Heat kernel of supercritical SDEs with unbounded drifts

Stéphane Menozzi, Zhang Xicheng

HEAT KERNEL OF SUPERCRITICAL NONLOCAL OPERATORS WITH UNBOUNDED DRIFTS

STEPHANE MENOZZI AND XICHENG ZHANG

ABSTRACT. Let $\alpha \in (0,2)$ and $d \in \mathbb{N}$. Consider the following SDE in $\mathbb{R}^d$:

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + a(t, X_{t-})\mathrm{d}L_t^{(\alpha)}, \quad X_0 = x,$$

where $L^{(\alpha)}$ is a d-dimensional rotationally invariant α-stable process, $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $a : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ are Hölder continuous functions in space, with respective order $\beta, \gamma \in (0,1)$ such that $(\beta \wedge \gamma) + \alpha > 1$, uniformly in t. Here b may be unbounded. When a is bounded and uniformly elliptic, we show that the unique solution $X_t(x)$ of the above SDE admits a continuous density, which enjoys sharp two-sided estimates. We also establish sharp upper-bound for the logarithmic derivative. In particular, we cover the whole *supercritical* range $\alpha \in (0,1)$. Our proof is based on *ad hoc* parametrix expansions and probabilistic techniques.

1. INTRODUCTION

Throughout this paper we fix $\alpha \in (0,2)$. Let $L^{(\alpha)}$ be a d-dimensional rotationally invariant α-stable process. We consider the following stochastic differential equation:

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + a(t, X_{t-})\mathrm{d}L_t^{(\alpha)}, \tag{1.1}$$

where $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $a : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ are Borel measurable functions and satisfy that for some $\beta \in ((1-\alpha)^+, 1]$ and $\kappa_0 \geqslant 1$,

$$|b(t,0)| \leqslant \kappa_0, \ |b(t,x) - b(t,y)| \leqslant \kappa_0(|x-y|^\beta \vee |x-y|), \tag{$\mathbf{H}_b^\beta$}$$

and for some $\gamma \in ((1-\alpha)^+, 1]$ and $\kappa_1 \geqslant 1$,

$$\kappa_1^{-1}\mathbb{I} \leqslant (aa^*)(t,x) \leqslant \kappa_1\mathbb{I}, \quad |a(t,x) - a(t,y)| \leqslant \kappa_1|x-y|^\gamma, \tag{$\mathbf{H}_a^\gamma$}$$

where a^* stands for the transpose of a and $\mathbb{I}$ is the identity matrix. Under $(\mathbf{H}_b^\beta)$ and $(\mathbf{H}_a^\gamma)$, it is well known that for each $(s,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, there is a unique weak solution $X_{s,t}(x)$ to SDE (1.1) starting from x at time s (see e.g. [12, Theorem 1.1]), and the generator of SDE (1.1) writes as

$$\mathscr{L}_s f(x) := \tfrac{1}{2}\mathcal{L}_s f(x) + b(s,x) \cdot \nabla f(x), \tag{1.2}$$

where $\mathcal{L}_s$ is given by

$$\mathcal{L}_s f(x) = \int_{\mathbb{R}^d} \delta_f^{(2)}(x; a(s,x)z)\frac{\mathrm{d}z}{|z|^{d+\alpha}} = \int_{\mathbb{R}^d} \delta_f^{(2)}(x; z)\frac{\kappa(s,x,z)}{|z|^{d+\alpha}}\mathrm{d}z \tag{1.3}$$

X. Zhang is supported by NNSF of China (No. 11731009) and the DFG through the CRC 1283 "Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications".

S. Menozzi has been funded by the Russian Science Foundation project (project No. 20-11-20119).

with

$$\delta_f^{(2)}(x;z) := f(x+z) + f(x-z) - 2f(x) \tag{1.4}$$

and

$$\kappa(s,x,z) := \det(a^{-1}(s,x))(|z|/|a^{-1}(s,x)z|)^{d+\alpha}. \tag{1.5}$$

Clearly, by $(\mathbf{H}_a^\gamma)$ we have for some $\bar\kappa_1 \geqslant 1$,

$$\bar\kappa_1^{-1} \leqslant \kappa(s,x,z) \leqslant \bar\kappa_1, \quad |\kappa(s,x,z) - \kappa(s,y,z)| \leqslant \bar\kappa_1 |x-y|^\gamma. \tag{1.6}$$

The operator $\mathscr{L}_s$ is called *supercritical* for $\alpha \in (0,1)$ since in this case, the drift term plays a dominant role. Namely, from the self-similarity properties of the driving process $L^{(\alpha)}$ in (1.1), it holds that for any $s>0$, $L_s^{(\alpha)} \overset{\text{(law)}}{=} s^{1/\alpha} L_1^{(\alpha)}$ and for $s \in (0,1)$, $\alpha \in (0,1)$, $s^{1/\alpha} < s$. This precisely means that the fluctuations induced by the noise are smaller than the typical order of the drift term in (1.1). For $\alpha \in (1,2)$, the converse phenomenon happens. Since for $s \in (0,1)$, $s^{1/\alpha} > s$, the fluctuations of the noise prevail in the SDE. From the operator viewpoint, $\mathcal{L}_s$ plays a dominant role and we say that $\mathscr{L}_s$ is *subcritical*. For the remaining case $\alpha = 1$, the noise and drift both have the same typical order and the operator $\mathscr{L}_s$ is called *critical*. Note that for $\alpha \in (0,1)$, since $z \mapsto \kappa(s,x,z)$ is symmetric, we have

$$\mathcal{L}_s f(x) = 2 \int_{\mathbb{R}^d} \delta_f^{(1)}(x;z) \frac{\kappa(s,x,z)\mathrm{d}z}{|z|^{d+\alpha}},$$

where

$$\delta_f^{(1)}(x;z) := f(x+z) - f(x).$$

Let us now indicate that there is a quite large literature concerning stable driven SDEs. We can first mention the seminal work of Kolokoltsov [20] who obtained, for an SDE driven by a symmetric stable process with smooth non-degenerate spectral measure, Lipschitz non-degenerate diffusion coefficient and non trivial Lipschitz bounded drifts when $\alpha > 1$, two sided estimates for the density of the type:

$$p(s,x,t,y) \asymp_C (t-s)^{-d/\alpha} \left(1 + \frac{|x-y|}{(t-s)^{1/\alpha}}\right)^{-(d+\alpha)}, \tag{1.7}$$

where $C \geqslant 1$ depends on the non-degeneracy and Lipschitz constants of the coefficients and the final considered time horizon T. Here and below, $Q_1 \asymp_C Q_2$ means that $C^{-1}Q_2 \leqslant Q_1 \leqslant CQ_2$.

Going to weaker regularity of the coefficients in (1.1) then first leads to investigate the well-posedness of the martingale problem associated with the formal generator associated with the dynamics (1.1). In [1], Bass and Chen showed the weak well-posedness for SDE (1.1) when a is only continuous and uniformly elliptic, b is Lipschitz and $L_t^{(\alpha)}$ is cylindrical α-stable process. In the subcritical case we can mention the work by Mikulevicius and Pragarauskas [26] who derived that weak uniqueness holds for equation (1.1) for bounded Hölder coefficients when $\alpha \geqslant 1$ and a non degenerate a. The martingale problem was in their framework studied from some related Schauder estimates established on the associated Integro Partial Differential Equation (IPDE).

In the super-critical case, the well-posedness of the martingale problem was recently investigated by Kulik *et al.* [18], [21] (see also [10, 12]). In [21], the authors consider SDEs of type (1.1) with bounded Hölder drift and non-degenerate

scalar diffusion coefficients under the natural condition $\alpha+\beta>1$[1] and obtain the existence of the heat kernel, a corresponding two-sided estimate of the form (1.7) as well as some estimates corresponding to the time derivative through parametrix type expansions. Let us emphasize that in the super-critical regime, the time derivative of the heat kernel roughly typically behaves as t^{-1} at time t whereas the spatial gradient is then more singular, as it is expected to have typical behavior of order $t^{-1/\alpha}\geqslant t^{-1}$ for $t\in(0,1]$.

Concerning other results related to (possibly non symmetric) stable heat kernel estimates we can refer e.g. to [6], [8], [9], [13] and for even more general stable like processes to the recent works [19], [11]. Let us also mention that sharp gradient estimates for driftless stable like operators have recently been obtained in [15]. We eventually refer to the work [22] and [7] for gradient estimates associated with additive and multiplicative cylindrical noises in (1.1), respectively.

In the current work we face two difficulties, we want to establish gradient estimates for all $\alpha\in(0,2)$ and for unbounded drifts. It is known, and somehow intuitive, that for unbounded drifts the heat kernel bounds must reflect somehow the transport induced by the drift. This was for instance observed for a Lipschitz drift in [14] for degenerate Kolmogorov SDEs which can be viewed as ODEs perturbed on some components by a Brownian noise propagating through the whole chain thanks to a weak type Hörmander condition on the drift. Before going further let us also mention the work by Huang [17] which establishes two-sided estimates for stable driven SDE with unbounded Lipschitz drift and $\alpha\in(1,2)$, which then reads as: for $s<t$ and $x,y\in\mathbb{R}^d$,

$$p(s,x,t,y)\asymp_C(t-s)^{-d/\alpha}\left(1+\frac{|\theta_{s,t}(x)-y|}{(t-s)^{1/\alpha}}\right)^{-(d+\alpha)}, \tag{1.8}$$

where $\theta_{s,t}(x)$ denotes the flow associated to the drift in (1.1). Namely,

$$\dot{\theta}_{s,t}(x)=b(t,\theta_{s,t}(x)),\theta_{s,s}(x)=x.$$

In the non-degenerate Brownian case, the type of heat kernel estimate in [14] has recently been extended to drifts satisfying a linear-growth without *a priori* smoothness assumptions on the drift, see [25]. In the quoted work, under additional Hölder continuity of the drift for the second derivatives, the estimates also extend to the derivatives up to order two with the corresponding additional parabolic singularity.

We here somehow follow the main line of that work but are faced with many additional difficulties. In particular, a common feature to both the Gaussian SDEs considered in [25] and the stable driven here is that we first need to establish the gradient estimates for smooth coefficients, with constants depending on the derivatives of the coefficients. In the current strictly stable framework we cannot rely on the Malliavin calculus arguments of [25] because of integrability issues. We thus establish here some direct bounds on the associated semi-group and its derivatives when the coefficients are smooth which serve as a starting point to derive the estimates concerning the logarithmic gradient of the density. This part is crucial and quite intricate (see Theorem 4.1 below). Then, in a second time, exploiting thoroughly the two-sided estimate, which holds independently of the smoothness

[1]it is indeed well known that from the seminal work of Tanaka *et al.* [28] that weak uniqueness may fail if this condition is not met.

of the coefficients, we establish that the gradient estimates hold also independently of such a smoothness. We eventually conclude through a compactness argument.

To state our main result, we introduce the regularized flow associated with drift b. For $\varepsilon > 0$, let $b_\varepsilon(t,x) := b(t,\cdot) * \rho_\varepsilon(x)$, where $\rho_\varepsilon(x) = \varepsilon^{-d}\rho(x/\varepsilon)$ and ρ is a smooth density function with support in the unit ball. Note that under $(\mathbf{H}_b^\beta)$,

$$|b_\varepsilon(t,x) - b(t,x)| \leqslant \kappa_0 \varepsilon^\beta, \quad \varepsilon \in (0,1), \tag{1.9}$$

and

$$|\nabla b_\varepsilon(t,x)| = \left|\int_{\mathbb{R}^d} (b(t,y) - b(t,x))\nabla\rho_\varepsilon(x-y)\mathrm{d}y\right| \leqslant \kappa_0(\varepsilon^{\beta-1}+1)\|\nabla\rho\|_{L^1}. \tag{1.10}$$

In particular, since $\alpha + \beta > 1$, for any $T > 0$, there is a $C > 0$ such that for any $0 \leqslant s < t \leqslant T$,

$$\int_s^t \|\nabla b_{|r-s|^{1/\alpha}}(r,\cdot)\|_\infty \mathrm{d}r \leqslant C\int_s^t \Big(|r-s|^{\frac{\beta-1}{\alpha}} + 1\Big)\mathrm{d}r \leqslant C(t-s)^{\frac{\alpha+\beta-1}{\alpha}}. \tag{1.11}$$

Thus, for fixed $s > 0$, the following ODE admits a unique solution $\theta_{s,t}(x)$:

$$\dot\theta_{s,t} = b_{|t-s|^{1/\alpha}}(t,\theta_{s,t}), \quad \theta_{s,s} = x, \quad t \geqslant 0. \tag{1.12}$$

Note that for $t > s$, $\theta_{s,t}(x)$ denotes the forward solution of the above ODE, while for $t < s$, it denotes the backward solution. We carefully mention that our main results will be stated w.r.t. to the flow θ in (1.12) which is precisely associated with a mollified drift with parameter corresponding to the typical scale of the driving process of the SDE (1.1) at the current considered time.

For notational simplicity, we introduce the following parameter set

$$\Theta := (\kappa_0, \kappa_1, d, \alpha, \beta, \gamma). \tag{1.13}$$

We also denote for $T \in (0,\infty]$,

$$\mathbb{D}_T := \{(s,x,t,y) : 0 \leqslant s < t < T, x, y \in \mathbb{R}^d\}.$$

We will frequently use from now on the notation $\lesssim$. For two quantities Q_1 and Q_2, we mean by $Q_1 \lesssim Q_2$ that there exists $C := C(T,\Theta)$ such that $Q_1 \leqslant CQ_2$. Other possible dependencies for the constants will be explicitly specified. Moreover, we also use the following notation

$$|\mathcal{D}^{(\alpha)} f|(x) := \int_{\mathbb{R}^d} \frac{|\delta_f^{(2)}(x;z)|}{|z|^{d+\alpha}}\mathrm{d}z. \tag{1.14}$$

The aim of this paper is to show the following result.

Theorem 1.1. *Under $(\mathbf{H}_b^\beta)$ and $(\mathbf{H}_a^\gamma)$, for each $0 \leqslant s < t < \infty$ and $x \in \mathbb{R}^d$, $X_{s,t}(x)$ admits a density $p(s,x,t,y)$ (called heat kernel of $\mathscr{L}_s$) that is continuous as a function of x, y, and such that for each $t > 0$ and $x, y \in \mathbb{R}^d$ and Lebesgue almost all $s \in [0,t)$,*

$$\partial_s p(s,x,t,y) = \mathscr{L}_s p(s,\cdot,t,y)(x), \quad p(s,x,t,\cdot) \to \delta_{\{x\}}(\cdot) \text{ weakly as } s \uparrow t,$$

where $\delta_{\{x\}}(\mathrm{d}y)$ denotes the Dirac measure concentrated at x. Moreover, we have

(i) **(Two-sides estimate)** *For any $T > 0$, there is a constant $C_1 = C_1(T,\Theta) \geqslant 1$ such that for all $(s,x,t,y) \in \mathbb{D}_T$,*

$$p(s,x,t,y) \asymp_{C_1} (t-s)((t-s)^{1/\alpha} + |\theta_{s,t}(x) - y|)^{-d-\alpha}, \tag{1.15}$$

where $\theta_{s,t}(x)$ is defined by ODE (1.12).

(ii) **(Fractional derivative estimate)** *For any $T>0$, there is a constant $C_2=C_2(T,\Theta)>0$ such that for all $(s,x,t,y)\in\mathbb{D}_T$,*

$$|\mathcal{D}^{(\alpha)}p(s,\cdot,t,y)|(x)\leqslant C_2((t-s)^{1/\alpha}+|\theta_{s,t}(x)-y|)^{-d-\alpha}. \tag{1.16}$$

(iii) **(Gradient estimate in x)** *For any $T>0$, there is a constant $C_3=C_3(T,\Theta)>0$ such that for all $(s,x,t,y)\in\mathbb{D}_T$,*

$$|\nabla_x\log p(s,x,t,y)|\leqslant C_3(t-s)^{-1/\alpha}. \tag{1.17}$$

Remark 1.2. If $|b(t,x)-b(t,y)|\leqslant\kappa_0|x-y|^\beta$ for any $x,y\in\mathbb{R}^d$ with $|x-y|\leqslant 1$ and $|b(0,t)|\leqslant\kappa_0$ for all $t\geqslant 0$, then $\mathbf{(H}_b^\beta\mathbf{)}$ holds. In particular, for $c(x)$ being a bounded β-Hölder continuous function, $b(x):=x+c(x)$ satisfies $\mathbf{(H}_b^\beta\mathbf{)}$.

Remark 1.3. For $\alpha\in[1,2)$, we can replace $\theta_{s,t}(x)$ in (1.15) by any *regularized* flow $\theta^{(\varepsilon)}_{s,t}(x)$ defined in (2.1) below. When $\alpha\in(0,1)$, we choose the regularizing parameter $\varepsilon=(t-s)^{1/\alpha}$ since we need to use ε to compensate the time singularity in the supercritical case. For $\alpha\in(0,1]$, since b is continuous in x, we can replace $\theta_{s,t}(x)$ in (1.15) by any measurable *Peano* flow $\vartheta_{s,t}(x)$ of the ODE $\dot\vartheta_{s,t}(x)=b(t,\vartheta_{s,t}(x))$.

Remark 1.4. When $b\equiv 0$ and $L^{(\alpha)}$ is a general α-stable-like process, it was proven in [24] that the gradient estimate (1.17) holds for $\alpha\in(\frac12,2)$. See also [11] for general cases. It seems that our gradient estimate (1.17) is the first result for SDE (1.1) driven by rotationally invariant α-stable process with $\alpha\in(0,\frac12]$.

The paper is organized as follows. We give in Section 2 some preliminary estimates needed for the main analysis. This concerns the mollified flow, some exit probabilities, convolution inequalities and the density of the *proxy process* involved in the parametrix (which has a dynamic similar to (1.1) with coefficient frozen along a suitable deterministic flow). Section 3 is then devoted to the derivation of the two-sided bound and the fractional derivative estimate under $\mathbf{(H}_b^\beta\mathbf{)}$ and $\mathbf{(H}_a^\gamma\mathbf{)}$ when the coefficients are additionally supposed to be smooth. We specifically address the gradient estimate under those same assumptions in Section 4. We eventually present in Section 5 some compactness arguments to derive the results of Theorem 1.1 under the sole conditions $\mathbf{(H}_b^\beta\mathbf{)}$, $\mathbf{(H}_a^\gamma\mathbf{)}$.

2. Preliminaries

2.1. **ODE flow.** We first present some basic properties about the solution $\theta_{s,t}(x)$ of the ODE (1.12). Since the drift coefficient therein depends on the initial time s, the following flow property does no longer hold:

$$\theta_{r,t}\circ\theta_{s,r}(x)=\theta_{s,t}(x),\quad s<r<t.$$

However, the above flow property holds for the following regularized ODE:

$$\dot\theta^{(\varepsilon)}_{s,t}(x)=b_\varepsilon(t,\theta^{(\varepsilon)}_{s,t}(x)),\quad \theta^{(\varepsilon)}_{s,s}(x)=x, \tag{2.1}$$

for any fixed regularizing parameter $\varepsilon>0$. Below we fix $\alpha\in(0,2)$ and always assume $(\mathbf{H}_b^\beta)$. The following lemma is easy.

Lemma 2.1. *(i) For each $\varepsilon>0$ and $s,t\geqslant 0$, $x\mapsto\theta^{(\varepsilon)}_{s,t}(x)$ is a C^1-diffeomorphism and*

$$(\theta^{(\varepsilon)}_{s,t})^{-1}(y)=\theta^{(\varepsilon)}_{t,s}(y). \tag{2.2}$$

Moreover, for all $s,r,t\geqslant 0$, it holds that

$$\theta^{(\varepsilon)}_{s,t}(x)=\theta^{(\varepsilon)}_{r,t}\circ\theta^{(\varepsilon)}_{s,r}(x). \tag{2.3}$$

(ii) For all $\varepsilon,\varepsilon'>0$ and $s,t\geqslant 0$, $x\in\mathbb{R}^d$, it holds that

$$|\theta^{(\varepsilon')}_{s,t}(x)-\theta^{(\varepsilon)}_{s,t}(x)|\leqslant 2\kappa_0(\varepsilon\vee\varepsilon')^\beta|t-s|\mathrm{e}^{\kappa_0\|\nabla\rho\|_{L^1}((\varepsilon\vee\varepsilon')^{\beta-1}+1)|t-s|}, \tag{2.4}$$

(iii) For any $T>0$, there is a constant $C=C(T,d,\kappa_0)>0$ such that for all $s,t\in[0,T]$, $x,y\in\mathbb{R}^d$ and $\varepsilon=|t-s|^{1/\alpha}$,

$$|\theta^{(\varepsilon)}_{s,t}(x)-y|\asymp_C|x-\theta^{(\varepsilon)}_{t,s}(y)|,\quad |\theta^{(\varepsilon)}_{s,t}(x)-\theta^{(\varepsilon)}_{s,t}(y)|\asymp_C|x-y|. \tag{2.5}$$

Proof. (i) Note that by (2.1), for $0\leqslant s<t$:

$$\theta^{(\varepsilon)}_{s,t}(x)=x+\int_s^t b(r,\theta^{(\varepsilon)}_{s,r}(x))\mathrm{d}r,\quad \theta^{(\varepsilon)}_{t,s}(y)=y-\int_s^t b(r,\theta^{(\varepsilon)}_{t,r}(y))\mathrm{d}r.$$

Let $y=\theta^{(\varepsilon)}_{s,t}(x)$. By the flow property, we have

$$y=\big(\theta^{(\varepsilon)}_{s,t}\big)^{-1}(y)+\int_s^t b(r,\theta^{(\varepsilon)}_{s,r}\circ\big(\theta^{(\varepsilon)}_{s,t}\big)^{-1}(y))\mathrm{d}r=(\theta^{(\varepsilon)}_{s,t})^{-1}(y)+\int_s^t b(r,\big(\theta^{(\varepsilon)}_{r,t}\big)^{-1}(y))\mathrm{d}r.$$

Since the ODE has a unique solution, we immediately have $\big(\theta^{(\varepsilon)}_{s,t}\big)^{-1}(y)=\theta^{(\varepsilon)}_{t,s}(y)$. As for (2.3), it follows from (2.2) and the flow property.

(ii) Without loss of generality, we assume $\varepsilon'<\varepsilon$. Since by (1.9) and (1.10),

$$|b_\varepsilon(t,x)-b_{\varepsilon'}(t,x)|\leqslant 2\kappa_0\varepsilon^\beta,\quad \|\nabla b_\varepsilon\|_\infty\leqslant\kappa_0\|\nabla\rho\|_{L^1}(\varepsilon^{\beta-1}+1), \tag{2.6}$$

by definition we have

$$\begin{aligned}|\theta^{(\varepsilon')}_{s,t}(x)-\theta^{(\varepsilon)}_{s,t}(x)|&\leqslant\int_s^t|b_{\varepsilon'}(r,\theta^{(\varepsilon')}_{s,r}(x))-b_\varepsilon(r,\theta^{(\varepsilon')}_{s,r}(x))|\mathrm{d}r\\&\quad+\int_s^t|b_\varepsilon(r,\theta^{(\varepsilon')}_{s,r}(x)-b_\varepsilon(r,\theta^{(\varepsilon)}_{s,r}(x)|\mathrm{d}r\\&\leqslant 2\kappa_0\varepsilon^\beta(t-s)+\kappa_0\|\nabla\rho\|_{L^1}(\varepsilon^{\beta-1}+1)\int_s^t|\theta^{(\varepsilon')}_{s,r}(x)-\theta^{(\varepsilon)}_{s,r}(x)|\mathrm{d}r.\end{aligned}$$

Using Gronwall's inequality, we obtain (2.4).

(iii) Without loss of generality, we assume $0\leqslant s<t\leqslant T$. Note that for $u\in[s,t]$,

$$|\theta^{(\varepsilon)}_{s,u}(x)-\theta^{(\varepsilon)}_{s,u}(y)|\leqslant|x-y|+\int_s^u\|\nabla b_\varepsilon(r,\cdot)\|_\infty|\theta^{(\varepsilon)}_{s,r}(x)-\theta^{(\varepsilon)}_{s,r}(y)|\mathrm{d}r.$$

For $\varepsilon=|t-s|^{1/\alpha}$, it follows from the Gronwall inequality and (1.10) that

$$|\theta^{(\varepsilon)}_{s,t}(x)-\theta^{(\varepsilon)}_{s,t}(y)|\leqslant\mathrm{e}^{\kappa_0\|\nabla\rho\|_{L^1}(|t-s|^{(\beta-1)/\alpha}+1)|t-s|}|x-y|.$$

We thus get (2.5) by (i) and symmetry. □

The following result is a consequence of the above lemma, which plays a crucial role below.

Lemma 2.2. *(i) For each $s,t\geqslant 0$, the map $x\mapsto\theta_{s,t}(x)$, where $(\theta_{s,u}(x))_{u\in[s,t]}$ solves (1.12), is a C^1-diffeomorphism and there is a constant $C_0=C_0(T,\Theta)>0$ such that*

$$|\det(\nabla\theta^{-1}_{s,t}(x))-1|\leqslant C_0|t-s|^{(\alpha+\beta-1)/\alpha}.$$

(ii) For any $T>0$, there is a constant $C_1=C_1(T,\Theta)\geqslant 1$ such that for all $s,t\in[0,T]$ and $x,y\in\mathbb{R}^d$,

$$|t-s|^{1/\alpha}+|\theta_{s,t}(x)-y|\asymp_{C_1}|t-s|^{1/\alpha}+|x-\theta_{t,s}(y)|. \tag{2.7}$$

(iii) For any $T > 0$, there is a constant $C_2 = C_2(T, \Theta) > 0$ such that for all $s, r, t \in [0, T]$ and $x \in \mathbb{R}^d$,

$$|\theta_{s,t}(x) - \theta_{r,t} \circ \theta_{s,r}(x)| \leqslant C_2 |(r \vee s \vee t) - (r \wedge s \wedge t)|^{1/\alpha}. \tag{2.8}$$

Proof. (i) It is well known that

$$\det(\nabla \theta_{s,t}(x)) = 1 + \int_s^t \mathrm{div} b_{|r-s|^{1/\alpha}}(r, \theta_{s,r}(x)) \det(\nabla \theta_{s,r}(x)) \mathrm{d}r.$$

Thus

$$\det(\nabla \theta_{s,t}(x)) = \exp\left\{ \int_s^t \mathrm{div} b_{|r-s|^{1/\alpha}}(r, \theta_{s,r}(x)) \mathrm{d}r \right\}.$$

The desired estimate follows by (1.11) and $\nabla \theta_{s,t}^{-1}(x) = (\nabla \theta_{s,t})^{-1}(\theta_{s,t}^{-1}(x))$.

(ii) Fix $s < t$. For $u \in [s, t]$, by definition we have

$$\theta_{s,u}(x) = x + \int_s^u b_{|r-s|^{1/\alpha}}(r, \theta_{s,r}(x)) \mathrm{d}r,$$

and for $\varepsilon = |t - s|^{1/\alpha}$,

$$\theta_{s,u}^{(\varepsilon)}(x) = x + \int_s^u b_\varepsilon(r, \theta_{s,r}^{(\varepsilon)}(x)) \mathrm{d}r.$$

By (2.6) with $\varepsilon' = |r - s|^{1/\alpha}$ and $\varepsilon = |t - s|^{1/\alpha}$, we have for all $u \geqslant s$,

$$\begin{aligned} |\theta_{s,u}(x) - \theta_{s,u}^{(\varepsilon)}(x)| &\leqslant \int_s^u \left| b_{|r-s|^{1/\alpha}}(r, \theta_{s,r}(x)) - b_\varepsilon(r, \theta_{s,r}^{(\varepsilon)}(x)) \right| \mathrm{d}r \\ &\lesssim |t - s|^{\beta/\alpha + 1} + (t - s)^{(\beta-1)/\alpha} \int_s^u |\theta_{s,r}(x) - \theta_{s,r}^{(\varepsilon)}(x)| \mathrm{d}r, \end{aligned}$$

which yields by Gronwall's inequality that

$$|\theta_{s,t}(x) - \theta_{s,t}^{(\varepsilon)}(x)| \lesssim \mathrm{e}^{(t-s)^{(\beta-1)/\alpha+1}} |t - s|^{\beta/\alpha+1} \lesssim |t - s|^{1/\alpha}, \tag{2.9}$$

where the second inequality is due to $\alpha + \beta > 1$. Thus, by (2.5) and (2.9), we have

$$\begin{aligned} |\theta_{s,t}(x) - y| &\leqslant |\theta_{s,t}^{(\varepsilon)}(x) - y| + |\theta_{s,t}(x) - \theta_{s,t}^{(\varepsilon)}(x)| \\ &\lesssim |x - \theta_{s,t}^{(\varepsilon)}(y)| + |t - s|^{1/\alpha} \lesssim |x - \theta_{t,s}(y)| + |t - s|^{1/\alpha}. \end{aligned}$$

The right hand side inequality of (2.7) follows. By symmetry, we also have the left hand side inequality.

(iii) Let $s, r, t \geqslant 0$ and $\varepsilon := |r \vee s \vee t - r \wedge s \wedge t|^{1/\alpha}$. By (2.3) we have

$$\begin{aligned} |\theta_{s,t}(x) - \theta_{r,t} \circ \theta_{s,r}(x)| &\leqslant |\theta_{s,t}(x) - \theta_{s,t}^{(\varepsilon)}(x)| + |\theta_{r,t}^{(\varepsilon)} \circ \theta_{s,r}^{(\varepsilon)}(x) - \theta_{r,t}^{(\varepsilon)} \circ \theta_{s,r}(x)| \\ &\quad + |\theta_{r,t}^{(\varepsilon)} \circ \theta_{s,r}(x) - \theta_{r,t} \circ \theta_{s,r}(x)|. \end{aligned}$$

The desired estimate (2.8) again follows by (2.5) and (2.9). □

Remark 2.3. By (2.8), we have

$$|x - \theta_{t,s} \circ \theta_{s,t}(x)| \lesssim_{C_2} |t - s|^{1/\alpha}, \tag{2.10}$$

and by (2.7),

$$|t - s|^{1/\alpha} + |x - y| \asymp_{C_1} |t - s|^{1/\alpha} + |\theta_{s,t}(x) - \theta_{t,s}(y)|. \tag{2.11}$$

Put it differently, we have an *approximate* flow property for the ODE (1.12). Namely, the flow property holds up to an additive time factor which has the same magnitude as the current typical time (self-similarity index of the driving process).

2.2. **Probability estimates.** We need the following master formula.

Lemma 2.4. *(Lévy system) Let $X_t := X_{0,t}$ be any solution of SDE (1.1). For any nonnegative measurable function $f : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ and finite stopping time τ,*

$$\mathbb{E}\sum_{r\in(0,\tau]} f(r, X_{r-}, \Delta X_r) = \mathbb{E}\int_0^\tau\!\!\int_{\mathbb{R}^d} f(r, X_{r-}, z)\frac{\kappa(r, X_{r-}, z)}{|z|^{d+\alpha}}\mathrm{d}z\mathrm{d}r,$$

where $\Delta X_r := X_r - X_{r-}$ and $\kappa(r, x, z)$ is defined in (1.5).

Proof. Let $N(\mathrm{d}t, \mathrm{d}z)$ be the counting measure associated with $L_t^{(\alpha)}$, i.e.,

$$N((0, t] \times E) := \sum_{s\in[0,t]} \mathbf{1}_E(\Delta L_s^{(\alpha)}),\ E \in \mathscr{B}(\mathbb{R}^d).$$

Noting that

$$\Delta X_t = a(t, X_{t-})\Delta L_t^{(\alpha)},$$

we have for any $\varepsilon > 0$,

$$\sum_{r\in(0,\tau]} f(r, X_{r-}, \Delta X_r)\mathbf{1}_{|\Delta X_r|\geqslant\varepsilon} = \int_0^\tau\!\!\int_{|a(r,X_{r-})z|\geqslant\varepsilon} f(r, X_{r-}, a(r, X_{r-})z)N(\mathrm{d}r, \mathrm{d}z).$$

Since the compensated measure of $N(\mathrm{d}t, \mathrm{d}z)$ is $\frac{\mathrm{d}z\mathrm{d}t}{|z|^{d+\alpha}}$, by the change of variable, we have

$$\begin{aligned}\mathbb{E}\sum_{r\in(0,\tau]} f(r, X_{r-}, \Delta X_r)\mathbf{1}_{|\Delta X_r|\geqslant\varepsilon} &= \mathbb{E}\int_0^\tau\!\!\int_{|a(r,X_{r-})z|\geqslant\varepsilon} f(r, X_{r-}, a(r, X_{r-})z)\frac{\mathrm{d}z\mathrm{d}r}{|z|^{d+\alpha}}\\ &= \mathbb{E}\int_0^\tau\!\!\int_{|z|\geqslant\varepsilon} f(r, X_{r-}, z)\frac{\kappa(r, X_{r-}, z)}{|z|^{d+\alpha}}\mathrm{d}z\mathrm{d}r,\end{aligned}$$

where $\kappa(r, x, z)$ is given in (1.5), which in turn gives the desired formula by the monotone convergence theorem. □

Fix $(s, x) \in \mathbb{R}_+ \times \mathbb{R}^d$. For $\eta > 0$, define the stopping time

$$\tau_{s,x}^\eta := \inf\Big\{t > s : |X_{s,t}(x) - \theta_{s,t}(x)| > \eta\Big\}, \tag{2.12}$$

which corresponds to the exit time of the diffusion from a tube around the deterministic ODE introduced in (1.12). We now give a *tube estimate* which roughly says that, for a given spatial threshold η, the probability that the difference between the process $X_{s,\cdot}(x)$ and the deterministic regularized flow $\theta_{s,\cdot}(x)$ leaves the tube of radius η before a certain fraction $\varepsilon\eta^\alpha$ of the corresponding typical time scale η^α is somehow *small*.

Lemma 2.5. *Under $(\mathbf{H}_b^\beta)$ and $(\mathbf{H}_a^0)$, for any $T > 0$, there is an $\varepsilon \in (0, 1)$ only depending on T, Θ such that for all $(s, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ and $\eta \in (0, T^{1/\alpha}]$,*

$$\mathbb{P}(\tau_{s,x}^\eta < s + \varepsilon\eta^\alpha) \leqslant 1/2.$$

Proof. Without loss of generality, we assume $(s, x) = (0, 0)$, and for simplicity write

$$X_t := X_{0,t}(0),\quad \theta_t := \theta_{0,t}(0),\quad \tau := \tau_{0,0}^\eta.$$

Let $f \in C_b^2(\mathbb{R}^d)$ with $f(0) = 0$ and $f(x) = 1$ for $|x| > 1$. For $\eta > 0$, set

$$f_\eta(x) := f(x/\eta), \quad u := \varepsilon\eta^\alpha.$$

Note that

$$Y_t := X_t - \theta_t = \int_0^t (b(r, X_r) - b_{r^{1/\alpha}}(r, \theta_r))\mathrm{d}r + \int_0^t a(r, X_{r-})\mathrm{d}L_r^{(\alpha)}.$$

By Itô's formula, we have

$$\mathbb{E}f_\eta(Y_{u\wedge\tau}) = \mathbb{E}\int_0^{u\wedge\tau} \Big[(b(r, X_r) - b_{r^{1/\alpha}}(r, \theta_r)) \cdot \nabla f_\eta(Y_r) + \tfrac{1}{2}\mathcal{L}_r f_\eta(Y_r)\Big]\mathrm{d}r.$$

Note that

$$\begin{aligned} |b(r, X_r) - b_{r^{1/\alpha}}(r, \theta_r)| &\leqslant |b(r, X_r) - b(r, \theta_r)| + |b(r, \theta_r) - b_{r^{1/\alpha}}(r, \theta_r)| \\ &\underset{(\mathbf{H}_b^\beta),(2.6)}{\leqslant} \kappa_0(|X_r - \theta_r|^\beta + |X_r - \theta_r|) + \kappa_0 r^{\beta/\alpha}. \end{aligned}$$

Also, with the notation of (1.4),

$$|\delta^{(2)}_{f_\eta}(x; z)| \leqslant (|z|^2\|\nabla^2 f_\eta\|_\infty) \wedge (4\|f_\eta\|_\infty) \lesssim (|z|^2/\eta^2) \wedge 1.$$

Hence,

$$\begin{aligned} \mathbb{E}f_\eta(Y_{u\wedge\tau}) &\lesssim \mathbb{E}\int_0^u \left[(|Y_r|^\beta + |Y_r| + r^{\beta/\alpha}) \cdot \frac{\mathbf{1}_{|Y_r|\leqslant\eta}}{\eta} + \int_{\mathbb{R}^d} \frac{(|z|^2/\eta^2) \wedge 1}{|z|^{d+\alpha}}\mathrm{d}z\right]\mathrm{d}r \\ &\lesssim \int_0^u (\eta^{\beta-1} + 1 + r^{\beta/\alpha}\eta^{-1} + \eta^{-\alpha})\mathrm{d}r \\ &\underset{u=\varepsilon\eta^\alpha}{\lesssim} \varepsilon(\eta^{\alpha+\beta-1} + \eta^\alpha + \varepsilon^{\beta/\alpha}\eta^{\alpha+\beta-1} + 1) \lesssim \varepsilon(T^{1+(\beta-1)/\alpha} + T + 1), \end{aligned}$$

where the last inequality is due to $\alpha + \beta > 1$ and $\eta \leqslant T^{1/\alpha}$. Importantly, the implicit constant is independent of ε. Note that

$$\mathbb{P}(\tau < u) = \mathbb{E}\mathbf{1}_{\tau<u} \leqslant \mathbb{E}f_\eta\big(Y_{u\wedge\tau}\big) \lesssim \varepsilon(T^{1+(\beta-1)/\alpha} + T + 1).$$

The desired estimate follows by choosing ε small enough. □

The following lemma will be used to show the lower bound estimate of the heat kernel. It gives a lower bound estimate for the probability that, considered two points x, y in the *off-diagonal regime* between times s and t, namely such that $|x - \theta_{t,s}(y)| \geqslant K(t-s)^{1/\alpha}$, after a time $\varepsilon(t-s)$ with ε as in Lemma 2.5, the stochastic forward transport of x by the SDE, i.e. $X_{s,s+\varepsilon(t-s)}(x)$ and the backward deterministic transport of y by the regularized flow $\theta_{t,s+\varepsilon(t-s)}(y)$ belong to a *diagonal* tube, with radius $K(t-s)^{1/\alpha}$, where again $(t-s)^{1/\alpha}$ corresponds to the current typical scale between times s and t.

Lemma 2.6. *Suppose that* $(\mathbf{H}_b^\beta)$ *and* $(\mathbf{H}_a^0)$ *hold. Let* $\varepsilon \in (0,1)$ *be as in Lemma 2.5. For any* $T > 0$*, there are constants* $c_0 \in (0,1), K \geqslant 1$ *depending only on* T, Θ *such that for all* $0 \leqslant s < t \leqslant T$ *and* $|x - \theta_{t,s}(y)| \geqslant K(t-s)^{1/\alpha}$*,*

$$\mathbb{P}\Big(|X_{s,s+\varepsilon(t-s)}(x) - \theta_{t,s+\varepsilon(t-s)}(y)| \leqslant K(t-s)^{1/\alpha}\Big) \geqslant \frac{c_0(t-s)^{1+d/\alpha}}{|x - \theta_{t,s}(y)|^{d+\alpha}}.$$

This Lemma will be crucial for the lower bound estimate of the heat kernel, since it precisely gives the control needed for a chaining argument, see Theorem 3.5 below. As opposed to the continuous case, for SDEs driven by pure jump processes, a single intermediate time, associated with a *large jump*, is needed for the chaining. Roughly speaking between times $s+\varepsilon(t-s)$ and t we will use the global

diagonal bound of order $(t-s)^{-d/\alpha}$, since ε is meant to be *small enough*, and the above Lemma controls the probability that the process enters a good neigborhood of the backward flow to do so. The lower bound is the sought one in the sense that when multiplying it by $(t-s)^{-d/\alpha}$ exactly makes the expression in (1.15), $(t-s)((t-s)^{1/\alpha}+|\theta_{s,t}(x)-y|)^{-(d+\alpha)} \asymp_C (t-s)|\theta_{s,t}(x)-y|^{-(d+\alpha)}$ appear since $|x-\theta_{s,t}(y)| \geqslant K(t-s)^{1/\alpha}$ (*off-diagonal regime*).

Proof of Lemma 2.6. Without loss of generality, we assume $s=0$ and for simplicity, we write

$$\eta := t^{1/\alpha},\ u := \varepsilon\eta^\alpha = \varepsilon t,\quad X_r(x) := X_{0,r}(x).$$

Define a stopping time

$$\sigma := \inf\Big\{r>0 : |X_r(x)-\theta_{t,r}(y)| \leqslant \eta\Big\}.$$

By the right continuity of $r \mapsto X_r(x)-\theta_{t,r}(y)$, one sees that

$$|X_\sigma(x)-\theta_{t,\sigma}(y)| \leqslant \eta,\ a.s.$$

In particular, for $\sigma \leqslant u$, by (2.11) and (2.8), there is a constant $C_0 = C_0(\Theta) \geqslant 1$ such that

$$\begin{aligned}|X_u(x)-\theta_{t,u}(y)| &\leqslant |X_{\sigma,u}(X_\sigma(x))-\theta_{\sigma,u}(X_\sigma(x))|\\ &\quad + |\theta_{\sigma,u}(X_\sigma(x))-\theta_{\sigma,u}(\theta_{t,\sigma}(y))|\\ &\quad + |\theta_{\sigma,u}(\theta_{t,\sigma}(y))-\theta_{t,u}(y)|\\ &\leqslant |X_{\sigma,u}(X_\sigma(x))-\theta_{\sigma,u}(X_\sigma(x))| + C_0\eta.\end{aligned}$$

Let $K \geqslant C_0+1$. Then

$$\big\{|X_{\sigma,u}(X_\sigma(x))-\theta_{\sigma,u}(X_\sigma(x))|<\eta\big\} \subset \big\{|X_u(x)-\theta_{t,u}(y)| \leqslant K\eta\big\}.$$

Thus, by the strong Markov property, we have

$$\begin{aligned}\mathbb{P}(|X_u(x)-\theta_{t,u}(y)| \leqslant K\eta) &\geqslant \mathbb{P}(\sigma\leqslant u; |X_u(x)-\theta_{t,u}(y)| \leqslant K\eta)\\ &\geqslant \mathbb{P}\left(\sigma\leqslant u; |X_{\sigma,u}(X_\sigma(x))-\theta_{\sigma,u}(X_\sigma(x))|<\eta\right)\\ &\geqslant \mathbb{P}(\sigma\leqslant u)\inf_{(s,z)\in[0,u]\times\mathbb{R}^d}\mathbb{P}\left(|X_{s,u}(z)-\theta_{s,u}(z)|<\eta\right).\end{aligned}$$

Let $\tau^\eta_{s,z}$ be defined by (2.12). By Lemma 2.5 we have

$$\mathbb{P}\left(|X_{s,u}(z)-\theta_{s,u}(z)| \geqslant \eta\right) \leqslant \mathbb{P}(\tau^\eta_{s,z}\leqslant u) \leqslant \mathbb{P}(\tau^\eta_{s,z}\leqslant s+\varepsilon\eta^\alpha) \leqslant 1/2, \tag{2.13}$$

which implies that

$$\inf_{(s,z)\in[0,u]\times\mathbb{R}^d}\mathbb{P}\left(|X_{s,u}(z)-\theta_{s,u}(z)|<\eta\right) \geqslant 1/2$$

and

$$\mathbb{P}(|X_u(x)-\theta_{t,u}(y)| \leqslant K\eta) \geqslant \mathbb{P}(\sigma\leqslant u)/2. \tag{2.14}$$

Next we need to obtain a lower bound estimate for $\mathbb{P}(\sigma\leqslant u)$. Let $\tau := \tau^\eta_{0,x}$. For $r < u\wedge\tau$, by (2.7), there are constants $c_0, C_1 > 0$ such that

$$\begin{aligned}|X_r(x)-\theta_{t,r}(y)| &\geqslant |\theta_{0,r}(x)-\theta_{t,r}(y)| - |X_r(x)-\theta_{0,r}(x)|\\ &\geqslant c_0|x-\theta_{t,0}(y)| - C_1 t^{1/\alpha} - \eta.\end{aligned}$$

In particular, if we choose $K \geqslant (C_0+1)\vee((C_1+2)/c_0)$, then since by assumption

$$|x-\theta_{t,0}(y)| \geqslant K\eta,$$

it holds that for $r < u \wedge \tau$,

$$|X_r(x) - \theta_{t,r}(y)| \geqslant c_0 K\eta - (C_1+1)\eta > \eta. \tag{2.15}$$

Thus we have

$$\mathbf{1}_{\{|X_{u\wedge\tau}(x)-\theta_{t,u\wedge\tau}(y)|\leqslant\eta\}} = \sum_{r\in(0,u\wedge\tau]} \mathbf{1}_{\{|X_r(x)-\theta_{t,r}(y)|\leqslant\eta\}},$$

i.e. we have at most one term in the above summand. We then derive from Lemma 2.4,

$$\begin{aligned}\mathbb{P}\{|X_{u\wedge\tau}(x) - \theta_{t,u\wedge\tau}(y)| \leqslant \eta\} &= \mathbb{E}\sum_{r\in(0,u\wedge\tau]} \mathbf{1}_{\{|X_r(x)-\theta_{t,r}(y)|\leqslant\eta\}} \\ &= \mathbb{E}\int_0^{u\wedge\tau}\int_{|z-\theta_{t,r}(y)|\leqslant\eta} \frac{\kappa(r, X_{r-}(x), z - X_{r-}(x))}{|z - X_{r-}(x)|^{d+\alpha}}\mathrm{d}z\mathrm{d}r.\end{aligned}$$

On the other hand, noting that for $r < u\wedge\tau$ and $|z-\theta_{t,r}(y)| \leqslant \eta$, we get

$$\begin{aligned}|z - X_{r-}(x)| &\leqslant |z-\theta_{t,r}(y)| + |\theta_{t,r}(y) - \theta_{0,r}(x)| + |\theta_{0,r}(x) - X_{r-}(x)| \\ &\leqslant \eta + C_2|x-\theta_{t,0}(y)| + C_2 t^{1/\alpha} + \eta \leqslant C_2|x-\theta_{t,0}(y)| + C_3\eta,\end{aligned}$$

using as well (2.7) for the last but one inequality. Since

$$\{|X_{u\wedge\tau}(x) - \theta_{t,u\wedge\tau}(y)| \leqslant \eta\} \subset \{\sigma \leqslant u\},$$

we further have

$$\begin{aligned}\mathbb{P}\{\sigma\leqslant u\} &\geqslant \mathbb{P}\{|X_{u\wedge\tau}(x) - \theta_{t,u\wedge\tau}(y)| \leqslant \eta\} \\ &\geqslant \mathbb{E}\int_0^{u\wedge\tau}\int_{|z-\theta_{t,r}(y)|\leqslant\eta} \frac{\kappa_1^{-1}}{(C_2|x-\theta_{t,0}(y)| + C_3\eta)^{d+\alpha}}\mathrm{d}z\mathrm{d}r \\ &= \frac{\mathbb{E}(u\wedge\tau)\kappa_1^{-1}\eta^d\cdot\mathrm{Vol}(B_1)}{(C_2|x-\theta_{t,0}(y)| + C_3\eta)^{d+\alpha}} \geqslant \frac{c_0 t^{1+d/\alpha}}{|x-\theta_{t,0}(y)|^{d+\alpha}},\end{aligned} \tag{2.16}$$

where the last step is due to $|x-\theta_{t,0}(y)| \geqslant K\eta$ and

$$\mathbb{E}(u\wedge\tau) \geqslant u\mathbb{P}(\tau > u) \overset{(2.13)}{\geqslant} u/2 = \varepsilon t/2.$$

Combining (2.14) and (2.16), we obtain the desired estimate. □

2.3. **Convolution inequalities.** This Section is dedicated to some useful convolution controls associated with functions that are known to be upper-bounds of the isotropic stable density and its gradient, see e.g. [20], [4]. Though a bit technical, these results will turn out to be crucial in order to control the parametrix series representation of the density and its gradient (see e.g. equation (3.10), Lemma 3.3 and Theorem 3.5 below).

For $\eta \in (0,2)$ and $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, let

$$\varrho^{(\eta)}(x) := (1+|x|)^{-d-\eta}, \quad \varrho^{(\eta)}(t,x) := t^{-d/\alpha}\varrho^{(\eta)}(t^{-1/\alpha}x).$$

For $\beta \geqslant 0$ and $\gamma \in \mathbb{R}$, we introduce the following functions for later use

$$\varrho^{(\eta)}_{\beta,\gamma}(t,x) := (1 \wedge (t^{1/\alpha} + |x|))^{\beta} t^{(\gamma-\eta)/\alpha}\varrho^{(\eta)}(t,x) \tag{2.17}$$

and

$$\phi^{(\eta)}_{\beta,\gamma}(s,x,t,y) := \varrho^{(\eta)}_{\beta,\gamma}(t-s, x-\theta_{t,s}(y)). \tag{2.18}$$

Note that

$$\varrho^{(\eta)}_{\beta,\gamma}(t,x) = \frac{(1\wedge(t^{1/\alpha}+|x|))^{\beta}t^{\gamma/\alpha}}{(t^{1/\alpha}+|x|)^{d+\eta}}.$$

For $T>0$, by (2.7) we have for $(s,x,t,y)\in\mathbb{D}_T$,

$$\phi^{(\eta)}_{\beta,\gamma}(s,x,t,y)\asymp\varrho^{(\eta)}_{\beta,\gamma}(t-s,\theta_{s,t}(x)-y), \tag{2.19}$$

and for $\beta\in[0,\eta]$,

$$\int_{\mathbb{R}^d}\phi^{(\eta)}_{\beta,\gamma}(s,x,t,y)\mathrm{d}y\lesssim\int_{\mathbb{R}^d}\varrho^{(\eta)}_{\beta,\gamma}(t-s,y)\mathrm{d}y\lesssim(t-s)^{\frac{\beta+\gamma-\eta}{\alpha}}. \tag{2.20}$$

For two functions f,g on $\mathbb{D}_\infty$, we write

$$(f\odot g)_r(s,x,t,y):=\int_{\mathbb{R}^d}f(s,x,r,z)g(r,z,t,y)\mathrm{d}z$$

and

$$(f\otimes g)(s,x,t,y):=\int_s^t(f\odot g)_r(s,x,t,y)\mathrm{d}r.$$

The following lemma is the same as in [8, Lemma 2.1].

Lemma 2.7. *Fix $\alpha\in(0,2)$. For any $\beta_1,\beta_2\in[0,\frac{\alpha}{4}]$ and $T>0$, there is a $C=C(T,\Theta,\beta_1,\beta_2)>0$ such that for all $\gamma_1>-\beta_1$ and $\gamma_2>-\beta_2$, $r\in[s,t]$ and $x,y\in\mathbb{R}^d$,*

$$\big(\phi^{(\alpha)}_{\beta_1,0}\odot\phi^{(\alpha)}_{\beta_2,0}\big)_r(s,x,t,y)\lesssim_C\big((r-s)^{\frac{\beta_1-\alpha}{\alpha}}+(t-r)^{\frac{\beta_2-\alpha}{\alpha}}\big)\phi^{(\alpha)}_{\beta_1\wedge\beta_2,0}(s,x,t,y) \tag{2.21}$$

and

$$\phi^{(\alpha)}_{\beta_1,\gamma_1}\otimes\phi^{(\alpha)}_{\beta_2,\gamma_2}(s,x,t,y)\lesssim_C\mathcal{B}(\tfrac{\beta_1+\gamma_1}{\alpha},\tfrac{\beta_2+\gamma_2}{\alpha})\phi^{(\alpha)}_{\beta_1\wedge\beta_2,\beta_1+\beta_2+\gamma_1+\gamma_2}(s,x,t,y), \tag{2.22}$$

where $\mathcal{B}(\gamma,\beta)$ is the usual Beta function defined by

$$\mathcal{B}(\gamma,\beta):=\int_0^1(1-s)^{\gamma-1}s^{\beta-1}\mathrm{d}s,\quad\gamma,\beta>0.$$

Proof. We follow the proof in [11]. Let $\ell(u):=\frac{u^{d+\alpha}}{1\wedge u^\beta}$. It is easy to see that, as soon as $d+\alpha>\beta$, ℓ is increasing on $\mathbb{R}_+$ and for any $\lambda\geqslant1$,

$$\ell(\lambda u)\leqslant\lambda^{d+\alpha}\ell(u). \tag{2.23}$$

Hence,

$$\ell(u+w)\leqslant\ell(2(u\vee w))\leqslant2^{d+\alpha}\ell(u\vee w)\leqslant2^{d+\alpha}(\ell(u)+\ell(w)). \tag{2.24}$$

Now for $r\in[s,t]$ and $x,y\in\mathbb{R}^d$, since

$$|t+s|^{1/\alpha}+|x+y|\leqslant2^{1/\alpha}\big(|s|^{1/\alpha}+|x|+|t|^{1/\alpha}+|y|\big),$$

by (2.23) and (2.24), we have

$$\ell(|t+s|^{1/\alpha}+|x+y|)\lesssim_C\ell(|s|^{1/\alpha}+|x|)+\ell(|t|^{1/\alpha}+|y|).$$

In particular,

$$\frac{((t+s)^{1/\alpha}+|x+y|)^{d+\alpha}}{1\wedge((t+s)^{1/\alpha}+|x+y|)^{\beta_1\wedge\beta_2}}\lesssim\frac{(s^{1/\alpha}+|x|)^{d+\alpha}}{1\wedge(s^{1/\alpha}+|x|)^{\beta_1}}+\frac{(t^{1/\alpha}+|y|)^{d+\alpha}}{1\wedge(t^{1/\alpha}+|y|)^{\beta_2}}.$$

Hence,

$$\begin{aligned}&\frac{1\wedge(s^{1/\alpha}+|x|)^{\beta_1}}{(s^{1/\alpha}+|x|)^{d+\alpha}}\times\frac{1\wedge(t^{1/\alpha}+|y|)^{\beta_2}}{(t^{1/\alpha}+|y|)^{d+\alpha}}\\&\lesssim_C\left[\frac{1\wedge(s^{1/\alpha}+|x|)^{\beta_1}}{(s^{1/\alpha}+|x|)^{d+\alpha}}+\frac{1\wedge(t^{1/\alpha}+|y|)^{\beta_2}}{(t^{1/\alpha}+|y|)^{d+\alpha}}\right]\end{aligned}$$

$$\times \frac{1\wedge((t+s)^{1/\alpha}+|x+y|)^{\beta_1\wedge\beta_2}}{((t+s)^{1/\alpha}+|x+y|)^{d+\alpha}}.$$

By this, the desired estimates follow by (2.19), (2.20) and (2.8). □

2.4. **Density estimate.** Let $a:\mathbb{R}_+\to\mathbb{R}^d\otimes\mathbb{R}^d$ be a measurable $d\times d$-matrix-valued function satisfying the non-degeneracy condition

$$\kappa_1^{-1}|\xi|^2\leqslant|a(s)\xi|^2\leqslant\kappa_1|\xi|^2. \tag{2.25}$$

Fix $\alpha\in(0,2)$ and consider the following jump process

$$X^a_{s,t}:=\int_s^t a(r)\mathrm{d}W_{S_r}, \tag{2.26}$$

where W is a d-dimensional Brownian motion and S is an $\alpha/2$-stable subordinator independent from W, both defined on some probability space $(\Omega,\mathcal{F},\mathbb{P})$. Note that

$$X^a_{s,t}\overset{(d)}{=}(t-s)^{1/\alpha}X^{\tilde{a}}_{0,1},$$

where

$$\tilde{a}(r):=a(s+r(t-s)).$$

We have the following lemma that can be derived from the approach initially used in [2] (see also [4]). We provide below a proof for completeness.

Lemma 2.8. *For any $0\leqslant s<t<\infty$, $X^a_{s,t}$ has a smooth density $p^a_{s,t}(x)$ with the scaling property*

$$p^a_{s,t}(x)=(t-s)^{-d/\alpha}p^{\tilde{a}}_{0,1}((t-s)^{-1/\alpha}x), \tag{2.27}$$

which enjoys the following estimates:

$$p^a_{s,t}(x)\asymp_{C_0}\varrho^{(\alpha)}_{0,\alpha}(t-s,x), \tag{2.28}$$

and for any $j\in\mathbb{N}$,

$$|\nabla^j p^a_{s,t}(x)|\lesssim_{C_j}\varrho^{(\alpha+j)}_{0,\alpha}(t-s,x), \tag{2.29}$$

where the constants only depend on κ_1,d,α. Moreover, suppose that the integrand in (2.26) writes as $a_\xi(r)$ and smoothly depends on the parameter $\xi\in\mathbb{R}^d$ so that (2.25) holds uniformly and $\sup_{r,\xi}|\nabla^k_\xi a_\xi(r)|<\infty$ for any $k\in\mathbb{N}$. Let $p^{a_\xi}_{s,t}$ be the density of the integral in (2.26) associated with a_ξ. Then we have for $k\in\mathbb{N}$ and $j\in\mathbb{N}_0$,

$$|\nabla^k_\xi\nabla^j_x p^{a_\xi}_{s,t}(x)|\lesssim_{C_{j,k}}\varrho^{(\alpha+j)}_{0,\alpha}(t-s,x). \tag{2.30}$$

Importantly, this last bound means that, the differentiation w.r.t. the parameter ξ appearing in the diffusion coefficient a_ξ does not yield an additional time singularity.

Proof. The two sided estimate (2.28) is well known (see e.g. [9]). We show (2.29). Without loss of generality, we assume $s=0$ and write

$$X_t:=\int_0^t a(r)\mathrm{d}W_{S_r}.$$

Fix a càdlàg path ℓ_s. Consider the following Gaussian random variable:

$$X^\ell_t:=\int_0^t a(r)\mathrm{d}W_{\ell_r}.$$

It has a density

$$g^{a,\ell}_t(x)=(2\pi)^{-d/2}\sqrt{\det\left((\mathcal{C}^{a,\ell}_t)^{-1}\right)}\exp\{-\langle\left(\mathcal{C}^{a,\ell}_t\right)^{-1}x,x\rangle/2\}, \tag{2.31}$$

where

$$\mathcal{C}_t^{a,\ell} := \int_0^t (aa^*)(r)\mathrm{d}\ell_r.$$

From the non-degeneracy assumption (2.25), we have

$$\langle \big(\mathcal{C}_t^{a,\ell}\big)^{-1} x, x\rangle \asymp |x|^2/\ell_t, \quad \det\big((\mathcal{C}_t^{a,\ell})^{-1}\big) \asymp \ell_t^{-d},$$

and

$$|\nabla g_t^{a,\ell}(x)| \lesssim |x|/\ell_t \exp\{-\lambda|x|^2/\ell_t\}.$$

The density $p_{0,t}^a(x) =: p_t^a(x)$ of X_t is given by

$$p_t^a(x) = \mathbb{E} g_t^{a,S}(x). \tag{2.32}$$

The bound of (2.29) is direct from the Fourier representation of the density when $|x| \leqslant t^{1/\alpha}$. On the other hand, for $|x| > t^{1/\alpha}$, from the global bound on the law of the subordinator

$$\mu_{S_t}(\mathrm{d}r) := \mathbb{P}\circ S_t^{-1}(\mathrm{d}r) \lesssim t\, r^{-\alpha/2-1}\mathrm{d}r,$$

it readily follows that

$$|\nabla p_t^a(x)| \leqslant \mathbb{E}|\nabla g_t^{a,S}(x)| \lesssim |x|\mathbb{E}(S_t^{-d/2-1}\exp\{-\lambda|x|^2/S_t\}) < +\infty.$$

Hence, from the bounded convergence theorem it holds that

$$|\nabla p_t^a(x)| \lesssim |x| \int_0^\infty r^{-(d+2)/2}\mathrm{e}^{-\lambda|x|^2/r}\mu_{S_t}(\mathrm{d}r),$$

and the integral expression in the r.h.s. precisely corresponds to the stable heat kernel in dimension $d+2$ at time t and point $\tilde{x} \in \mathbb{R}^{d+2}$ s.t. $|\tilde{x}| = \sqrt{\lambda}|x|$. Thus, from (2.28),

$$\begin{aligned}|\nabla p_t^a(x)| &\lesssim |x| t^{-(d+2)/\alpha}\frac{1}{(1+t^{-1/\alpha}|\tilde{x}|)^{d+2+\alpha}}\\ &\lesssim t(t^{1/\alpha}+|x|)^{-d-\alpha-1} = \varrho_{0,\alpha}^{(\alpha+1)}(t,x).\end{aligned}$$

The approach is similar for higher order derivatives. This is also the case for (2.30) recalling that differentiating a Gaussian density w.r.t. the variance does not induce additional singularities. The proof is complete. □

Remark 2.9. We would like to emphasize that the gradient estimate (2.29) plays a crucial role for two-sided estimates due to the fact that for any $\beta \in [0,1]$,

$$|x|^\beta \varrho_{0,\alpha}^{(\alpha+1)}(t,x) = \frac{t|x|^\beta}{(t^{1/\alpha}+|x|)^{d+\alpha+1}} \leqslant \frac{t^{(\alpha+\beta-1)/\alpha}}{(t^{1/\alpha}+|x|)^{d+\alpha}} = \varrho_{0,\beta+\alpha-1}^{(\alpha)}(t,x).$$

In particular, for any $\beta \in [0,1]$,

$$|x-\theta_{t,s}(x)|^\beta \phi_{0,\alpha}^{(\alpha+1)}(s,x,t,y) \leqslant \phi_{0,\beta+\alpha-1}^{(\alpha)}(s,x,t,y). \tag{2.33}$$

We carefully point out that the Gradient estimate (2.29), which remarkably emphasizes a concentration gain, does not hold for a general α-stable like process [15]. This is also why, for the driving process in (1.1), we limit ourselves to the rotationally invariant, and thus symmetric, α-stable process and do not handle general α-stable like processes.

The following lemma is taken from [9, Lemmas 3.2 and 3.3].

Lemma 2.10. *Under (2.25), there is a constant $C = C(d, \alpha, \kappa_1) > 0$ such that*

$$|\nabla p^a_{s,t} - \nabla p^{\bar{a}}_{s,t}|(x) \lesssim_C \|a - \bar{a}\|_\infty \varrho^{(\alpha+1)}_{0,\alpha}(t-s, x). \tag{2.34}$$

Also,

$$|\mathcal{D}^{(\alpha)} p^a_{s,t}|(x) \lesssim_C \varrho^{(\alpha)}_{0,0}(t-s, x), \tag{2.35}$$

and

$$|\mathcal{D}^{(\alpha)}(p^a_{s,t} - p^{\bar{a}}_{s,t})|(x) \lesssim_C \|a - \bar{a}\|_\infty \varrho^{(\alpha)}_{0,0}(t-s, x). \tag{2.36}$$

Moreover, we also have

$$\begin{aligned} &\int_{\mathbb{R}^d} |\delta^{(2)}_{p^a_{s,t}}(x_1; z) - \delta^{(2)}_{p^a_{s,t}}(x_2; z)| \frac{\mathrm{d}z}{|z|^{d+\alpha}} \\ &\quad \lesssim_C \left(\frac{|x_1 - x_2|}{(t-s)^{1/\alpha}} \wedge 1 \right) \left(\sum_{i=1,2} \varrho^{(\alpha)}_{0,0}(t-s, x_i) \right). \end{aligned} \tag{2.37}$$

Proof. From the scaling property (2.27), it suffices to consider $s = 0$ and $t = 1$. Note that

$$\begin{aligned} |\delta^{(2)}_{p^a_1}(x; z)| &= |p^a_1(x+z) + p^a_1(x-z) - 2p^a_1(x)| \\ &\lesssim (|z|^2 \wedge 1)(\varrho^{(\alpha)}(x+z) + \varrho^{(\alpha)}(x-z) + \varrho^{(\alpha)}(x)). \end{aligned}$$

By elementary calculations, one sees that

$$\int_{\mathbb{R}^d} \varrho^{(\alpha)}(x+z) \frac{(|z|^2 \wedge 1)\mathrm{d}z}{|z|^{d+\alpha}} \lesssim_C \varrho^{(\alpha)}(x). \tag{2.38}$$

Thus (2.35) follows. As for (2.36) and (2.37), they can be derived similarly to [11, Lemma 2.7 and Lemma 2.8]. The statement (2.34) can also be derived from the arguments developed therein. We omit the details. □

3. Heat kernel of nonlocal operators with smooth coefficients

In this section we assume that $\mathbf{(H}^\beta_b\mathbf{)}$ and $\mathbf{(H}^\gamma_a\mathbf{)}$ hold and additionally that for any $j \in \mathbb{N}$,

$$\|\nabla^j b\|_\infty + \|\nabla^j a\|_\infty < \infty. \tag{3.1}$$

We shall denote

$$\mathscr{C} := \big\{(b, a) : \text{ satisfying } \mathbf{(H}^\beta_b\mathbf{)}, \mathbf{(H}^\gamma_a\mathbf{)} \text{ with common bounds } \kappa_0, \kappa_1 \text{ and (3.1).}\big\}$$

Under $\mathbf{(H}^\gamma_a\mathbf{)}$ and (3.1), for each $(s, x) \in \mathbb{R}_+ \times \mathbb{R}^d$, it is well known that there is a unique solution $(X_{s,t}(x))_{t \geqslant s}$ to SDE (1.1), and $X_{s,t}(x)$ has for $t > s$ a density $p(s, x, t, y)$ so that (cf. [13, 7])

$$P_{s,t} f(x) := \mathbb{E} f(X_{s,t}(x)) = \int_{\mathbb{R}^d} f(y) p(s, x, t, y) \mathrm{d}y, \quad f \in L^\infty(\mathbb{R}^d).$$

The density is also a mild solution of the Kolmogorov equation in the sense that for all $\varphi \in C^2_0(\mathbb{R}^d)$

$$P_{s,t}\varphi(x) = \varphi(x) + \int_s^t P_{s,r} \mathscr{L}_r \varphi(x) \mathrm{d}r. \tag{3.2}$$

Fix $(\tau, \xi) \in [s, t] \times \mathbb{R}^d$. Consider the following freezing process

$$X^{(\tau,\xi)}_{s,t} := x + \int_s^t b_{|r-\tau|^{1/\alpha}}(r, \theta_{\tau,r}(\xi)) \mathrm{d}r + \int_s^t a(r, \theta_{\tau,r}(\xi)) \mathrm{d}L^{(\alpha)}_r.$$

By Lemma 2.8, the density of $X^{(\tau,\xi)}_{s,t}$ is given by

$$\tilde{p}^{(\tau,\xi)}(s,x,t,y)=p^{a^{(\tau,\xi)}}_{s,t}\left(x-y+\int_s^t b_{|r-\tau|^{1/\alpha}}(r,\theta_{\tau,r}(\xi))\mathrm{d}r\right), \tag{3.3}$$

where $a^{(\tau,\xi)}(r):=a(r,\theta_{\tau,r}(\xi))$ and $p^{a^{(\tau,\xi)}}_{s,t}$ is the density of $\int_s^t a^{(\tau,\xi)}(r)\mathrm{d}L^{(\alpha)}_r$ given in Lemma 2.8. In particular,

$$\partial_s\tilde{p}^{(\tau,\xi)}(s,x,t,y)+\tilde{\mathscr{L}}^{(\tau,\xi)}_s\tilde{p}^{(\tau,\xi)}(s,\cdot,t,y)(x)=0, \tag{3.4}$$

where

$$\tilde{\mathscr{L}}^{(\tau,\xi)}_s f(x):=\tfrac{1}{2}\tilde{\mathcal{L}}^{(\tau,\xi)}_s f(x)+b_{|s-\tau|^{1/\alpha}}(s,\theta_{\tau,s}(\xi))\cdot\nabla f(x)$$

and

$$\tilde{\mathcal{L}}^{(\tau,\xi)}_s f(x)=\int_{\mathbb{R}^d}\delta^{(2)}_f(x;z)\frac{\kappa(s,\theta_{\tau,s}(\xi),z)}{|z|^{d+\alpha}}\mathrm{d}z$$

with

$$\kappa(s,\theta_{s,\tau}(\xi),z):=\frac{\det(a^{-1}(s,\theta_{\tau,s}(\xi))|z|^{d+\alpha}}{|a^{-1}(s,\theta_{\tau,s}(\xi))z|^{d+\alpha}}.$$

For simplicity, we shall write

$$\mathscr{A}^{(\tau,\xi)}_s f(x):=(\mathscr{L}_s-\tilde{\mathscr{L}}^{(\tau,\xi)}_s)f(x)=\mathscr{K}^{(\tau,\xi)}_s f(x)+\mathscr{B}^{(\tau,\xi)}_s f(x), \tag{3.5}$$

where

$$\mathscr{K}^{(\tau,\xi)}_s f(x):=\tfrac{1}{2}(\mathcal{L}_s-\tilde{\mathcal{L}}^{(\tau,\xi)}_s)f(x),$$

and

$$\mathscr{B}^{(\tau,\xi)}_s f(x):=\big(b(s,x)-b_{|s-\tau|^{1/\alpha}}(s,\theta_{\tau,s}(\xi))\big)\cdot\nabla f(x).$$

Let us introduce the corresponding frozen semi-group:

$$\widetilde{P}^{(\tau,\xi)}_{s,t}f(x):=\mathbb{E}f(X^{(\tau,\xi)}_{s,t}(x)). \tag{3.6}$$

We have the following Duhamel type representation formula:

Lemma 3.1. *For any $f\in C^\infty_b(\mathbb{R}^d)$ and $(\tau,\xi)\in[s,t]\times\mathbb{R}^d$, it holds that*

$$P_{s,t}f=\widetilde{P}^{(\tau,\xi)}_{s,t}f+\int_s^t P_{s,r}\mathscr{A}^{(\tau,\xi)}_r\widetilde{P}^{(\tau,\xi)}_{r,t}f\mathrm{d}r=\widetilde{P}^{(\tau,\xi)}_{s,t}f+\int_s^t\widetilde{P}^{(\tau,\xi)}_{s,r}\mathscr{A}^{(\tau,\xi)}_r P_{r,t}f\mathrm{d}r.$$

Proof. We drop for the proof the superscript (τ,ξ) for notational simplicity. Since from (3.2) and (3.4),

$$\partial_t P_{s,t}f=P_{s,t}\mathscr{L}_t f,\quad \partial_s\widetilde{P}_{s,t}f=-\widetilde{\mathscr{L}_s}\widetilde{P}_{s,t}f,$$

by the chain rule, we have

$$\partial_r(P_{s,r}\widetilde{P}_{r,t}f)=P_{s,r}\mathscr{L}_r\widetilde{P}_{r,t}f-P_{s,r}\widetilde{\mathscr{L}_r}\widetilde{P}_{r,t}f=P_{s,r}\mathscr{A}_r\widetilde{P}_{r,t}f.$$

Integrating both sides from s to t with respect to r yields

$$P_{s,t}f=\widetilde{P}_{s,t}f+\int_s^t P_{s,r}\mathscr{A}_r\widetilde{P}_{r,t}f\mathrm{d}r.$$

Similarly, one can show that

$$\widetilde{P}_{s,t}f=P_{s,t}f-\int_s^t\widetilde{P}_{s,r}\mathscr{A}_r P_{r,t}f\mathrm{d}r.$$

The proof is complete. □

By Lemma 3.1, we have for each $(\tau,\xi)\in[s,t]\times\mathbb{R}^d$ and $x,y\in\mathbb{R}^d$,

$$p(s,x,t,y)=\widetilde{p}^{(\tau,\xi)}(s,x,t,y)+\int_s^t\int_{\mathbb{R}^d}p(s,x,r,z)\mathscr{A}_r^{(\tau,\xi)}\widetilde{p}^{(\tau,\xi)}(r,\cdot,t,y)(z)\mathrm{d}z\mathrm{d}r.$$

In particular, if we take $(\tau,\xi)=(t,y)$ and define

$$p_0(s,x,t,y):=\tilde{p}^{(t,y)}(s,x,t,y)=p_{s,t}^{a^{(t,y)}}\left(x-\theta_{t,s}(y)\right), \tag{3.7}$$

then we obtain the *forward* representation,

$$p(s,x,t,y)=p_0(s,x,t,y)+\int_s^t\int_{\mathbb{R}^d}p(s,x,r,z)\mathscr{A}_r^{(t,y)}p_0(r,\cdot,t,y)(z)\mathrm{d}z\mathrm{d}r. \tag{3.8}$$

Let

$$q_0(s,x,t,y):=\mathscr{A}_s^{(t,y)}p_0(s,\cdot,t,y)(x),$$

and define recursively for $n\geqslant 1$,

$$q_n:=q_0\otimes q_{n-1},\quad q=\sum_{n=0}^{\infty}q_n. \tag{3.9}$$

By iteration, we *formally* obtain from (3.8) and (3.9),

$$p=p_0+p\otimes q_0=p_0+\sum_{n=0}^{\infty}p_0\otimes q_n=p_0+p_0\otimes q. \tag{3.10}$$

The following lemma is a direct consequence of (3.7), (2.35) and (2.29).

Lemma 3.2. *For any $\alpha\in(0,2)$ and $j=0,1,\cdots$, we have*

$$|\nabla^j p_0(s,\cdot,t,y)|(x)\lesssim\phi_{0,\alpha}^{(\alpha+j)}(s,x,t,y) \tag{3.11}$$

and

$$|\mathcal{D}^{(\alpha)}p_0(s,\cdot,t,y)|(x)\lesssim\phi_{0,0}^{(\alpha)}(s,x,t,y). \tag{3.12}$$

The following lemma corresponds to [8, Theorem 3.1].

Lemma 3.3. *The series $q=\sum_{n=0}^{\infty}q_n$ is absolutely convergent, and for each $s<t$, $(x,y)\mapsto q(s,x,t,y)$ is equi-continuous in $(b,a)\in\mathscr{C}$. Moreover, for any $T>0$, there is a constant $C=C(T,\Theta)>0$ such that for all $(s,x,t,y)\in\mathbb{D}_T$,*

$$|q(s,x,t,y)|\lesssim_C\left(\phi_{\gamma_0,0}^{(\alpha)}+\phi_{0,\gamma_0}^{(\alpha)}\right)(s,x,t,y), \tag{3.13}$$

where $\gamma_0:=(\alpha+\beta-1)\wedge\gamma$, and for any $\gamma_1\in(0,\gamma_0)$,

$$\begin{aligned}&|q(s,x,t,y)-q(s,x',t,y)|\lesssim_C(|x-x'|^{\gamma_1}\wedge 1)\\&\quad\times\left(\left(\phi_{\gamma_0,-\gamma_1}^{(\alpha)}+\phi_{0,\gamma_0-\gamma_1}^{(\alpha)}\right)(s,x,t,y)+\left(\phi_{\gamma_0,-\gamma_1}^{(\alpha)}+\phi_{0,\gamma_0-\gamma_1}^{(\alpha)}\right)(s,x',t,y)\right).\end{aligned} \tag{3.14}$$

Proof. (i) First of all, note that by (1.6),

$$|\kappa(s,x,z)-\kappa(s,\theta_{t,s}(y),z)|\lesssim(|x-\theta_{t,s}(y)|^{\gamma}\wedge 1)$$

and by $(\mathbf{H}_b^{\beta})$,

$$|b(s,x)-b_{|s-t|^{1/\alpha}}(s,\theta_{t,s}(y))|\lesssim|x-\theta_{t,s}(y)|^{\beta}+|x-\theta_{t,s}(y)|+|t-s|^{\beta/\alpha}.$$

Thus, we have by (3.12),

$$|\mathscr{K}_s^{(t,y)}p_0(s,\cdot,t,y)(x)|\lesssim\phi_{\gamma,0}^{(\alpha)}(s,x,t,y),$$

and by (3.11) and (2.33),

$$|\mathscr{B}_s^{(t,y)}p_0(s,\cdot,t,y)(x)|\lesssim\phi_{0,\alpha+\beta-1}^{(\alpha)}(s,x,t,y).$$

So, for $\gamma_0 = \gamma \wedge (\alpha+\beta-1)$,

$$|q_0(s,x,t,y)| \lesssim \Big(\phi^{(\alpha)}_{\gamma,0} + \phi^{(\alpha)}_{0,\alpha+\beta-1}\Big)(s,x,t,y) \lesssim \Big(\phi^{(\alpha)}_{\gamma_0,0} + \phi^{(\alpha)}_{0,\gamma_0}\Big)(s,x,t,y).$$

Suppose now that for some $k \in \mathbb{N}$,

$$|q_{k-1}(s,x,t,y)| \leqslant C_k \Big(\phi^{(\alpha)}_{\gamma_0,(k-1)\gamma_0} + \phi^{(\alpha)}_{0,k\gamma_0}\Big)(s,x,t,y).$$

By Lemma 2.7, we have

$$\begin{aligned}|q_k(s,x,t,y)| &\leqslant CC_k\big(\phi^{(\alpha)}_{\gamma_0,0} + \phi^{(\alpha)}_{0,\gamma_0}\big) \otimes \big(\phi^{(\alpha)}_{\gamma_0,(k-1)\gamma_0} + \phi^{(\alpha)}_{0,k\gamma_0}\big)(s,x,t,y)\\ &\leqslant C_0C_k\mathcal{B}(\tfrac{\gamma_0}{\alpha}, \tfrac{k\gamma_0}{\alpha})\big(\phi^{(\alpha)}_{\gamma_0,k\gamma_0} + \phi^{(\alpha)}_{0,(k+1)\gamma_0}\big)(s,x,t,y). \qquad (3.15)\end{aligned}$$

Hence,

$$C_{k+1} = C_0C_k\mathcal{B}(\tfrac{\gamma_0}{\alpha}, \tfrac{k\gamma_0}{\alpha}).$$

From the relation $\mathcal{B}(\gamma,\beta) = \frac{\Gamma(\gamma)\Gamma(\beta)}{\Gamma(\gamma+\beta)}$, where Γ is the usual Gamma function, we obtain

$$C_k = C_0^k \prod_{i=1}^{k-1} \mathcal{B}(\tfrac{\gamma_0}{\alpha}, \tfrac{(k-1)\gamma_0}{\alpha}) = \frac{(C_0\Gamma(\gamma_0/\alpha))^k}{\Gamma(k\gamma_0/\alpha)},$$

with the usual convention that $\prod_{i=1}^0 = 1$. Thus

$$\begin{aligned}\sum_{k=0}^\infty |q_k(s,x,t,y)| &\leqslant \sum_{k=0}^\infty \frac{(C_0\Gamma(\gamma_0/\alpha))^k}{\Gamma(k\gamma_0/\alpha)}\Big(\phi^{(\alpha)}_{\gamma_0,k\gamma_0} + \phi^{(\alpha)}_{0,(k+1)\gamma_0}\Big)(s,x,t,y)\\ &\leqslant \sum_{k=0}^\infty \frac{(C_0\Gamma(\gamma_0/\alpha))^k}{\Gamma(k\gamma_0/\alpha)}\Big(\phi^{(\alpha)}_{\gamma_0,0} + \phi^{(\alpha)}_{0,\gamma_0}\Big)(s,x,t,y).\end{aligned}$$

This gives (3.13).

(ii) For fixed $s<t$, by Lemma 2.8 and the definition of q_0, one sees that $(x,y) \mapsto q_0(s,x,t,y)$ is equi-continuous in $(b,a) \in \mathscr{C}$. Furthermore, it follows by induction that, for each $k \in \mathbb{N}$, $(x,y) \mapsto q_k(s,x,t,y)$ is also equi-continuous in $(b,a) \in \mathscr{C}$. Hence, $(x,y) \mapsto q(s,x,t,y)$ is equi-continuous in $(b,a) \in \mathscr{C}$.

(iii) If $|x-x'| \geqslant (t-s)^{1/\alpha}$, then we have

$$\begin{aligned}|q_0(s,x,t,y)| &\lesssim (|x-x'|^{\gamma_1} \wedge 1)(t-s)^{-\gamma_1/\alpha}\big(\phi^{(\alpha)}_{\gamma_0,0} + \phi^{(\alpha)}_{0,\gamma_0}\big)(s,x,t,y)\\ &= (|x-x'|^{\gamma_1} \wedge 1)\big(\phi^{(\alpha)}_{\gamma_0,-\gamma_1} + \phi^{(\alpha)}_{0,\gamma_0-\gamma_1}\big)(s,x,t,y).\end{aligned}$$

Next we assume $|x-x'| \leqslant (t-s)^{1/\alpha}$. In this case, it is easy to see from (2.17)-(2.18), that

$$\phi^{(\eta)}_{0,0}(s,x,t,y) \asymp \phi^{(\eta)}_{0,0}(s,x',t,y), \quad \eta \geqslant 0. \qquad (3.16)$$

By (2.35), (2.37) and (3.16), we have

$$\begin{aligned}&|\mathscr{K}_s^{(t,y)}p_0(s,\cdot,t,y)(x) - \mathscr{K}_s^{(t,y)}p_0(s,\cdot,t,y)(x')|\\ &\leqslant \|\kappa(\cdot,x,\cdot) - \kappa(\cdot,x',\cdot)\|_\infty \int_{\mathbb{R}^d} |\delta^{(2)}_{p_0(s,\cdot,t,y)}(x;z)|\frac{\mathrm{d}z}{|z|^{d+\alpha}}\\ &\quad + \|\kappa(\cdot,x,\cdot) - \kappa(\cdot,\theta_{t,s}(y),\cdot)\|_\infty\\ &\qquad \times \int_{\mathbb{R}^d} |\delta^{(2)}_{p_0(s,\cdot,t,y)}(x;z) - \delta^{(2)}_{p_0(s,\cdot,t,y)}(x';z)|\frac{\mathrm{d}z}{|z|^{d+\alpha}}\\ &\leqslant (|x-x'|^\gamma \wedge 1)\phi^{(\alpha)}_{0,0}(s,x,t,y) + (|x-\theta_{t,s}(y)|^\gamma \wedge 1)\\ &\quad \times \left(\frac{|x-x'|}{(t-s)^{1/\alpha}} \wedge 1\right)\Big(\phi^{(\alpha)}_{0,0}(s,x,t,y) + \phi^{(\alpha)}_{0,0}(s,x',t,y)\Big)\end{aligned}$$

$$\lesssim (|x-x'|^{\gamma_1}\wedge 1)\Big(\phi^{(\alpha)}_{0,\gamma-\gamma_1}(s,x,t,y)+\phi^{(\alpha)}_{\gamma,-\gamma_1}(s,x,t,y)\Big).$$

Moreover, by (3.11), (3.16) and (2.33), we also have

$$\begin{aligned}
&|\mathscr{B}^{(t,y)}_s p_0(s,\cdot,t,y)(x)-\mathscr{B}^{(t,y)}_s p_0(s,\cdot,t,y)(x')|\leqslant |b(s,x)-b(s,x')|\cdot|\nabla p_0(s,\cdot,t,y)|(x')\\
&\quad+\big|b(s,x)-b_{|s-t|^{1/\alpha}}(s,\theta_{t,s}(y))\big|\cdot|\nabla p_0(s,\cdot,t,y)(x')-\nabla p_0(s,\cdot,t,y)(x)|\\
&\lesssim |x-x'|^\beta\phi^{(\alpha+1)}_{0,\alpha}(s,x,t,y)+(|x-\theta_{t,s}(y)|^\beta+|t-s|^{\beta/\alpha})|x-x'|\phi^{(\alpha+2)}_{0,\alpha}(s,x,t,y)\\
&\lesssim (|x-x'|^{\gamma_1}\wedge 1)\phi^{(\alpha)}_{0,\alpha+\beta-1-\gamma_1}(s,x,t,y).
\end{aligned}$$

Combining the above calculations and recalling $\gamma_0=\gamma\wedge(\alpha+\beta-1)$, we obtain

$$\begin{aligned}
&|q_0(s,x,t,y)-q_0(s,x',t,y)|\lesssim_C (|x-x'|^{\gamma_1}\wedge 1)\\
&\quad\times\Big((\phi^{(\alpha)}_{\gamma_0,-\gamma_1}+\phi^{(\alpha)}_{0,\gamma_0-\gamma_1})(s,x,t,y)+(\phi^{(\alpha)}_{\gamma_0,-\gamma_1}+\phi^{(\alpha)}_{0,\gamma_0-\gamma_1})(s,x',t,y)\Big).
\end{aligned}$$

Using this last estimate, equation (3.14) follows from the same iterative argument as in (i). □

Remark 3.4. This lemma allows to iterate the representation (3.8) which leads to the representation (3.10) of the density.

We now aim at proving the following *a priori* estimate about $p(s,x,t,y)$.

Theorem 3.5. *Under* $(\mathbf{H}^\gamma_a)$, $(\mathbf{H}^\beta_b)$ *and* (3.1), *for each* $0\leqslant s<t<\infty$, $X_{s,t}(x)$ *admits a density* $p(s,x,t,y)$ *that is equi-continuous in* $(b,a)\in\mathscr{C}$ *as a function of* $x,y\in\mathbb{R}^d$, *and there is a constant* $C=C(T,\Theta)>0$ *so that for all* $(s,x,t,y)\in\mathbb{D}_T$,

$$p(s,x,t,y)\asymp_C \phi^{(\alpha)}_{0,\alpha}(s,x,t,y). \tag{3.17}$$

Proof. Note that by (3.7), (2.28) and (2.7),

$$p_0(s,x,t,y)\asymp_C \phi^{(\alpha)}_{0,\alpha}(s,x,t,y).$$

By Lemma 2.7, we have

$$|p_0\otimes q|(s,x,t,y)\lesssim_C (\phi^{(\alpha)}_{0,\alpha+\gamma_0}+\phi^{(\alpha)}_{\gamma_0,\alpha})(s,x,t,y).$$

The upper bound follows from (3.10).

Next we use Lemma 2.6 to show the lower bound estimate. Let K be as in Lemma 2.6. Suppose that $|x-\theta_{t,s}(y)|\leqslant 2K(t-s)^{1/\alpha}$. Then we have

$$\begin{aligned}
p(s,x,t,y)&\geqslant p_0(s,x,t,y)-|p_0\otimes q(s,x,t,y)|\\
&\geqslant c_0\phi^{(\alpha)}_{0,\alpha}(s,x,t,y)-(\phi^{(\alpha)}_{0,\alpha+\gamma_0}+\phi^{(\alpha)}_{\gamma_0,\alpha})(s,x,t,y)\\
&\geqslant (c_0-C_1(t-s)^{\frac{\gamma_0}{\alpha}})\phi^{(\alpha)}_{0,\alpha}(s,x,t,y).
\end{aligned}$$

In particular, if $t-s\leqslant\ell$ with ℓ small enough and $|x-\theta_{t,s}(y)|\leqslant 2K(t-s)^{1/\alpha}$, then

$$p(s,x,t,y)\geqslant \tfrac{c_0}{2}\phi^{(\alpha)}_{0,\alpha}(s,x,t,y)\geqslant c_1(t-s)^{-d/\alpha}. \tag{3.18}$$

Next we prove the above estimate still holds for

$$|x-\theta_{t,s}(y)|\geqslant 2K(t-s)^{1/\alpha}.$$

Let $\varepsilon\in(0,1/2)$ be as in Lemma 2.6 and small enough so that $2(1-\varepsilon)^{1/\alpha}\geqslant 1$. Let

$$r:=s+\varepsilon(t-s),\quad B:=\{z:|z-\theta_{t,r}(y)|\leqslant 2K(t-r)^{1/\alpha}\}.$$

Since $2(1-\varepsilon)^{1/\alpha}\geqslant 1$, we clearly have

$$B\supset\{z:|z-\theta_{t,r}(y)|\leqslant K(t-s)^{1/\alpha}\}=:B'.$$

Now from the Chapman-Kolmogorov equation, we have for $t-s\leqslant\ell$,

$$\begin{aligned}p(s,x,t,y)&=\int_{\mathbb{R}^d}p(s,x,r,z)p(r,z,t,y)\mathrm{d}z\\&\geqslant\int_B p(s,x,r,z)p(r,z,t,y)\mathrm{d}z\\&\geqslant\inf_{z\in B}p(r,z,t,y)\int_B p(s,x,r,z)\mathrm{d}z\\&\overset{(3.18)}{\geqslant}c_1(t-r)^{-d/\alpha}\mathbb{P}(X_{s,r}(x)\in B)\\&\geqslant c_2(t-s)^{-d/\alpha}\mathbb{P}(X_{s,r}(x)\in B')\\&\geqslant c_3(t-s)|x-\theta_{t,s}(y)|^{-d-\alpha},\end{aligned}$$

where the last step is due to Lemma 2.6. Thus we obtain that for some $c_4>0$ and all $s,t\in[0,T]$,

$$p(s,x,t,y)\geqslant c_4\phi^{(\alpha)}_{0,\alpha}(s,x,t,y),\ \ t-s\leqslant\ell,\ \ x,y\in\mathbb{R}^d.$$

For $t-s>\ell$, the bound follows iteratively from the Chapman-Kolmogorov equation. The proof is complete. □

For the fractional derivative estimates, we need the following lemma.

Lemma 3.6. *For $s<t$, let $h_{s,t}(x):=\int_{\mathbb{R}^d}p_0(s,x,t,y)\mathrm{d}y$. We have for some $C>0$,*

$$|\mathcal{D}^{(\alpha)}h_{s,t}|(x)\lesssim_C(t-s)^{\gamma_0/\alpha-1},\ \ \gamma_0:=\gamma\wedge(\alpha+\beta-1).$$

Proof. By definition we have

$$\begin{aligned}|\mathcal{D}^{(\alpha)}h_{s,t}|(x)&=\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}\delta^{(2)}_{\tilde{p}^{(t,y)}(s,\cdot,t,y)}(x;z)\mathrm{d}y\right|\frac{\mathrm{d}z}{|z|^{d+\alpha}}\\&=\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}\delta^{(2)}_{p^{a^{(t,y)}}_{s,t}}(x-\theta_{t,s}(y);z)\mathrm{d}y\right|\frac{\mathrm{d}z}{|z|^{d+\alpha}}\\&\leqslant\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}\delta^{(2)}_{p^{a^{(t,y)}}_{s,t}-p^{a^{(s,x)}}_{s,t}}(x-\theta_{t,s}(y);z)\mathrm{d}y\right|\frac{\mathrm{d}z}{|z|^{d+\alpha}}\\&\quad+\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}\delta^{(2)}_{p^{a^{(s,x)}}_{s,t}}(x-\theta_{t,s}(y);z)\mathrm{d}y\right|\frac{\mathrm{d}z}{|z|^{d+\alpha}}=:I_1+I_2.\end{aligned}$$

For I_1, noting that by $(\mathbf{H}^\gamma_a)$ and Lemma 2.2,

$$|a(r,\theta_{s,r}(x))-a(r,\theta_{t,r}(y))|\lesssim 1\wedge|x-\theta_{t,s}(y)|^\gamma+|t-s|^{\gamma/\alpha},$$

we have

$$\begin{aligned}I_1&\leqslant\int_{\mathbb{R}^d}|\mathcal{D}^{(\alpha)}(p^{a^{(t,y)}}_{s,t}-p^{a^{(s,x)}}_{s,t})|(x-\theta_{t,s}(y))\mathrm{d}y\\&\overset{(2.36)}{\lesssim}\int_{\mathbb{R}^d}\Big(\phi^{(\alpha)}_{\gamma,0}+\phi^{(\alpha)}_{0,\gamma}\Big)(s,x,t,y)\mathrm{d}y\overset{(2.20)}{\lesssim}(t-s)^{\gamma/\alpha-1}.\end{aligned}$$

For I_2, by the change of variable we have

$$\begin{aligned}I_2&=\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}\delta^{(2)}_{p^{a^{(s,x)}}_{s,t}}(x-y;z)\det(\nabla\theta^{-1}_{s,t}(y))\mathrm{d}y\right|\frac{\mathrm{d}z}{|z|^{d+\alpha}}\\&=\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}\delta^{(2)}_{p^{a^{(s,x)}}_{s,t}}(x-y;z)\Big(\det(\nabla\theta^{-1}_{s,t}(y))-1\Big)\mathrm{d}y\right|\frac{\mathrm{d}z}{|z|^{d+\alpha}},\end{aligned}$$

where we have used that

$$\int_{\mathbb{R}^d} p_{s,t}^{a^{(s,x)}}(x-y)\mathrm{d}y = 1 \Rightarrow \int_{\mathbb{R}^d} \delta^{(2)}_{p_{s,t}^{a^{(s,x)}}}(x-y)\mathrm{d}y = 0.$$

Thus by (i) of Lemma 2.2 and (2.35), we have

$$\begin{aligned} I_2 &\lesssim (t-s)^{\frac{\beta+\alpha-1}{\alpha}} \int_{\mathbb{R}^d} |\mathcal{D}^{(\alpha)} p_{s,t}^{a^{(s,x)}}|(x-y)\mathrm{d}y \\ &\lesssim (t-s)^{\frac{\beta+\alpha-1}{\alpha}} \int_{\mathbb{R}^d} \varrho^{(\alpha)}_{0,0}(t-s, x-y)\mathrm{d}y \\ &\lesssim (t-s)^{\frac{\beta-1}{\alpha}} = (t-s)^{-1+\frac{\alpha+\beta-1}{\alpha}}. \end{aligned}$$

The proof is complete. □

Lemma 3.7. *(Fractional derivative estimate) For any $T > 0$, we have for some $C = C(T, \Theta) > 0$,*

$$|\mathcal{D}^{(\alpha)} p(s, \cdot, t, y)|(x) \lesssim_C \phi^{(\alpha)}_{0,0}(s, x, t, y).$$

Proof. Let $u = (s+t)/2$. By (3.10) and the definition of $\delta^{(2)}$, we have

$$\begin{aligned} \delta^{(2)}_{p(s,\cdot,t,y)}(x;\bar{z}) &= \delta^{(2)}_{p_0(s,\cdot,t,y)}(x;\bar{z}) + \int_s^t\int_{\mathbb{R}^d} \delta^{(2)}_{p_0(s,\cdot,r,z)}(x;\bar{z}) q(r,z,t,y)\mathrm{d}z\mathrm{d}r \\ &= \delta^{(2)}_{p_0(s,\cdot,t,y)}(x;\bar{z}) + \int_s^u \left(\int_{\mathbb{R}^d} \delta^{(2)}_{p_0(s,\cdot,r,z)}(x;\bar{z})\mathrm{d}z\right) q(r,\theta_{s,r}(x),t,y)\mathrm{d}r \\ &+ \int_s^u\int_{\mathbb{R}^d} \delta^{(2)}_{p_0(s,\cdot,r,z)}(x;\bar{z})(q(r,z,t,y) - q(r,\theta_{s,r}(x),t,y))\mathrm{d}z\mathrm{d}r \\ &+ \int_u^t\int_{\mathbb{R}^d} \delta^{(2)}_{p_0(s,\cdot,r,z)}(x;\bar{z}) q(r,z,t,y)\mathrm{d}z\mathrm{d}r. \end{aligned}$$

With the notations of Lemma 3.6, set $h_{s,r}(x) = \int_{\mathbb{R}^d} p_0(s,x,r,z)\mathrm{d}z$. By (1.14) and the Fubini theorem, we have

$$\begin{aligned} |\mathcal{D}^{(\alpha)} p(s,\cdot,t,y)|(x) &\leqslant |\mathcal{D}^{(\alpha)} p_0(s,\cdot,t,y)|(x) + \int_s^u |\mathcal{D}^{(\alpha)} h_{s,r}|(x)|q(r,\theta_{s,r}(x),t,y)|\mathrm{d}r \\ &+ \int_s^u\int_{\mathbb{R}^d} |\mathcal{D}^{(\alpha)} p_0(s,\cdot,r,z)|(x)|q(r,z,t,y) - q(r,\theta_{s,r}(x),t,y)|\mathrm{d}z\mathrm{d}r \\ &+ \int_u^t\int_{\mathbb{R}^d} |\mathcal{D}^{(\alpha)} p_0(s,\cdot,r,z)|(x)|q(r,z,t,y)|\mathrm{d}z\mathrm{d}r \\ &=: I_1(x) + I_2(x) + I_3(x) + I_4(x). \end{aligned}$$

For I_1, by (3.12) we have

$$I_1(x) \lesssim \phi^{(\alpha)}_{0,0}(s,x,t,y).$$

Recall

$$\gamma_0 = (\alpha+\beta-1)\wedge\gamma, \quad \gamma_1 \in (0,\gamma_0).$$

For I_2, by Lemma 3.6, (3.13), (2.19) and (2.8), we have

$$\begin{aligned} I_2(x) &\lesssim \int_s^u (r-s)^{\frac{\gamma_0}{\alpha}-1}\big(\phi^{(\alpha)}_{\gamma_0,0} + \phi^{(\alpha)}_{0,\gamma_0}\big)(r,\theta_{s,r}(x),t,y)\mathrm{d}r \\ &\lesssim \left(\int_s^u (r-s)^{\frac{\gamma_0}{\alpha}-1}\mathrm{d}r\right)\big(\phi^{(\alpha)}_{\gamma_0,0} + \phi^{(\alpha)}_{0,\gamma_0}\big)(s,x,t,y) \\ &\lesssim \big(\phi^{(\alpha)}_{\gamma_0,\gamma_0} + \phi^{(\alpha)}_{0,2\gamma_0}\big)(s,x,t,y). \end{aligned}$$

For I_3, by (3.12), (3.14) and (2.22), we have

$$\begin{aligned} I_3(x) &\lesssim \int_s^u \int_{\mathbb{R}^d} \phi^{(\alpha)}_{\gamma_1,0}(s,x,r,z)\big(\phi^{(\alpha)}_{\gamma_0,-\gamma_1}+\phi^{(\alpha)}_{0,\gamma_0-\gamma_1}\big)(r,z,t,y)\mathrm{d}z\mathrm{d}r \\ &+ \int_s^u \int_{\mathbb{R}^d} \phi^{(\alpha)}_{\gamma_1,0}(s,x,r,z)\mathrm{d}z\big(\phi^{(\alpha)}_{\gamma_0,-\gamma_1}+\phi^{(\alpha)}_{0,\gamma_0-\gamma_1}\big)(s,x,t,y)\mathrm{d}r \\ &\lesssim \big(\phi^{(\alpha)}_{\gamma_0,0}+\phi^{(\alpha)}_{0,\gamma_0}\big)(s,x,t,y). \end{aligned}$$

For I_4, by (3.12), (3.13) and (2.21), we have

$$I_4(x) \lesssim \int_u^t \big(\phi^{(\alpha)}_{0,0}\odot(\phi^{(\alpha)}_{\gamma_0,0}+\phi^{(\alpha)}_{0,\gamma_0})\big)_r(s,x,t,y)\mathrm{d}r \lesssim \phi^{(\alpha)}_{0,0}(s,x,t,y).$$

Combining the above estimates, we complete the proof. □

4. A priori gradient estimates

The aim of this section is to show the following a priori gradient estimate.

Theorem 4.1. *Under* $(\mathbf{H}^\beta_b)$, $(\mathbf{H}^\gamma_a)$ *and* (3.1), *for any* $T>0$, *there is a constant* $C=C(T,\Theta)>0$ *such that for all* $f\in\mathcal{B}_b(\mathbb{R}^d)$, $0\leqslant s<t\leqslant T$ *and* $x\in\mathbb{R}^d$,

$$|\nabla P_{s,t}f(x)| \lesssim_C (t-s)^{-1/\alpha}P_{s,t}|f|(x). \tag{4.1}$$

Moreover, $x\mapsto\nabla P_{s,t}f(x)$ *is equi-continuous in* $(b,a)\in\mathscr{C}$.

We shall prove this theorem for $\alpha\in[1,2)$ and $\alpha\in(0,1)$ separately by different methods.

4.1. Critical and Subcritical cases: $\alpha\in[1,2)$.

In this subsection we start from the series expansion (3.10) for the density to derive the estimate

$$|\nabla_x p(s,x,t,y)| \lesssim_{C_3} \phi^{(\alpha)}_{0,\alpha-1}(s,x,t,y), \tag{4.2}$$

when $(\mathbf{H}^\beta_b)$, $(\mathbf{H}^\gamma_a)$ and (3.1) are in force and $\alpha\in[1,2)$. This precisely gives (4.1). We recall that, with the notations of Section 3:

$$p(s,x,t,y)=p_0(s,x,t,y)+(p_0\otimes q)(s,x,t,y).$$

Therefore, for $u=\frac{s+t}{2}$ and $\xi=\theta_{s,r}(x)$,

$$\begin{aligned} \nabla_x p(s,x,t,y) &= \nabla_x p_0(s,x,t,y)+\int_u^t\int_{\mathbb{R}^d}\nabla_x p_0(s,x,r,z)q(r,z,t,y)\mathrm{d}z\mathrm{d}r \\ &+ \int_s^u\int_{\mathbb{R}^d}(\nabla_x p_0-\nabla_x\tilde{p}^{(r,\xi)})(s,x,r,z)q(r,z,t,y)\mathrm{d}z\mathrm{d}r \\ &+ \int_s^u\int_{\mathbb{R}^d}\nabla_x\tilde{p}^{(r,\xi)}(s,x,r,z)(q(r,z,t,y)-q(r,\xi,t,y))\mathrm{d}z\mathrm{d}r \\ &=: G_1(s,x,t,y)+G_2(s,x,t,y)+G_3(s,x,t,y)+G_4(s,x,t,y), \end{aligned}$$

where for the last term, we have used precisely the cancellation property

$$\int_{\mathbb{R}^d}\nabla_x\tilde{p}^{(r,\xi)}(s,x,r,z)\mathrm{d}z=0.$$

For G_1, by (3.11) we clearly have

$$|G_1(s,x,t,y)| \lesssim \phi^{(\alpha+1)}_{0,\alpha}(s,x,t,y)\leqslant\phi^{(\alpha)}_{0,\alpha-1}(s,x,t,y),$$

using Remark 2.9, equation (2.33), for the last inequality. For G_2, by (3.11), (3.13) and (2.22), we have

$$\begin{aligned}|G_2(s,x,t,y)| &\leqslant \int_u^t \int_{\mathbb{R}^d} \phi_{0,\alpha}^{(\alpha+1)}(s,x,r,z)|q(r,z,t,y)|\mathrm{d}z\mathrm{d}r\\ &\lesssim (t-s)^{-\frac{1}{\alpha}}\phi_{0,\alpha}^{(\alpha)} \otimes \big(\phi_{\gamma_0,0}^{(\alpha)} + \phi_{0,\gamma_0}^{(\alpha)}\big)(s,x,t,y)\\ &\lesssim (t-s)^{-\frac{1}{\alpha}}\phi_{0,\alpha+\gamma_0}^{(\alpha)}(s,x,t,y) = \phi_{0,\alpha+\gamma_0-1}^{(\alpha)}(s,x,t,y).\end{aligned}$$

For G_3, noting that by (3.3),

$$\begin{aligned}\nabla_x p_0(s,x,r,z) &= \nabla_x p_{s,r}^{a^{(r,z)}}\left(x - z + \int_s^r b_{|r'-r|^{1/\alpha}}(r', \theta_{r,r'}(z))\mathrm{d}r'\right),\\ \nabla_x \tilde{p}^{(r,\xi)}(s,x,r,z) &= \nabla_x p_{s,r}^{a^{(r,\xi)}}\left(x - z + \int_s^r b_{|r'-r|^{1/\alpha}}(r', \theta_{r,r'}(\xi))\mathrm{d}r'\right),\end{aligned}$$

by (2.34), (2.29), (1.11) and (2.5), one finds that

$$\begin{aligned}|\nabla_x p_0 - \nabla_x \tilde{p}^{(r,\xi)}|(s,x,r,z) &\lesssim \phi_{0,\alpha}^{(\alpha+1)}(s,x,r,z)(1\wedge|z-\theta_{s,r}(x)|^{\gamma})\\ &\quad + \phi_{0,\alpha}^{(\alpha+2)}(s,x,r,z)\big(|z-\theta_{s,r}(x)|^{\beta} + (r-s)^{\frac{\beta}{\alpha}}\big)(r-s)\\ &\underset{(2.33)}{\lesssim} \big(\phi_{0,\alpha+\gamma-1}^{(\alpha)} + \phi_{0,2\alpha+\beta-2}^{(\alpha)}\big)(s,x,r,z)\\ &\lesssim \phi_{0,\alpha-1+\gamma_0}^{(\alpha)}(s,x,r,z),\end{aligned} \tag{4.3}$$

where $\gamma_0 = \gamma\wedge(\alpha+\beta-1)$. Therefore, due to $\alpha\in[1,2)$, by (2.22),

$$\begin{aligned}|G_3(s,x,t,y)| &\lesssim \phi_{0,\alpha-1+\gamma_0}^{(\alpha)} \otimes \big(\phi_{\gamma_0,0}^{(\alpha)} + \phi_{0,\gamma_0}^{(\alpha)}\big)(s,x,t,y)\\ &\lesssim \phi_{0,\alpha-1+2\gamma_0}^{(\alpha)}(s,x,t,y) \leqslant \phi_{0,\alpha-1}^{(\alpha)}(s,x,t,y).\end{aligned} \tag{4.4}$$

For G_4, by (3.3), (2.29) and (3.14) we have for $\gamma_1\in(0,\gamma_0)$,

$$\begin{aligned}|G_4(s,x,t,y)| &\leqslant \int_s^u \int_{\mathbb{R}^d} |\nabla_x \tilde{p}^{(r,\xi)}(s,x,r,z)|\ |(q(r,z,t,y) - q(r,\xi,t,y))|\mathrm{d}z\mathrm{d}r\\ &\lesssim \int_s^u \mathrm{d}r \int_{\mathbb{R}^d} \mathrm{d}z \phi_{0,\alpha-1}^{(\alpha)}(s,x,r,z)(1\wedge|z-\xi|^{\gamma_1})\\ &\quad\times\Big[\big(\phi_{\gamma_0,-\gamma_1}^{(\alpha)} + \phi_{0,\gamma_0-\gamma_1}^{(\alpha)}\big)(r,z,t,y) + \big(\phi_{\gamma_0,-\gamma_1}^{(\alpha)} + \phi_{0,\gamma_0-\gamma_1}^{(\alpha)}\big)(r,\xi,t,y)\Big].\end{aligned}$$

Since $t-r\asymp t-s$ for $r\in[s,u]$ and $\xi=\theta_{s,r}(x)$, from (2.7) in Lemma 2.2, it holds

$$\big(\phi_{\gamma_0,-\gamma_1}^{(\alpha)} + \phi_{0,\gamma_0-\gamma_1}^{(\alpha)}\big)(r,\xi,t,y) \lesssim \big(\phi_{\gamma_0,-\gamma_1}^{(\alpha)} + \phi_{0,\gamma_0-\gamma_1}^{(\alpha)}\big)(s,x,t,y).$$

Thus by (2.22), we eventually have

$$\begin{aligned}|G_4(s,x,t,y)| &\lesssim \int_s^u \mathrm{d}r \int_{\mathbb{R}^d} \mathrm{d}z \phi_{0,\alpha+\gamma_1-1}^{(\alpha)}(s,x,r,z)\\ &\quad\times\Big[\big(\phi_{\gamma_0,-\gamma_1}^{(\alpha)} + \phi_{0,\gamma_0-\gamma_1}^{(\alpha)}\big)(r,z,t,y) + \big(\phi_{\gamma_0,-\gamma_1}^{(\alpha)} + \phi_{0,\gamma_0-\gamma_1}^{(\alpha)}\big)(s,x,t,y)\Big]\\ &\lesssim \big(\phi_{0,\gamma_0-\gamma_1}^{(\alpha)} + \phi_{\gamma_0,\alpha-1}^{(\alpha)}\big)(s,x,t,y).\end{aligned}$$

Combining the above calculations, we obtain (4.2). Moreover, by the dominated convergence theorem, from the above calculations, it is easy to see that

$$\lim_{x\to x_0}\sup_{(b,a)\in\mathscr{C}} |G_i^{b,a}(s,x,t,y) - G_i^{b,a}(s,x_0,t,y)| = 0,\ \ i=1,2,3,4,$$

where $G_i^{b,a}$ are defined as above through the coefficients b, a. For instance,

$$\lim_{x\to x_0}\sup_{(b,a)\in\mathscr{C}}|G_2^{b,a}(s,x,t,y)-G_2^{b,a}(s,x_0,t,y)|$$
$$\leqslant\int_u^t\lim_{x\to x_0}\sup_{(b,a)\in\mathscr{C}}\int_{\mathbb{R}^d}|\nabla_x p_0(s,x,r,z)-\nabla_x p_0(s,x_0,r,z)|\,|q(r,z,t,y)|\mathrm{d}z\mathrm{d}r,$$

and for each $r\in(u,t)$, by (3.13) and (3.11),

$$\lim_{x\to x_0}\sup_{(b,a)\in\mathscr{C}}\int_{\mathbb{R}^d}|\nabla_x p_0(s,x,r,z)-\nabla_x p_0(s,x_0,r,z)|\,|q(r,z,t,y)|\mathrm{d}z$$
$$\lesssim\int_{\mathbb{R}^d}\lim_{x\to x_0}\sup_{(b,a)\in\mathscr{C}}|\nabla_x p_0(s,x,r,z)-\nabla_x p_0(s,x_0,r,z)|\phi_{0,0}^{(\alpha)}(r,z,t,y)\mathrm{d}z=0.$$

In particular, Theorem 4.1 holds for $\alpha\in[1,2)$.

Remark 4.2. We remark that for $\alpha\in(0,1)$, under $\alpha+\beta>1$, the second inequality in (4.4) may not hold since $\alpha+\gamma_0-1$ may be less than zero. This is also the reason that we have to make a different treatment for supercritical case. Let us mention that this proof anyhow works even in the super-critical case under the most stringent condition $\alpha+\frac{\beta}{2}>1$. Eventually, we also point out that the previous arguments can be simplified if $\alpha\in(1,2)$ for which the full parametrix expansion (3.10) of the density can actually be directly differentiated since the induced singularity in time remains integrable.

4.2. **Supercritical case $\alpha\in(0,1)$.** The following gradient estimate comes in [29].

Theorem 4.3. *(Gradient estimate) Under $(\mathbf{H}_b^\beta)$, $(\mathbf{H}_a^\gamma)$ and (3.1), for any $T>0$, there is a constant $C>0$ such that for all $f\in\mathcal{B}_b(\mathbb{R}^d)$ and $0\leqslant s<t\leqslant T$,*

$$|\nabla P_{s,t}f(x)|\lesssim_C(t-s)^{-1/\alpha}\|f\|_\infty,$$

where the constant C may depend on $\|\nabla b\|_\infty$ and $\|\nabla a\|_\infty$.

Below we fix $s<t$ and $x\in\mathbb{R}^d$ and divide the proof into six steps.

(Step 1). For notational simplicity, we shall write for $r\in[s,t]$,

$$\widetilde{\mathscr{A}_r}:=\mathscr{A}_r^{(s,x)}=\mathscr{K}_r^{(s,x)}+\mathscr{B}_r^{(s,x)}=:\widetilde{\mathscr{K}_r}+\widetilde{\mathscr{B}}_r,$$

and

$$h(s,x,t,y):=\big(\nabla\tilde{p}^{(\tau,\xi)}(s,\cdot,t,y)(x)\big)_{(\tau,\xi)=(s,x)}\overset{(3.3)}{=}-\nabla_y g_{s,t}^{(s,x)}(\theta_{s,t}(x)-y),\tag{4.5}$$

and for a function f,

$$H_{s,t}f(x):=\int_{\mathbb{R}^d}h(s,x,t,y)f(y)\mathrm{d}y.$$

By Lemma 3.3 we have

$$\nabla P_{s,t}f(x)=\nabla\widetilde{P}_{s,t}^{(\tau,\xi)}f(x)+\int_s^t\nabla\widetilde{P}_{s,t}^{(\tau,\xi)}\mathscr{A}_r^{(\tau,\xi)}P_{r,t}f(x)\mathrm{d}r.$$

Taking $(\tau,\xi)=(s,x)$ and using the above notations, we can write

$$\nabla P_{s,t}f(x)=H_{s,t}f(x)+\int_s^t H_{s,r}\widetilde{\mathscr{A}_r}P_{r,t}f(x)\mathrm{d}r=H_{s,t}f(x)+\sum_{i=1}^4 I_{s,t}^{(i)}f(x),\tag{4.6}$$

where for $u:=\frac{s+t}{2}$,

$$I_{s,t}^{(1)}f(x):=\int_s^u H_{s,r}\widetilde{\mathscr{K}_r}P_{r,t}f(x)\mathrm{d}r,$$

$$I^{(2)}_{s,t}f(x) := \int_s^u H_{s,r}\widetilde{\mathscr{B}}_r P_{r,t}f(x)\mathrm{d}r,$$
$$I^{(3)}_{s,t}f(x) := \int_u^t H_{s,r}\widetilde{\mathscr{K}_r} P_{r,t}f(x)\mathrm{d}r,$$
$$I^{(4)}_{s,t}f(x) := \int_u^t H_{s,r}\widetilde{\mathscr{B}}_r P_{r,t}f(x)\mathrm{d}r.$$

(Step 2). Note that for $j\in\mathbb{N}$,

$$|\nabla_y^j h(s,x,t,y)| \overset{(4.5)}{=} |\nabla_y^{j+1} g^{(s,x)}_{s,t}(\theta_{s,t}(x)-y)| \overset{(2.29)}{\lesssim} \phi^{(\alpha+j+1)}_{0,\alpha}(s,x,t,y). \tag{4.7}$$

Thus we have

$$\begin{aligned}|H_{s,t}f(x)| &\lesssim \int_{\mathbb{R}^d}\phi^{(\alpha+1)}_{0,\alpha}(s,x,t,y)|f(y)|\mathrm{d}y\\ &\leqslant (t-s)^{-\frac{1}{\alpha}}\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,\alpha}(s,x,t,y)|f(y)|\mathrm{d}y.\end{aligned}$$

For $I^{(1)}_{s,t}f(x)$, noting that by Lemma 3.7,

$$\begin{aligned}|\widetilde{\mathscr{K}_r}P_{r,t}f(z)| &\lesssim (1\wedge|z-\theta_{s,r}(x)|^\gamma)|\mathcal{D}^{(\alpha)}P_{r,t}f|(z)\\ &\lesssim (1\wedge|z-\theta_{s,r}(x)|^\gamma)\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,0}(r,z,t,y)|f(y)|\mathrm{d}y,\end{aligned}$$

and using (2.33) and Lemma 2.7, we have

$$\begin{aligned}|I^{(1)}_{s,t}f(x)| &\overset{(4.7)}{\lesssim} \int_s^u\int_{\mathbb{R}^d}\phi^{(\alpha+1)}_{\gamma,\alpha}(s,x,r,z)\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,0}(r,z,t,y)|f(y)|\mathrm{d}y\mathrm{d}z\mathrm{d}r\\ &\overset{(2.33)}{\lesssim} \int_s^u\int_{\mathbb{R}^d}\big(\phi^{(\alpha)}_{0,\alpha+\gamma-1}\odot\phi^{(\alpha)}_{0,0}\big)_r(s,x,t,y)|f(y)|\mathrm{d}y\mathrm{d}r\\ &\overset{(2.21)}{\lesssim} \int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,\alpha+\gamma-1}(s,x,t,y)|f(y)|\mathrm{d}y.\end{aligned}$$

For $I^{(2)}_{s,t}f(x)$, noting that

$$|\widetilde{\mathscr{B}}_r P_{r,t}f(z)| \lesssim (|\theta_{s,r}(x)-z|^\beta + |\theta_{s,r}(x)-z| + (r-s)^{\beta/\alpha})|\nabla P_{r,t}f(z)|,$$

using (4.7) and (2.33), we have

$$\begin{aligned}|I^{(2)}_{s,t}f(x)| &\lesssim \int_s^u\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,\alpha+\beta-1}(s,x,r,z)|\nabla P_{r,t}f(z)|\mathrm{d}z\mathrm{d}r\\ &\lesssim (t-s)^{-\frac{1}{\alpha}}\int_s^t\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,\alpha+\beta-1}(s,x,r,z)(t-r)^{\frac{1}{\alpha}}|\nabla P_{r,t}f(z)|\mathrm{d}z\mathrm{d}r.\end{aligned}$$

(Step 3). In this step we treat the hard term $I^{(3)}_{s,t}f(x)$. Let $\varepsilon := (t-r)^{1/\alpha}$ and

$$\kappa_\varepsilon(r,z,z') := \kappa(r,\cdot,z')*\rho_\varepsilon(z),\quad \bar\kappa_\varepsilon(r,z,z') := \kappa_\varepsilon(r,z,z') - \kappa_\varepsilon(r,\theta_{r,s}(x),z')$$

and

$$\widetilde{\mathscr{K}^{(\varepsilon)}_r}f(z) = 2\int_{\mathbb{R}^d}\delta^{(1)}_f(z;z')\frac{\bar\kappa_\varepsilon(r,z,z')}{|z'|^{d+\alpha}}\mathrm{d}z'.$$

Let us write

$$\begin{aligned}I^{(3)}_{s,t}f(x) &= \int_u^t\Big(H_{s,r}(\widetilde{\mathscr{K}_r}-\widetilde{\mathscr{K}^{(\varepsilon)}_r})P_{r,t}f(x) + H_{s,r}\widetilde{\mathscr{K}^{(\varepsilon)}_r}P_{r,t}f(x)\Big)\mathrm{d}r\\ &=: \int_u^t\Big(J^{(\varepsilon)}_{1,r}(s,x,t) + J^{(\varepsilon)}_{2,r}(s,x,t)\Big)\mathrm{d}r.\end{aligned}$$

Let $\gamma_1 \in (0,\gamma)$. Noting that

$$|(\kappa-\kappa_\varepsilon)(r,z,z') - (\kappa-\kappa_\varepsilon)(r,\theta_{r,s}(x),z')| \lesssim_C (|z-\theta_{r,s}(x)|^{\gamma_1}\wedge 1)\varepsilon^{\gamma-\gamma_1},$$

by definition and Lemma 3.7, we have

$$\begin{aligned}|(\widetilde{\mathscr{K}_r}-\widetilde{\mathscr{K}_r^{(\varepsilon)}})P_{r,t}f(z)| &\lesssim (|z-\theta_{r,s}(x)|^{\gamma_1}\wedge 1)\varepsilon^{\gamma-\gamma_1}|\mathcal{D}^{(\alpha)}P_{r,t}f|(z)\\ &\lesssim (|z-\theta_{r,s}(x)|^{\gamma_1}\wedge 1)\varepsilon^{\gamma-\gamma_1}\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,0}(r,z,t,y)|f(y)|\mathrm{d}y.\end{aligned}$$

For $J^{(\varepsilon)}_{1,r}$, recalling $\varepsilon=(t-r)^{1/\alpha}$, we have

$$\begin{aligned}\int_u^t |J^{(\varepsilon)}_{1,r}(s,x,t)|\mathrm{d}r &\overset{(4.7)}{\lesssim} \int_u^t\int_{\mathbb{R}^d}\big(\phi^{(\alpha)}_{0,\alpha+\gamma_1-1}\odot\phi^{(\alpha)}_{0,\gamma-\gamma_1}\big)_r(s,x,t,y)|f(y)|\mathrm{d}y\mathrm{d}r\\ &\overset{(2.21)}{\lesssim}\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,\alpha+\gamma-1}(s,x,t,y)|f(y)|\mathrm{d}y.\end{aligned}$$

For $J^{(\varepsilon)}_{2,r}$, by the change of variables and Fubini's theorem, we have

$$\begin{aligned}J^{(\varepsilon)}_{2,r}(s,x,t) &= \int_{\mathbb{R}^d}h(s,x,r,z)\int_{\mathbb{R}^d}\delta^{(1)}_{P_{r,t}f}(z;z')\frac{\bar\kappa_\varepsilon(r,z,z')}{|z'|^{d+\alpha}}\mathrm{d}z'\mathrm{d}z\\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\delta^{(1)}_{h(s,x,r,\cdot)\bar\kappa_\varepsilon(r,\cdot,z')}(z;z')\frac{\mathrm{d}z'}{|z'|^{d+\alpha}}P_{r,t}f(z)\mathrm{d}z\\ &= \int_{\mathbb{R}^d}h(s,x,r,z)\int_{\mathbb{R}^d}\delta^{(1)}_{\bar\kappa_\varepsilon(r,\cdot,z')}(z;z')\frac{\mathrm{d}z'}{|z'|^{d+\alpha}}P_{r,t}f(z)\mathrm{d}z\\ &+\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\delta^{(1)}_{h(s,x,r,\cdot)}(z;z')\bar\kappa_\varepsilon(r,z+z',z')\frac{\mathrm{d}z'}{|z'|^{d+\alpha}}P_{r,t}f(z)\mathrm{d}z.\end{aligned}$$

Noting that by $\mathbf{(H}^\gamma_a\mathbf{)}$,

$$|\delta^{(1)}_{\bar\kappa_\varepsilon(r,\cdot,z')}(z;z')|\lesssim(\varepsilon^{\gamma-1}|z'|)\wedge|z'|^\gamma\wedge 1,$$

we have

$$\int_{\mathbb{R}^d}|\delta^{(1)}_{\bar\kappa_\varepsilon(r,\cdot,z')}(z;z')|\frac{\mathrm{d}z'}{|z'|^{d+\alpha}}\lesssim\int_{\mathbb{R}^d}((\varepsilon^{\gamma-1}|z'|)\wedge|z'|^\gamma\wedge 1)\frac{\mathrm{d}z'}{|z'|^{d+\alpha}}\lesssim\varepsilon^{(\gamma-\alpha)\wedge 0}.$$

On the other hand, by (4.5) and (2.29),

$$|\delta^{(1)}_{h(s,x,r,\cdot)}(z;z')|\lesssim\big(((r-s)^{-\frac{1}{\alpha}}|z'|)\wedge 1\big)\big(\phi^{(\alpha)}_{0,\alpha-1}(s,x,r,z+z')+\phi^{(\alpha)}_{0,\alpha-1}(s,x,r,z)\big).$$

Thus, as in (2.38) we have

$$\begin{aligned}\int_{\mathbb{R}^d}\frac{|\delta^{(1)}_{h(s,x,r,\cdot)}(z;z')|\mathrm{d}z'}{|z'|^{d+\alpha}} &\lesssim\int_{\mathbb{R}^d}\big(((r-s)^{-\frac{1}{\alpha}}|z'|)\wedge 1\big)\phi^{(\alpha)}_{0,\alpha-1}(s,x,r,z+z')\frac{\mathrm{d}z'}{|z'|^{d+\alpha}}\\ &+\phi^{(\alpha)}_{0,\alpha-1}(s,x,r,z)\int_{\mathbb{R}^d}\big(((r-s)^{-\frac{1}{\alpha}}|z'|)\wedge 1\big)\frac{\mathrm{d}z'}{|z'|^{d+\alpha}}\\ &\lesssim\phi^{(\alpha)}_{0,\alpha-1}(s,x,r,z)(r-s)^{-1}=\phi^{(\alpha)}_{0,-1}(s,x,r,z).\end{aligned}$$

Therefore,

$$|J^{(\varepsilon)}_{2,r}(s,x,t)|\lesssim\int_{\mathbb{R}^d}\Big[\varepsilon^{(\gamma-\alpha)\wedge 0}\phi^{(\alpha)}_{0,\alpha-1}(s,x,r,z)+\phi^{(\alpha)}_{0,-1}(s,x,r,z)\Big]P_{r,t}|f|(z)\mathrm{d}z.$$

Recall $\varepsilon=(t-r)^{\frac{1}{\alpha}}$. By (2.33), we obtain

$$\int_u^t|J^{(\varepsilon)}_{2,r}(s,x,t)|\mathrm{d}r\lesssim(t-s)^{-\frac{1}{\alpha}}\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,\alpha}(s,x,t,y)|f(y)|\mathrm{d}y.$$

(Step 4). For $\varepsilon=(t-r)^{1/\alpha}$, we define

$$\bar{b}_\varepsilon(r,z):=(b*\rho_\varepsilon)(r,z)-(b*\rho_\varepsilon*\rho_{|r-s|^{1/\alpha}})(r,\theta_{r,s}(x))$$

and

$$\widetilde{\mathscr{B}}_r^{(\varepsilon)}f(z):=\bar{b}_\varepsilon(r,z)\cdot\nabla f(z).$$

For $I_{s,t}^{(4)}$, we similarly write

$$\begin{aligned}I_{s,t}^{(4)}f(x)&=\int_u^t\Big(H_{s,r}(\widetilde{\mathscr{B}}_r-\widetilde{\mathscr{B}}_r^{(\varepsilon)})P_{r,t}f(x)+H_{s,r}\widetilde{\mathscr{B}}_r^{(\varepsilon)}P_{r,t}f(x)\Big)\mathrm{d}r\\&=:\int_u^t\Big(J_{3,r}^{(\varepsilon)}(s,x,t)+J_{4,r}^{(\varepsilon)}(s,x,t)\Big)\mathrm{d}r.\end{aligned}$$

For $J_{3,r}^{(\varepsilon)}$, since

$$|\bar{b}_0-\bar{b}_\varepsilon|(r,z)\leqslant\kappa_0\varepsilon^\beta=\kappa_0(t-s)^{\beta/\alpha},$$

by (4.7) we have

$$\begin{aligned}|J_{3,r}^{(\varepsilon)}(s,x,t)|&=\left|\int_{\mathbb{R}^d}h(s,x,r,z)(\bar{b}_0(r,z)-\bar{b}_\varepsilon(r,z))\cdot\nabla P_{r,t}f(z)\mathrm{d}z\right|\\&\lesssim\int_{\mathbb{R}^d}\phi_{0,\alpha}^{(\alpha+1)}(s,x,r,z)(t-r)^{\frac{\beta}{\alpha}}|\nabla P_{r,t}f(z)|\mathrm{d}z,\end{aligned}$$

and

$$\int_u^t|J_{3,r}^{(\varepsilon)}(s,x,t)|\mathrm{d}r\lesssim(t-s)^{-\frac{1}{\alpha}}\int_s^t\int_{\mathbb{R}^d}\phi_{0,\alpha}^{(\alpha)}(s,x,r,z)(t-r)^{\frac{\beta}{\alpha}}|\nabla P_{r,t}f(z)|\mathrm{d}z\mathrm{d}r.$$

For $J_{4,r}^{(\varepsilon)}$, we derive integrating by parts that

$$\begin{aligned}|J_{4,r}^{(\varepsilon)}(s,x,t)|&=\left|\int_{\mathbb{R}^d}h(s,x,r,z)\,\bar{b}_\varepsilon(r,z)\cdot\nabla_zP_{r,t}f(z)\mathrm{d}z\right|\\&\leqslant\left|\int_{\mathbb{R}^d}h(s,x,r,z)\,\mathrm{div}\bar{b}_\varepsilon(r,z)P_{r,t}f(z)\mathrm{d}z\right|\\&\quad+\left|\int_{\mathbb{R}^d}\bar{b}_\varepsilon(r,z)\cdot\nabla_zh(s,x,r,z)P_{r,t}f(z)\mathrm{d}z\right|.\end{aligned}$$

Since

$$|\mathrm{div}\bar{b}_\varepsilon(r,z)|=|\mathrm{div}b_\varepsilon(r,z)|\leqslant\kappa_0\varepsilon^{\beta-1}=\kappa_0(t-r)^{(\beta-1)/\alpha}$$

and

$$|\bar{b}_\varepsilon|(r,z)\lesssim|z-\theta_{r,s}(x)|^\beta+(r-s)^{\beta/\alpha},$$

by (4.7) and (2.33) we have

$$\begin{aligned}|J_{4,r}^{(\varepsilon)}(s,x,t)|&\lesssim\int_{\mathbb{R}^d}\phi_{0,\alpha-1}^{(\alpha)}(s,x,r,z)(t-r)^{(\beta-1)/\alpha}|P_{r,t}f(z)|\mathrm{d}z\\&\quad+\int_{\mathbb{R}^d}\phi_{0,\alpha+\beta-2}^{(\alpha)}(s,x,r,z)|P_{r,t}f(z)|\mathrm{d}z.\end{aligned}$$

Thus,

$$\begin{aligned}\int_u^t|J_{4,r}^{(\varepsilon)}(s,x,t)|\mathrm{d}r&\lesssim\int_u^t\int_{\mathbb{R}^d}\big(\phi_{0,\alpha-1}^{(\alpha)}\odot\phi_{0,\alpha+\beta-1}^{(\alpha)}\big)_r(s,x,t,y)|f(y)|\mathrm{d}y\mathrm{d}r\\&\quad+\int_u^t\int_{\mathbb{R}^d}\big(\phi_{0,\alpha+\beta-2}^{(\alpha)}\odot\phi_{0,\alpha}^{(\alpha)}\big)_r(s,x,t,y)|f(y)|\mathrm{d}y\mathrm{d}r\\&\overset{(2.21)}{\lesssim}\int_u^t\big[(r-s)^{\frac{\alpha-1}{\alpha}}(t-r)^{\frac{\alpha+\beta-1}{\alpha}}+(r-s)^{\frac{\alpha+\beta-2}{\alpha}}(t-r)\big]\end{aligned}$$

$$\times \big[(r-s)^{-1}+(t-r)^{-1}\big]\mathrm{d}r\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,0}(s,x,t,y)|f(y)|\mathrm{d}y$$
$$\lesssim \int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,2\alpha+\beta-2}(s,x,t,y)|f(y)|\mathrm{d}y$$
$$\lesssim (t-s)^{-\frac{1}{\alpha}}\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,\alpha}(s,x,t,y)|f(y)|\mathrm{d}y,$$

recalling that $\alpha+\beta>1$ for the last inequality. Hence,

$$|I^{(4)}_{s,t}f(x)|\lesssim (t-s)^{-\frac{1}{\alpha}}\Big(\int_s^t\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,\alpha}(s,x,r,z)(t-r)^{\frac{\beta}{\alpha}}|\nabla P_{r,t}f(z)|\mathrm{d}z\mathrm{d}r$$
$$+\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,\alpha}(s,x,t,y)|f(y)|\mathrm{d}y\Big).$$

(Step 5). Combining the above calculations, we obtain

$$|\nabla P_{s,t}f(x)|\lesssim (t-s)^{-\frac{1}{\alpha}}\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,\alpha}(s,x,t,y)|f(y)|\mathrm{d}y$$
$$+(t-s)^{-\frac{1}{\alpha}}\int_s^t\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,\alpha}(s,x,r,z)(t-r)^{\frac{\beta}{\alpha}}|\nabla P_{r,t}f(z)|\mathrm{d}z\mathrm{d}r$$
$$+(t-s)^{-\frac{1}{\alpha}}\int_s^t\int_{\mathbb{R}^d}\phi^{(\alpha)}_{0,\alpha+\beta-1}(s,x,r,z)(t-r)^{\frac{1}{\alpha}}|\nabla P_{r,t}f(z)|\mathrm{d}z\mathrm{d}r.$$

By the lower bound estimate, we further have

$$(t-s)^{\frac{1}{\alpha}}|\nabla P_{s,t}f(x)|\lesssim P_{s,t}|f|(x)+\int_s^t(t-r)^{\frac{\beta}{\alpha}}P_{s,r}|\nabla P_{r,t}f|(x)\mathrm{d}r$$
$$+\int_s^t(r-s)^{\frac{\beta-1}{\alpha}}(t-r)^{\frac{1}{\alpha}}P_{s,r}|\nabla P_{r,t}f|(x)\mathrm{d}r. \tag{4.8}$$

For fixed $0\leqslant u<t\leqslant T$ and $s\in(u,t)$, we let

$$\Gamma^t_u(s,x):=(t-s)^{\frac{1}{\alpha}}P_{u,s}|\nabla P_{s,t}f|(x).$$

Using $P_{u,s}$ act on both sides of (4.8) and by $P_{u,s}P_{s,r}=P_{u,r}$, we derive that

$$\Gamma^t_u(s,x)\lesssim P_{u,t}|f|(x)+\int_s^t\Big[(r-s)^{\frac{\beta-1}{\alpha}}+(t-r)^{\frac{\beta-1}{\alpha}}\Big]\Gamma^t_u(r,x)\mathrm{d}r.$$

Note that by Theorem 4.3,

$$\sup_{s\in[u,t]}\|\Gamma^t_u(s,\cdot)\|_\infty<\infty.$$

Since $\alpha+\beta>1$, from the Volterra-Gronwall inequality, we obtain that for all $s\in(u,t)$,

$$\Gamma^t_u(s,x)\lesssim P_{u,t}|f|(x).$$

Taking limit $u\uparrow s$, we obtain

$$(t-s)^{\frac{1}{\alpha}}|\nabla P_{s,t}f|(x)\lesssim P_{s,t}|f|(x),$$

which eventually yields the desired gradient estimate.

(Step 6). Finally, by (4.6) and the dominated convergence theorem, one can show that

$$\lim_{x\to x_0}\sup_{(b,a)\in\mathscr{C}}|\nabla P^{b,a}_{s,t}f(x)-\nabla P^{b,a}_{s,t}f(x_0)|=0.$$

Indeed, from the above proof, it suffices to show that

$$\lim_{x\to x_0}\sup_{(b,a)\in\mathscr{C}}|H^{b,a}_{s,t}f(x)-H^{b,a}_{s,t}f(x_0)|=0.$$

This follows by Lemma 2.8.

5. Proof of Theorem 1.1

The point is here to prove Theorem 1.1. Namely, we want to extend the bounds of Theorem 3.5 and Lemmas 3.7 and (4.2) under the sole assumptions $\mathbf{(H}_a^\gamma\mathbf{)}$, $\mathbf{(H}_b^\beta\mathbf{)}$. Let a_ε and b_ε be the smooth approximations of a and b, respectively. Hence, assumptions $\mathbf{(H}_a^\gamma\mathbf{)}$, $\mathbf{(H}_b^\beta\mathbf{)}$ and (3.1) are met by $a_\varepsilon, b_\varepsilon$ for the SDE

$$\mathrm{d}X_t^\varepsilon = b_\varepsilon(t, X_t^\varepsilon)\mathrm{d}t + a_\varepsilon(t, X_{t-}^\varepsilon)\mathrm{d}L_t^{(\alpha)}. \tag{5.1}$$

The following convergence in law result was established in [7], see Theorem 1.1 therein.

Theorem 5.1. *Let $X_{s,t}^\varepsilon(x)$ be the unique solution of SDE (5.1). Then $X_{s,t}^\varepsilon(x)$ weakly converges to $X_{s,t}(x)$.*

Proof. For fixed $(s,x)\in\mathbb{R}_+\times\mathbb{R}^d$, since the coefficients b, a are linear growth, it is by now standard to show that the law of $X_{s,\cdot}^\varepsilon(x)$ is tight in the space of all cádlág functions. By a standard way, one can show that any weak accumulation point of the law of $X_{s,\cdot}^\varepsilon(x)$ is a weak solution of SDE (1.1). Finally, by the weak uniqueness, one sees that $X_{s,t}^\varepsilon(x)$ weakly converges to $X_{s,t}(x)$. □

Denoting by p_ε the associated density, it therefore holds from Theorem 3.5, Lemma 3.7 and Theorem 4.1 that

(i) **(Two-sides estimate)** For any $T>0$, there is a constant $C_1=C_1(T,\Theta)>0$ such that for all $0\leqslant s<t\leqslant T$ and $x,y\in\mathbb{R}^d$,

$$p_\varepsilon(s,x,t,y)\asymp_{C_1}\phi_{0,\alpha}^{(\alpha)}(s,x,t,y). \tag{5.2}$$

(ii) **(Fractional derivative estimate)** For any $T>0$, there is a constant $C_2=C_2(T,\Theta)>0$ such that for all $0\leqslant s<t\leqslant T$ and $x,y\in\mathbb{R}^d$,

$$|\mathcal{D}^{(\alpha)}p_\varepsilon(s,\cdot,t,y)|(x)\lesssim_{C_2}\phi_{0,0}^{(\alpha)}(s,x,t,y). \tag{5.3}$$

(iii) **(Gradient estimate in x)** For any $T>0$, there is a constant $C_3=C_3(T,\Theta)>0$ such that for all $0\leqslant s<t\leqslant T$ and $x,y\in\mathbb{R}^d$,

$$|\nabla P_{s,t}^\varepsilon f(x)|\lesssim_{C_3}(t-s)^{-1/\alpha}P_{s,t}^\varepsilon|f|(x). \tag{5.4}$$

where the constants in the above controls only depend on $\mathbf{(H}_a^\gamma\mathbf{)}$, $\mathbf{(H}_b^\beta\mathbf{)}$ through Θ (see precisely (1.13)).

By Theorem 5.1, we have for any $f\in C_b(\mathbb{R}^d)$,

$$\lim_{\varepsilon\to0}P_{s,t}^\varepsilon f(x):=\lim_{\varepsilon\to0}\mathbb{E}f(X_{s,t}^\varepsilon(x))=\mathbb{E}f(X_{s,t}(x))=:P_{s,t}f(x). \tag{5.5}$$

(i) (Two-sided estimates) For nonnegative measurable functions f, we get from (5.2)

$$C_1^{-1}\int_{\mathbb{R}^d}\phi_{0,\alpha}^{(\alpha)}(s,x,t,y)f(y)\mathrm{d}y\leqslant\mathbb{E}f(X_{s,t}(x))\leqslant C_1\int_{\mathbb{R}^d}\phi_{0,\alpha}^{(\alpha)}(s,x,t,y)f(y)\mathrm{d}y,$$

which implies that $X_{s,t}(x)$ has a density $p(s,x,t,y)$ having lower and upper bound as in (1.15). On the other hand, for fixed $s<t$, by Theorem 3.5 we have

$$(x,y)\mapsto p_\varepsilon(s,x,t,y)\text{ is equi-continuous in }\varepsilon\in(0,1).$$

From the Ascoli-Arzelà theorem, there are a subsequence ε_k and a continuous function $\bar{p}(s,x,t,y)$ as a function of $x,y\in\mathbb{R}^d$ such that

$$p_{\varepsilon_k}(s,x,t,y)\to\bar{p}(s,x,t,y)\text{ locally uniformly in }x,y\in\mathbb{R}^d, \tag{5.6}$$

which together with (5.5) yields that

$$p(s,x,t,y) = \bar{p}(s,x,t,y) \text{ is continuous as a function of } x,y \in \mathbb{R}^d. \tag{5.7}$$

(ii) (Fractional derivative estimates) It follows by (5.3), (5.6), (5.7) and Fatou's lemma that

$$\begin{aligned}|\mathcal{D}^{(\alpha)}p(s,\cdot,t,y)|(x) &= \int_{\mathbb{R}^d} \lim_{k\to\infty} |\delta^{(2)}_{p_{\varepsilon_k}(s,\cdot,t,y)}(x;z)| \frac{\mathrm{d}z}{|z|^{d+\alpha}} \\ &\leqslant \varliminf_{k\to\infty} \int_{\mathbb{R}^d} |\delta^{(2)}_{p_{\varepsilon_k}(s,\cdot,t,y)}(x;z)| \frac{\mathrm{d}z}{|z|^{d+\alpha}} \\ &= \varliminf_{k\to\infty} |\mathcal{D}^{(\alpha)}p_{\varepsilon_k}(s,\cdot,t,y)|(x) \lesssim_{C_2} \phi^{(\alpha)}_{0,0}(s,x,t,y).\end{aligned}$$

(iii) (Gradient estimates) For fixed $f \in C_b(\mathbb{R}^d)$, by (5.4),

$$x \mapsto \nabla P^{\varepsilon}_{s,t} f(x) \text{ is equi-continuous in } \varepsilon,$$

which together with (5.5) implies that $x \mapsto P_{s,t}(x)$ is continuous differentiable. By taking limits along a subsequence ε_k for (5.4), we obtain

$$|\nabla P_{s,t} f(x)| \lesssim_{C_3} (t-s)^{-1/\alpha} P_{s,t}|f|(x).$$

Finally, for fixed $t' > t$ and $y \in \mathbb{R}^d$, we let $f(x) := p(t,x,t',y)$, then by the Chapman-Kolmogorov equation, we obtain

$$|\nabla p(s,\cdot,t',y)(x)| \lesssim_{C_3} (t-s)^{-1/\alpha} p(s,x,t',y).$$

This then readily gives estimate (4.2) (logarithmic derivative) of Theorem 1.1.

References

[1] R. F. Bass and Z.-Q. Chen, Systems of equations driven by stable processes. *Probab. Theory Relat. Fields*, 134 (2006), 175-214.

[2] Bendikov A., Asymptotic formulas for symmetric stable semigroups. *Expositiones Mathematicae*, 13(1994), 381-384.

[3] R. M. Blumenthal and R. K. Getoor, Some theorems on stable processes. *Trans. Amer. Math. Soc.* **95** (1960), 263-273.

[4] K. Bogdan and T. Jakubowski. Estimates of heat kernel of fractional Laplacian perturbed by gradient operators. *Comm. Math. Phys.*, 271(1):179–198, 2007.

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

[6] Z.-Q. Chen E. Hu, L. Xie and X. Zhang, Heat kernels for non-symmetric diffusions operators with jumps. *J. Differential Equations* **263** (2017), 6576-6634.

[7] Z.-Q. Chen Z. Hao and X. Zhang. Hölder regularity and gradient estimates for SDEs driven by cylindrical α-stable processes. *Electron. J. Probab.* 25 (2020), article no. 137, 1-23.

[8] Z.-Q. Chen and X. Zhang. Heat kernels and analyticity of non-symmetric jump diffusion semigroups. *Probab. Theory Related Fields*, 165(1-2):267–312, 2016.

[9] Z.-Q. Chen and X. Zhang. Heat kernels for time-dependent non-symmetric stable-like operators *J. Math. Anal. Appl.* , 465 (2018) 1-21(2019).

[10] Z.-Q. Chen and X. Zhang. Uniqueness of stable-like processes. arXiv preprint arXiv:1604.02681(2016).

[11] Z.-Q. Chen and X. Zhang. Heat kernels for time-dependent non-symmetric mixed Lévy-type operators. arXiv preprint arXiv:1604.02681(2020).

[12] Z.-Q. Chen, X. Zhang and G. Zhao, Supercritical SDEs driven by multiplicative stable-like Lévy processes. To appear in *Trans. Amer. Math. Soc.*, (2021+).

[13] A. Debussche and N. Fournier. Existence of densities for stable-like driven SDE's with Hölder continuous coefficients, *J. Funct. Anal.*, **264**, No. 4, 1757–1778(2013).

[14] F. Delarue and S. Menozzi. Density estimates for a random noise propagating through a chain of differential equations. *Journal of Functional Analysis*, 259–6:1577–1630, 2010.

[15] K. Du and X. Zhang. Optimal gradient estimates of heat kernels of stabe-like operators. *Proc. of Amer. Math. Soci.*. Vol. 147, No 8, 3559-3565(2019).
[16] A. Friedman. Partial differential equations of parabolic type. Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.
[17] L. Huang. Density estimates for SDEs driven by tempered stable processes. *arXiv:1504.04183*, 2015.
[18] V. Knopova and A.M. Kulik. Parametrix construction of the transition probability density of the solution to an SDE driven by α-stable noise, *A*nn. Inst. Henri Poincaré Probab. Stat., 54(1), 100–140 (2018).
[19] V. Knopova, A. Kulik, R. Schilling. Construction and heat kernel estimates of general stable-like Markov processes. *arXiv:2005.08491*, (2020).
[20] V. N. Kolokoltsov. Symmetric Stable Laws and Stable-Like Jump-Diffusions. *Proceedings of the London Mathematical Society*, 80(3):725–768, May 2000.
[21] A. M. Kulik. On weak uniqueness and distributional properties of a solution to an SDE with α-stable noise. *Stochastic Process. Appl.*, 129(2):473–506, 2019.
[22] A. Kulik, S. Peszat, E. Priola.Gradient formula for transition semigroup corresponding to stochastic equation driven by a system of independent Lévy processes.*a*rXiv:2006.09133 (2020).
[23] R. Léandre. Régularité de processus de sauts dégénérés. *Ann. Inst. H. Poincaré Probab. Statist.*, 21(2), 125–146, 1985.
[24] W. Liu, R. Song and L. Xie, Gradient estimates for the fundamental solution of Lévy type operator. *Adv. Nonlinear Anal.* 9 (2020), 1453-1462.
[25] S. Menozzi, A. Pesce ad X. Zhang. Density and gradient estimates for non degenerate Brownian SDEs with unbounded measurable drift, *J. Differential Equations*, **272**, 330–369(2021)
[26] R. Mikulevicius and H. Pragarauskas. On the Cauchy problem for integro-differential operators in Hölder classes and the uniqueness of the martingale problem. *Potential Anal.*, 40(4):539–563, 2014.
[27] K. Sato: Lévy processes and infinitely divisible distributions. Cambridge University Press, London 1999.
[28] H. Tanaka, M. Tsuchiya, and S. Watanabe. Perturbation of drift-type for Lévy processes. *J. Math. Kyoto Univ.*, 14:73–92, 1974.
[29] F.Y. Wang, L. Xu and X. Zhang: Gradient estimates for SDEs driven by multiplicative Lévy noise. *Journal of Functional Analysis*, 269 (2015) 3195-3219.
[30] F.Y. Wang, X. Zhang. Heat kernel for fractional diffusion operators with perturbations. *Forum Math.* 27 (2015) 973-994.

Stephane Menozzi: Laboratoire de Modélisation Mathématique d'Evry (LaMME), UMR CNRS 8071, Université d'Evry Val d'Essonne (Université Paris Saclay), 23 Boulevard de France 91037 Evry, France and Laboratory of Stochastic Analysis, HSE, Pokrovsky Blvd, 11, Moscow, Russian Federation, Email: stephane.menozzi@univ-evry.fr

Xicheng Zhang: School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, P.R.China, Email: XichengZhang@gmail.com