HEAT KERNELS FOR NON-SYMMETRIC DIFFUSION OPERATORS WITH JUMPS

ZHEN-QING CHEN, ERYAN HU, LONGJIE XIE AND XICHENG ZHANG

Abstract. For $d \geqslant 2$, we prove the existence and uniqueness of heat kernels to the following time-dependent second order diffusion operator with jumps:

$$\mathscr{L}_t := \frac{1}{2}\sum_{i,j=1}^d a_{ij}(t,x)\partial^2_{ij} + \sum_{i=1}^d b_i(t,x)\partial_i + \mathscr{L}^\kappa_t,$$

where $a = (a_{ij})$ is a uniformly bounded, elliptic, and Hölder continuous matrix-valued function, b belongs to some suitable Kato's class, and $\mathscr{L}^\kappa_t$ is a non-local α-stable-type operator with bounded kernel κ. Moreover, we establish sharp two-sided estimates, gradient estimate and fractional derivative estimate for the heat kernel under some mild conditions.

AMS 2010 Mathematics Subject Classification: Primary 35K05, 60J35, 47G20; Secondary 47D07

Keywords and Phrases: Heat kernel, transition density, non-local operator, Kato class, Lévy system, gradient estimate

1. Introduction

Let $C_0(\mathbb{R}^d)$ be the Banach space of all continuous functions on $\mathbb{R}^d$ vanishing at infinity equipped with uniform norm, and $C_c(\mathbb{R}^d)$ the space of all continuous functions on $\mathbb{R}^d$ with compact support. Let $\mathscr{L}$ be a linear operator on $C_0(\mathbb{R}^d)$ with domain Dom($\mathscr{L}$). Suppose that $C^\infty_c(\mathbb{R}^d) \subset \mathrm{Dom}(\mathscr{L})$. We say $\mathscr{L}$ satisfies a positive maximum principle if for all $f \in C^\infty_c(\mathbb{R}^d)$ reaching a positive maximum at point $x_0 \in \mathbb{R}^d$, then $\mathscr{L} f(x_0) \leqslant 0$. The well-known Courrège theorem states that $\mathscr{L}$ satisfies the positive maximum principle if and only if $\mathscr{L}$ takes the following form

$$\begin{aligned}\mathscr{L} f(x) = \frac{1}{2}\sum_{i,j=1}^d a_{ij}(x)\partial^2_{ij} f(x) + \sum_{i=1}^d b_i(x)\partial_i f(x) + c(x)f(x)\\ + \int_{\mathbb{R}^d}\big(f(x+z) - f(x) - 1_{\{|z|\leqslant 1\}} z\cdot\nabla f(x)\big)\,\mu_x(\mathrm{d}z),\end{aligned} \tag{1.1}$$

where $a = (a_{ij}(x))_{1\leqslant i,j\leqslant d}$ is a $d\times d$-symmetric positive definite matrix-valued measurable function on $\mathbb{R}^d$, $b(x) : \mathbb{R}^d \to \mathbb{R}^d$, $c : \mathbb{R}^d \to (-\infty, 0]$ are measurable functions and $\mu_x(\mathrm{d}z)$ is a family of Lévy measures, with that a, b, c, μ enjoy some continuity with respect to x (see [21]). On the other hand, from the probabilistic viewpoint, consider the following SDE with jumps:

$$\begin{aligned}\mathrm{d}X_t = \sigma(X_t)\mathrm{d}W_t + b(X_t)\mathrm{d}t + \int_{|z|\leqslant 1} g(X_{t-}, z)\tilde{N}(\mathrm{d}t, \mathrm{d}z)\\ + \int_{|z|>1} g(X_{t-}, z)N(\mathrm{d}t, \mathrm{d}z), \quad X_0 = x,\end{aligned} \tag{1.2}$$

where $\sigma(x) = \sqrt{a(x)}$, $g(x,z) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, W is a d-dimensional standard Brownian motion, while N is a Poisson random measure with intensity measure ν, and $\tilde{N}$ is the associated

Research of XZ is partially supported by NNSFC grant of China (Nos. 11271294, 11325105).

compensated Poisson random measure. Under some Lipschitz assumptions in x-variable on $\sigma(x)$, $b(x)$ and $g(x,z)$, it is well knownn that the above SDE admits a unique strong solution, which defines a strong Markov process whose infinitesimal generator $\mathscr{L}$ is of the form (1.1) with $\mu_x(\mathrm{d}z) = \nu \circ g^{-1}(x,\cdot)(\mathrm{d}z)$ (see [20]). A natural question is whether SDE (1.2) has a (weak) solution without Lipschitz assumption on $\sigma(x)$, $b(x)$ and $g(x,z)$, and how about its density.

In this work we are concerned with the existence, uniqueness, and estimates of fundamental solutions of time-dependent version of the operator $\mathscr{L}$ in (1.1), with minimal regularity assumptions on $a(t,x)$, $b(t,x)$ and $\kappa(t,x,z)$, where $\kappa(t,x,z) := |z|^{d+\alpha}\mu_{t,x}(\mathrm{d}z)/\mathrm{d}z$. More precisely, we shall consider the following time-inhomogeneous and non-symmetric non-local operators:

$$\mathscr{L}_t f(x) := \mathscr{L}_t^a f(x) + b_t \cdot \nabla f(x) + \mathscr{L}_t^\kappa f(x), \tag{1.3}$$

where

$$\mathscr{L}_t^a f(x) := \frac{1}{2}\sum_{i,j=1}^d a_{ij}(t,x)\partial_{ij}^2 f(x),\ b_t \cdot \nabla f(x) := \sum_{i=1}^d b_i(t,x)\partial_i f(x),$$

$$\mathscr{L}_t^\kappa f(x) := \int_{\mathbb{R}^d} \left(f(x+z) - f(x) - 1_{\{|z|\leqslant 1\}} z \cdot \nabla f(x)\right) \frac{\kappa(t,x,z)}{|z|^{d+\alpha}}\mathrm{d}z.$$

Here $a(t,x) := (a_{ij}(t,x))_{1\leqslant i,j\leqslant d}$ is a $d \times d$-symmetric matrix-valued measurable function on $[0,\infty)\times\mathbb{R}^d$, $b(t,x) : [0,\infty)\times\mathbb{R}^d \to \mathbb{R}^d$ and $\kappa(t,x,z) : [0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}$ are measurable functions, and $\alpha \in (0,2)$.

With different choices of a, b and κ, we get different types of operators $\mathscr{L}_t$. For example, when $a = \mathbb{I}_{d\times d}$, $b = 0$ and $\kappa(t,x,z) = \mathcal{A}(d,-\alpha)\kappa$ for some $\kappa > 0$, $\mathscr{L}_t = \frac{1}{2}\Delta + \kappa\Delta^{\alpha/2}$ is the generator of independent sum of Brownian motion and rotational α-stable process with weight κ. Here $\mathcal{A}(d,-\alpha)$ is a positive constant: $\mathcal{A}(d,-\alpha) := \alpha 2^{\alpha-1}\pi^{-d/2}\Gamma((d+\alpha)/2)\Gamma(1-\alpha/2)^{-1}$ and Γ is the Gamma function defined by $\Gamma(\lambda) := \int_0^\infty t^{\lambda-1}e^{-t}\mathrm{d}t$, $\lambda > 0$. Moreover, the heat kernel of $\frac{1}{2}\Delta + \kappa\Delta^{\alpha/2}$ exists, denoted by $p^\kappa(t,x,y) = p^\kappa(t,|y-x|)$. It is shown in [13, Theorem 1.4] (see also [24, Theorem 2.13] and [9, Corollary 1.2]) that , there are constants $C, \lambda \geqslant 1$ depending only on d, α such that for all $t > 0$ and $x \in \mathbb{R}^d$,

$$\begin{aligned} C^{-1}\left(t^{-d/2}\wedge(\kappa t)^{-d/\alpha}\right)\wedge\left(t^{-d/2}\mathrm{e}^{-\lambda|x|^2/t} + \tfrac{\kappa t}{|x|^{d+\alpha}}\right) &\leqslant p^\kappa(t,x) \\ \leqslant C\left(t^{-d/2}\wedge(\kappa t)^{-d/\alpha}\right)\wedge\left(t^{-d/2}\mathrm{e}^{-\lambda^{-1}|x|^2/t} + \tfrac{\kappa t}{|x|^{d+\alpha}}\right)&. \end{aligned} \tag{1.4}$$

The above in particular implies that for each $T, M > 0$, all $\kappa \in [0,M]$, $t \in (0,T]$ and $x \in \mathbb{R}^d$,

$$\widetilde{C}^{-1}\left(t^{-d/2}\mathrm{e}^{-\lambda|x|^2/t} + t^{-d/2}\wedge\tfrac{\kappa t}{|x|^{d+\alpha}}\right) \leqslant p^\kappa(t,x) \leqslant \widetilde{C}\left(t^{-d/2}\mathrm{e}^{-\lambda^{-1}|x|^2/t} + t^{-d/2}\wedge\tfrac{\kappa t}{|x|^{d+\alpha}}\right), \tag{1.5}$$

where $\widetilde{C} \geqslant 1$ depends on T, M, d, α. For notational convenience, define for $\gamma, \lambda \in \mathbb{R}$, $t > 0$ and $x \in \mathbb{R}^d$,

$$\xi_{\lambda,\gamma}(t,x) := t^{(\gamma-d)/2}\mathrm{e}^{-\lambda|x|^2/t} \quad \text{and} \quad \eta_{\alpha,\gamma}(t,x) := t^{\gamma/2}(|x| + t^{1/2})^{-d-\alpha}. \tag{1.6}$$

It is easy to check that we can rewrite (1.5) as

$$\widehat{C}^{-1}\left(\xi_{\lambda,0}(t,x) + \kappa\eta_{\alpha,2}(t,x)\right) \leqslant p^\kappa(t,x) \leqslant \widehat{C}\left(\xi_{\lambda^{-1},0}(t,x) + \kappa\eta_{\alpha,2}(t,x)\right) \tag{1.7}$$

for some $\widehat{C} \geqslant 1$ depending on T, M, d, α.

When $\kappa(t,x,z) = \mathcal{A}(d,-\alpha)1_{|z|\leqslant 1}$, $\mathscr{L}^\kappa$ is just the truncated fractional Laplacian operator $\bar{\Delta}^{\alpha/2}$:

$$\bar{\Delta}^{\alpha/2} f(x) = \int_{\{|z|\leqslant 1\}} \left(f(x+z) - f(x) - z \cdot \nabla f(z)\right) \frac{\mathcal{A}(d,-\alpha)}{|z|^{d+\alpha}}\mathrm{d}z.$$

It follows from [8] that the heat kernel of $\bar{\Delta}^{\alpha/2}$, denoted by $\bar{p}_\alpha(t,x,y) = \bar{p}_\alpha(t,x-y)$, exists and it is jointly continuous and has the following estimates: there are constants $C_i = C_i(d,\alpha) > 1,\ i = 1,2$ such that

$$
\begin{aligned}
C_1^{-1}\left(\left(\tfrac{t}{|x|}\right)^{C_2|x|}1_{|x|>1} + \left(t^{-\frac{d}{\alpha}} \wedge \tfrac{t}{|x|^{d+\alpha}}\right)1_{|x|\leqslant 1}\right) &\leqslant \bar{p}_\alpha(t,x) \\
\leqslant C_1\left(\left(\tfrac{t}{|x|}\right)^{C_2^{-1}|x|}1_{|x|>1} + \left(t^{-\frac{d}{\alpha}} \wedge \tfrac{t}{|x|^{d+\alpha}}\right)1_{|x|\leqslant 1}\right), &\quad t \in (0,1], x \in \mathbb{R}^d.
\end{aligned} \tag{1.8}
$$

Throughout this paper, we assume $d \geqslant 2$ and make the following assumptions on a and κ:

($\mathbf{H}^a$) There are $c_1 > 0$ and $\beta \in (0,1)$ such that for all $t > 0$ and $x,y \in \mathbb{R}^d$,

$$|a(t,y) - a(t,x)| \leqslant c_1|y-x|^\beta, \tag{1.9}$$

and for some $c_2 \geqslant 1$,

$$c_2^{-1}\mathbb{I}_{d\times d} \leqslant a(t,x) \leqslant c_2\mathbb{I}_{d\times d}. \tag{1.10}$$

Here $\mathbb{I}_{d\times d}$ denotes the $d\times d$ identity matrix.

($\mathbf{H}^\kappa$) $\kappa(t,x,z)$ is a bounded measurable function and if $\alpha = 1$, we require for any $0 < r < R < \infty$,

$$\int_{r<|z|\leqslant R} z\kappa(t,x,z)|z|^{-d-1}\mathrm{d}z = 0. \tag{1.11}$$

Let $Z(t,x;s,y)$ be the fundamental solution of $\{\mathscr{L}_t^a; t \geqslant 0\}$; see Theorem 2.3 below for details. Since $\mathscr{L}_t$ can be viewed as a perturbation of $\mathscr{L}_t^a$ by $\mathscr{L}_t^{b,\kappa} := b\cdot\nabla + \mathscr{L}_t^\kappa$, heuristically the fundamental solution (or heat kernel) $p(t,x;s,y)$ of $\mathscr{L}_t$ should satisfy the following Duhamel's formula: for all $0 \leqslant t < s < \infty$ and $x,y \in \mathbb{R}^d$,

$$p(t,x;s,y) = Z(t,x;s,y) + \int_t^s\int_{\mathbb{R}^d} p(t,x;r,z)\mathscr{L}_r^{b,\kappa}Z(r,\cdot;s,y)(z)\mathrm{d}z\mathrm{d}r, \tag{1.12}$$

or

$$p(t,x;s,y) = Z(t,x;s,y) + \int_t^s\int_{\mathbb{R}^d} Z(t,x;r,z)\mathscr{L}_r^{b,\kappa}p(r,\cdot;s,y)(z)\mathrm{d}z\mathrm{d}r. \tag{1.13}$$

For any $T \in (0,\infty]$ and $\varepsilon \in [0,T)$, we write

$$\mathbb{D}_\varepsilon^T := \Big\{(t,x;s,y) : x,y \in \mathbb{R}^d \text{ and } s,t \geqslant 0 \text{ with } \varepsilon < s-t < T\Big\}.$$

The following are the main results of this paper. See (2.28) below for the definition of space-time Kato class $\mathbb{K}_2$ of functions on $\mathbb{R}\times\mathbb{R}^d$. We will see from Proposition 2.7 below that $\mathbb{K}_2$ contains $L^q(\mathbb{R}; L^p(\mathbb{R}^d))$ for any $p,q \in [1,\infty]$ with $\frac{d}{p} + \frac{2}{q} < 1$.

Theorem 1.1. *Let $\alpha \in (0,2)$. Under* **($\mathbf{H}^a$)**, **($\mathbf{H}^\kappa$)** *and $b \in \mathbb{K}_2$, there is a unique continuous function $p(t,x;s,y)$ on $\mathbb{D}_0^\infty$ that satisfies* (1.12)*, and*

(1) *(Upper-bound estimate) For any $T > 0$, there exist constants $C_0, \lambda_0 > 0$ such that on $\mathbb{D}_0^T$,*

$$|p(t,x;s,y)| \leqslant C_0(\xi_{\lambda_0,0} + \|\kappa\|_\infty\eta_{\alpha,2})(s-t,y-x). \tag{1.14}$$

Moreover, the following hold.

(2) *(C-K equation) For all $0 \leqslant t < r < s < \infty$ and $x,y \in \mathbb{R}^d$,*

$$\int_{\mathbb{R}^d} p(t,x;r,z)p(r,z;s,y)\mathrm{d}z = p(t,x;s,y). \tag{1.15}$$

(3) *(Gradient estimate) For any $T > 0$, there exist constants $C_1, \lambda_1 > 0$ such that on $\mathbb{D}_0^T$,*

$$|\nabla_x p(t,x;s,y)| \leqslant C_1(\xi_{\lambda_1,-1} + \|\kappa\|_\infty\eta_{\alpha,1})(s-t,y-x). \tag{1.16}$$

(4) *(Fractional derivative estimate) If in addition for* $\alpha \in (0,1]$, $b \in \mathbb{K}_1$ *and for* $\alpha \in (1,2)$, $b \in \overline{\mathbb{K}}_\alpha$ *(see* (2.29) *below for a definition), then for any* $T > 0$*, there exists a constant* $C_2 > 0$ *such that on* $\mathbb{D}_0^T$*,*

$$|\Delta^{\alpha/2} p(t,\cdot;s,y)(x)| \leqslant C_2 \eta_{\alpha,0}(s-t, y-x). \tag{1.17}$$

Meanwhile, equation (1.13) *holds on* $\mathbb{D}_0^\infty$*.*

(5) *(Conservativeness) For any* $0 \leqslant t < s < \infty$ *and* $x \in \mathbb{R}^d$*,*

$$\int_{\mathbb{R}^d} p(t,x;s,y)\mathrm{d}y = 1. \tag{1.18}$$

(6) *(Generator) For any* $f \in C_b^2(\mathbb{R}^d)$*, we have*

$$P_{t,s} f(x) - f(x) = \int_t^s P_{t,r} \mathscr{L}_r f(x) \mathrm{d}r, \tag{1.19}$$

where $P_{t,s} f(x) := \int_{\mathbb{R}^d} p(t,x;s,y) f(y)\mathrm{d}y$*.*

(7) *(Continuity) For any bounded and uniformly continuous function* $f(x)$*, we have*

$$\lim_{|t-s|\to 0} \|P_{t,s} f - f\|_\infty = 0. \tag{1.20}$$

Remark 1.2. *Estimate* (1.17) *is new even for* $\kappa \equiv 0$*.*

Note that in Theorem 1.1, we do not assume $\kappa(t,x,z) \geqslant 0$ and so the fundamental solution $p(t,x;s,y)$ can take negative values; see Remark 1.4 below. The following theorem gives the lower bound estimate.

Theorem 1.3. *Under the same assumptions of Theorem 1.1, if for each* $t > 0$ *and* $x \in \mathbb{R}^d$*,*

$$\kappa(t,x,z) \geqslant 0, \quad a.e.\ z \in \mathbb{R}^d, \tag{1.21}$$

then $p(t,x;s,y) \geqslant 0$ *on* $\mathbb{D}_0^\infty$*. Moreover, for any* $T > 0$*, there are constants* $C_3, \lambda_3 > 0$ *such that*

$$p(t,x;s,y) \geqslant C_3(\xi_{\lambda_3,0} + m_\kappa \eta_{\alpha,2})(s-t, y-x) \text{ on } \mathbb{D}_0^T, \tag{1.22}$$

where $m_\kappa := \inf_{(t,x)} \operatorname{essinf}_{z\in\mathbb{R}^d} \kappa(t,x,z)$*.*

Remark 1.4. *Under the hypothesis of Theorem 1.3, if in addition,* κ *satisfies that for each* $t \geqslant 0$*,*

$$x \mapsto \kappa(t,x,z) \text{ is continuous} \quad a.e.\ z \in \mathbb{R}^d,$$

then, we can prove that (1.21) is also a necessary condition to the positivity of $p(t,x;s,y)$*. For example, see the proof of* [14, Theorem 1.2]*, or* [7, Lemma 4.5] *or* [25].

The following corollary follows immediately from Theorems 1.1 and 1.3.

Corollary 1.5. *Let* $\alpha \in (0,2)$*. Under* **(H^a)**, **(H$^\kappa$)**, $b \in \mathbb{K}_2$ *and* (1.21)*, for every* $T > 0$*, there are positive constants* $C, \lambda \geqslant 1$ *such that on* $\mathbb{D}_0^T$*,*

$$C^{-1}\Big(\xi_{\lambda,0} + m_\kappa \eta_{\alpha,2}\Big)(s-t, y-x) \leqslant p(t,x;s,y) \leqslant C\Big(\xi_{\lambda^{-1},0} + \|\kappa\|_\infty \eta_{\alpha,2}\Big)(s-t, y-x).$$

In the truncated case, we consider the following two conditions on κ:

(HU$^\kappa$) $0 \leqslant \kappa(t,x,z) \leqslant \kappa_0 1_{|z|\leqslant 1}(z)$ for some $\kappa_0 > 0$.

(HL$^\kappa$) $\kappa(t,x,z) \geqslant \kappa_0 1_{|z|\leqslant 1}(z)$ for some $\kappa_0 > 0$.

Theorem 1.6. *Suppose that* **(H^a)**, **(H$^\kappa$)** *and* $b \in \mathbb{K}_2$ *hold. Let* $T > 0$*, and for* $\lambda > 0$*, set*

$$\bar{\eta}_{\alpha,\lambda}(t,x) := t(|x| + t^{1/2})^{-(d+\alpha)} 1_{|x|\leqslant 1/2} + (t/|x|)^{\lambda|x|} 1_{|x|>1/2}. \tag{1.23}$$

(i) *If in addition κ satisfies* **(HU$^\kappa$)**, *then there are constants $C_1, \lambda_1 > 0$ such that*

$$p(t,x;s,y) \leqslant C_1\,(\xi_{\lambda_1,0} + \overline{\eta}_{\alpha,1/8})\,(s-t, y-x) \quad \text{on } \mathbb{D}_0^T.$$

(ii) *If in addition κ satisfies* **(HL$^\kappa$)**, *then there are constants $C_2, \lambda_2 > 0$ such that*

$$p(t,x;s,y) \geqslant C_2\,(\xi_{\lambda_2,0} + \overline{\eta}_{\alpha,8})\,(s-t, y-x) \quad \text{on } \mathbb{D}_0^T.$$

Heat kernel analysis takes an important place in PDE and in probability theory, as heat kernel encodes all the information about the corresponding generator and the corresponding Markov processes. Since explicit formula can only be derived in some very special and limited cases, the main focus of the heat kernel analysis is on its sharp estimates. While it is relatively easy to get some crude bounds, obtaining sharp two-sided bounds on the heat kernel is typically quite delicate and challenging. It requires deep understanding of the corresponding generator. For second order elliptic operators and diffusion process, a lot is known and there are many beautiful results . For instance, the celebrated Aronson's estimate [1] asserts that the heat kernel for uniformly elliptic operators of divergence form with measurable coefficients has two-sided Gaussian-type bounds. Aronson's estimate also holds for non-divergence form elliptic operators with Hölder continuous coefficients; see Theorem 2.3 below.

The study of heat kernel for non-local operators is relatively recent, propelled by interest in discontinuous Markov processes, as many physical, engineering and social phenomena can be successfully modeled by using discontinuous Markov processes including Lévy processes. The infinitesimal generators of discontinuous Markov processes are non-local operators. During the past several years there is also many interest from the theory of PDE (such as singular obstacle problems) to study non-local operators; see, for example, [4, 23] and the references therein. Quite many progress has been made in the last fifteen years on the development of the DeGiorgi-Nash-Moser-Aronson type theory for symmetric non-local operators. For example, Kolokoltsov [22] obtained two-sided heat kernel estimates for certain stable-like processes in $\mathbb{R}^d$, whose infinitesimal generators are a class of pseudo-differential operators having smooth symbols. Bass and Levin [2] used a completely different approach to obtain similar estimates for discrete time Markov chain on $\mathbb{Z}^d$, where the conductance between x and y is comparable to $|x-y|^{-n-\alpha}$ for $\alpha \in (0,2)$. In Chen and Kumagai [11], two-sided heat kernel estimates and a scale-invariant parabolic Harnack inequality (PHI in abbreviation) for symmetric α-stable-like processes on d-sets are obtained. Recently in [12], two-sided heat kernel estimates and PHI are established for symmetric non-local operators of variable order. The DeGiorgi-Nash-Moser-Aronson type theory is studied very recently in Chen and Kumagai [13] for symmetric diffusions with jumps. We refer the reader to the survey articles [5, 19] and the references therein on the study of heat kernels for symmetric non-local operators. However, for non-symmetric non-local operators, much less is known. In [3], Bogdan and Jakubowski considered a fundamental solution to the non-local operator $\Delta^{\alpha/2}+b(x)\cdot\nabla$ with $\alpha \in (1,2)$ and b belonging to some Kato's class, and obtained its sharp two-sided estimates. The uniqueness of fundamental solution to $\Delta^{\alpha/2} + b(x)\cdot\nabla$ and its connection to stable processes with drifts are settled in Chen and Wang [15]. In [26], Xie and Zhang studied the critical case $a(t,x)\Delta^{1/2} + b(t,x)\cdot\nabla$. Heat kernels for subordinate Brownian motions with drifts have been studied in [7] and [6]. Chen and Wang [14] studied heat kernel estimates for $\Delta^{\alpha/2}$ under non-local perturbation, while Wang [25] investigated heat kernel for Δ perturbed by non-local operators. Recently, Chen and Zhang [16] obtained sharp two-sided estimates, gradient estimate and fractional derivative estimate of the heat kernel for general non-local and non-symmetric operator $\mathscr{L}^\kappa$ with $\kappa(t,x,z) = \kappa(x,z)$ by using Levi's parametrix method.

In this paper, we concentrate on the study of heat kernel for non-symmetric operators $\mathscr{L}$ of type (1.3), which have both diffusive and non-local parts . When $\kappa(t,x,y) \geqslant 0$, its fundamental solution $p(t,x;s,y)$ becomes a family of transition density and so it determines a Feller process X having strong Feller property. Clearly, the law of X is a solution to the martingale problem for $(\mathscr{L}, C_c^2(\mathbb{R}^d))$. Is the solution to the martingale problem for $(\mathscr{L}, C_c^2(\mathbb{R}^d))$ unique? It is also tempting to ask that when $\kappa(t,x,z)/|z|^{d+\alpha}$ is of the form $\nu \circ g^{-1}(t,x,\cdot)(\mathrm{d}z)$ for some $g(t,x,z) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ and a σ-finite measure ν on $\mathbb{R}^d \setminus \{0\}$, whether this Feller process X satisfies the following SDE:

$$\begin{aligned}\mathrm{d}X_t = \sigma(t,X_t)\mathrm{d}W_t + b(t,X_t)\mathrm{d}t + \int_{|z|\leqslant 1} g(t,X_{t-},z)\tilde{N}(\mathrm{d}t,\mathrm{d}z)\\ + \int_{|z|>1} g(t,X_{t-},z)N(\mathrm{d}t,\mathrm{d}z), \quad X_0 = x,\end{aligned} \tag{1.24}$$

where $\sigma(t,x) = \sqrt{a(t,x)}$, W is a d-dimensional standard Brownian motion, N is a Poisson random measure with intensity measure ν, and $\tilde{N}$ is the associated compensated Poisson random measure? We plan to address these questions in a separate work.

The rest of the paper is organized as follows. In Section 2, we present some key estimates that will be used later. In Section 3, we prove our main result Theorem 1.1. The main crux of work is on various gradient and fractional derivative estimates, which is crucial for the iteration procedure and rigorously establishing the Duhamel's formula. In Section 4, we first show the positivity of $p(t,x;s,y)$ by the maximum principle under the non-negativeness of κ. We then derive the lower bound estimate by a probabilistic approach after obtaining the on-diagonal estimate of $p(t,x;s,y)$. In Section 5, we consider the truncated case. In the Appendix, we show a maximum principle and derive two-sided Aronson-type Gaussian estimates for heat kernels of time-dependent second-order elliptic differential operators.

We conclude this introduction by mentioning some conventions that will be used throughout this paper. The letter C or c with or without subscripts will denote an unimportant constant. For two quantities f and g, $f \asymp g$ means that $C^{-1}f \leqslant f \leqslant Cg$ for some $C \geqslant 1$, and $f \preceq g$ means that $f \leqslant Cg$ for some $C \geqslant 1$. The letter $\mathbb{N}$ will denote the collection of positive integers, and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

2. Preliminaries

2.1. **Basic estimates.** We first prove the following elementary but important estimates (which can also be called $3P$-inequalities) for later use. Recall that the functions $\xi_{\lambda,\gamma}$ and $\eta_{\alpha,\gamma}$ are defined in (1.6).

Lemma 2.1. (i) *For any $\alpha \in (0,+\infty)$ and $\lambda > 0$, there exist positive constants $C_1 = C_1(d,\alpha,\lambda)$ and $C_2 = C_2(d,\alpha)$ such that for all $t > 0$ and $x \in \mathbb{R}^d$,*

$$\xi_{\lambda,0}(t,x) \leqslant C_1 \eta_{\alpha,\alpha}(t,x), \tag{2.1}$$

and for all $\gamma \geqslant 0$ and $t > 0$,

$$\int_{\mathbb{R}^d} \eta_{\alpha,\gamma}(t,x)\mathrm{d}x \leqslant C_2 t^{(\gamma-\alpha)/2}, \tag{2.2}$$

and for some $C_3 = C_3(d,\alpha) > 0$ and all $t,s > 0$, $x \in \mathbb{R}^d$ and $\gamma \in [0,\alpha]$,

$$\int_{\mathbb{R}^d} \eta_{\alpha,\gamma}(t,x-z)\eta_{\alpha,\alpha}(s,z)\mathrm{d}z \geqslant C_3 \eta_{\alpha,\gamma}(t+s,x). \tag{2.3}$$

(ii) *For any $0 < \alpha \leqslant \beta$ and for all $t,s > 0$, $x,y \in \mathbb{R}^d$, we have*

$$\eta_{\alpha,0}(t,x)\eta_{\beta,0}(s,y) \leqslant 2^{d+\alpha}\Big(\eta_{\beta,0}(t,x) + \eta_{\beta,0}(s,y)\Big)\eta_{\alpha,0}(t+s,x+y). \tag{2.4}$$

Moreover, there is a constant $C_4 = C_4(d,\alpha,\beta) > 0$ *such that for all* $\gamma_1, \gamma_2 > \beta - 2$,

$$\begin{aligned}&\int_t^s\int_{\mathbb{R}^d} \eta_{\alpha,\gamma_1}(s-r,y-z)\eta_{\beta,\gamma_2}(r-t,z-x)\mathrm{d}z\mathrm{d}r\\&\qquad\leqslant C_4\mathcal{B}(\tfrac{\gamma_1-\beta}{2}+1,\tfrac{\gamma_2-\beta}{2}+1)\eta_{\alpha,2+\gamma_1+\gamma_2-\beta}(s-t,y-x),\end{aligned}\tag{2.5}$$

where $\mathcal{B}(\beta,\gamma) := \int_0^1 (1-s)^{\beta-1}s^{\gamma-1}\mathrm{d}s$ *is the usual Beta function.*

(iii) *For any* $\alpha \in (0,2)$*, there exists a constant* $C_5 = C_5(d,\alpha,\lambda) > 0$ *such that for all* $\gamma_1 > -2$ *and* $\gamma_2 > \alpha - 2$,

$$\begin{aligned}&\int_t^s\int_{\mathbb{R}^d} \xi_{\lambda,\gamma_1}(r-t,z-x)\eta_{\alpha,\gamma_2}(s-r,y-z)\mathrm{d}z\mathrm{d}r\\&\qquad\leqslant C_5\mathcal{B}(\tfrac{\gamma_1}{2}+1,\tfrac{\gamma_2-\alpha}{2}+1)\eta_{\alpha,2+\gamma_1+\gamma_2}(s-t,y-x).\end{aligned}\tag{2.6}$$

(iv) *For any* $\lambda > 0$*, we have*

$$\int_{\mathbb{R}^d} \xi_{\lambda,0}(t,x-y)\xi_{\lambda,0}(s,y)\mathrm{d}y = (\pi\lambda^{-1})^{d/2}\xi_{\lambda,0}(t+s,x),\tag{2.7}$$

and for all $\gamma_1,\gamma_2 > -2$,

$$\begin{aligned}&\int_t^s\int_{\mathbb{R}^d} \xi_{\lambda,\gamma_1}(r-t,z-x)\xi_{\lambda,\gamma_2}(s-r,y-z)\mathrm{d}z\mathrm{d}r\\&\qquad= (\pi\lambda^{-1})^{d/2}\mathcal{B}(\tfrac{\gamma_1}{2}+1,\tfrac{\gamma_2}{2}+1)\xi_{\lambda,2+\gamma_1+\gamma_2}(s-t,y-x).\end{aligned}\tag{2.8}$$

Proof. (i) If $|x| \leqslant t^{1/2}$, then

$$\xi_{\lambda,0}(t,x) \leqslant t^{-d/2} \leqslant 2^{d+\alpha}\eta_{\alpha,\alpha}(t,x).$$

If $|x| > t^{1/2}$, then

$$\xi_{\lambda,0}(t,x) = t^{\alpha/2}|x|^{-(d+\alpha)}\left(|x|^2/t\right)^{(d+\alpha)/2}\mathrm{e}^{-\lambda|x|^2/t} \leq t^{\alpha/2}|x|^{-(d+\alpha)} \leq \eta_{\alpha,\alpha}(t,x).$$

Moreover, we have

$$\int_{\mathbb{R}^d}\eta_{\alpha,\gamma}(t,x)\mathrm{d}x \leqslant t^{\gamma/2}\left(\int_{|x|\leqslant t^{1/2}} t^{-\frac{d+\alpha}{2}}\mathrm{d}x + \int_{|x|>t^{1/2}}|x|^{-d-\alpha}\mathrm{d}x\right) \leq t^{(\gamma-\alpha)/2}.$$

To prove (2.3), it suffices to show it for $\gamma = \alpha$. Thus, by symmetry we may assume $s \leqslant t$. Noticing that for $|z| \leqslant s^{1/2}$,

$$|x-z| + t^{1/2} \leqslant |x| + |z| + t^{1/2} \leqslant |x| + 2(t+s)^{1/2},$$

we have

$$\begin{aligned}&\int_{\mathbb{R}^d}\eta_{\alpha,\alpha}(t,x-z)\eta_{\alpha,\alpha}(s,z)\mathrm{d}z \geqslant \int_{|z|\leqslant s^{1/2}}\frac{t^{\alpha/2}}{(|x-z|+t^{1/2})^{d+\alpha}}\eta_{\alpha,\alpha}(s,z)\mathrm{d}z\\&\geqslant \frac{2^{-\alpha/2}(t+s)^{\alpha/2}}{(|x|+2(t+s)^{1/2})^{d+\alpha}}\int_{|z|\leqslant s^{1/2}}\eta_{\alpha,\alpha}(s,z)\mathrm{d}z \geqslant \frac{2^{-\alpha/2}}{2^{d+\alpha}}\eta_{\alpha,\alpha}(t+s,x)\int_{|z|\leqslant 1}\eta_{\alpha,\alpha}(1,z)\mathrm{d}z.\end{aligned}$$

(ii) Estimate (2.4) follows by the following easy inequality:

$$\begin{aligned}(|x+y|+(t+s)^{1/2})^{d+\alpha} &\leqslant 2^{d+\alpha-1}\Big\{(|x|+t^{1/2})^{d+\alpha} + (|y|+s^{1/2})^{d+\alpha}\Big\}\\&\leqslant 2^{d+\alpha}\Big\{(|x|+t^{1/2})^{d+\alpha} + (|x|+t^{1/2})^{\alpha-\beta}(|y|+s^{1/2})^{d+\beta}\Big\},\end{aligned}$$

where the second inequality is due to $b^\alpha \leqslant a^\alpha + a^{\alpha-\beta}b^\beta$ for $0 \leqslant \alpha \leqslant \beta$ and $a,b \geqslant 0$. Moreover, by (2.4) and (2.2), we have

$$\int_t^s\int_{\mathbb{R}^d}\eta_{\alpha,\gamma_1}(s-r,y-z)\eta_{\beta,\gamma_2}(r-t,z-x)\mathrm{d}z\mathrm{d}r \leqslant 2^{d+\alpha}\eta_{\alpha,0}(s-t,y-x)$$

$$\times \int_t^s (s-r)^{\gamma_1/2}(r-t)^{\gamma_2/2}\int_{\mathbb{R}^d}\Big(\eta_{\beta,0}(s-r,y-z)+\eta_{\beta,0}(r-t,z-x)\Big)\mathrm{d}z\mathrm{d}r$$
$$\leq \eta_{\alpha,0}(s-t,y-x)\int_t^s \Big((s-r)^{(\gamma_1-\beta)/2}(r-t)^{\gamma_2/2}+(s-r)^{\gamma_1/2}(r-t)^{(\gamma_2-\beta)/2}\Big)\mathrm{d}r$$
$$\leq \eta_{\alpha,2+\gamma_1+\gamma_2-\beta}(s-t,y-x)\Big(\mathcal{B}(\tfrac{\gamma_1-\beta}{2}+1,\tfrac{\gamma_2}{2}+1)+\mathcal{B}(\tfrac{\gamma_1}{2}+1,\tfrac{\gamma_2-\beta}{2}+1)\Big)$$
$$\leq \mathcal{B}(\tfrac{\gamma_1-\beta}{2}+1,\tfrac{\gamma_2-\beta}{2}+1)\eta_{\alpha,2+\gamma_1+\gamma_2-\beta}(s-t,y-x).$$

(iii) It follows by (2.1) with $\xi_{\lambda,\gamma_1}(t,x)\leqslant C_1\eta_{\alpha,\alpha+\gamma_1}(t,x)$ and (2.5) with $\beta=\alpha$.

(iv) It follows by Chapman-Kolmogorov's equation for Brownian transition density function. □

2.2. **Fractional derivative estimates of Gaussian kernel.** For $\alpha\in(0,2)$, set

$$z^{(\alpha)} := z1_{\alpha\in(1,2)} + z1_{|z|\leqslant 1}1_{\alpha=1}.$$

Let $J:\mathbb{R}^d\to\mathbb{R}$ be a bounded measurable function. For a function $f(x)$ on $\mathbb{R}^d$, define

$$\widetilde{\mathscr{L}}^J f(x) := \int_{\mathbb{R}^d}\delta_f^{(\alpha)}(x;z)J(z)|z|^{-d-\alpha}\mathrm{d}z, \tag{2.9}$$

where

$$\delta_f^{(\alpha)}(x;z) := f(x+z)-f(x)-z^{(\alpha)}\cdot\nabla f(x). \tag{2.10}$$

The following lemma will play an important role in the sequel.

Lemma 2.2. *Given $\alpha\in(0,2)$, let $J:\mathbb{R}^d\to\mathbb{R}$ be a bounded measurable function with*

$$\int_{r<|z|\leqslant R} z\cdot J(z)|z|^{-d-1}\mathrm{d}z = 0,\quad 0<r<R<\infty. \tag{2.11}$$

Let $T>0$ and $G_t(x):(0,T)\times\mathbb{R}^d\to\mathbb{R}$ be a C^2 function in x. Suppose that for each $j=0,1,2$, there are $C_j>0$ and $\beta_j\geqslant 0$ such that for $t\in(0,T)$ and $x\in\mathbb{R}^d$,

$$|\nabla^j G_t(x)|\leqslant C_j\eta_{\alpha,\alpha-\beta_j-j}(t,x). \tag{2.12}$$

Then for any $\gamma\in[0,(2-\alpha)\wedge 1)$, there exists a constant $C=C(\gamma,d,\alpha)>0$ such that

$$\Big|\widetilde{\mathscr{L}}^J G_t(x)\Big|\leqslant C\|J\|_\infty\Big(C_0t^{-\frac{\beta_0}{2}}+C_1t^{-\frac{\beta_1}{2}}+C_1^\gamma C_2^{1-\gamma}t^{-\frac{\gamma\beta_1+(1-\gamma)\beta_2}{2}}\Big)\eta_{\alpha,0}(t,x). \tag{2.13}$$

Proof. By definition (2.9)-(2.10), (2.11) and (1.11), we have

$$\Big|\widetilde{\mathscr{L}}^J G_t(x)\Big|\leqslant\|J\|_\infty\Bigg[\int_{|z|\leqslant t^{1/2}}|\delta_{G_t}^{(\alpha)}(x;z)|\cdot|z|^{-d-\alpha}\mathrm{d}z+\int_{|z|>t^{1/2}}|G_t(x+z)|\cdot|z|^{-d-\alpha}\mathrm{d}z$$
$$+|G_t(x)|\int_{|z|>t^{1/2}}|z|^{-d-\alpha}\mathrm{d}z+1_{\alpha\in(1,2)}|\nabla G_t(x)|\int_{|z|>t^{1/2}}|z|\cdot|z|^{-d-\alpha}\mathrm{d}z\Bigg]$$
$$=:\|J\|_\infty\Big[I_1+I_2+I_3+I_4\Big].$$

Notice that for $\alpha\in(0,1)$,

$$\delta_{G_t}^{(\alpha)}(x;z)=\int_0^1\langle z,\nabla G_t(x+\theta z)\rangle\mathrm{d}\theta,$$

and for $\alpha\in[1,2)$,

$$\delta_{G_t}^{(\alpha)}(x;z)=\int_0^1\int_0^1\theta\langle z\otimes z,\nabla^2G_t(x+\theta'\theta z)\rangle\mathrm{d}\theta'\mathrm{d}\theta.$$

By (2.12), we have for all $|z| \leqslant t^{1/2}$,

$$
\begin{aligned}
|\delta^{(\alpha)}_{G_t}(x;z)| &\leqslant |z| \int_0^1 |\nabla G_t(x+\theta z)| \mathrm{d}\theta + 1_{\alpha\in[1,2)} |z| \cdot |\nabla G_t(x)| \\
&\leqslant C_1 \left(\int_0^1 \eta_{\alpha,\alpha-\beta_1-1}(t, x+\theta z) \mathrm{d}\theta + 1_{\alpha\in[1,2)} \eta_{\alpha,\alpha-\beta_1-1}(t,x) \right) |z| \\
&\leqslant 2^{d+\alpha} C_1\, \eta_{\alpha,\alpha-\beta_1-1}(t,x)\, |z|,
\end{aligned}
$$

and if $\alpha \in [1,2)$, we alternatively have

$$
|\delta^{(\alpha)}_{G_t}(x;z)| \leqslant 2^{d+\alpha} C_2\, \eta_{\alpha,\alpha-\beta_2-2}(t,x)\, |z|^2.
$$

Thus for I_1, if $\alpha \in (0,1)$, then

$$
I_1 \leqslant 2^{d+\alpha} C_1\, \eta_{\alpha,\alpha-\beta_1-1}(t,x) \int_{|z|\leqslant t^{1/2}} |z|^{1-d-\alpha} \mathrm{d}z \preceq C_1 t^{-\beta_1/2} \eta_{\alpha,0}(t,x);
$$

if $\alpha \in [1,2)$, then by interpolation, we have for all $\gamma \in [0, 2-\alpha)$,

$$
\begin{aligned}
I_1 &\leqslant 2^{d+\alpha} \int_{|z|\leqslant t^{1/2}} \Big(C_1\, \eta_{\alpha,\alpha-\beta_1-1}(t,x)\, |z| \Big)^{\gamma} \Big(C_2\, \eta_{\alpha,\alpha-\beta_2-2}(t,x)\, |z|^2 \Big)^{1-\gamma} |z|^{-d-\alpha} \mathrm{d}z \\
&\preceq C_1^{\gamma} C_2^{1-\gamma} t^{-\gamma\beta_1/2-(1-\gamma)\beta_2/2} \eta_{\alpha,0}(t,x).
\end{aligned}
$$

For I_2, by (2.4) and (2.2), we have

$$
\begin{aligned}
I_2 &\leqslant C_0 t^{\alpha-\beta_0/2} \int_{|z|>t^{1/2}} \eta_{\alpha,0}(t, x+z) |z|^{-d-\alpha} \mathrm{d}z \\
&\leqslant 2^{d+\alpha} C_0 t^{\alpha-\beta_0/2} \int_{|z|>t^{1/2}} \eta_{\alpha,0}(t, x+z) \eta_{\alpha}(t,-z) \mathrm{d}z \\
&\leqslant 4^{d+\alpha} C_0 t^{\alpha-\beta_0/2} \eta_{\alpha,0}(2t, x) \int_{\mathbb{R}^d} [\eta_{\alpha,0}(t, x+z) + \eta_{\alpha}(t,-z)] \mathrm{d}z \\
&\preceq C_0 t^{-\beta_0/2} \eta_{\alpha,0}(t,x).
\end{aligned}
$$

For I_3 and I_4, it is easy to see that

$$
I_3 + I_4 \preceq \Big(C_0 t^{-\beta_0/2} + C_1 t^{-\beta_1/2} \Big) \eta_{\alpha,0}(t,x).
$$

Combing the above calculations, we get (2.13). □

Under **($\mathbf{H}^a$)**, it is more or less well known that there exists a fundamental solution to the operator $\partial_t - \mathscr{L}^a_t$ (cf. [18]). However, to the best of our knowledge, most of the proofs also require the Hölder continuity of a with respect to the time variable t. For the readers' convenience, a proof of the following result is provided in Appendix 6.2.

Theorem 2.3. *Under* **($\mathbf{H}^a$)**, *there is a unique continuous function Z on $\mathbb{D}^{\infty}_0$ such that for a.e. $t \in (0,s)$, and every $x, y \in \mathbb{R}^d$,*

$$
\partial_t Z(t,x;s,y) + \mathscr{L}^a_t Z(t,\cdot;s,y)(x) = 0, \tag{2.14}
$$

and

(1) *(Upper and gradient estimate) For $j = 0,1,2$ and $T > 0$, there exist constants $C, \lambda > 0$ such that on $\mathbb{D}^T_0$,*

$$
|\nabla^j_x Z(t,x;s,y)| \leqslant C \xi_{\lambda,-j}(s-t, y-x). \tag{2.15}
$$

Moreover, $Z(t,x;s,y)$ enjoys the following properties.

(2) *(Hölder estimate in y) For* $j = 0, 1$, $\beta' \in (0, \beta)$ *and* $T > 0$, *there exist constants* $C, \lambda > 0$ *such that on* $\mathbb{D}_0^T$,

$$|\nabla_x^j Z(t, x; s, y_1) - \nabla_x^j Z(t, x; s, y_2)| \leqslant C|y_1 - y_2|^{\beta'} \sum_{i=1,2} \xi_{\lambda, -\beta'-j}(s - t, y_i - x). \tag{2.16}$$

(3) *(Continuity) For any bounded and uniformly continuous function* $f(x)$,

$$\lim_{|t-s|\to 0} \|P_{t,s}^{(Z)} f - f\|_\infty = 0, \tag{2.17}$$

where $P_{t,s}^{(Z)} f(x) := \int_{\mathbb{R}^d} Z(t, x; s, y) f(y) \mathrm{d}y$.

(4) *(C-K equation) For all* $0 \leqslant t < r < s < \infty$, *we have*

$$P_{t,r}^{(Z)} P_{r,s}^{(Z)} f = P_{t,s}^{(Z)} f. \tag{2.18}$$

(5) *(Conservativeness) For all* $0 \leqslant t < s < \infty$, *we have*

$$P_{t,s}^{(Z)} 1 = 1. \tag{2.19}$$

(6) *(Generator) For any* $f \in C_b^2(\mathbb{R}^d)$, *we have*

$$P_{t,s}^{(Z)} f(x) - f(x) = \int_t^s P_{t,r}^{(Z)} \mathscr{L}_r^a f(x) \mathrm{d}r = \int_t^s \mathscr{L}_r^a P_{r,s}^{(Z)} f(x) \mathrm{d}r. \tag{2.20}$$

(7) *(Two-sided estimates) For any* $T > 0$, *there exist constants* $C, \lambda \geqslant 1$ *such that on* $\mathbb{D}_0^T$,

$$C^{-1} \xi_{\lambda,0}(s - t, y - x) \leqslant Z(t, x; s, y) \leqslant C \xi_{\lambda^{-1},0}(s - t, y - x). \tag{2.21}$$

We call $Z(t, x; s, y)$ the fundamental solution or heat kernel of $\mathscr{L}^a$. The following corollary gives fractional derivative estimates of Gaussian kernels.

Corollary 2.4. *Let* $\alpha \in (0, 2)$ *and* $J : \mathbb{R}^d \to \mathbb{R}$ *be a bounded measurable function satisfying* (2.11). *Let* $Z(t, x; s, y)$ *be as in Theorem 2.3,* $\beta \in (0, 1)$ *be as in* (1.9), *and* $T > 0$.

(i) *There is a constant* $C > 0$ *such that on* $\mathbb{D}_0^T$,

$$\left| \widetilde{\mathscr{L}^J} Z(t, \cdot; s, y)(x) \right| \leqslant C \|J\|_\infty \eta_{\alpha,0}(s - t, y - x). \tag{2.22}$$

(ii) *For any* $\gamma \in [0, (2 - \alpha) \wedge 1)$, *there is a constant* $C > 0$ *such that on* $\mathbb{D}_0^T$,

$$\left| \widetilde{\mathscr{L}^J} Z(t, \cdot; s, y)(x_1) - \widetilde{\mathscr{L}^J} Z(t, \cdot; s, y)(x_2) \right| \leqslant C \|J\|_\infty |x_1 - x_2|^\gamma \sum_{i=1,2} \eta_{\alpha,-\gamma}(s - t, y - x_i). \tag{2.23}$$

(iii) *For any* $\gamma \in [0, (2 - \alpha) \wedge 1)$ *and* $\beta' \in (0, \beta)$, *there is a constant* $C > 0$ *such that on* $\mathbb{D}_0^T$,

$$\left| \widetilde{\mathscr{L}^J} Z(t, \cdot; s, y_1)(x) - \widetilde{\mathscr{L}^J} Z(t, \cdot; s, y_2)(x) \right| \leqslant C \|J\|_\infty |y_1 - y_2|^{\beta'\gamma} \sum_{i=1,2} \eta_{\alpha,-\beta'\gamma}(s - t, y_i - x). \tag{2.24}$$

(iv) *For any* $\gamma \in (0, 1]$, *there is a constant* $C > 0$ *such that on* $\mathbb{D}_0^T$,

$$\left| \nabla_x Z(t, x_1; s, y) - \nabla_x Z(t, x_2; s, y) \right| \leqslant C|x_1 - x_2|^\gamma \sum_{i=1,2} \xi_{\lambda/2,-\gamma-1}(s - t, y - x_i). \tag{2.25}$$

Proof. (i) By (2.15) and (2.1), estimate (2.22) follows by applying Lemma 2.2 to function

$$(r, x) \mapsto Z(t, x + y; t + r, y)$$

with $\beta_j = 0$, $j = 0, 1, 2$, and letting $r = s - t$.

(ii) For fixed $t < s$ and $x_1, x_2, y \in \mathbb{R}^d$, let us define

$$G_r(z) := Z(t, z + y; t + r, y) - Z(t, x_2 - x_1 + z + y; t + r, y).$$

Clearly,

$$I := \widetilde{\mathscr{L}^J} Z(t,\cdot;s,y)(x_1) - \widetilde{\mathscr{L}^J} Z(t,\cdot;s,y)(x_2) = \widetilde{\mathscr{L}^J} G_{s-t}(x_1-y).$$

If $|x_1-x_2| > \sqrt{s-t}$, then by (i), we have

$$\begin{aligned}|I| &\leqslant |\widetilde{\mathscr{L}^J} Z(t,\cdot;s,y)(x_1)| + |\widetilde{\mathscr{L}^J} Z(t,\cdot;s,y)(x_2)| \\ &\preceq C\|J\|_\infty\Big(\eta_{\alpha,0}(s-t,y-x_1)+\eta_{\alpha,0}(s-t,y-x_2)\Big) \\ &\leqslant C\|J\|_\infty|x_1-x_2|^\gamma\Big(\eta_{\alpha,-\gamma}(s-t,y-x_1)+\eta_{\alpha,-\gamma}(s-t,y-x_2)\Big).\end{aligned} \tag{2.26}$$

If $|x_1-x_2| \leqslant \sqrt{s-t}$, then by (2.15), we have for $j=0,1$,

$$\begin{aligned}|\nabla^j G_{s-t}(z)| &\leqslant |x_1-x_2|\int_0^1 |\nabla_x^{j+1} Z(t,z+y+\theta(x_2-x_1);s,y)|\mathrm{d}\theta \\ &\preceq |x_1-x_2|\int_0^1 \xi_{\lambda,-j-1}(s-t,-z-\theta(x_2-x_1))\mathrm{d}\theta \\ &\preceq |x_1-x_2|\cdot \xi_{\lambda/2,-j-1}(s-t,z),\end{aligned}$$

and

$$|\nabla^2 G_{s-t}(z)| \leqslant |\nabla_x^2 Z(t,z+y;s,y)| + |\nabla_x^2 Z(t,x_2-x_1+z+y;s,y)| \preceq \xi_{\lambda,-2}(s-t,z).$$

Hence, by (2.1) and Lemma 2.2 with $\beta_0=\beta_1=1$ and $\beta_2=0$, we obtain that for $|x_1-x_2| \leqslant \sqrt{s-t}$,

$$\begin{aligned}|I| &\leqslant C\|J\|_\infty\Big(|x_1-x_2|(s-t)^{-1/2}+|x_1-x_2|^\gamma(s-t)^{-\frac{\gamma}{2}}\Big)\eta_{\alpha,0}(s-t,x_1-y) \\ &\leqslant C\|J\|_\infty|x_1-x_2|^\gamma\eta_{\alpha,-\gamma}(s-t,x_1-y).\end{aligned}$$

Combining this with (2.26), we obtain (2.23).

(iii) As above, for fixed $t<s$ and $x,y_1,y_2\in\mathbb{R}^d$, let us define

$$G_r(z) := Z(t,z+y_1;t+r,y_1) - Z(t,z+y_1;t+r,y_2).$$

Clearly,

$$I := \widetilde{\mathscr{L}^J} Z(t,\cdot;s,y_1)(x) - \widetilde{\mathscr{L}^J} Z(t,\cdot;s,y_2)(x) = \widetilde{\mathscr{L}^J} G_{s-t}(x-y_1).$$

If $|y_1-y_2| > \sqrt{s-t}$, then by (i), we have

$$\begin{aligned}|I| &\leqslant |\widetilde{\mathscr{L}^J} Z(t,\cdot;s,y_1)(x)| + |\widetilde{\mathscr{L}^J} Z(t,\cdot;s,y_2)(x)| \\ &\preceq \|J\|_\infty\Big(\eta_{\alpha,0}(s-t,y_1-x)+\eta_{\alpha,0}(s-t,y_2-x)\Big) \\ &\preceq \|J\|_\infty|y_1-y_2|^{\beta'\gamma}\Big(\eta_{\alpha,-\beta'\gamma}(s-t,y_1-x)+\eta_{\alpha,-\beta'\gamma}(s-t,y_2-x)\Big).\end{aligned} \tag{2.27}$$

If $|y_1-y_2| \leqslant \sqrt{s-t}$, then by (2.16), we have for $j=0,1$,

$$\begin{aligned}|\nabla^j G_{s-t}(z)| &\preceq |y_1-y_2|^{\beta'}\Big(\xi_{\lambda,-\beta'-j}(s-t,-z)+\xi_{\lambda,-\beta'-j}(s-t,y_2-y_1-z)\Big) \\ &\preceq |y_1-y_2|^{\beta'}\xi_{\lambda/2,-\beta'-j}(s-t,z),\end{aligned}$$

and by (2.15),

$$|\nabla^2 G_{s-t}(z)| \leqslant |\nabla_x^2 Z(t,z+y_1;s,y_1)| + |\nabla_x^2 Z(t,z+y_1;s,y_2)| \preceq \xi_{\lambda,-2}(s-t,z).$$

Hence, by (2.1) and Lemma 2.2 with $\beta_0=\beta_1=\beta'$ and $\beta_2=0$, we obtain that for $|y_1-y_2| \leqslant \sqrt{s-t}$,

$$\begin{aligned}|I| &\leqslant C\|J\|_\infty\Big(|y_1-y_2|^{\beta'}(s-t)^{-\frac{\beta'}{2}}+|y_1-y_2|^{\beta'\gamma}(s-t)^{-\frac{\beta'\gamma}{2}}\Big)\eta_{\alpha,0}(s-t,x-y_1) \\ &\leqslant C\|J\|_\infty|y_1-y_2|^{\beta'\gamma}\eta_{\alpha,-\beta'\gamma}(s-t,x-y_1).\end{aligned}$$

Combining this with (2.27), we obtain (2.24).

(iv) It follows from (2.15) and the same argument as above. □

2.3. **Kato's class.** We introduce the following Kato's class of space-time functions. Recall that functions $\xi_{\lambda,\gamma}(t,x)$ and $\eta_{\alpha,\gamma}(t,x)$ are defined in (1.6). For a function $g(t,x)$, we will use $g(t\pm s, x\pm y)$ as an abbreviation for $\sum_{j,k=0}^{1} g(t+(-1)^j s, x+(-1)^k y)$.

Definition 2.5. *(Generalized Kato's class) Let $\eta_{\alpha,\gamma}$ be given by* (1.6). *For $\alpha\in[1,+\infty)$, define*

$$\mathbb{K}_\alpha := \Big\{ f:\mathbb{R}\times\mathbb{R}^d\to\mathbb{R} \text{ satisfies } \lim_{\delta\to 0} K^f_\alpha(\delta)=0\Big\}, \tag{2.28}$$

$$\bar{\mathbb{K}}_\alpha := \Big\{ f:\mathbb{R}\times\mathbb{R}^d\to\mathbb{R} \text{ satisfies } \bar{K}^f_\alpha(1)<\infty\Big\}, \tag{2.29}$$

where

$$K^f_\alpha(\delta) := \sup_{(t,x)} \int_0^\delta\int_{\mathbb{R}^d} |f|(t\pm s, x\pm y)\eta_{\alpha,\alpha-1}(s,y)\mathrm{d}y\mathrm{d}s, \tag{2.30}$$

$$\bar{K}^f_\alpha(\delta) := \sup_{(t,x)} \int_0^\delta\int_{\mathbb{R}^d} |f|(t\pm s, x\pm y)\eta_{\alpha,0}(s,y)\mathrm{d}y\mathrm{d}s. \tag{2.31}$$

For $\alpha=+\infty$, we define

$$\mathbb{K}_\infty := \Big\{ f:\mathbb{R}\times\mathbb{R}^d\to\mathbb{R} \text{ satisfies } \lim_{\delta\to 0} N^f_\lambda(\delta)=0 \text{ for every } \lambda>0\Big\},$$

where

$$N^f_\lambda(\delta) := \sup_{(t,x)} \int_0^\delta\int_{\mathbb{R}^d} |f|(t\pm s, x\pm y)\xi_{\lambda,-1}(s,y)\mathrm{d}y\mathrm{d}s.$$

Remark 2.6. *$\mathbb{K}_\infty$ is the same as the Kato class defined in* [27]. *For any $\lambda>0$ and $\alpha\in[1,\infty)$, by (2.1), it is easy to see that there exists a constant $C=C(d,\alpha,\lambda)>0$ such that for all $\delta\in(0,1)$,*

$$N^f_\lambda(\delta)\leqslant CK^f_\alpha(\delta). \tag{2.32}$$

Moreover, for any time-independent function $f(t,x)=f(x)$, we have $f\in\mathbb{K}_\alpha$ if and only if

$$M^\alpha_f(\delta) := \sup_{x\in\mathbb{R}^d}\int_{\mathbb{R}^d} |f(x+y)|\cdot\frac{1}{|y|^{d-1}}\left(1\wedge\frac{\delta}{|y|^2}\right)^{(1+\alpha)/2}\mathrm{d}y\to 0 \quad \text{as } \delta\to 0.$$

Indeed, it follows by noticing that

$$\int_0^\delta \eta_{\alpha,\alpha-1}(s,y)\mathrm{d}s = \int_0^\delta \frac{s^{(\alpha-1)/2}\mathrm{d}s}{(|y|+s^{1/2})^{d+\alpha}} \asymp \frac{(|y|^2\wedge\delta)^{(1+\alpha)/2}}{|y|^{d+\alpha}} = \frac{1}{|y|^{d-1}}\left(1\wedge\frac{\delta}{|y|^2}\right)^{(1+\alpha)/2}.$$

We have the following results about the above Kato classes.

Proposition 2.7. *Let $\alpha\in[1,\infty)$, $p,q\in[1,\infty]$ and $f:\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}$.*

(i) *There is a constant $C=C(d,\alpha)>0$ such that for all $\delta>0$ and $j\in\mathbb{N}$,*

$$K^f_\alpha(j\delta)\leqslant CjK^f_\alpha(\delta),\quad \bar{K}^f_\alpha(j\delta)\leqslant Cj\bar{K}^f_\alpha(\delta) \tag{2.33}$$

and

$$\sup_{(t,x)}\int_0^\delta\int_{\mathbb{R}^d}|f|(t\pm s,x\pm y)\mathrm{d}y\mathrm{d}s\leqslant C\Big((\delta^{-(d+1)/2}K^f_\alpha(\delta))\wedge(\delta^{-(d+\alpha)/2}\bar{K}^f_\alpha(\delta))\Big). \tag{2.34}$$

(ii) *If* $\frac{d}{p}+\frac{2}{q}<1$, *then*

$$L^q(\mathbb{R};L^p(\mathbb{R}^d))\subset\mathbb{K}_1\subset\mathbb{K}_\alpha\subset\mathbb{K}_\infty;$$

and if $\alpha\in[1,2)$ *and* $\frac{d}{p}+\frac{2}{q}<2-\alpha$, *then*

$$L^q(\mathbb{R};L^p(\mathbb{R}^d))\subset\overline{\mathbb{K}}_\alpha\subset\mathbb{K}_\alpha.$$

Proof. (i) By definition (2.30), we have

$$\begin{aligned}K^f_\alpha(j\delta)&=\sup_{(t,x)}\sum_{k=0}^{j-1}\int_{k\delta}^{(k+1)\delta}\int_{\mathbb{R}^d}|f|(t\pm s,x\pm y)\eta_{\alpha,\alpha-1}(s,y)\mathrm{d}y\mathrm{d}s\\&\leqslant K^f_\alpha(\delta)+\sum_{k=1}^{j-1}\sup_{(t,x)}\int_0^\delta\int_{\mathbb{R}^d}|f|((t\pm k\delta)\pm s,x\pm y)\eta_{\alpha,\alpha-1}(s+k\delta,y)\mathrm{d}y\mathrm{d}s.\end{aligned}\tag{2.35}$$

Denoting the term in the above sum by I_k, by (2.3) and (2.2), we have for each $k=1,\cdots,j-1$,

$$\begin{aligned}I_k&\le\sup_{(t,x)}\int_0^\delta\int_{\mathbb{R}^d}|f|((t\pm k\delta)\pm s,x\pm y)\int_{\mathbb{R}^d}\eta_{\alpha,\alpha-1}(s,y-z)\eta_{\alpha,\alpha}(k\delta,z)\mathrm{d}z\mathrm{d}y\mathrm{d}s\\&=\sup_{(t,x)}\int_{\mathbb{R}^d}\left(\int_0^\delta\int_{\mathbb{R}^d}|f|((t\pm k\delta)\pm s,(x\pm z)\pm y)\eta_{\alpha,\alpha-1}(s,y)\mathrm{d}y\mathrm{d}s\right)\eta_{\alpha,\alpha}(k\delta,z)\mathrm{d}z\\&\leqslant K^f_\alpha(\delta)\int_{\mathbb{R}^d}\eta_{\alpha,\alpha}(k\delta,z)\mathrm{d}z=K^f_\alpha(\delta)\left(\int_{\mathbb{R}^d}\eta_{\alpha,\alpha}(1,z)\mathrm{d}z\right).\end{aligned}$$

Substituting this into (2.35), we get the first inequality in (2.33). Similarly, we can prove the second inequality in (2.33). On the other hand, for any $(t,x)\in\mathbb{R}\times\mathbb{R}^d$,

$$\begin{aligned}\int_0^\delta\int_{\mathbb{R}^d}|f|(t\pm s,x\pm y)\mathrm{d}y\mathrm{d}s&=\int_\delta^{2\delta}\int_{\mathbb{R}^d}|f|((t\pm\delta)\pm s,x\pm y)\frac{\eta_{\alpha,\alpha-1}(s,y)}{\eta_{\alpha,\alpha-1}(s,y)}\mathrm{d}y\mathrm{d}s\\&\leqslant\frac{(2\delta)^{(\alpha-1)/2}}{\delta^{(d+\alpha)/2}}\sup_{(t,x)}\int_\delta^{2\delta}\int_{\mathbb{R}^d}|f|((t\pm\delta)\pm s,x\pm y)\eta_{\alpha,\alpha-1}(s,y)\mathrm{d}y\mathrm{d}s\\&\leqslant2^{(\alpha-1)/2}\delta^{-(d+1)/2}K^f_\alpha(2\delta)\lesssim\delta^{-(d+1)/2}K^f_\alpha(\delta).\end{aligned}$$

Similar, we have

$$\int_0^\delta\int_{\mathbb{R}^d}|f|(t\pm s,x\pm y)\mathrm{d}y\mathrm{d}s\lesssim\delta^{-(d+\alpha)/2}\bar{K}^f_\alpha(\delta).$$

(ii) The inclusions of $\mathbb{K}_1\subset\mathbb{K}_\alpha\subset\mathbb{K}_\infty$ and $\overline{\mathbb{K}}_\alpha\subset\mathbb{K}_\alpha$ follow by definitions and (2.32). Let us prove $L^q(\mathbb{R};L^p(\mathbb{R}^d))\subset\cap_{\alpha\geqslant1}\mathbb{K}_\alpha$. Let $f\in L^q(\mathbb{R};L^p(\mathbb{R}^d))$. By Hölder's inequality, we have

$$K^f_\alpha(\delta)\leqslant\left(\int_{\mathbb{R}}\left(\int_{\mathbb{R}^d}|f(s,y)|^p\mathrm{d}y\right)^{\frac{q}{p}}\mathrm{d}s\right)^{\frac{1}{q}}I_\alpha(\delta),$$

where

$$I_\alpha(\delta):=\left(\int_0^\delta\left(\int_{\mathbb{R}^d}\frac{s^{(\alpha-1)p^*/2}\mathrm{d}y}{(|y|+s^{1/2})^{(d+\alpha)p^*}}\right)^{\frac{q^*}{p^*}}\mathrm{d}s\right)^{\frac{1}{q^*}},$$

with $q^*:=\frac{q}{q-1}$ and $p^*:=\frac{p}{p-1}$. Noticing that

$$\int_{\mathbb{R}^d}\frac{s^{(\alpha-1)p^*/2}\mathrm{d}y}{(|y|+s^{1/2})^{(d+\alpha)p^*}}\leqslant s^{(\alpha-1)p^*/2}\left(\int_{|y|\leqslant s^{1/2}}s^{-\frac{(d+\alpha)p^*}{2}}\mathrm{d}y+\int_{|y|>s^{1/2}}\frac{\mathrm{d}y}{|y|^{(d+\alpha)p^*}}\right)\lesssim s^{\frac{d-(d+1)p^*}{2}},$$

we have

$$I_\alpha(\delta) \le \left(\int_0^\delta s^{\frac{dq^*}{2p^*}-\frac{(d+1)q^*}{2}}\mathrm{d}s\right)^{\frac{1}{q^*}}.$$

Thus $I_\alpha(\delta)$ converges to zero as $\delta \to 0$ provided that

$$\frac{dq^*}{2p^*} - \frac{d+1}{2}q^* + 1 > 0 \Leftrightarrow \frac{d}{p} + \frac{2}{q} < 1.$$

Similarly, we can show that $L^q(\mathbb{R}; L^p(\mathbb{R}^d)) \subset \overline{\mathbb{K}}_\alpha$ provided $\frac{d}{p} + \frac{2}{q} < 2 - \alpha$. □

Next we study the mollifying approximation of $f \in \mathbb{K}_\alpha$. Let $\rho(t,x) : \mathbb{R}^{d+1} \to [0,1]$ be a smooth function with support in the unit ball and $\int \rho = 1$. For $\varepsilon \in (0,1)$, define a family of mollifiers ρ_ε as follows:

$$\rho_\varepsilon(t,x) := \varepsilon^{-d-1}\rho(\varepsilon^{-1}t, \varepsilon^{-1}x).$$

For $f \in \mathbb{K}_\alpha$, we define

$$f_\varepsilon(t,x) := f * \rho_\varepsilon(t,x) = \int_{\mathbb{R}^{d+1}} f(s,y)\rho_\varepsilon(t-s, x-y)\mathrm{d}y\mathrm{d}s. \tag{2.36}$$

By Fubini's theorem, it is easy to see that

$$K_\alpha^{f_\varepsilon}(\delta) \leqslant K_\alpha^f(\delta). \tag{2.37}$$

Lemma 2.8. *For $\alpha \in [1,\infty)$ and $f \in \mathbb{K}_\alpha$, we have*

$$\lim_{\varepsilon\to 0}\int_0^1\int_{\mathbb{R}^d} |f_\varepsilon - f|(t \pm s, x \pm y)\eta_{\alpha,\alpha-1}(s,y)\mathrm{d}y\mathrm{d}s = 0.$$

Proof. First of all, notice that

$$\lim_{\delta\to 0}\sup_{\varepsilon\in(0,1)}\int_0^\delta\int_{\mathbb{R}^d} |f_\varepsilon - f|(t \pm s, x \pm y)\eta_{\alpha,\alpha-1}(s,y)\mathrm{d}y\mathrm{d}s \overset{(2.37)}{\leqslant} 2\lim_{\delta\to 0} K_\alpha^f(\delta) = 0.$$

So, it suffices to prove that for fixed $\delta \in (0,1)$,

$$\lim_{\varepsilon\to 0}\int_\delta^1\int_{\mathbb{R}^d} \frac{|f_\varepsilon - f|(t \pm s, x \pm y)s^{(\alpha-1)/2}}{(|y| + s^{1/2})^{d+\alpha}}\mathrm{d}y\mathrm{d}s = 0. \tag{2.38}$$

Let $f^n(t,x) := (-n) \vee f(t,x) \wedge n$ and $f^n_\varepsilon := f^n * \rho_\varepsilon$. Since ρ_ε has support in $\{(t,x) : |(t,x)| \leqslant \varepsilon\}$, by the definition of convolution, we have

$$\begin{aligned}
&\sup_{\varepsilon\in(0,\delta/4)}\int_\delta^1\int_{\mathbb{R}^d} \frac{|f^n_\varepsilon - f_\varepsilon|(t \pm s, x \pm y)s^{(\alpha-1)/2}}{(|y| + s^{1/2})^{d+\alpha}}\mathrm{d}y\mathrm{d}s \\
&\quad\leqslant \int_{3\delta/4}^{1+\delta/4}\int_{\mathbb{R}^d} \frac{|f^n - f|(t \pm s, x \pm y)(s+\delta/4)^{(\alpha-1)/2}}{(|y| - \delta/4 + (s-\delta/4)^{1/2})^{d+\alpha}}\mathrm{d}y\mathrm{d}s \\
&\quad\leqslant \int_{3\delta/4}^{1+\delta/4}\int_{\mathbb{R}^d} \frac{|f^n - f|(t \pm s, x \pm y)(2s)^{(\alpha-1)/2}}{(|y| + s^{1/2}/4)^{d+\alpha}}\mathrm{d}y\mathrm{d}s,
\end{aligned} \tag{2.39}$$

which converges to zero as $n \to \infty$ by the dominated convergence theorem. On the other hand, for fixed $n \in \mathbb{N}$, since $\lim_{\varepsilon\to 0} f^n_\varepsilon = f^n$ a.e., by the bounded convergence theorem, we have

$$\lim_{\varepsilon\to 0}\int_\delta^1\int_{\mathbb{R}^d} \frac{|f^n_\varepsilon - f^n|(t \pm s, x \pm y)s^{(\alpha-1)/2}}{(|y| + s^{1/2})^{d+\alpha}}\mathrm{d}y\mathrm{d}s = 0.$$

Combining this with (2.39), we obtain (2.38). □

3. Proof of Theorem 1.1

In the remaining part of this paper, we shall fix $\alpha \in (0,2)$ and assume **(H^a)**, **(H$^\kappa$)** and $b \in \mathbb{K}_2$. Below, a function $f(t,x)$ on $[0,\infty)\times\mathbb{R}^d$ will be automatically extended to $\mathbb{R}\times\mathbb{R}^d$ by letting $f(t,\cdot)=0$ for $t<0$. Notice that

$$\mathscr{L}_t^{b,\kappa} := b_t\cdot\nabla + \mathscr{L}_t^\kappa = \tilde{b}_t\cdot\nabla + \widetilde{\mathscr{L}}^{\kappa(t,\cdot)}, \tag{3.1}$$

where $\widetilde{\mathscr{L}}^{\kappa(t,\cdot)}$ is defined by (2.9), and

$$\tilde{b}(t,x) := b(t,x) + 1_{\alpha\in(1,2)}\int_{|z|>1} z\kappa(t,x,z)|z|^{-d-\alpha}\mathrm{d}z - 1_{\alpha\in(0,1)}\int_{|z|\leqslant 1} z\kappa(t,x,z)|z|^{-d-\alpha}\mathrm{d}z. \tag{3.2}$$

By definition, it is easy to see that for some $c=c(d,\alpha)>0$,

$$K_2^{\tilde{b}}(r) \leqslant K_2^b(r) + c\|\kappa\|_\infty r^{1/2}, \quad r>0. \tag{3.3}$$

Let $Z(t,x;s,y)$ be the heat kernel of $\mathscr{L}_t^a$ constructed in Theorem 2.3. We will construct the fundamental solution $p(t,x;s,y)$ of $\mathscr{L}_t$ by using Duhamel's formula (1.12). To solve that integral equation, let $p_0(t,x;s,y):=Z(t,x;s,y)$, and for $n\in\mathbb{N}$, define

$$p_n(t,x;s,y) := \int_t^s\int_{\mathbb{R}^d} p_{n-1}(t,x;r,z)\mathscr{L}_r^{b,\kappa}Z(r,\cdot;s,y)(z)\mathrm{d}z\mathrm{d}r. \tag{3.4}$$

We first prepare the following lemma for later use.

Lemma 3.1. *For any $\lambda>0$ and $j=0,1$, there exists a constant $C_j=C_j(d,\alpha,\lambda)>0$ such that for all $(t,x;s,y)\in\mathbb{D}_0^\infty$,*

$$\begin{aligned}&\int_t^s\int_{\mathbb{R}^d}\eta_{\alpha,2-j}(r-t,z-x)|b(r,z)|\xi_{\lambda,-1}(s-r,y-z)\mathrm{d}z\mathrm{d}r\\ &\qquad\leqslant C_jK_2^b(s-t)\eta_{\alpha,2-j}(s-t,y-x),\end{aligned}\tag{3.5}$$

and

$$\begin{aligned}&\int_t^s\int_{\mathbb{R}^d}\xi_{\lambda,-j}(r-t,z-x)|b(r,z)|\xi_{2\lambda,-1}(s-r,y-z)\mathrm{d}z\mathrm{d}r\\ &\qquad\leqslant C_jK_2^b(s-t)\xi_{\lambda,-j}(s-t,y-x),\end{aligned}\tag{3.6}$$

where $K_2^b(s-t)$ is defined by (2.30).

Proof. Notice that by (2.1),

$$\xi_{\lambda,-j}(s-t,y-x)\leq\eta_{2,2-j}(s-t,y-x),\quad j=0,1.$$

Hence, by (2.4), we have

$$\begin{aligned}&\int_t^s\int_{\mathbb{R}^d}\eta_{\alpha,2-j}(r-t,z-x)\,|b(r,z)|\,\xi_{\lambda,-1}(s-r,y-z)\mathrm{d}z\mathrm{d}r\\
&\leq\eta_{\alpha,0}(s-t,y-x)\int_t^s\int_{\mathbb{R}^d}(r-t)^{(2-j)/2}(s-r)^{1/2}|b(r,z)|\\
&\qquad\times\Big(\eta_{2,0}(r-t,z-x)+\eta_{2,0}(s-r,y-z)\Big)\mathrm{d}z\mathrm{d}r\\
&\leq\eta_{\alpha,2-j}(s-t,y-x)\int_t^s\int_{\mathbb{R}^d}|b(r,z)|\\
&\qquad\times\Big(\eta_{2,1}(r-t,z-x)+\eta_{2,1}(s-r,y-z)\Big)\mathrm{d}z\mathrm{d}r\\
&\leqslant C_jK_2^b(s-t)\eta_{\alpha,2-j}(s-t,y-x),\end{aligned}$$

where the last step is due to the change of variables and the definition of K_2^b. Thus (3.5) is proved. Estimate (3.6) follows from [27, Lemma 3.1] and (2.32). $\square$

Lemma 3.2. *For each $n \in \mathbb{N}$ and $j = 0, 1$, $\nabla_x^j p_n(t, x; s, y)$ is a jointly continuous function on $\mathbb{D}_0^\infty$, and for any $T > 0$, there exist constants $c, \lambda > 0$ such that for all $n \in \mathbb{N}$ and $(t, x; s, y) \in \mathbb{D}_0^T$,*

$$
\begin{aligned}
|\nabla_x^j p_n(t, x; s, y)| \leqslant\ & c(c\ell_{b,\kappa}(s-t))^{n-1}\|\kappa\|_\infty \eta_{\alpha,2-j}(s-t, y-x) \\
&+ (c\ell_{b,\kappa}(s-t))^n \xi_{\lambda,-j}(s-t, y-x),
\end{aligned}
\tag{3.7}
$$

where $\ell_{b,\kappa}(r) := \|\kappa\|_\infty \left(r^{1-\frac{\alpha}{2}} + r^{1/2}\right) + K_2^b(r)$.

Proof. (1) First of all, by definition, **(H$^\kappa$)**, (2.22) and (2.15), there is a $\lambda > 0$ such that

$$
|\mathscr{L}^{\kappa(t,\cdot)} Z(t, \cdot; s, y)(x)| \leq \|\kappa\|_\infty \eta_{\alpha,0}(s-t, y-x), \quad |\nabla_x Z(t, x; s, y)| \leq \xi_{2\lambda,-1}(s-t, y-x). \tag{3.8}
$$

For $r > 0$, let

$$
\widetilde{\ell}_{b,\kappa}(r) := \|\kappa\|_\infty r^{1-\frac{\alpha}{2}} + K_2^{\tilde b}(r),
$$

where $\tilde b$ is defined by (3.2). In view of (3.1) and (3.3), it is enough to prove (3.7) with $\widetilde{\ell}_{b,\kappa}$ instead of $\ell_{b,\kappa}$. For $n = 1$, by (2.6) and (3.6) we have

$$
\begin{aligned}
|\nabla_x^j p_1(t, x; s, y)| \leq & \int_t^s \int_{\mathbb{R}^d} \xi_{\lambda,-j}(r-t, z-x) \Big(|\tilde b(r, z)| \cdot \xi_{2\lambda,-1}(s-r, y-z) \\
& \qquad + \|\kappa\|_\infty \eta_{\alpha,0}(s-r, y-z) \Big) \mathrm{d}z\mathrm{d}r \\
\leqslant\ & c_0 \widetilde{\ell}_{b,\kappa}(s-t) \cdot \xi_{\lambda,-j}(s-t, y-x) + c_1 \|\kappa\|_\infty \eta_{\alpha,2-j}(s-t, y-x).
\end{aligned}
$$

Suppose that (3.7) holds for $\widetilde{\ell}_{b,\kappa}$ and for some $n \in \mathbb{N}$. By (3.8) and the induction hypothesis, we have

$$
\begin{aligned}
& |\nabla_x^j p_{n+1}(t, x; s, y)| \\
\leq & \int_t^s \int_{\mathbb{R}^d} \Big(c(c\widetilde{\ell}_{b,\kappa}(r-t))^{n-1} \|\kappa\|_\infty \eta_{\alpha,2-j}(r-t, z-x) + (c\widetilde{\ell}_{b,\kappa}(r-t))^n \xi_{\lambda,-j}(r-t, z-x) \Big) \\
& \quad \times \Big(|\tilde b(r, z)| \cdot \xi_{2\lambda,-1}(s-r, y-z) + \|\kappa\|_\infty \eta_{\alpha,0}(s-r, y-z) \Big) \mathrm{d}z\mathrm{d}r \\
\leq\ & c(c\widetilde{\ell}_{b,\kappa}(s-t))^{n-1} \|\kappa\|_\infty (I_1 + I_2) + (c\widetilde{\ell}_{b,\kappa}(s-t))^n (I_3 + I_4),
\end{aligned}
$$

where

$$
\begin{aligned}
I_1 &:= \int_t^s \int_{\mathbb{R}^d} \eta_{\alpha,2-j}(r-t, z-x) \cdot |\tilde b(r, z)| \cdot \xi_{2\lambda,-1}(s-r, y-z) \mathrm{d}z\mathrm{d}r, \\
I_2 &:= \|\kappa\|_\infty \int_t^s \int_{\mathbb{R}^d} \eta_{\alpha,2-j}(r-t, z-x) \cdot \eta_{\alpha,0}(r-t, z-x) \mathrm{d}z\mathrm{d}r, \\
I_3 &:= \int_t^s \int_{\mathbb{R}^d} \xi_{\lambda,-j}(r-t, z-x) \cdot |\tilde b(r, z)| \cdot \xi_{2\lambda,-1}(s-r, y-z) \mathrm{d}z\mathrm{d}r, \\
I_4 &:= \|\kappa\|_\infty \int_t^s \int_{\mathbb{R}^d} \xi_{\lambda,-j}(r-t, z-x) \cdot \eta_{\alpha,0}(r-t, z-x) \mathrm{d}z\mathrm{d}r.
\end{aligned}
$$

By (3.5) and (2.5), one sees that

$$
I_1 + I_2 \preceq \widetilde{\ell}_{b,\kappa}(s-t) \cdot \eta_{\alpha,2-j}(s-t, y-x),
$$

and by (3.6) and (2.6),

$$
I_3 \preceq \widetilde{\ell}_{b,\kappa}(s-t) \cdot \xi_{\lambda,-j}(s-t, y-x), \quad I_4 \preceq \|\kappa\|_\infty \eta_{\alpha,2-j}(s-t, y-x).
$$

Therefore,

$$
\begin{aligned}
|\nabla_x^j p_{n+1}(t,x;s,y)| \leqslant\ & c_2 c(c\widetilde{\ell}_{b,\kappa}(s-t))^{n-1}\widetilde{\ell}_{b,\kappa}(s-t)\cdot\|\kappa\|_\infty \eta_{\alpha,2-j}(s-t,y-x)\\
&+(c\widetilde{\ell}_{b,\kappa}(s-t))^n\Big(c_3\widetilde{\ell}_{b,\kappa}(s-t)\cdot\xi_{\lambda,-j}(s-t,y-x)+c_4\|\kappa\|_\infty\eta_{\alpha,2-j}(s-t,y-x)\Big)\\
\leqslant\ & (c_2+c_4)(c\widetilde{\ell}_{b,\kappa}(s-t))^n\cdot\|\kappa\|_\infty\eta_{\alpha,2-j}(s-t,y-x)\\
&+c_3(c\widetilde{\ell}_{b,\kappa}(s-t))^n\widetilde{\ell}_{b,\kappa}(s-t)\cdot\xi_{\lambda,-j}(s-t,y-x).
\end{aligned}
$$

Finally, by choosing $c=c_0\vee c_1\vee(c_2+c_4)\vee c_3$, we obtain the desired result.

(2) We use induction to show the joint continuity of $\nabla_x^j p_n$. First of all, by Theorem 2.3, $\nabla_x^j p_0$ is jointly continuous. Suppose now that $\nabla_x^j p_{n-1}$ is jointly continuous for some $n\in\mathbb{N}$. For fixed $\varepsilon>0$ and $\delta\in(0,\varepsilon/2)$, we write

$$
\begin{aligned}
\nabla_x^j p_n(t,x;s,y) &= \left(\int_{s-\delta}^{s}+\int_t^{t+\delta}\right)\int_{\mathbb{R}^d}\nabla_x^j p_{n-1}(t,x;r,z)\mathscr{L}_r^{b,\kappa}Z(r,\cdot;s,y)(z)\mathrm{d}z\mathrm{d}r\\
&\quad+\int_{t+\delta}^{s-\delta}\int_{\mathbb{R}^d}\nabla_x^j p_{n-1}(t,x;r,z)\mathscr{L}_r^{b,\kappa}Z(r,\cdot;s,y)(z)\mathrm{d}z\mathrm{d}r\\
&=: I_1(\delta,t,x;s,y)+I_2(\delta,t,x;s,y).
\end{aligned}
$$

For I_1, as in step (1), there is a $C_\varepsilon>0$ such that for all $(t,x;s,y)\in\mathbb{D}_\varepsilon^T$ and $\delta\in(0,\varepsilon/2)$,

$$
|I_1(\delta,t,x;s,y)|\leqslant C_\varepsilon\ell_{b,\kappa}(\delta). \tag{3.9}
$$

For I_2, by the dominated convergence theorem and induction hypothesis, one sees that for fixed $\delta\in(0,\varepsilon/2)$,

$$
(t,x;s,y)\mapsto I_2(\delta,t,x;s,y)\text{ is continuous on }\mathbb{D}_\varepsilon^T.
$$

Combining this with (3.9), we obtain the continuity of $\nabla_x^j p_n$ on $\mathbb{D}_\varepsilon^T$. Since $\varepsilon,T>0$ are arbitrary, $\nabla_x^j p_n$ is jointly continuous on $\mathbb{D}_0^\infty$. The proof is complete. □

Lemma 3.3. *Let $J:\mathbb{R}^d\to\mathbb{R}$ be a bounded measurable function satisfying* (2.11).

(i) *If $\alpha\in(0,1]$ and $b\in\mathbb{K}_1$, then for any $T>0$, there is a constant $c>0$ such that for all $n\in\mathbb{N}_0$ and $(t,x;s,y)\in\mathbb{D}_0^T$,*

$$
|\widetilde{\mathscr{L}^J} p_n(t,\cdot;s,y)(x)|\leqslant c(c\ell_{b,\kappa}(s-t))^n\|J\|_\infty\eta_{\alpha,0}(s-t,y-x), \tag{3.10}
$$

where $\ell_{b,\kappa}(r)=\|\kappa\|_\infty\left(r^{1-\frac{\alpha}{2}}+r^{1/2}\right)+K_1^b(r)$.

(ii) *If $\alpha\in(1,2)$ and $b\in\overline{\mathbb{K}}_\alpha$, then* (3.10) *still holds with $\ell_{b,\kappa}(r):=\|\kappa\|_\infty\left(r^{1-\frac{\alpha}{2}}+r^{1/2}\right)+r^{\frac{\alpha-1}{2}}\overline{K}_\alpha^b(1)$.*

Moreover, for the above two classes of Kato's functions b, we have

$$
p_{n+1}(t,x;s,y)=\int_t^s\int_{\mathbb{R}^d}Z(t,x;r,z)\mathscr{L}_r^{b,\kappa}p_n(r,\cdot;s,y)(z)\mathrm{d}z\mathrm{d}r. \tag{3.11}
$$

Proof. As before, we set for $r>0$

$$
\widetilde{\ell}_{b,\kappa}(r):=\|\kappa\|_\infty r^{1-\frac{\alpha}{2}}+K_2^{\tilde{b}}(r).
$$

By (3.1) and (3.3), we only need to prove (3.10) with $\widetilde{\ell}_{b,\kappa}$ instead of $\ell_{b,\kappa}$. By (2.22), one sees that (3.10) and (3.11) hold for $n=0$. Now suppose that (3.10) and (3.11) hold for $\widetilde{\ell}_{b,\kappa}$ and for some $n\in\mathbb{N}_0$. By Fubini's theorem, (2.22), (3.8) and (2.5), we have

$$
|\widetilde{\mathscr{L}_t^J} p_{n+1}(t,\cdot;s,y)(x)|=\left|\int_t^s\int_{\mathbb{R}^d}\widetilde{\mathscr{L}_t^J}p_n(t,\cdot;r,z)(x)\mathscr{L}_r^{b,\kappa}Z(r,\cdot;s,y)(z)\mathrm{d}z\mathrm{d}r\right|
$$

$$
\begin{aligned}
&\leq \int_t^s\int_{\mathbb{R}^d} (c\widetilde{\ell}_{b,\kappa}(r-t))^n\|J\|_\infty\eta_{\alpha,0}(r-t,z-x)\\
&\quad\times\Big(|\tilde{b}(r,z)|\cdot\xi_{2\lambda,-1}(s-r,y-z)+\|\kappa\|_\infty\eta_{\alpha,0}(s-r,y-z)\Big)\mathrm{d}z\mathrm{d}r\\
&=(c\widetilde{\ell}_{b,\kappa}(s-t))^n\|J\|_\infty(I_1+I_2),
\end{aligned}
$$

where

$$
\begin{aligned}
I_1 &:= \int_t^s\int_{\mathbb{R}^d}\eta_{\alpha,0}(r-t,z-x)\cdot|\tilde{b}(r,z)|\cdot\xi_{2\lambda,-1}(s-r,y-z)\mathrm{d}z\mathrm{d}r,\\
I_2 &:= \|\kappa\|_\infty\int_t^s\int_{\mathbb{R}^d}\eta_{\alpha,0}(r-t,z-x)\cdot\eta_{\alpha,0}(s-r,y-z)\mathrm{d}z\mathrm{d}r.
\end{aligned}
$$

For I_1, by (2.1) and (2.4), we have

$$
\begin{aligned}
I_1 &\leq (s-t)^{(\alpha\vee 1-1)/2}\int_t^s\int_{\mathbb{R}^d}\eta_{\alpha,0}(r-t,z-x)\cdot|\tilde{b}(r,z)|\cdot\eta_{\alpha\vee 1,0}(s-r,y-z)\mathrm{d}z\mathrm{d}r\\
&\leq \eta_{\alpha,\alpha\vee 1-1}(s-t,y-x)\int_t^s\int_{\mathbb{R}^d}|\tilde{b}(r,z)|\Big(\eta_{\alpha\vee 1,0}(r-t,z-x)+\eta_{\alpha\vee 1,0}(s-r,y-z)\Big)\mathrm{d}z\mathrm{d}r.
\end{aligned}
$$

If $\alpha\in(0,1]$, then

$$
I_1\leq\eta_{\alpha,0}(s-t,y-x)K_1^{\tilde{b}}(s-t).
$$

If $\alpha\in(1,2)$, then

$$
I_1\leq\eta_{\alpha,\alpha-1}(s-t,y-x)\overline{K}_\alpha^{\tilde{b}}(s-t)=(s-t)^{\frac{\alpha-1}{2}}\overline{K}_\alpha^{\tilde{b}}(s-t)\eta_{\alpha,0}(s-t,y-x).
$$

For I_2, by (2.5) we have

$$
I_2\leq\|\kappa\|_\infty\eta_{\alpha,2-\alpha}(s-t,y-x)=\|\kappa\|_\infty(s-t)^{1-\frac{\alpha}{2}}\eta_{\alpha,0}(s-t,y-x).
$$

Combining the above calculations, we obtain (3.10).

Moreover, by Fubini's theorem again and the induction hypothesis, we have

$$
\begin{aligned}
p_{n+2}(t,x;s,y) &= \int_t^s\int_{\mathbb{R}^d}p_{n+1}(t,x;r,z)\mathscr{L}_r^{b,\kappa}Z(r,\cdot;s,y)(z)\mathrm{d}z\mathrm{d}r\\
&=\int_t^s\int_{\mathbb{R}^d}\int_t^r\int_{\mathbb{R}^d}Z(t,x;r',z')\mathscr{L}_{r'}^{b,\kappa}p_n(r',\cdot;r,z)(z')\mathrm{d}z'\mathrm{d}r'\\
&\qquad\qquad\times\mathscr{L}_r^{b,\kappa}Z(r,\cdot;s,y)(z)\mathrm{d}z\mathrm{d}r\\
&=\int_t^s\int_{\mathbb{R}^d}Z(t,x;r',z')\int_{r'}^s\int_{\mathbb{R}^d}\mathscr{L}_{r'}^{b,\kappa}p_n(r',\cdot;r,z)(z')\\
&\qquad\qquad\times\mathscr{L}_r^{b,\kappa}Z(r,\cdot;s,y)(z)\mathrm{d}z\mathrm{d}r\mathrm{d}z'\mathrm{d}r'\\
&=\int_t^s\int_{\mathbb{R}^d}Z(t,x;r',z')\mathscr{L}_r^{b,\kappa}p_{n+1}(r',\cdot;s,y)(z')\mathrm{d}z'\mathrm{d}r'.
\end{aligned}
$$

The proof is complete. □

Under additional regularity assumptions on b and κ, we can show further regularity of $p_n(t,x;s,y)$ as given in the following lemma.

Lemma 3.4. *If b and κ are bounded measurable and for some $C_0>0$ and $\gamma\in(0,1)$,*

$$
|b(t,y)-b(t,x)|+|\kappa(t,y,z)-\kappa(t,x,z)|\leqslant C_0|y-x|^\gamma,\quad t\in\mathbb{R}_+,\ x,y,z\in\mathbb{R}^d,\tag{3.12}
$$

then $\nabla^2_x p_n$ is continuous on $\mathbb{D}_0^\infty$, and for any $T > 0$, there are two constants $\lambda > 0$ and $C_1 > 0$ such that for all $n \in \mathbb{N}$ and $(t, x; s, y) \in \mathbb{D}_0^T$,

$$|\nabla^2_x p_n(t, x; s, y)| \leqslant C_1\Big(C_1(s-t)^{\frac{(2-\alpha)\wedge 1}{2}}\Big)^{n-1}(\eta_{\alpha,0} + \xi_{\lambda/2,-1})(s-t, y-x). \tag{3.13}$$

Proof. Below, we always assume that $0 < s - t \leqslant T$ and $x_1, x_2, x, y \in \mathbb{R}^d$.
(1) By (2.15), (2.25) and (3.12), we have

$$|\tilde{b}_t \cdot \nabla Z(t,\cdot; s, y)(x_1) - \tilde{b}_t \cdot \nabla Z(t,\cdot; s, y)(x_2)| \preceq |x_1 - x_2|^\gamma \Big(\sum_i \xi_{\lambda/2,-\gamma-1}(s-t, y-x_i)\Big),$$

and by (2.23), (2.22) and (3.12),

$$|\widetilde{\mathscr{L}}^{\kappa(t,\cdot)} Z(t,\cdot; s, y)(x_1) - \widetilde{\mathscr{L}}^{\kappa(t,\cdot)} Z(t,\cdot; s, y)(x_2)| \preceq |x_1 - x_2|^\gamma \Big(\sum_i \eta_{\alpha,-\gamma}(s-t, y-x_i)\Big).$$

Let $H_{s,y}(t, x) := \mathscr{L}^{b,\kappa}_t Z(t,\cdot; s, y)(x)$. By the above two estimates, we have

$$|H_{s,y}(t, x_1) - H_{s,y}(t, x_2)| \preceq |x_1 - x_2|^\gamma \Big(\sum_i (\eta_{\alpha,-\gamma} + \xi_{\lambda/2,-\gamma-1})(s-t, y-x_i)\Big). \tag{3.14}$$

Moreover, by (2.15) and (2.22), we also have

$$|H_{s,y}(t, x)| \preceq (\eta_{\alpha,0} + \xi_{\lambda,-1})(s-t, y-x). \tag{3.15}$$

(2) We use induction to prove (3.13). First of all, for $n = 1$, by (6.27) below, we have

$$\nabla^2_x p_1(t, x; s, y) = \int_t^s \int_{\mathbb{R}^d} \nabla^2_x p_0(t, x; r, z) H_{s,y}(r, z)\mathrm{d}z\mathrm{d}r = I_1 + I_2 + I_3,$$

where

$$\begin{aligned}
I_1 &:= \int_{\frac{s+t}{2}}^s \int_{\mathbb{R}^d} \nabla^2_x Z(t, x; r, z) H_{s,y}(r, z)\mathrm{d}z\mathrm{d}r,\\
I_2 &:= \int_t^{\frac{s+t}{2}} \int_{\mathbb{R}^d} \nabla^2_x Z(t, x; r, z)(H_{s,y}(r, z) - H_{s,y}(r, x))\mathrm{d}z\mathrm{d}r,\\
I_3 &:= \int_t^{\frac{s+t}{2}} \left(\int_{\mathbb{R}^d} \nabla^2_x Z(t, x; r, z)\mathrm{d}z\right) H_{s,y}(r, x)\mathrm{d}r.
\end{aligned}$$

For I_1, by (2.15), (3.15), (2.6) and (2.8), we have

$$\begin{aligned}
|I_1| &\preceq \int_{\frac{s+t}{2}}^s \int_{\mathbb{R}^d} \xi_{\lambda,-2}(r-t, z-x)(\eta_{\alpha,0} + \xi_{\lambda,-1})(s-r, y-z)\mathrm{d}z\mathrm{d}r\\
&\preceq (s-t)^{-1} \int_t^s \int_{\mathbb{R}^d} \xi_{\lambda,0}(r-t, z-x)(\eta_{\alpha,0} + \xi_{\lambda,-1})(s-r, y-z)\mathrm{d}z\mathrm{d}r\\
&\preceq \eta_{\alpha,0}(s-t, y-x) + \xi_{\lambda,-1}(s-t, y-x).
\end{aligned}$$

For I_2, by (2.15), (3.14), (2.6) and (2.8), we similarly have

$$\begin{aligned}
|I_2| \preceq & \int_t^{\frac{s+t}{2}} \int_{\mathbb{R}^d} \xi_{\lambda,-2}(r-t, z-x)|x-z|^\gamma \cdot \Big((\eta_{\alpha,-\gamma} + \xi_{\lambda,-\gamma-1})(s-r, y-z)\\
&\qquad + (\eta_{\alpha,-\gamma} + \xi_{\lambda,-\gamma-1})(s-r, y-x)\Big)\mathrm{d}z\mathrm{d}r\\
\preceq & \int_t^{\frac{s+t}{2}} \int_{\mathbb{R}^d} \xi_{\lambda/2,\gamma-2}(r-t, z-x) \cdot \Big((\eta_{\alpha,-\gamma} + \xi_{\lambda,-\gamma-1})(s-r, y-z)\\
&\qquad + (\eta_{\alpha,-\gamma} + \xi_{\lambda,-\gamma-1})(s-r, y-x)\Big)\mathrm{d}z\mathrm{d}r\\
\preceq & \ \eta_{\alpha,0}(s-t, y-x) + \xi_{\lambda/2,-1}(s-t, y-x).
\end{aligned}$$

For I_3, by (6.31) below and (3.15), we have

$$|I_3| \preceq \left(\int_t^{\frac{s+t}{2}} (r-t)^{\frac{\beta}{2}-1}\mathrm{d}r\right)(\eta_{\alpha,0}+\xi_{\lambda,-1})(s-t,y-x) \preceq (\eta_{\alpha,0}+\xi_{\lambda,-1})(s-t,y-x).$$

Combining the above calculations, we obtain (3.13) for $n=1$.

(3) Suppose (3.13) holds for some $n\in\mathbb{N}$. By the induction hypothesis, (3.15) and Lemma 2.1,

$$\begin{aligned}|\nabla_x^2 p_{n+1}(t,x;s,y)| &\leqslant C_1(C_1(s-t)^{\frac{(2-\alpha)\wedge 1}{2}})^{n-1}\\ &\times \int_t^s\int_{\mathbb{R}^d}(\eta_{\alpha,0}+\xi_{\lambda/2,-1})(r-t,z-x)(\eta_{\alpha,0}+\xi_{\lambda,-1})(s-r,y-z)\mathrm{d}z\mathrm{d}r\\ &\leqslant C_1(C_1(s-t)^{\frac{(2-\alpha)\wedge 1}{2}})^{n}(\eta_{\alpha,0}+\xi_{\lambda/2,-1})(s-t,y-x).\end{aligned}$$

Thus we obtain (3.13).

(4) The joint continuity of $\nabla_x^2 p_n$ follows by the same argument as in Lemma 3.2. The proof is complete. □

Now we can prove the solvability of the integral equation (1.12).

Theorem 3.5. *Under* **(H^a)**, **(H$^\kappa$)** *and $b\in\mathbb{K}_2$, there exists a $\delta>0$ so that* (1.12) *has a unique continuous solution $p(t,x;s,y)$ on $\mathbb{D}_0^\delta$ such that*

$$|p(t,x;s,y)| \leqslant C_1(\xi_{\lambda,0}+\|\kappa\|_\infty\eta_{\alpha,2})(s-t,y-x) \quad \text{on } \mathbb{D}_0^\delta \tag{3.16}$$

for some constant $C_1>0$. Moreover, the following hold.

(i) *(Gradient estimate) $\nabla_x p$ is continuous on $\mathbb{D}_0^\delta$ and for some $C_2>0$ and all $(t,x;s,y)\in\mathbb{D}_0^\delta$,*

$$|\nabla_x p(t,x;s,y)| \leqslant C_2(\xi_{\lambda,-1}+\|\kappa\|_\infty\eta_{\alpha,1})(s-t,y-x). \tag{3.17}$$

(ii) *(On-diagonal lower bound estimate) There is a constant $C_3>0$ such that for all $|y-x|\leqslant\sqrt{s-t}<\delta$,*

$$p(t,x;s,y) \geqslant C_3(s-t)^{-d/2}. \tag{3.18}$$

(iii) *(C-K equation) For all $(t,x;s,y)\in\mathbb{D}_0^\delta$ and $r\in(t,s)$, the following Chapman-Kolmogorov equation holds:*

$$\int_{\mathbb{R}^d} p(t,x;r,z)p(r,z;s,y)\mathrm{d}z = p(t,x;s,y). \tag{3.19}$$

(iv) *(Fractional derivative estimate) Let $J:\mathbb{R}^d\to\mathbb{R}$ be a bounded measurable function satisfying* (2.11). *If in addition for $\alpha\in(0,1]$, $b\in\mathbb{K}_1$ and for $\alpha\in(1,2)$, $b\in\overline{\mathbb{K}}_\alpha$, then for some $C_4>0$ and all $(t,x;s,y)\in\mathbb{D}_0^\delta$,*

$$|\widetilde{\mathscr{L}}^J p(t,\cdot;s,y)(x)| \leqslant C_4\|J\|_\infty\eta_{\alpha,0}(s-t,y-x) \tag{3.20}$$

and

$$p(t,x;s,y) = Z(t,x;s,y) + \int_t^s\int_{\mathbb{R}^d} Z(t,x;r,z)\mathscr{L}_r^{b,\kappa}p(r,\cdot;s,y)(z)\mathrm{d}z\mathrm{d}r. \tag{3.21}$$

(v) *(Second order derivative estimate) If b and κ are bounded measurable and satisfy* (3.12), *then $\nabla_x^2 p$ is continuous on $\mathbb{D}_0^\delta$ and for some $C_5>0$,*

$$|\nabla_x^2 p(t,x;s,y)| \leqslant C_5(\xi_{\lambda,-2}+\|\kappa\|_\infty\eta_{\alpha,0})(s-t,y-x). \tag{3.22}$$

Proof. (i) Let $\ell_{b,\kappa}(r)$ and the constant c be as in Lemma 3.2 with $T = 1$. In view of $\lim_{\delta\to0}\ell_{b,\kappa}(\delta) = 0$, one can choose a $\delta_1 > 0$ such that $c\ell_{b,\kappa}(\delta_1) \leqslant 1/2$. Thus by (3.7), the series $p(t,x;s,y) := \sum_{n=0}^\infty p_n(t,x;s,y)$ and $G(t,x;s,y) := \sum_{n=0}^\infty \nabla_x p_n(t,x;s,y)$ are locally uniformly absolutely convergent on $\mathbb{D}_0^{\delta_1}$. In particular, p, G are continuous on $\mathbb{D}_0^{\delta_1}$ and

$$\nabla_x p(t,x;s,y) = G(t,x;s,y).$$

On the other hand, due to (3.4) we have

$$\sum_{n=0}^m p_n(t,x;s,y) = p_0(t,x;s,y) + \int_t^s\int_{\mathbb{R}^d}\sum_{n=0}^{m-1} p_n(t,x;r,z)\mathscr{L}_r^{b,\kappa}Z(r,\cdot;s,y)(z)\mathrm{d}z\mathrm{d}r.$$

By taking limits and the dominated convergence theorem, we obtain (1.12). Moreover, by (3.7), we have for $j = 0,1$,

$$\begin{aligned}|\nabla_x^j p(t,x;s,y) - \nabla_x^j p_0(t,x;s,y)| &\leqslant \sum_{n=1}^\infty |\nabla_x^j p_n(t,x;s,y)|\\ &\leqslant 2c(\|\kappa\|_\infty\eta_{\alpha,2-j} + \ell_{b,\kappa}\xi_{\lambda,-j})(s-t,y-x),\end{aligned}\tag{3.23}$$

which in turn implies (3.16) and (3.17).

Now let $\tilde p(t,x;s,y)$ be another solution to (1.12) satisfying (3.16). As in the proof of (3.7), we can show that for all $n\in\mathbb{N}$,

$$\begin{aligned}|p(t,x;s,y) - \tilde p(t,x;s,y)| \leqslant C_1\left(c(c\ell_{b,\kappa}(s-t))^{n-1} + (c\ell_{b,\kappa}(s-t))^n\right)\|\kappa\|_\infty\eta_{\alpha,2}(s-t,y-x)\\ + C_1(c\ell_{b,\kappa}(s-t))^n\xi_{\lambda,0}(s-t,y-x).\end{aligned}$$

Since $c\ell_b(s-t)\leqslant 1/2$, letting $n\to\infty$, we obtain the uniqueness.

(ii) By (3.23), if $|x-y|\leqslant\sqrt{s-t}$, then we have

$$\begin{aligned}p(t,x;s,y) &\geqslant p_0(t,x;s,y) - 2c(\|\kappa\|_\infty\eta_{\alpha,2} + \ell_{b,\kappa}\xi_{\lambda,0})(s-t,y-x)\\ &\geqslant (c_1 - c_2\ell_{b,\kappa}(s-t))(s-t)^{-d/2} \geqslant c_1(s-t)^{-d/2}/2,\end{aligned}$$

provided $s-t\leqslant\delta_2$ with δ_2 being small enough so that $\ell_{b,\kappa}(\delta_2)\leqslant\frac{c_1}{2c_2}$.

(iii) By Fubini's theorem, we have

$$\int_{\mathbb{R}^d} p(t,x;r,z)p(r,z;s,y)\mathrm{d}z = \sum_{n=0}^\infty\sum_{m=0}^n\int_{\mathbb{R}^d} p_m(t,x;r,z)p_{n-m}(r,z;s,y)\mathrm{d}z.$$

For proving (3.19), it suffices to prove that for each $n\in\mathbb{N}_0$,

$$\sum_{m=0}^n\int_{\mathbb{R}^d} p_m(t,x;r,z)p_{n-m}(r,z;s,y)\mathrm{d}z = p_n(t,x;s,y).\tag{3.24}$$

For $n = 0$, it is clearly true by (2.18). Now suppose (3.24) holds for some $n\in\mathbb{N}$. Write

$$\sum_{m=0}^{n+1}\int_{\mathbb{R}^d} p_m(t,x;r,z)p_{n+1-m}(r,z;s,y)\mathrm{d}z = I_1 + I_2,$$

where

$$\begin{aligned}I_1 &:= \int_{\mathbb{R}^d} p_{n+1}(t,x;r,z)p_0(r,z;s,y)\mathrm{d}z,\\ I_2 &:= \sum_{m=0}^n\int_{\mathbb{R}^d} p_m(t,x;r,z)p_{n+1-m}(r,z;s,y)\mathrm{d}z.\end{aligned}$$

Observing that

$$\int_{\mathbb{R}^d} \mathscr{L}_t^{b,\kappa} p_0(t,\cdot;r,z)(x) p_0(r,z;s,y)\mathrm{d}z = \mathscr{L}_t^{b,\kappa} p_0(t,\cdot;s,y)(x), \tag{3.25}$$

by (3.4) and Fubini's theorem, we have

$$\begin{aligned} I_1 &= \int_{\mathbb{R}^d} \left(\int_t^r \int_{\mathbb{R}^d} p_n(t,x;r',z') \mathscr{L}_{r'}^{b,\kappa} p_0(r',\cdot;r,z)(z')\mathrm{d}z'\mathrm{d}r' \right) p_0(r,z;s,y)\mathrm{d}z \\ &= \int_t^r \int_{\mathbb{R}^d} p_n(t,x;r',z') \left(\int_{\mathbb{R}^d} \mathscr{L}_{r'}^{b,\kappa} p_0(r',\cdot;r,z)(z') p_0(r,z;s,y)\mathrm{d}z \right) \mathrm{d}z'\mathrm{d}r' \\ &= \int_t^r \int_{\mathbb{R}^d} p_n(t,x;r',z') \mathscr{L}_{r'}^{b,\kappa} p_0(r',\cdot;s,y)(z')\mathrm{d}z'\mathrm{d}r'. \end{aligned}$$

Similarly, by (3.4) and the induction hypothesis, we have

$$I_2 = \int_r^s \int_{\mathbb{R}^d} p_n(t,x;r',z') \mathscr{L}_{r'}^{b,\kappa} p_0(r',\cdot;s,y)(z')\mathrm{d}z'\mathrm{d}r'.$$

Hence,

$$I_1 + I_2 = \int_t^s \int_{\mathbb{R}^d} p_n(t,x;r',z') \mathscr{L}_{r'}^{b,\kappa} p_0(r',\cdot;s,y)(z')\mathrm{d}z'\mathrm{d}r' = p_{n+1}(t,x;s,y),$$

which gives (3.24).

(iv) If in addition for $\alpha \in (0,1]$, $b \in \mathbb{K}_1$ and for $\alpha \in (1,2)$, $b \in \overline{\mathbb{K}}_\alpha$, then by (3.10), and since $\lim_{\delta\to} \ell_{b,\kappa}(\delta) = 0$, where $\ell_{b,\kappa}(r)$ is the same as in Lemma 3.3 with $T = 1$, as above there is a $\delta_3 > 0$ such that the series $\sum_{n=0}^\infty |\widetilde{\mathscr{L}^\kappa} p_n(t,\cdot;s,y)(x)|$ is locally uniformly convergent on $\mathbb{D}_0^{\delta_3}$. In particular, we have on $\mathbb{D}_0^{\delta_1\wedge\delta_3}$,

$$\widetilde{\mathscr{L}^\kappa} p(t,\cdot;s,y)(x) = \sum_{n=0}^\infty \widetilde{\mathscr{L}^\kappa} p_n(t,\cdot;s,y)(x),$$

and so (3.20) holds. Moreover, by (3.11), we also have (3.21).

(v) Let C_5 be the constant C_1 in (3.13) with $T = 1$. As above, it follows from (3.13) with $T = 1$ that there is a $\delta_4 > 0$ such that $C_5 \delta_4^{\frac{(2-\alpha)\wedge 1}{2}} = 1/2$.

Finally, we just need to set $\delta := \delta_1 \wedge \delta_2 \wedge \delta_3 \wedge \delta_4$. □

Using (3.19), we can extend the definition of $p(t,x;s,y)$ to $\mathbb{D}_0^\infty$.

Proof of Theorem 1.1. We shall show that $p(t,x;s,y)$ in Theorem 3.5 has all the properties in Theorem 1.1. First of all, we need to extend the definition of $p(t,x;s,y)$ from $\mathbb{D}_0^\delta$ to $\mathbb{D}_0^\infty$ by C-K equation as follows: If $\delta < s - t \leqslant 2\delta$, we define

$$p(t,x;s,y) = \int_{\mathbb{R}^d} p(t,x;\tfrac{t+s}{2},z) p(\tfrac{t+s}{2},z;s,y)\mathrm{d}z. \tag{3.26}$$

Proceeding this procedure, we can extend p to $\mathbb{D}_0^\infty$.

(1), (2), (3) and (4) follow from (3.16), (3.17), (3.19), (3.20), (3.26) and Lemma 2.1. As for (1.12) and (1.13) on $\mathbb{D}_0^\infty$, it follows by (1.12) and (1.13) on $\mathbb{D}_0^\delta$ and C-K equation.

(5) By the construction of $p(t,x;s,y)$ (see the proof of Theorem 3.5(i)), there is a $\delta_1 > 0$ such

that $p(t,x;s,y) = \sum_{n=0}^\infty p_n(t,x;s,y)$ on $\mathbb{D}_0^{\delta_1}$. Then, by the dominated convergence theorem and Fubini's theorem, for any $s,t \geqslant 0$ with $0 < s-t < \delta_1$ and $x \in \mathbb{R}^d$,

$$\int_{\mathbb{R}^d} p(t,x;s,y)\mathrm{d}y = \sum_{n=0}^\infty \int_{\mathbb{R}^d} p_n(t,x;s,y)\mathrm{d}y = \int_{\mathbb{R}^d} Z(t,x;s,y)\mathrm{d}y$$
$$+ \sum_{n=1}^\infty \int_t^s \int_{\mathbb{R}^d} p_{n-1}(t,x;r,z)\left(\int_{\mathbb{R}^d} \mathscr{L}_r^{b,\kappa} Z(r,\cdot;s,y)(z)\mathrm{d}y\right)\mathrm{d}z\mathrm{d}r = 1 + 0 = 1.$$

Hence, the conservativeness (1.18) follows from the above equality and (3.26).

(6) Let $P_{t,s}f(x) := \int_{\mathbb{R}^d} p(t,x;s,y)f(y)\mathrm{d}y$. By (1.12), we have for any bounded measurable f,

$$P_{t,s}f(x) = P_{t,s}^{(Z)}f(x) + \int_t^s P_{t,r}\mathscr{L}_r^{b,\kappa}P_{r,s}^{(Z)}f(x)\mathrm{d}r. \tag{3.27}$$

Hence, by (2.20), for $f \in C_b^2(\mathbb{R}^d)$, we have

$$\begin{aligned} P_{t,s}f(x) - f(x) &= P_{t,s}^{(Z)}f(x) - f(x) + \int_t^s P_{t,r}\mathscr{L}_r^{b,\kappa}P_{r,s}^{(Z)}f(x)\mathrm{d}r \\ &= \int_t^s P_{t,r}^{(Z)}\mathscr{L}_r^a f(x)\mathrm{d}r + \int_t^s P_{t,r}\mathscr{L}_r^{b,\kappa}P_{r,s}^{(Z)}f(x)\mathrm{d}r, \end{aligned} \tag{3.28}$$

and, by (3.27) and Fubini's theorem,

$$\begin{aligned} \int_t^s P_{t,r}\mathscr{L}_r^a f(x)\mathrm{d}r - \int_t^s P_{t,r}^{(Z)}\mathscr{L}_r^a f(x)\mathrm{d}r &= \int_t^s\int_t^r P_{t,u}\mathscr{L}_u^{b,\kappa}P_{u,r}^{(Z)}\mathscr{L}_r^a f(x)\mathrm{d}u\mathrm{d}r \\ &= \int_t^s P_{t,u}\mathscr{L}_u^{b,\kappa}\left(\int_u^s P_{u,r}^{(Z)}\mathscr{L}_r^a f(x)\mathrm{d}r\right)\mathrm{d}u \\ &= \int_t^s P_{t,u}\mathscr{L}_u^{b,\kappa}\Big(P_{u,s}^{(Z)}f(x) - f(x)\Big)\mathrm{d}u. \end{aligned}$$

Combining this with (3.28), we obtain

$$P_{t,s}f(x) - f(x) = \int_t^s P_{t,r}\Big(\mathscr{L}_r^a + \mathscr{L}_r^{b,\kappa}\Big)f(x)\mathrm{d}r = \int_t^s P_{t,r}\mathscr{L}_r f(x)\mathrm{d}r.$$

(7) By (1.12) and (2.17), we only need to show that

$$\lim_{|t-s|\to 0}\sup_{x\in\mathbb{R}^d}\left|\int_t^s P_{t,r}\mathscr{L}_r^{b,\kappa}P_{r,s}^{(Z)}f(x)\mathrm{d}r\right| = 0.$$

Notice that by (3.16), (3.8) and Lemma 3.1,

$$\begin{aligned} \left|\int_t^s P_{t,r}\mathscr{L}_r^{b,\kappa}P_{r,s}^{(Z)}f(x)\mathrm{d}r\right| &\le \|f\|_\infty \int_t^s\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(\xi_{\lambda,0} + \|\kappa\|_\infty\eta_{\alpha,2})(r-t,z-x) \\ &\quad \times \Big(|\tilde{b}(r,z)|\cdot\xi_{\lambda,-1}(s-r,y-z) + \|\kappa\|_\infty\eta_{\alpha,0}(s-r,y-z)\Big)\mathrm{d}z\mathrm{d}y\mathrm{d}r \\ &\le \|f\|_\infty\Bigg(\int_{\mathbb{R}^d}\Big(\Big(\|\kappa\|_\infty(s-t)^{1-\frac{\alpha}{2}} + K_2^{\tilde{b}}(s-t)\Big)(\xi_{\lambda,0} + \|\kappa\|_\infty\eta_{\alpha,2})(s-t,y-x) \\ &\quad + \|\kappa\|_\infty(\eta_{\alpha,2} + \|\kappa\|_\infty\eta_{\alpha,4-\alpha})(s-t,y-x)\Big)\mathrm{d}y\Bigg) \\ &\leqslant C\|f\|_\infty\Big(\ell_{b,\kappa}(s-t) + (\ell_{b,\kappa}(s-t))^2\Big), \end{aligned}$$

where $\ell_{b,\kappa}(r)$ is the same as in Lemma 3.2, and the constant C is independent of x and s,t. Since $b \in \mathbb{K}_2$, one derives the desired limit. □

4. Proof of the lower bound

4.1. **Positivity.** In this subsection, we show that if $\kappa \geqslant 0$, then the continuous kernel $p(t, x; s, y)$ constructed in Theorem 3.5 is non-negative.

Theorem 4.1. *Under* **(H^a)**, **(H$^\kappa$)** *and* $b \in \mathbb{K}_2$, *if* $\kappa \geqslant 0$, *then the heat kernel* $p(t, x; s, y)$ *constructed in Theorem 3.5 is non-negative.*

Proof. We divide the proof into three steps.

(i) Let $b_\varepsilon := b * \rho^{(1)}_\varepsilon$ and $\kappa_\varepsilon := \kappa * \rho^{(2)}_\varepsilon$, where $\rho^{(1)}_\varepsilon \in C^\infty_c(\mathbb{R}^{d+1})$ is supported in $B_\varepsilon \subset \mathbb{R}^{d+1}$ satisfying $\int \rho^{(1)}_\varepsilon = 1$ and $\rho^{(2)}_\varepsilon \in C^\infty_c(\mathbb{R}^{2d+1})$ is supported in $B_\varepsilon \subset \mathbb{R}^{2d+1}$ satisfying $\int \rho^{(2)}_\varepsilon = 1$. For example,

$$\kappa_\varepsilon(t, x, z) := \int_{\mathbb{R}^{2d+1}} \kappa(s, y_1, y_2)\rho^{(2)}_\varepsilon(t - s, x - y_1, z - y_2)\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}s.$$

Let $p_\varepsilon(t, x; s, y)$ be the corresponding heat kernel constructed in Theorem 3.5. We claim that

$$p_\varepsilon(t, x; s, y) \geqslant 0 \quad \text{on } \mathbb{D}^\infty_0. \tag{4.1}$$

While it is possible to use Hille-Yosida-Ray theorem and Courrége's first theorem to prove the claim, as it was done in [14, Theorem 1.2] (see also [25, Lemma 4.1] and [7, Lemma 4.9]), we present here a self-contained proof based on the maximum principle established in Theorem 6.1 in the Appendix. Notice that for any $\theta > 0$ and $\varepsilon \in (0, 1)$,

$$K^{b_\varepsilon}_\theta(\delta) \leqslant K^b_\theta(\delta), \quad \|\kappa_\varepsilon\|_\infty \leqslant \|\kappa\|_\infty, \tag{4.2}$$

and by (2.34), each pair of b_ε and κ_ε satisfies (3.12). Hence, by (3.16) and (3.17), we have the following uniform estimate:

$$\sup_{\varepsilon\in(0,1)} |\nabla^j_x p_\varepsilon(t, x; s, y)| \leqslant C(\xi_{\lambda,-j} + \|\kappa\|_\infty \eta_{\alpha,2-j})(s - t, y - x), \quad j = 0, 1. \tag{4.3}$$

Let $f \in C^2_b(\mathbb{R}^d)$ be non-negative. Fix $s > 0$ and set

$$u_\varepsilon(t, x) := \int_{\mathbb{R}^d} p_\varepsilon(t, x; s, y) f(y)\mathrm{d}y, \quad t < s.$$

In order to show (4.1), it suffices to check that the conditions of Theorem 6.1 are satisfied for u_ε. Since $b_\varepsilon, \kappa_\varepsilon \in C^\infty_b(\mathbb{R}_+ \times \mathbb{R}^d)$, by Theorem 3.5, one sees that $u_\varepsilon \in C_b([0, s] \times \mathbb{R}^d)$ and $(t, x) \mapsto \nabla^j u_\varepsilon(t, x)$ is continuous for $j = 0, 1, 2$, and

$$u_\varepsilon(t, x) = P^{(Z)}_{t,s} f(x) + \int_t^s P^{(Z)}_{t,r} \mathscr{L}^{b_\varepsilon,\kappa_\varepsilon}_r u_\varepsilon(r, x)\mathrm{d}r,$$

where $P^{(Z)}_{t,s} f(x) := \int_{\mathbb{R}^d} Z(t, x; s, y) f(y)\mathrm{d}y$. Moreover, by (3.16), (3.17), (3.22) and Lemma 2.2, as in the proof of (2.23), we have for any $\gamma \in (0, (2 - \alpha) \wedge 1)$,

$$|\mathscr{L}^{b_\varepsilon,\kappa_\varepsilon}_r u_\varepsilon(r, x) - \mathscr{L}^{b_\varepsilon,\kappa_\varepsilon}_r u_\varepsilon(r, x')| \leqslant C_\varepsilon (s - r)^{-\frac{\alpha+\gamma}{2}} |x - x'|^\gamma.$$

Hence, for Lebesgue almost all $t \in [0, s]$ (see Lemma 6.4 below),

$$\partial_t u_\varepsilon + \mathscr{L}^a_t u_\varepsilon + \mathscr{L}^{b_\varepsilon,\kappa_\varepsilon}_t u_\varepsilon = 0.$$

Thus we can use Theorem 6.1 to conclude (4.1).

(ii) In this step, we show that

$$p_\varepsilon \text{ is locally equi-continuous on } \mathbb{D}^\infty_0. \tag{4.4}$$

Recall that

$$p_\varepsilon(t, x; s, y) = Z(t, x; s, y) + \int_t^s \int_{\mathbb{R}^d} p_\varepsilon(t, x; r, z) \mathscr{L}^{b_\varepsilon,\kappa_\varepsilon}_r Z(r, \cdot; s, y)(z)\mathrm{d}z\mathrm{d}r. \tag{4.5}$$

Let D be a compact subset of $\mathbb{D}_{t_0}^{T_0} \subset \mathbb{D}_0^\infty$, where $0 < t_0 < T_0 < \infty$. We first show that

$$\lim_{|h|\to 0} \sup_{\varepsilon\in(0,1)} \sup_{(t,x,s,y)\in D} |p_\varepsilon(t,x;s,y) - p_\varepsilon(t,x;s,y+h)| = 0. \tag{4.6}$$

Notice that for any $\delta < (s-t)/4$, by (4.3), (3.8), (2.1) and (2.2),

$$\begin{aligned}
&\left|\int_{s-\delta}^{s}\int_{\mathbb{R}^d} p_\varepsilon(t,x;r,z)\mathscr{L}_r^{b_\varepsilon,\kappa_\varepsilon} Z(r,\cdot;s,y)(z)\mathrm{d}z\mathrm{d}r\right| \\
&\leq \int_{s-\delta}^{s}\int_{\mathbb{R}^d} (\xi_{\lambda,0} + \|\kappa\|_\infty \eta_{\alpha,2})(r-t, z-x) \\
&\quad \times \Big((|b_\varepsilon(r,z)| + \|\kappa\|_\infty)\cdot \xi_{\lambda,-1}(s-r, y-z) + \|\kappa\|_\infty \eta_{\alpha,0}(s-r,y-z)\Big)\mathrm{d}z\mathrm{d}r \\
&\leq \Big((s-t-\delta)^{-d/2} + (s-t-\delta)^{1-(d+\alpha)/2}\Big) \\
&\quad \times \int_0^\delta\int_{\mathbb{R}^d} \Big(|b_\varepsilon(s-r,y-z)|\cdot \eta_{\alpha,\alpha-1}(r,z) + \|\kappa\|_\infty \eta_{\alpha,0}(r,z)\Big)\mathrm{d}z\mathrm{d}r \\
&\leq (s-t)^{-d/2}\Big(K_2^b(\delta) + \|\kappa\|_\infty \delta^{1/2} + \|\kappa\|_\infty \delta^{1-\alpha/2}\Big).
\end{aligned} \tag{4.7}$$

By (4.3), (2.16), (2.24), Lemmas 3.1 and 2.1, for any $\beta' \in (0,\beta)$ and $\gamma \in (0, (2-\alpha)\wedge 1)$, we have

$$\begin{aligned}
&\left|\int_{t}^{s-\delta}\int_{\mathbb{R}^d} p_\varepsilon(t,x;r,z)\mathscr{L}_r^{b_\varepsilon,\kappa_\varepsilon}(Z(r,\cdot;s,y) - Z(r,\cdot;s,y+h))(z)\mathrm{d}z\mathrm{d}r\right| \\
&\leq |h|^{\beta'}\Bigg[\int_t^{s-\delta}\int_{\mathbb{R}^d} (\xi_{\lambda,0} + \|\kappa\|_\infty\eta_{\alpha,2})(r-t,z-x)\,(|b_\varepsilon(r,z)| + \|\kappa_\varepsilon\|_\infty) \\
&\qquad \times \Big(\xi_{\lambda,-\beta'-1}(s-r,y-z) + \xi_{\lambda,-\beta'-1}(s-r,y+h-z)\Big)\mathrm{d}z\mathrm{d}r\Bigg] \\
&\quad + |h|^{\beta'\gamma}\|\kappa\|_\infty\Bigg[\int_t^s\int_{\mathbb{R}^d} (\xi_{\lambda,0} + \|\kappa\|_\infty\eta_{\alpha,2})(r-t,z-x) \\
&\qquad \times \Big(\eta_{\alpha,-\beta'\gamma}(s-r,y-z) + \eta_{\alpha,-\beta'\gamma}(s-r,y+h-z)\Big)\mathrm{d}z\mathrm{d}r\Bigg] \\
&\leq |h|^{\beta'}\delta^{-\beta'/2}\Bigg[\int_t^s\int_{\mathbb{R}^d} (\xi_{\lambda,0} + \|\kappa\|_\infty\eta_{\alpha,2})(r-t,z-x)\,(|b_\varepsilon(r,z)| + \|\kappa_\varepsilon\|_\infty) \\
&\quad \times \Big(\xi_{\lambda,-1}(s-r,y-z) + \xi_{\lambda,-1}(s-r,y+h-z)\Big)\mathrm{d}z\mathrm{d}r\Bigg] \\
&\quad + |h|^{\beta'\gamma}\Big((s-t)^{2-\beta'\gamma-(d+\alpha)} + \|\kappa\|_\infty(s-t)^{4-\alpha-\beta'\gamma-(d+\alpha)}\Big) \\
&\leq |h|^{\beta'}\delta^{-\beta'/2}K_2^b(s-t)\,(1+\|\kappa\|_\infty)\,(s-t)^{-d/2} \\
&\quad + |h|^{\beta'\gamma}(s-t)^{2-\beta'\gamma-(d+\alpha)}\Big(1 + \|\kappa\|_\infty(s-t)^{2-\alpha}\Big),
\end{aligned}$$

which together with (4.7) gives (4.6). Similarly, we can show

$$\begin{aligned}
&\lim_{|h|\to 0}\sup_{\varepsilon\in(0,1)}\sup_{(t,x;s,y)\in D} |p_\varepsilon(t,x;s,y) - p_\varepsilon(t,x+h;s,y)| = 0, \\
&\lim_{|\delta|\to 0}\sup_{\varepsilon\in(0,1)}\sup_{(t,x;s,y)\in D} |p_\varepsilon(t,x;s,y) - p_\varepsilon(t+\delta,x;s,y)| = 0, \\
&\lim_{|\delta|\to 0}\sup_{\varepsilon\in(0,1)}\sup_{(t,x;s,y)\in D} |p_\varepsilon(t,x;s,y) - p_\varepsilon(t,x;s+\delta,y)| = 0.
\end{aligned}$$

Thus we obtain (4.4).

(iii) By (4.4), Ascoli-Arzela's lemma and a diagonalization argument, there exist a subsequence ε_k (still denoted by ε for simplicity) and a continuous function $\bar{p}$ such that

$$p_\varepsilon(t, x; s, y) \to \bar{p}(t, x; s, y) \text{ for all } (t, x; s, y) \in \mathbb{D}_0^\infty. \tag{4.8}$$

Now we want to take limits on both sides of (4.5). First, by (4.2), (4.3), (4.8) and the dominated convergence theorem, we have

$$\int_t^s\int_{\mathbb{R}^d} p_\varepsilon(t, x; r, z)\widetilde{\mathscr{L}_r^{\kappa_\varepsilon}} Z(r, \cdot; s, y)(z)\mathrm{d}z\mathrm{d}r \overset{\varepsilon\to 0}{\to} \int_t^s\int_{\mathbb{R}^d} \bar{p}(t, x; r, z)\widetilde{\mathscr{L}_r^{\kappa}} Z(r, \cdot; s, y)(z)\mathrm{d}z\mathrm{d}r.$$

Next, for the term containing $\widetilde{b_\varepsilon}(r, \cdot)\cdot\nabla$, by (4.3) and (2.15), we have

$$\begin{aligned}
&\Big|\int_t^s\int_{\mathbb{R}^d} p_\varepsilon(t, x; r, z)\widetilde{b_\varepsilon}(r, z)\nabla_z Z(r, z; s, y)\mathrm{d}z\mathrm{d}r - \int_t^s\int_{\mathbb{R}^d} \bar{p}(t, x; r, z)\tilde{b}(r, z)\nabla_z Z(r, z; s, y)\mathrm{d}z\mathrm{d}r\Big| \\
&\quad \leqslant \int_t^s\int_{\mathbb{R}^d}(\xi_{\lambda,0} + \|\kappa_\varepsilon\|_\infty\eta_{\alpha,2})(r-t, z-x)\cdot|\widetilde{b_\varepsilon} - \tilde{b}|(r, z)\cdot\xi_{\lambda_1,-1}(s-r, y-z)\mathrm{d}z\mathrm{d}r \\
&\quad + \int_t^s\int_{\mathbb{R}^d}|p_\varepsilon - \bar{p}|(t, x; r, z)\cdot|\tilde{b}(r, z)|\cdot\xi_{\lambda_1,-1}(s-r, y-z)\mathrm{d}z\mathrm{d}r =: I_1(\varepsilon) + I_2(\varepsilon),
\end{aligned}$$

where $\tilde{b}$ is defined in (3.2) and $\widetilde{b_\varepsilon}$ is defined by

$$\widetilde{b_\varepsilon}(t, x) := b_\varepsilon(t, x) + 1_{\alpha\in(1,2)}\int_{|z|>1} z\kappa_\varepsilon(t, x, z)|z|^{-d-\alpha}\mathrm{d}z - 1_{\alpha\in(0,1)}\int_{|z|\leqslant 1} z\kappa_\varepsilon(t, x, z)|z|^{-d-\alpha}\mathrm{d}z.$$

For $I_1(\varepsilon)$, by (2.1), (2.4), Lemma 2.8 and the dominated convergence theorem, we have

$$\begin{aligned}
I_1(\varepsilon) &\leqslant \int_t^s\int_{\mathbb{R}^d}(\eta_{2,2} + \|\kappa_\varepsilon\|_\infty\eta_{\alpha,2})(r-t, z-x)\cdot|\widetilde{b_\varepsilon} - \tilde{b}|(r, z)\cdot\eta_{2,1}(s-r, y-z)\mathrm{d}z\mathrm{d}r \\
&\leqslant (\eta_{2,0} + \|\kappa\|_\infty\eta_{\alpha,0})(s-t, y-x)\int_t^s\int_{\mathbb{R}^d}|\widetilde{b_\varepsilon} - \tilde{b}|(r, z)\cdot(r-t)(s-r)^{1/2} \\
&\quad \times\Big(\eta_{2,0}(r-t, z-x) + \eta_{2,0}(s-r, y-z)\Big)\mathrm{d}z\mathrm{d}r \\
&\leqslant (\eta_{2,1} + \|\kappa\|_\infty\eta_{\alpha,1})(s-t, y-x)\Big[\int_t^s\int_{\mathbb{R}^d}|\widetilde{b_\varepsilon} - \tilde{b}|(r, z)\cdot\eta_{2,1}(r-t, z-x)\mathrm{d}z\mathrm{d}r \\
&\quad + \int_t^s\int_{\mathbb{R}^d}|\widetilde{b_\varepsilon} - \tilde{b}|(r, z)\cdot\eta_{2,1}(s-r, y-z)\mathrm{d}z\mathrm{d}r\Big] \overset{\varepsilon\to 0}{\to} 0.
\end{aligned}$$

For $I_2(\varepsilon)$, by (4.3), (4.8) and the dominated convergence theorem again, we have

$$I_2(\varepsilon) \overset{\varepsilon\to 0}{\to} 0.$$

Combining the above limits, one sees that $\bar{p}(t, x; s, y)$ satisfies (1.12). Similarly, we can show $\bar{p}$ also satisfies (1.14). Thus by the uniqueness, we obtain $\bar{p} = p$ and so, $p \geqslant 0$. □

4.2. **Lower bound estimate.** Throughout this subsection, we assume $\kappa \geqslant 0$. By Theorem 4.1 and (1.15), $\{p(t, x; s, y) : (t, x; s, y) \in \mathbb{D}_0^\infty\}$ is a family of transition probability density functions. It determines a Feller process

$$\Big(\Omega, \mathscr{F}, (\mathbb{P}_{t,x})_{(t,x)\in\mathbb{R}_+\times\mathbb{R}^d}; (X_s)_{s\geqslant 0}\Big),$$

with the property that

$$\mathbb{P}_{t,x}(X_s = x,\ 0 \leqslant s \leqslant t) = 1,$$

and for $r \in [t,s]$ and $E \in \mathscr{B}(\mathbb{R}^d)$,

$$\mathbb{E}_{t,x}(X_s \in E \mid X_r) = \int_E p(r, X_r; s, y)\mathrm{d}y. \tag{4.9}$$

Moreover, for any $f \in C_b^2(\mathbb{R}^d)$, it follows from (1.19) and the Markov property of X that under $\mathbb{P}_{t,x}$, with respect to the filtration $\mathscr{F}_s := \sigma\{X_r, r \leqslant s\}$,

$$M_s^f := f(X_s) - f(X_t) - \int_t^s \mathscr{L}_r f(X_r)\mathrm{d}r \text{ is a martingale.} \tag{4.10}$$

In other words, $\mathbb{P}_{t,x}$ solves the martingale problem for $(\mathscr{L}_t, C_b^2(\mathbb{R}^d))$.

For any Borel set E, let

$$\sigma_E := \inf\{s \geqslant 0 : X_s \in E\}, \qquad \tau_E := \inf\{s \geqslant 0 : X_s \notin E\},$$

be the first hitting and exit time, respectively, of E.

Below for simplicity, we write

$$\mathcal{J}_\alpha(t,x,z) := \kappa(t,x,z-x)|z-x|^{-d-\alpha}.$$

We now determine the Lévy system of the Feller process X, which in particular is a Hunt process. The proof of the following result is similar to that of [10]. For completeness, we give a detailed proof here.

Lemma 4.2. *Suppose that E and F are two disjoint open sets in $\mathbb{R}^d$. Then*

$$\sum_{t<r\leqslant s} 1_{\{X_{r-}\in E, X_r \in F\}} - \int_t^s 1_E(X_r)\int_F \mathcal{J}_\alpha(r, X_r, z)\mathrm{d}z\mathrm{d}r$$

is a $\mathbb{P}_{t,x}$-martingale for every $t \geqslant 0$ and $x \in \mathbb{R}^d$.

Proof. First of all, by (4.10), $\{X_s, s \geqslant 0\}$ is a semi-martingale under $\mathbb{P}_{t,x}$. Let $f \in C_b^2(\mathbb{R}^d)$ with $f = 0$ on E and $f = 1$ on F. By Itô's formula, we have

$$f(X_s) - f(X_t) = \sum_{i=1}^d \int_t^s \partial_i f(X_{r-})\mathrm{d}X_r + \sum_{t<r\leqslant s}\beta_r(f) + \frac{1}{2}\sum_{i,j=1}^d \int_t^s \partial_i\partial_j f(X_{r-})\mathrm{d}\langle X^c, X^c\rangle_r,$$

where

$$\beta_r(f) := f(X_r) - f(X_{r-}) - \sum_{i=1}^d \partial_i f(X_{r-})(X_r - X_{r-}).$$

Hence,

$$\begin{aligned} N_s := \int_t^s 1_E(X_{r-})\mathrm{d}M_r^f &= \sum_{i=1}^d \int_t^s 1_E(X_{r-})\partial_i f(X_{r-})\mathrm{d}X_r + \sum_{t<r\leqslant s} 1_E(X_{r-})\beta_r(f) \\ &\quad + \frac{1}{2}\sum_{i,j=1}^d \int_t^s 1_E(X_{r-})\partial_{ij}^2 f(X_{r-})\mathrm{d}\langle X^c, X^c\rangle_r - \int_t^s 1_E(X_r)\mathscr{L}_r f(X_r)\mathrm{d}r \end{aligned}$$

is a martingale. Since $f(x) = \partial_i f(x) = \partial_{ij}^2 f(x) = 0$ for $x \in E$, we further have

$$\begin{aligned} N_s &= \sum_{t<r\leqslant s} 1_E(X_{r-}) f(X_r) - \int_t^s 1_E(X_r)\mathscr{L}_r^\kappa f(X_r)\mathrm{d}r \\ &= \sum_{t<r\leqslant s} 1_E(X_{r-}) f(X_r) - \int_t^s 1_E(X_r)\int_{\mathbb{R}^d} f(z)\mathcal{J}_\alpha(r, X_r, z)\mathrm{d}z\mathrm{d}r. \end{aligned}$$

By choosing $f_n \in C_b^2(\mathbb{R}^d)$ with $f_n|_E = 0$, $f_n|_F = 1$ and $f_n \to 1_F$, then taking limits, we obtain the desired result. ◻

In particular, Lemma 4.2 implies that

$$\mathbb{E}_{t,x}\left[\sum_{t<r\leqslant s} 1_E(X_{r-})1_F(X_r)\right] = \mathbb{E}_{t,x}\left[\int_t^s\int_{\mathbb{R}^d} 1_E(X_r)1_F(z)\mathcal{J}_\alpha(r,X_r,z)\mathrm{d}z\mathrm{d}r\right].$$

Let f be a non-negative measurable function on $\mathbb{R}_+\times\mathbb{R}^d\times\mathbb{R}^d$ that vanishes along the diagonal. By a routine measure theoretic argument, we get

$$\mathbb{E}_{t,x}\left[\sum_{t<r\leqslant s} f(r,X_{r-},X_r)\right] = \mathbb{E}_{t,x}\left[\int_t^s\int_{\mathbb{R}^d} f(r,X_r,z)\mathcal{J}_\alpha(r,X_r,z)\mathrm{d}z\mathrm{d}r\right].$$

Finally, we can follow the same method as in [13] to get the following Lévy system.

Lemma 4.3. *Let f be a non-negative measurable function on $\mathbb{R}_+\times\mathbb{R}^d\times\mathbb{R}^d$ that vanishes along the diagonal. Then for every stopping time T (with respect to the filtration of X), we have*

$$\mathbb{E}_{t,x}\left[\sum_{t<r\leqslant T} f(r,X_{r-},X_r)\right] = \mathbb{E}_{t,x}\left[\int_t^T\int_{\mathbb{R}^d} f(r,X_r,z)\mathcal{J}_\alpha(r,X_r,z)\mathrm{d}z\mathrm{d}r\right]. \tag{4.11}$$

We need the following two lemmas.

Lemma 4.4. *For any $M>0$, there is a constant $\gamma_0\in(0,1)$ depending only on M and the constants in* (1.14) *with $T=1$ such that for all $\delta\in(0,M)$,*

$$\sup_{(t,x)\in\mathbb{R}_+\times\mathbb{R}^d}\mathbb{P}_{t,x}\Big(\tau_{B(x,\delta)}\leqslant t+\gamma_0\delta^2\Big)\leqslant\tfrac{1}{2}. \tag{4.12}$$

Proof. For simplicity, write $\tau:=\tau_{B(x,\delta)}$. By the strong Markov property of X, we have

$$\begin{aligned}\mathbb{P}_{t,x}\Big(\tau\leqslant t+r\Big) &\leqslant \mathbb{P}_{t,x}\Big(\tau\leqslant t+r; X_{t+r}\in B(x,\tfrac{\delta}{2})\Big)+\mathbb{P}_{t,x}\Big(X_{t+r}\notin B(x,\tfrac{\delta}{2})\Big)\\ &= \mathbb{P}_{t,x}\left(\mathbb{P}_{\tau,X_\tau}\Big(X_{t+r}\in B(x,\tfrac{\delta}{2})\Big);\tau\leqslant t+r\right)+\mathbb{P}_{t,x}\Big(X_{t+r}\notin B(x,\tfrac{\delta}{2})\Big)\\ &\leqslant \mathbb{P}_{t,x}\left(\mathbb{P}_{\tau,X_\tau}\Big(|X_{t+r}-X_\tau|\geqslant\tfrac{\delta}{2}\Big);\tau\leqslant t+r\right)+\mathbb{P}_{t,x}\Big(X_{t+r}\notin B(x,\tfrac{\delta}{2})\Big)\\ &\leqslant 2\sup_{t\leqslant s\leqslant t+r}\sup_{x\in\mathbb{R}^d}\mathbb{P}_{s,x}\Big(|X_{t+r}-x|\geqslant\tfrac{\delta}{2}\Big).\end{aligned} \tag{4.13}$$

On the other hand, by (4.9) and (1.14), there is a constant $C>0$ depending only the constants in (1.14) with $T=1$ such that for all $r\in(0,1)$, $t\leqslant s\leqslant t+r$ and $x\in\mathbb{R}^d$,

$$\begin{aligned}\mathbb{P}_{s,x}\Big(|X_{t+r}-x|\geqslant\tfrac{\delta}{2}\Big) &= \int_{|y-x|\geqslant\frac{\delta}{2}} p(s,x;t+r,y)\mathrm{d}y\leqslant C\int_{|y-x|\geqslant\frac{\delta}{2}}(\eta_{\alpha,2}+\xi_{\lambda,0})(t+r-s,y-x)\mathrm{d}y\\ &= C\int_{|y|\geqslant\frac{\delta}{2}}(\eta_{\alpha,2}+\xi_{\lambda,0})(t+r-s,y)\mathrm{d}y\leqslant C(t+r-s)(\delta^{-\alpha}+\delta^{-2}),\end{aligned}$$

which together with (4.13) yields

$$\mathbb{P}_{t,x}\Big(\tau_{B(x,\delta)}\leqslant t+r\Big)\leqslant Cr(\delta^{-\alpha}+\delta^{-2}). \tag{4.14}$$

By letting $r=\gamma_0\delta^2$ with γ_0 being small enough, we obtain (4.12). □

For a number $\theta>0$, define

$$m_\kappa^{(\theta)}=\inf_{(t,x)\in\mathbb{R}_+\times\mathbb{R}^d}\operatorname*{essinf}_{|z|<\theta}\kappa(t,x,z).$$

Lemma 4.5. *Let $M > 0$ and γ_0 be the same as in Lemma 4.4. For all $\gamma \in (0, \gamma_0]$, there exists a constant $c_1 > 0$ such that for all $\delta \in (0, M)$, $t > 0$, $\theta > 4\delta$ and $x, y \in \mathbb{R}^d$ with $\theta/2 \geqslant |x - y| \geqslant 2\delta$,*

$$\mathbb{P}_{t,x}\Big(\sigma_{B(y,\delta)} < t + \gamma\delta^2\Big) \geqslant c_1 \frac{\delta^{d+2} \cdot m_\kappa^{(\theta)}}{|y - x|^{d+\alpha}}. \tag{4.15}$$

Proof. For $\delta \in (0, \theta/4)$ and $\gamma \in (0, \gamma_0]$, by (4.12) we have

$$\mathbb{E}_{t,x}\bigg(\int_t^{(t+\gamma\delta^2)\wedge\tau_{B(x,\delta)}} \mathrm{d}r \bigg) \geqslant \gamma\delta^2 \mathbb{P}_{t,x}\Big(\tau_{B(x,\delta)} \geqslant t + \gamma\delta^2\Big) \geqslant \gamma\delta^2/2. \tag{4.16}$$

Noticing that

$$X_r \notin B(y, \delta) \ \text{ when } \ t < r < (t + \gamma\delta^2) \wedge \tau_{B(x,\delta)},$$

we have

$$\mathbf{1}_{X_{(t+\gamma\delta^2)\wedge\tau_{B(x,\delta)}} \in B(y,\delta)} = \sum_{t<r\leqslant(t+\gamma\delta^2)\wedge\tau_{B(x,\delta)}} \mathbf{1}_{X_r \in B(y,\delta)}.$$

By (4.11) and the definition of $\mathcal{J}_\alpha$, we have

$$\begin{aligned}\mathbb{P}_{t,x}\Big(\sigma_{B(y,\delta)} < t + \gamma\delta^2\Big) &\geqslant \mathbb{P}_{t,x}\Big(X_{(t+\gamma\delta^2)\wedge\tau_{B(x,\delta)}} \in B(y, \delta)\Big) \\ &= \mathbb{E}_{t,x} \int_t^{(t+\gamma\delta^2)\wedge\tau_{B(x,\delta)}} \int_{B(y,\delta)} \frac{\kappa(t, X_r, z - X_r)}{|z - X_r|^{d+\alpha}} \mathrm{d}z\mathrm{d}r. \end{aligned} \tag{4.17}$$

Since $\theta/2 \geqslant |y - x| \geqslant 2\delta$, we have for all $z \in B(y, \delta)$ and $X_r \in B(x, \delta)$,

$$|z - X_r| \leqslant |y - z| + |y - x| + |X_r - x| < 2|y - x| \leqslant \theta.$$

Thus by (4.17) and (4.16), we have

$$\mathbb{P}_{t,x}\Big(\sigma_{B(y,\delta)} < t + \gamma\delta^2\Big) \geqslant \frac{\gamma\delta^2}{2} \int_{B(y,\delta)} \frac{m_\kappa^{(\theta)}}{(2|y - x|)^{d+\alpha}} \mathrm{d}z \geqslant c_2 \frac{\delta^{d+2} \cdot m_\kappa^{(\theta)}}{|y - x|^{d+\alpha}}.$$

The proof is complete. □

Now we can give

Proof of lower bound (1.22). First of all, by the on-diagonal estimate (3.18) and (3.19), for any $T > 0$, there is a constant $C > 0$ such that

$$p(t, x; s, y) \geqslant C(s - t)^{-d/2}, \ \ |y - x| \leqslant \sqrt{s - t} \leqslant \sqrt{T}. \tag{4.18}$$

In fact, let δ be as in Theorem 3.5. If $s - t \leqslant 2\delta$ and $|y - x| \leqslant \sqrt{s - t}$, then by (3.26), we have

$$\begin{aligned} p(t, x; s, y) &= \int_{\mathbb{R}^d} p(t, x; \tfrac{t+s}{2}, z) p(\tfrac{t+s}{2}, z; s, y) \mathrm{d}z \\ &\geqslant \int_{B(\frac{x+y}{2}, \frac{\sqrt{s-t}}{2})} p(t, x; \tfrac{t+s}{2}, z) p(\tfrac{t+s}{2}, z; s, y) \mathrm{d}z \\ &\overset{(3.18)}{\geqslant} C_3^2 (s - t)^{-d} \mathrm{Vol}\Big(B(\tfrac{x+y}{2}, \tfrac{\sqrt{s-t}}{2})\Big) \geq (s - t)^{-d/2}. \end{aligned}$$

Using the above estimate repeatedly, we obtain (4.18).

Now by (4.18) and a standard chain argument (see [17]), for any $T > 0$, there are positive constants $C, \lambda_2 > 0$ such that

$$p(t, x; s, y) \geqslant C\xi_{\lambda_2,0}(s - t, y - x) \ \text{ on } \mathbb{D}_0^T. \tag{4.19}$$

Thus to prove (1.22), it remains to show that there is a $C' > 0$ such that

$$p(t, x; s, y) \geqslant C' m_k \eta_{\alpha,2}(s - t, y - x) \ \text{ on } \mathbb{D}_0^T,$$

where $m_\kappa := \inf_{(t,x)} \operatorname{essinf}_{z\in\mathbb{R}^d} \kappa(t,x,z)$. If $|y-x| \leqslant \sqrt{s-t}$, due to (4.19), there is nothing to prove. Below we assume

$$|y-x| > \sqrt{s-t} =: 3\delta.$$

Let $\gamma_0 \in (0,1)$ be the same as in Lemma 4.4 and $\theta \in (0,\infty]$ be an arbitrary fixed number. By the strong Markov property of X and Lemma 4.5, we have for any $\theta/2 > |y-x| \geqslant 3\delta$,

$$\begin{aligned}
\mathbb{P}_{t,x}\Big(X_{t+2\gamma_0\delta^2} \in B(y,2\delta)\Big) &\geqslant \mathbb{P}_{t,x}\left(\sigma := \sigma_{B(y,\delta)} < t+\gamma_0\delta^2; \sup_{s\in[\sigma,\sigma+\gamma_0\delta^2]} |X_s - X_\sigma| < \delta\right)\\
&= \mathbb{E}_{t,x}\left(\mathbb{P}_{\sigma,X_\sigma}\left(\sup_{s\in[\sigma,\sigma+\gamma_0\delta^2]} |X_s - X_\sigma| < \delta\right); \sigma_{B(y,\delta)} < t+\gamma_0\delta^2\right)\\
&\geqslant \inf_{r,z} \mathbb{P}_{r,z}\left(\tau_{B(z,\delta)} > r+\gamma_0\delta^2\right)\mathbb{P}_{t,x}\left(\sigma_{B(y,\delta)} < t+\gamma_0\delta^2\right)\\
&\overset{(4.12)}{\geqslant} \frac{1}{2}\mathbb{P}_{t,x}\left(\sigma_{B(y,\delta)} < t+\gamma_0\delta^2\right) \overset{(4.15)}{\geqslant} \frac{c_1\delta^{d+2}\cdot m_\kappa^{(\theta)}}{2|y-x|^{d+\alpha}}.
\end{aligned}$$

Hence, by (4.18), we have for any $\theta/2 > |y-x| \geqslant 3\delta$,

$$\begin{aligned}
p(t,x;s,y) &\geqslant \int_{B(y,2\delta)} p(t,x;t+2\gamma_0\delta^2,z)p(t+2\gamma_0\delta^2,z;s,y)\mathrm{d}z\\
&\geqslant \inf_{z\in B(y,2\delta)} p(t+2\gamma_0\delta^2,z;s,y)\mathbb{P}_{t,x}\Big(X_{t+2\gamma_0\delta^2} \in B(y,2\delta)\Big)\\
&\geqslant C(s-t)^{-d/2}\cdot\frac{c_1\delta^{d+2}\cdot m_\kappa^{(\theta)}}{2|y-x|^{d+\alpha}} \geqslant C' m_\kappa^{(\theta)}\eta_{\alpha,2}(s-t,y-x).
\end{aligned}\tag{4.20}$$

The proof is complete by setting $\theta = \infty$ in the above inequality. □

5. The truncated case

Unlike the upper and lower bound in Corollary 1.5 (see also (1.5)), in the truncated case, the heat kernel $p(t,x;s,y)$ decays exponentially as $|x-y| \to 0$, In this section, we prove Theorem 1.6 by establishing the following two lemmas.

Lemma 5.1. *Under* **(H^a)**, **(H$^\kappa$)**, **(HU$^\kappa$)** *and $b \in \mathbb{K}_2$, for any $T > 0$, there are constants $C_1, \lambda_1 > 0$ such that on $\mathbb{D}_0^T$,*

$$p(t,x;s,y) \leqslant C_1\,(\xi_{\lambda_1,0} + \bar\eta_{\alpha,1/8})\,(s-t,y-x),$$

where $\bar\eta_{\alpha,\lambda}$ is defined by (1.23).

Proof. By (1.14), we already know that

$$p(t,x;s,y) \leqslant C_0\,(\xi_{\lambda_0,0} + \|\kappa\|_\infty\eta_{\alpha,2})\,(s-t,y-x) \text{ on } \mathbb{D}_0^T. \tag{5.1}$$

However, the term $\eta_{\alpha,2}(s-t,y-x)$ is too large when $|y-x|$ is large. So, we need to establish a proper upper bound for this case. We use induction method to show that there is a constant $c_1 \geqslant 1$ such that for all $n \geqslant 1$,

$$p(t,x;s,y) \leqslant \left(\frac{c_1(s-t)}{n}\right)^n, \quad 0 < s-t \leqslant T,\ |y-x| \geqslant 2n, \tag{5.2}$$

First of all, by (5.1), we have for all $s-t \in (0,T]$ and $|y-x| \geqslant 2$,

$$p(t,x;s,y) \leqslant C_0\left((s-t)^{-d/2}\mathrm{e}^{-2\lambda_0/(s-t)} + \|\kappa\|_\infty 2^{-d-\alpha}(s-t)\right) \leqslant c_2(s-t).$$

Hence, (5.2) is true for $n=1$ as long as $c_1 \geqslant c_2$.

Next we assume that (5.2) holds for all $n \leqslant N$. We want to show (5.2) for

$$s-t \in (0,T], \quad |y-x| \geqslant 2(N+1).$$

Fix such x, y and let $\tau := \tau_{B(x,1)}$ be the first exit time of X from ball $B(x,1)$. By the Lévy system of X (see Lemma 4.3) and **(HU$^\kappa$)**, we have

$$\mathbb{P}_{t,x}\left(X_\tau \in B(x,2)^c\right) = \mathbb{E}_{t,x}\left[\int_t^\tau \int_{B(x,2)^c} \kappa(r, X_r, z - X_r) \mathrm{d}z\mathrm{d}r\right] = 0,$$

where we have used the fact that for $r < \tau$ and $z \in B(x,2)^c$,

$$|z - X_r| \geqslant |z-x| - |X_r - x| > 1.$$

Let $t_n := t + \frac{n(s-t)}{N+1}$ for $n = 0, 1, \cdots, N+1$. By the strong Markov property of X,

$$\begin{aligned} p(t,x;s,y) &= \mathbb{E}_{t,x}\left[p(\tau, X_\tau; s,y); \tau < s\right] \\ &= \mathbb{E}_{t,x}\left[p(\tau, X_\tau; s,y); \tau < s, X_\tau \in B(x,2)\right] \\ &= \sum_{n=0}^N \mathbb{E}_{t,x}\left[p(\tau, X_\tau; s,y); \tau \in [t_n, t_{n+1}), X_\tau \in B(x,2)\right] \\ &\leqslant \sum_{n=0}^N \sup_{(r,z)\in[t_n,t_{n+1})\times B(x,2)} p(r,z;s,y) \cdot \mathbb{P}_{t,x}\left(\tau \leqslant t_{n+1}\right). \end{aligned} \tag{5.3}$$

Noting that for all $z \in B(x,2)$ and $|y-x| \geqslant 2(N+1)$,

$$|y-z| \geqslant |y-x| - |z-x| \geqslant 2N,$$

by the induction hypothesis, we have

$$\sup_{(r,z)\in[t_n,t_{n+1})\times B(x,2)} p(r,z;s,y) \leqslant \left(\frac{c_1(s-t_n)}{N}\right)^N = c_1^N \left(\frac{(N+1-n)(s-t)}{N(N+1)}\right)^N. \tag{5.4}$$

On the other hand, by (4.14) with $r = (n+1)(s-t)/(N+1)$ and $\delta = 1$,

$$\mathbb{P}_{t,x}\left(\tau \leqslant t_{n+1}\right) \leqslant C\frac{(n+1)(s-t)}{N+1}, \quad n \leqslant N. \tag{5.5}$$

Combining (5.3)-(5.5), we get for all $s-t \in (0,T]$ and $|y-x| \geqslant 2(N+1)$,

$$\begin{aligned} p(t,x;s,y) &\leqslant c_1^N \cdot C \sum_{n=0}^N \left(\frac{(N+1-n)(s-t)}{N(N+1)}\right)^N \frac{(n+1)(s-t)}{N+1} \\ &= c_1^N \cdot C \left(\frac{s-t}{N+1}\right)^{N+1} \sum_{n=1}^{N+1} (N+2-n)\left(\frac{n}{N}\right)^N, \end{aligned}$$

where $C \geqslant 1$. Since $s \mapsto (N+2-s)(s/N)^N$ is increasing on $[0,N]$, we have

$$\begin{aligned} \sum_{n=1}^{N+1} (N+2-n)\left(\frac{n}{N}\right)^N &= \left(\frac{N+1}{N}\right)^N + 2 + \sum_{n=1}^{N-1}(N+2-n)\left(\frac{n}{N}\right)^N \\ &\leqslant \mathrm{e} + 2 + \int_0^N (N+2-s)(s/N)^N \mathrm{d}s \\ &= \mathrm{e} + 2 + \frac{1}{N^N}\left(\frac{(N+2)N^{N+1}}{N+1} - \frac{N^{N+2}}{N+2}\right) \\ &= \mathrm{e} + 2 + \frac{N}{N+1}\cdot\frac{3N+4}{N+2} \leqslant 10. \end{aligned}$$

Therefore,

$$p(t,x;s,y) \leqslant c_1^N \cdot 10C \left(\frac{s-t}{N+1}\right)^{N+1}.$$

Thus (5.2) is proven for $c_1 = c_2 \vee (10C)$.

Finally, for $|y-x| \geqslant 2$, choosing $n \in \mathbb{N}$ so that $2n \leqslant |y-x| < 2(n+1)$, by (5.2), we have

$$\begin{aligned} p(t,x;s,y) &\leqslant \left(\frac{c_1(s-t)}{n}\right)^n \preceq \left(\frac{2(n+1)}{n} \cdot \frac{c_1(s-t)}{|y-x|}\right)^{\frac{n}{2(n+1)}\cdot|y-x|} \\ &\leqslant \left(\frac{4c_1(s-t)}{|y-x|}\right)^{\frac{|y-x|}{4}} \preceq \left(\frac{s-t}{|y-x|}\right)^{\frac{|y-x|}{8}}, \end{aligned}$$

which together with (5.1) gives the desired estimate. □

Lemma 5.2. *Under* **(H^a)**, **(H$^\kappa$)**, **(HL$^\kappa$)** *and $b \in \mathbb{K}_2$, for any $T > 0$, there are constants $C_2, \lambda_2 > 0$ such that on $\mathbb{D}_0^T$,*

$$p(t,x;s,y) \geqslant C_2 \left(\xi_{\lambda_2,0} + \overline{\eta}_{\alpha,8}\right)(s-t, y-x).$$

Proof. First of all, by (4.19), we have

$$p(t,x;s,y) \geqslant C\xi_{\lambda_2,0}(s-t,y-x) \text{ on } \mathbb{D}_0^T, \tag{5.6}$$

and, by **(HL$^\kappa$)** and (4.20), for $s-t \in (0,T]$ and $|y-x| \leqslant 1/2$,

$$p(t,x;s,y) \geqslant c_1\eta_{\alpha,2}(s-t,y-x). \tag{5.7}$$

Thus it remains to prove this lemma for $s-t \in (0,T]$ and $|y-x| > 1/2$. Let n be the least integer greater than $4|y-x|$, that is,

$$2 \leqslant n-1 \leqslant 4|y-x| < n. \tag{5.8}$$

For $i = 0, 1, \cdots, n$, let us define

$$x_i = x + (y-x)i/n, \quad B_i := B(x_i, 1/8) \quad \text{and} \quad t_i = t + (s-t)i/n.$$

Noticing that for all $i = 0, 1, \cdots, n-1$ and $z_i \in B_i$,

$$|z_i - z_{i+1}| \leqslant |z_i - x_i| + |x_i - x_{i+1}| + |x_{i+1} - z_{i+1}| \leqslant \tfrac{1}{8} + \tfrac{|y-x|}{n} + \tfrac{1}{8} < \tfrac{1}{2},$$

by (5.7), we have

$$p(t_i,z_i;t_{i+1},z_{i+1}) \geqslant c_1\eta_{\alpha,2}(t_{i+1}-t_i, z_{i+1}-z_i) \geqslant c_1\left(\frac{1}{2} + \sqrt{\frac{T}{3}}\right)^{-(d+\alpha)}\left(\frac{s-t}{n}\right) =: c_2\left(\frac{s-t}{n}\right).$$

Hence, by C-K equation (3.19) and (5.8),

$$\begin{aligned} p(t,x;s,y) \geqslant & \int_{B_{n-1}} \cdots \int_{B_1} p(t_0,x;t_1,z_1) \cdots p(t_{n-1},z_{n-1};s,y) \mathrm{d}z_1 \cdots \mathrm{d}z_{n-1} \\ \geqslant & (\mathrm{Vol}(B_1))^{n-1}\left(\frac{c_2(s-t)}{n}\right)^n \geqslant c_3\left(\frac{s-t}{|y-x|}\right)^{8|y-x|}, \end{aligned}$$

which together with (5.7) and (5.6) yields the desired estimate. □

Theorem 1.6 follows directly from the above two lemmas.

6. Appendix

6.1. **A maximum principle.** In this subsection we show a maximum principle for operator $\mathscr{L}$, which has been used to show the uniqueness and positivity of heat kernels in this paper.

Theorem 6.1. *Let $a(t,x), b(t,x)$ and $\kappa(t,x,z)$ be bounded measurable with matrix $a(t,x) \geqslant 0$ and $\kappa(t,x,z) \geqslant 0$. For $T > 0$, let $u(t,x) \in C_b([0,T) \times \mathbb{R}^d)$ satisfy the following equation: for Lebesgue almost all $t \in [0,T)$,*

$$\partial_t u + \mathscr{L}_t u \leqslant 0, \quad \varliminf_{t \uparrow T} u(t,x) \geqslant 0.$$

Assume that for each $t \in [0,T)$, $x \in \mathbb{R}^d$ and $j = 0,1,2$, the mappings $x \mapsto \nabla_x^j u(t,x)$ and $t \mapsto \nabla_x^j u(t,x)$ are continuous. Then we have

$$u(t,x) \geqslant 0, \quad (t,x) \in [0,T) \times \mathbb{R}^d. \tag{6.1}$$

Proof. First of all, we assume that for each $(t,x) \in [0,T) \times \mathbb{R}^d$,

$$\partial_t u(t,x) + \mathscr{L}_t u(t,x) \leqslant \delta < 0 \text{ and } \lim_{|x| \to \infty} u(t,x) = \infty. \tag{6.2}$$

Suppose that (6.1) is not true. Since $\lim_{|x|\to\infty} u(t,x) = \infty$ and $\varliminf_{t\uparrow T} u(t,x) \geqslant 0$, there must be a point $(t_0, x_0) \in [0,T) \times \mathbb{R}^d$ such that

$$u(t_0, x_0) = \inf_{(t,x) \in [0,T) \times \mathbb{R}^d} u(t,x) < 0.$$

Since $x \mapsto \nabla^2 u(t_0, x)$ is continuous and x_0 is an infimum point of $x \mapsto u(t_0, x)$, we have

$$\nabla_x u(t_0, x_0) = 0 \text{ and } (\partial_i \partial_j u(t_0, x_0))_{ij} \text{ is positive definite and symmetric.}$$

Therefore,

$$\mathscr{L}_s u(t_0, \cdot)(x_0) \geqslant 0,$$

and by (6.2),

$$\begin{aligned} u(t,x_0) - u(t_0,x_0) &\leqslant (t-t_0)\delta - \int_{t_0}^t \mathscr{L}_s u(s,\cdot)(x_0) \mathrm{d}s \\ &\leqslant (t-t_0)\delta - \int_{t_0}^t \mathscr{L}_s u(s,\cdot)(x_0) - \mathscr{L}_s u(t_0,\cdot)(x_0) \mathrm{d}s. \end{aligned} \tag{6.3}$$

Notice that by definition and the assumptions,

$$\begin{aligned} &|\mathscr{L}_s u(s,\cdot)(x_0) - \mathscr{L}_s u(t_0,\cdot)(x_0)| \\ &\quad \leqslant \|a\|_\infty \|\nabla_x^2 u(s,x_0) - \nabla_x^2 u(t_0,x_0)\| + \|b\|_\infty |\nabla_x u(s,x_0) - \nabla_x u(t_0,x_0)| \\ &\quad + \|\kappa\|_\infty \int_{|z|\leqslant 1} \left(\int_0^1 \|\nabla_x^2 u(s, x_0 + \theta z) - \nabla_x^2 u(t_0, x_0 + \theta z)\| \mathrm{d}\theta \right) |z|^{2-d-\alpha} \mathrm{d}z \\ &\quad + \|\kappa\|_\infty \int_{|z|>1} |u(s, x_0 + z) - u(s,x_0) - u(t_0, x_0 + z) + u(t_0, x_0)| \cdot |z|^{-d-\alpha} \mathrm{d}z. \end{aligned}$$

Since $s \mapsto \nabla_x^j u(s,x)$, $j = 0,1,2$ are continuous, by dividing both sides of (6.3) by $t - t_0$ and letting $t \downarrow t_0$, we obtain

$$0 \leqslant \delta + \lim_{t \to t_0} \frac{1}{t - t_0} \int_{t_0}^t |\mathscr{L}_s u(s,\cdot)(x_0) - \mathscr{L}_s u(t_0,\cdot)(x_0)| \mathrm{d}s = \delta < 0,$$

which is impossible. In other words, the infimum is achieved at terminal time T, and (6.1) holds.

Next, we need to drop the restriction (6.2). Let

$$f(x) := (1+|x|^2)^\beta, \ \ \beta \in (0, \alpha/2).$$

For $\varepsilon, \delta > 0$, define

$$u_{\varepsilon,\delta}(t,x) := u(t,x) + \delta(T-t) + \varepsilon \mathrm{e}^{-t} f(x).$$

By easy calculations, one sees that for some $C > 0$,

$$|\mathscr{L}_t f(x)| \leqslant C(1+|x|^\beta),$$

and

$$\partial_t u_{\delta,\varepsilon}(t,x) + \mathscr{L}_t u_{\delta,\varepsilon}(t,x) \leqslant -\delta + \varepsilon \mathrm{e}^{-t}(\mathscr{L}_t f(x) - f(x)) \leqslant -\delta/2 < 0,$$

provided ε being small enough so that $\varepsilon \mathrm{e}^{-t}(\mathscr{L}_t f(x) - f(x)) < \delta/2$. Moreover, clearly

$$\lim_{x\to\infty} |u_{\varepsilon,\delta}(t,x)| = \infty.$$

Hence, by what we have proved,

$$u_{\varepsilon,\delta}(t,x) \geqslant 0.$$

By letting $\varepsilon \to 0$ and then $\delta \to 0$, we obtain (6.1). $\square$

6.2. **Proof of Theorem 2.3.** Let A be a $d\times d$ positive definite matrix and $Z_A(x)$ the d-dimensional Gaussian density function with covariance matrix A, i.e.,

$$Z_A(x) := \frac{\mathrm{e}^{-\langle A^{-1}x,x\rangle/2}}{\sqrt{(2\pi)^d \det(A)}}, \tag{6.4}$$

where $\det(A)$ denotes the determinant of A. For $x, y, z \in \mathbb{R}^d$ and $t < s$, define

$$A_{t,s}(y) := \int_t^s a_r(y)\mathrm{d}r \quad \text{and} \quad Z_y(t,x;s,z) := Z_{A_{t,s}(y)}(z-x).$$

Clearly, for each fixed y, $Z_y(t,x;s,z)$ is smooth in (x,z) and Lipschitz continuous in t and s. By definition and (1.10), for each $j \in \mathbb{N}_0$, there are constants $C_j, \lambda_j > 0$ only depending on d and c_2 such that for all $x, y, z \in \mathbb{R}^d$ and $t < s$,

$$|\nabla_x^j Z_y(t,x;s,z)| \leqslant C_j \xi_{\lambda_j,-j}(s-t, z-x). \tag{6.5}$$

Moreover, it is easy to see that $Z_y(t,x;s,z)$ satisfies the following equation: for fixed $s > 0$ and Lebesgue almost all $t \in [0,s]$,

$$\partial_t Z_y(t,x;s,z) + \mathscr{L}_t^{a.(y)} Z_y(t,\cdot;s,z)(x) = 0, \quad x,y,z \in \mathbb{R}^d. \tag{6.6}$$

Now, let us define

$$Z_0(t,x;s,y) := Z_y(t,x;s,y). \tag{6.7}$$

The following lemma establishes the Hölder continuity of $\nabla_x^j Z_0(t,x;s,y)$ in y for every $j \geqslant 0$.

Lemma 6.2. *Under* $(\mathbf{H}^a)$*, for each* $T > 0$*,* $j \in \mathbb{N}_0$ *and* $\beta' \in (0,\beta)$*, there exist constants* $C_j, \lambda_j > 0$ *depending on* T, d, c_1 *and* c_2 *such that for all* $x, y \in \mathbb{R}^d$ *and* $0 \leqslant t < s \leqslant T$*,*

$$\begin{aligned} &|\nabla_x^j Z_0(t,x;s,y) - \nabla_x^j Z_0(t,x;s,y')| \\ &\quad \leqslant C_j |y-y'|^{\beta'} \Big(\xi_{\lambda_j,-\beta'-j}(s-t, y-x) + \xi_{\lambda_j,-\beta'-j}(s-t, y'-x)\Big). \end{aligned} \tag{6.8}$$

Proof. Observe that by the chain rule,

$$\nabla^j Z_A(x) = H_j(A^{-1}, x) Z_A(x),$$

where $H_j(A, x)$ is a vector valued polynomial of A, x, and has the following property

$$H_j(\ell^2 A, \ell^{-1} x) = \ell^j H_j(A, x), \quad \ell > 0. \tag{6.9}$$

Thus by the definition of Z_0, we have

$$\nabla_x^j Z_0(t, x; s, y) = H_j(A_{t,s}^{-1}(y), y - x) Z_{A_{t,s}(y)}(y - x). \tag{6.10}$$

Due to (1.10), there is a constant $C > 0$ such that for all $0 < t < s$,

$$C^{-1}(s - t)^d \leqslant \det(A_{t,s}(y)) \leqslant C(s - t)^d, \tag{6.11}$$

and

$$\langle A_{t,s}^{-1}(y) z, z\rangle \geqslant \lambda |z|^2/(s - t), \quad |A_{t,s}^{-1}(y) z| \leqslant C|z|/(s - t). \tag{6.12}$$

Let us denote the left hand side of (6.8) by $\mathscr{I}$. Let $\delta > 0$ be a small number, whose value will be determined below. We consider two cases:

(Case $|y - y'| > \delta \wedge \sqrt{s - t}$). In this case, by (6.5) we have

$$\begin{aligned}\mathscr{I} &\preceq \xi_{\lambda_j, -j}(s - t, y - x) + \xi_{\lambda_j, -j}(s - t, y' - x) \\ &\preceq |y - y'|^{\beta'} \Big(\xi_{\lambda_j, -\beta' - j}(s - t, y - x) + \xi_{\lambda_j, -\beta' - j}(s - t, y' - x)\Big).\end{aligned}$$

(Case $|y - y'| \leqslant \delta \wedge \sqrt{s - t}$). In this case, by (6.10) we have

$$\begin{aligned}\mathscr{I} \leqslant\ & \frac{|H_j(A_{t,s}^{-1}(y), y - x)|}{\sqrt{\det(A_{t,s}(y))}} \Big| \mathrm{e}^{-\langle A_{t,s}^{-1}(y)(y - x), y - x\rangle} - \mathrm{e}^{-\langle A_{t,s}^{-1}(y')(y' - x), y' - x\rangle} \Big| \\ &+ \left| \frac{H_j(A_{t,s}^{-1}(y), y - x)}{\sqrt{\det(A_{t,s}(y))}} - \frac{H_j(A_{t,s}^{-1}(y'), y' - x)}{\sqrt{\det(A_{t,s}(y'))}} \right| \mathrm{e}^{-\langle A_{t,s}^{-1}(y')(y' - x), y' - x\rangle} =: \mathscr{I}_1 + \mathscr{I}_2.\end{aligned}$$

For $\mathscr{I}_1$, notice that

$$\begin{aligned}R &:= |\langle A_{t,s}^{-1}(y)(y - x), y - x\rangle - \langle A_{t,s}^{-1}(y')(y' - x), y' - x\rangle| \\ &\leqslant C_0(s - t)^{-1}(|y - y'|^{\beta} |y - x|^2 + |y - y'|(|y - x| + |y' - x|)) \\ &\leqslant (C_0 \delta^{\beta} + \tfrac{\lambda}{4})|y - x|^2/(s - t) + C,\end{aligned}$$

where λ is the same as in (6.12), and by (6.9) and (6.12),

$$|H_j(A_{t,s}^{-1}(y), z)| \leqslant C(s - t)^{-j/2} h_j(|z|/\sqrt{s - t}),$$

where h_j is a j-order real polynomial. Thus in view of $\mathrm{e}^R - 1 \leqslant R\mathrm{e}^R$ for all $R > 0$, by (6.11) and choosing δ small enough so that $C_0 \delta^{\beta} = \frac{\lambda}{4}$, we have

$$\begin{aligned}\mathscr{I}_1 &\preceq (s - t)^{-j/2} h_j(|y - x|/\sqrt{s - t}) \cdot \xi_{\lambda, 0}(s - t, y - x) \cdot (\mathrm{e}^R - 1) \\ &\preceq (s - t)^{-j/2} h_j(|y - x|/\sqrt{s - t}) \cdot \xi_{\lambda/2, 0}(s - t, y - x) \\ &\quad \times (s - t)^{-1}\Big(|y - y'|^{\beta} |y - x|^2 + |y - y'|(|y - x| + |y' - x|)\Big) \\ &\preceq |y - y'|^{\beta'} \xi_{\lambda/3, -\beta' - j}(s - t, y - x),\end{aligned}$$

where we have used the fact that $|y - y'| \leqslant \sqrt{s - t}$. Similarly, one can show

$$\mathscr{I}_2 \preceq |y - y'|^{\beta'} \xi_{\lambda/3, -\beta' - j}(s - t, y' - x).$$

Combining the above calculations, we get (6.8). □

The classical Levi's freezing coefficients method suggests that the heat kernel Z of $\mathscr{L}_t^a$ takes the following form:

$$Z(t,x;s,y) = Z_0(t,x;s,y) + \int_t^s \int_{\mathbb{R}^d} Z_0(t,x;r,z) Q(r,z;s,y) \mathrm{d}z \mathrm{d}r, \tag{6.13}$$

where Q satisfies the following integro-equation:

$$Q(t,x;s,y) = Q_0(t,x;s,y) + \int_t^s \int_{\mathbb{R}^d} Q_0(t,x;r,z) Q(r,z;s,y) \mathrm{d}z \mathrm{d}r \tag{6.14}$$

with

$$Q_0(t,x;s,y) := (\mathscr{L}_t^a - \mathscr{L}_t^{a.(y)}) Z_0(t,\cdot;s,y)(x) = \sum_{i,j=1}^d \Big(a_t^{ij}(x) - a_t^{ij}(y)\Big)\partial^2_{x_i x_j} Z_0(t,x;s,y). \tag{6.15}$$

Let us first solve the integral equation (6.14).

Lemma 6.3. *For each $n \in \mathbb{N}$, define $Q_n(t,x;s,y)$ recursively by*

$$Q_n(t,x;s,y) := \int_t^s \int_{\mathbb{R}^d} Q_0(t,x;r,z) Q_{n-1}(r,z;s,y) \mathrm{d}z \mathrm{d}r. \tag{6.16}$$

Under **(H^a)**, *the series $Q(t,x;s,y) := \sum_{n=0}^\infty Q_n(t,x;s,y)$ is locally uniformly and absolutely convergent, and solves the integral equation* (6.14). *Moreover,*

$$Q(t,x;s,y) = Q_0(t,x;s,y) + \int_t^s \int_{\mathbb{R}^d} Q(t,x;r,z) Q_0(r,z;s,y) \mathrm{d}z \mathrm{d}r, \tag{6.17}$$

and for any $T > 0$, on $\mathbb{D}_0^T$, we have

$$|Q(t,x;s,y)| \leqslant C \xi_{\lambda,\beta-2}(s-t;y-x), \tag{6.18}$$

and, for any $\beta' \in (0,\beta)$,

$$|Q(t,x_1;s,y) - Q(t,x_2;s,y)| \leqslant C|x_1 - x_2|^{\beta'} \sum_{i=1,2} \xi_{\lambda,\beta-\beta'-2}(s-t;y-x_i), \tag{6.19}$$

$$|Q(t,x;s,y_1) - Q(t,x;s,y_2)| \leqslant C|y_1 - y_2|^{\beta'} \sum_{i=1,2} \xi_{\lambda,\beta-\beta'-2}(s-t,y_i-x). \tag{6.20}$$

Proof. (i) First of all, by (6.5) and **(H^a)**, there exist $C_0, \lambda > 0$ such that

$$|Q_0(t,x;s,y)| \leqslant C_0 \xi_{\lambda,\beta-2}(s-t,y-x). \tag{6.21}$$

We use induction to prove that for all $n \in \mathbb{N}$, $0 \leqslant t < s$ and $x, y \in \mathbb{R}^d$,

$$|Q_{n-1}(t,x;s,y)| \leqslant \frac{(C_0\Gamma(\beta/2))^n}{(\lambda/\pi)^{d(n-1)/2}\Gamma(n\beta/2)} \xi_{\lambda,n\beta-2}(s-t,y-x), \tag{6.22}$$

where Γ is the usual Gamma function, and

$$Q_n(t,x;s,y) = \int_t^s \int_{\mathbb{R}^d} Q_{n-1}(t,x;r,z) Q_0(r,z;s,y) \mathrm{d}z \mathrm{d}r. \tag{6.23}$$

Suppose that (6.22) and (6.23) hold for some $n \in \mathbb{N}$. Let $\gamma_n := \frac{(C_0\Gamma(\beta/2))^n}{(\lambda/\pi)^{d(n-1)/2}\Gamma(n\beta/2)}$. By (6.21), (6.22) and (2.7), we have

$$\begin{aligned}
|Q_n(t,x;s,y)| &\leqslant C_0\gamma_n \int_t^s (r-t)^{\frac{\beta}{2}-1}(s-r)^{\frac{n\beta}{2}-1}\left(\int_{\mathbb{R}^d} \xi_{\lambda,0}(r-t,z-x)\xi_{\lambda,0}(s-r,y-z)\mathrm{d}z\right)\mathrm{d}r \\
&= C_0(\pi\lambda^{-1})^{d/2}\gamma_n \xi_{\lambda,0}(s-t,y-x)\int_t^s (r-t)^{\frac{\beta}{2}-1}(s-r)^{\frac{n\beta}{2}-1}\mathrm{d}r \\
&= C_0(\pi\lambda^{-1})^{d/2}\gamma_n \xi_{\lambda,(n+1)\beta-2}(s-t,y-x)\mathcal{B}(\tfrac{n\beta}{2},\tfrac{\beta}{2}) = \gamma_{n+1}\xi_{\lambda,(n+1)\beta-2}(s-t,y-x),
\end{aligned}$$

where in the last step we have used $\mathcal{B}(\frac{n\beta}{2}, \frac{\beta}{2}) = \Gamma(\frac{\beta}{2})\Gamma(\frac{n\beta}{2})/\Gamma(\frac{(n+1)\beta}{2})$. Moreover, by the induction hypothesis and Fubini's theorem, we have

$$\begin{aligned}
Q_{n+1}(t,x;s,y) &= \int_t^s\int_{\mathbb{R}^d} Q_0(t,x;r,z)\int_r^s\int_{\mathbb{R}^d} Q_{n-1}(r,z;r',z')Q_0(r',z';s,y)\mathrm{d}z'\mathrm{d}r'\mathrm{d}z\mathrm{d}r\\
&= \int_t^s\int_{\mathbb{R}^d}\int_t^{r'}\int_{\mathbb{R}^d} Q_0(t,x;r,z)Q_{n-1}(r,z;r',z')\mathrm{d}z\mathrm{d}rQ_0(r',z';s,y)\mathrm{d}z'\mathrm{d}r'\\
&= \int_t^s\int_{\mathbb{R}^d} Q_n(t,x;r',z')Q_0(r',z';s,y)\mathrm{d}z'\mathrm{d}r'.
\end{aligned}$$

Thus by (6.22), the series $Q = \sum_{n=0}^\infty Q_n$ is locally uniformly and absolutely convergent, and solves (6.14) and (6.17). Moreover, we also have the estimate (6.18).

(ii) Next, we prove (6.19). We first show that

$$|Q_0(t,x_1;s,y) - Q_0(t,x_2;s,y)| \preceq |x_1-x_2|^{\beta'}\sum_{i=1,2}\xi_{\lambda,\beta-\beta'-2}(s-t;y-x_i). \tag{6.24}$$

If $|x_1-x_2| > \sqrt{s-t}$, then by (6.21), one sees that (6.24) holds. If $|x_1-x_2| \leqslant \sqrt{s-t}$, then by the definition of Q_0 and (6.5), we have for some $\theta \in [0,1]$,

$$\begin{aligned}
|Q_0(t,x_1;s,y) - Q_0(t,x_2;s,y)| &\leqslant |a_t(x_1)-a_t(x_2)|\cdot|\nabla_x^2 Z_0(t,x_1;s,y)|\\
&\quad + |a_t(x_2)-a_t(y)|\cdot|\nabla_x^2 Z_0(t,x_1;s,y) - \nabla_x^2 Z_0(t,x_2;s,y)|\\
&\preceq |x_1-x_2|^\beta\xi_{\lambda_2,-2}(s-t,y-x_1) + |y-x_2|^\beta|x_1-x_2|\xi_{\lambda_3,-3}(s-t,y-x_2-\theta(x_1-x_2))\\
&\preceq |x_1-x_2|^{\beta'}\xi_{\lambda,\beta-\beta'-2}(s-t,y-x_1) + |y-x_2|^\beta|x_1-x_2|^{\beta'}\xi_{\lambda,-\beta'-2}(s-t,y-x_2)\\
&\preceq |x_1-x_2|^{\beta'}\sum_{i=1,2}\xi_{\lambda,\beta-\beta'-2}(s-t,y-x_i).
\end{aligned}$$

Write

$$G(t,x;s,y) := \int_t^s\int_{\mathbb{R}^d} Q_0(t,x;r,z)Q(r,z;s,y)\mathrm{d}z\mathrm{d}r.$$

By (6.18) and (6.24), we have

$$\begin{aligned}
&|G(t,x_1;s,y) - G(t,x_2;s,y)|\\
&\preceq |x_1-x_2|^{\beta'}\int_t^s\int_{\mathbb{R}^d}\xi_{\lambda,\beta-2}(s-r,y-z)\sum_{i=1,2}\xi_{\lambda,\beta-\beta'-2}(r-t,z-x_i)\mathrm{d}z\mathrm{d}r\\
&\preceq |x_1-x_2|^{\beta'}\sum_{i=1,2}\xi_{\lambda,\beta-\beta'-2}(s-t,y-x_i).
\end{aligned} \tag{6.25}$$

Combining this with (6.24), we obtain (6.19).

(iii) We now show that on $\mathbb{D}_0^T$,

$$|Q_0(t,x;s,y_1) - Q_0(t,x;s,y_2)| \preceq |y_1-y_2|^{\beta'}\sum_{i=1,2}\xi_{\lambda,\beta-\beta'-2}(s-t,y_i-x). \tag{6.26}$$

If $|y_1-y_2| > \sqrt{s-t}$, then by (6.21) again, one sees that (6.26) holds. If $|y_1-y_2| \leqslant \sqrt{s-t}$, we write

$$\begin{aligned}
|Q_0(t,x;s,y_1) - Q_0(t,x;s,y_2)| &\leqslant \big|a_t(y_1)-a_t(y_2)\big|\Big(|\nabla_x^2 Z_0(t,x;s,y_1)| + |\nabla_x^2 Z_0(t,x;s,y_2)|\Big)\\
&\quad + \Big(|a_t(x)-a_t(y_1)|\wedge|a_t(x)-a_t(y_2)|\Big)\Big|\nabla_x^2 Z_0(t,x;s,y_1) - \nabla_x^2 Z_0(t,x;s,y_2)\Big| =: I_1 + I_2.
\end{aligned}$$

For I_1, we have

$$I_1 \preceq |y_1-y_2|^\beta \sum_{i=1,2} \xi_{\lambda,-2}(s-t, y_i-x) \preceq |y_1-y_2|^{\beta'} \sum_{i=1,2} \xi_{\lambda,\beta-\beta'-2}(s-t, y_i-x).$$

For I_2, by (6.8) we have

$$\begin{aligned} I_2 &\preceq |y_1-y_2|^{\beta'}(|y_1-x|^\beta \wedge |y_2-x|^\beta) \sum_{i=1,2} \xi_{\lambda,-\beta'-2}(s-t, y_i-x) \\ &\preceq |y_1-y_2|^{\beta'} \sum_{i=1,2} \xi_{\lambda,\beta-\beta'-2}(s-t, y_i-x). \end{aligned}$$

Thus (6.26) is proven. Using (6.26) and as in (ii), we have (6.20). □

The following result can be derived in the same way as that in the proof of [18, Theorem 6, p.13] and so its proof is omitted here.

Lemma 6.4. *Let $f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ be a measurable function and satisfy that for some $\gamma \in (0,1)$*

$$|f(t,x) - f(t,x')| \leqslant C|x-x'|^\gamma.$$

Fix $s > 0$ and define $V(t,x) := \int_t^s \int_{\mathbb{R}^d} Z_0(t,x;r,z) f(r,z)\mathrm{d}z\mathrm{d}r$. Then we have the following conclusions:

(i) The mapping $(t,x) \mapsto \nabla_x^2 V(t,x)$ is continuous on $[0,s) \times \mathbb{R}^d$ and

$$\nabla_x^2 V(t,x) = \int_t^s \int_{\mathbb{R}^d} \nabla_x^2 Z_0(t,x;r,z) f(r,z)\mathrm{d}z\mathrm{d}r, \tag{6.27}$$

where the integral in the right hand side is understood in the sense of double integral.

(ii) For Lebesgue-almost all $t \in [0,s]$ and $x \in \mathbb{R}^d$,

$$\partial_t V(t,x) + \mathscr{L}_t^{a.(y)} V(t,x) + f(t,x) = 0. \tag{6.28}$$

Now we are ready to give

Proof of Theorem 2.3. We need to check $Z(t,x;s,y)$ defined by (6.13) has all the stated properties. Let

$$\Phi(t,x;s,y) := \int_t^s \int_{\mathbb{R}^d} Z_0(t,x;r,z) Q(r,z;s,y)\mathrm{d}z\mathrm{d}r.$$

(1) By (6.5) and (6.18), we have

$$\begin{aligned} |\Phi(t,x;s,y)| &\preceq \int_t^s \int_{\mathbb{R}^d} \xi_{\lambda,0}(r-t, z-x)\xi_{\lambda,\beta-2}(s-r, y-z)\mathrm{d}z\mathrm{d}r \\ &\preceq \left(\int_t^s (s-r)^{\frac{\beta-1}{2}}\mathrm{d}r\right)\xi_{\lambda,0}(s-t, y-x) \preceq \xi_{\lambda,\beta+1}(s-t, y-x), \end{aligned} \tag{6.29}$$

which together with (6.5) gives the upper bound of (2.21).

In view of (6.19), by (6.27), for $j = 1,2$, we may write

$$\begin{aligned} \nabla_x^j \Phi(t,x;s,y) &= \int_{\frac{s+t}{2}}^s \int_{\mathbb{R}^d} \nabla_x^j Z_0(t,x;r,z) Q(r,z;s,y)\mathrm{d}z\mathrm{d}r \\ &\quad + \int_t^{\frac{s+t}{2}} \int_{\mathbb{R}^d} \nabla_x^j Z_0(t,x;r,z)(Q(r,z;s,y) - Q(r,x;s,y))\mathrm{d}z\mathrm{d}r \\ &\quad + \int_t^{\frac{s+t}{2}} \left(\int_{\mathbb{R}^d} \nabla_x^j Z_0(t,x;r,z)\mathrm{d}z\right) Q(r,x;s,y)\mathrm{d}r =: I_1 + I_2 + I_3. \end{aligned} \tag{6.30}$$

For I_1, by (6.5) and (6.18), we have

$$\begin{aligned}|I_1| &\leq \int_{\frac{s+t}{2}}^{s}\int_{\mathbb{R}^d} |\nabla_x^j Z_0(t,x;r,z)|\cdot|Q(r,z;s,y)|\mathrm{d}z\mathrm{d}r\\ &\leq \int_{\frac{s+t}{2}}^{s}\int_{\mathbb{R}^d} \xi_{\lambda_j,-j}(r-t,z-x)\xi_{\lambda_j,\beta-2}(s-r,y-z)\mathrm{d}z\mathrm{d}r\\ &\leq \left(\int_{\frac{s+t}{2}}^{s}(r-t)^{-\frac{j}{2}}(s-r)^{\frac{\beta}{2}-1}\mathrm{d}r\right)\xi_{\lambda_j,0}(s-t,y-x) \leq \xi_{\lambda_j,\beta-j}(s-t,y-x).\end{aligned}$$

For I_2, by (6.5) and (6.19), we have

$$\begin{aligned}|I_2| &\leq \int_t^{\frac{s+t}{2}}\int_{\mathbb{R}^d} \xi_{\lambda_j,-j}(r-t,z-x)|z-x|^{\beta'}\Big(\xi_{\lambda_j,\beta-\beta'-2}(s-r,y-z)+\xi_{\lambda_j,\beta-\beta'-2}(s-r,y-x)\Big)\mathrm{d}z\mathrm{d}r\\ &\leq \int_t^{\frac{s+t}{2}}\int_{\mathbb{R}^d} \xi_{\lambda_j/2,\beta'-j}(r-t,z-x)\Big(\xi_{\lambda_j/2,\beta-\beta'-2}(s-r,y-z)+\xi_{\lambda_j/2,\beta-\beta'-2}(s-r,y-x)\Big)\mathrm{d}z\mathrm{d}r\\ &\leq \left(\int_t^{\frac{s+t}{2}}(r-t)^{\frac{\beta'-j}{2}}\mathrm{d}r\right)\xi_{\lambda_j/2,\beta-\beta'-2}(s-t,y-x) \leq \xi_{\lambda_j/2,\beta-j}(s-t,y-x).\end{aligned}$$

For I_3, noticing that for each $y\in\mathbb{R}^d$,

$$\int_{\mathbb{R}^d}\nabla_x^j Z_y(t,x;r,z)\mathrm{d}z = \nabla_x^j\int_{\mathbb{R}^d} Z_{A_{t,r}(y)}(r-t,z-x)\mathrm{d}z = 0,$$

by calculations as in Lemma 6.2, we have

$$\begin{aligned}\left|\int_{\mathbb{R}^d}\nabla_x^j Z_z(t,x;r,z)\mathrm{d}z\right| &= \left|\int_{\mathbb{R}^d}\nabla_x^j Z_{A_{t,r}(z)}(r-t,z-x)-\nabla_x^j Z_{A_{t,r}(x)}(r-t,z-x)\mathrm{d}z\right|\\ &\leq \int_{\mathbb{R}^d}|z-x|^{\beta}\xi_{\lambda_j,-j}(r-t,z-x)\mathrm{d}z \leq (r-t)^{\frac{\beta-j}{2}}.\end{aligned}\tag{6.31}$$

Therefore,

$$|I_3| \leq \int_t^{\frac{s+t}{2}}(r-t)^{\frac{\beta-j}{2}}\xi_{\lambda_j,\beta-2}(r-s,y-x)\mathrm{d}r \leq \xi_{\lambda_j,\beta-j}(s-t;y-x).$$

Combining the above calculations, we obtain

$$|\nabla_x^j\Phi(t,x;s,y)| \leq \xi_{\lambda_j,\beta-j}(s-t;y-x),$$

which together with (6.5) yields (2.15).

(2) By (6.5), (6.20) and (2.8), we have for $j=0,1$,

$$\begin{aligned}&|\nabla_x^j\Phi(t,x;s,y)-\nabla_x^j\Phi(t,x;s,y')|\\ &\leq |y-y'|^{\beta'}\int_t^s\int_{\mathbb{R}^d}\xi_{\lambda,-j}(r-t,z-x)\Big(\xi_{\lambda,\beta-\beta'-2}(s-r,y'-z)+\xi_{\lambda,\beta-\beta'-2}(s-r,y-z)\Big)\mathrm{d}z\mathrm{d}r\\ &\leq |y-y'|^{\beta'}\Big(\xi_{\lambda,\beta-\beta'-j}(s-t,y-x)+\xi_{\lambda,\beta-\beta'-j}(s-t,y'-x)\Big),\end{aligned}$$

which together with (6.13) and (6.8) implies (2.16).

(3) In view of (6.29), it suffices to show that

$$\lim_{|t-s|\to 0}\sup_{x\in\mathbb{R}^d}\left|\int_{\mathbb{R}^d}Z_0(t,x;s,y)f(y)\mathrm{d}y - f(x)\right| = 0.\tag{6.32}$$

Notice that (see (6.31))

$$\left|\int_{\mathbb{R}^d} Z_0(t,x;s,y)\mathrm{d}y - 1\right| = \left|\int_{\mathbb{R}^d} Z_{A_{t,s}(y)}(t,x;s,y) - Z_{A_{t,s}(x)}(t,x;s,y)\mathrm{d}y\right|$$
$$\leq \int_{\mathbb{R}^d} |y-x|^\beta \xi_{\lambda,0}(s-t,y-x)\mathrm{d}y \leq (s-t)^{\frac{\beta}{2}}.$$

Thus to prove (6.32), it reduces to show

$$\lim_{|t-s|\to 0}\sup_{x\in\mathbb{R}^d}\left|\int_{\mathbb{R}^d} Z_0(t,x;s,y)(f(y)-f(x))\mathrm{d}y\right| = 0.$$

Since f is uniformly continuous, for any $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $|y-x| \leqslant \delta$,

$$|f(y)-f(x)| \leqslant \varepsilon.$$

Therefore, by (6.5), we have

$$\left|\int_{\mathbb{R}^d} Z_0(t,x;s,y)(f(y)-f(x))\mathrm{d}y\right| \leqslant \left(\int_{|y-x|\leqslant\delta} + \int_{|y-x|>\delta}\right) Z_0(t,x;s,y)|f(y)-f(x)|\mathrm{d}y$$
$$\leq \varepsilon\int_{|y-x|\leqslant\delta} Z_0(t,x;s,y)\mathrm{d}y + 2\|f\|_\infty \int_{|y-x|>\delta} Z_0(t,x;s,y)\mathrm{d}y$$
$$\leq \varepsilon\int_{|y-x|\leqslant\delta} \xi_{\lambda,0}(s-t,y-x)\mathrm{d}y + 2\|f\|_\infty \int_{|y-x|>\delta} \xi_{\lambda,0}(s-t,y-x)\mathrm{d}y$$
$$\leq \varepsilon\int_{\mathbb{R}^d} \xi_{\lambda,0}(s-t,y-x)\mathrm{d}y + 2\|f\|_\infty (s-t)^{\alpha/2}\int_{|y-x|>\delta} |y-x|^{-d-\alpha}\mathrm{d}y$$
$$\leq \varepsilon + 2\|f\|_\infty (s-t)^{\alpha/2}\delta^{-\alpha}.$$

Letting $|t-s| \to 0$ and then $\varepsilon \to 0$, we get the desired limit.

(4) and (5). By (6.6) and (6.28), it is easy to see that for fixed $s > 0$ and Lebesgue almost all $t \in [0,s]$,

$$\partial_t Z(t,x;s,y) + \mathscr{L}^a_t Z(t,\cdot;s,y)(x) = 0, \quad x,y\in\mathbb{R}^d,$$

that is, equation (2.14) holds. In particular, if we let $P^{(Z)}_{t,s}f(x) := \int_{\mathbb{R}^d} Z(t,x;s,y)f(y)\mathrm{d}y$, then for bounded continuous function f on $\mathbb{R}^d$, $(t,x)\mapsto \nabla^2_x P^{(Z)}_{t,s}f(x)$ is continuous and

$$\partial_t P^{(Z)}_{t,s}f(x) + \mathscr{L}^a_t P^{(Z)}_{t,s}f(x) = 0, \quad \lim_{t\uparrow r} P^{(Z)}_{t,s}f(x) = P^{(Z)}_{r,s}f(x) \text{ for } r\in(0,s].$$

On the other hand, since $t\mapsto P^{(Z)}_{t,r}P^{(Z)}_{r,s}f(x)$ satisfies the same equation with the same final value, by Theorem 6.1, we get (2.18). Moreover, if we take $f \equiv 1$, then we get (2.19). The same reason yields the non-negativity of $Z(t,x;s,y)$.

(6) Fix $s > 0$. For $f \in C^2_b(\mathbb{R}^d)$, set

$$u(t,x) := f(x) + \int_t^s P^{(Z)}_{t,r}\mathscr{L}^a_r f(x)\mathrm{d}r, \quad t\in[0,s].$$

Then we have

$$\mathscr{L}^a_t u(t,x) = \mathscr{L}^a_t f(x) + \int_t^s \mathscr{L}^a_t P^{(Z)}_{t,r}\mathscr{L}^a_r f(x)\mathrm{d}r.$$

Integrating both sides from t_0 to s with respect to t and by Fubini's theorem, we obtain

$$\int_{t_0}^s \mathscr{L}^a_t u(t,x)\mathrm{d}t = \int_{t_0}^s \mathscr{L}^a_t f(x)\mathrm{d}t + \int_{t_0}^s\int_t^s \mathscr{L}^a_t P^{(Z)}_{t,r}\mathscr{L}^a_r f(x)\mathrm{d}r\mathrm{d}t$$

$$
\begin{aligned}
&= \int_{t_0}^{s} \mathscr{L}^a_t f(x)\mathrm{d}t + \int_{t_0}^{s}\int_{t_0}^{r} \mathscr{L}^a_t P^{(Z)}_{t,r} \mathscr{L}^a_r f(x)\mathrm{d}t\mathrm{d}r \\
&= \int_{t_0}^{s} P^{(Z)}_{t_0,r} \mathscr{L}^a_r f(x)\mathrm{d}r = u(t_0, x) - f(x).
\end{aligned}
$$

In particular, for almost all $t \in [0, s]$ and $x \in \mathbb{R}^d$,

$$\partial_t u(t, x) + \mathscr{L}^a_t u(t, x) = 0, \quad \lim_{t \uparrow s} u(t, x) = f(x).$$

Using Theorem 6.1 once again (uniqueness), we obtain

$$u(t, x) = P^{(Z)}_{t,s} f(x).$$

(7) The upper bound estimate has been shown in (1). We only need to prove the lower bound estimate. By definition, one sees that for $|x - y| \leqslant (s - t)^{1/2}$,

$$Z_0(t, x; s, y) \geqslant C_0(s - t)^{-d/2},$$

and

$$Z(t, x; s, y) \geqslant Z_0(t, x; s, y) - |\Phi(t, x; s, y)| \geqslant C_0(s - t)^{-d/2} - C_1(s - t)^{(\beta+1-d)/2}, \tag{6.33}$$

which has a lower bound $\frac{C_0}{2}(s - t)^{-d/2}$ provided $C_1(s - t)^{(\beta+1)/2} \leqslant \frac{C_0}{2}$. Since $Z(t, x; s, y)$ is non-negative, such an on-diagonal lower bound estimate together with C-K equation and a standard chain argument yields the lower bound estimate (2.21) (see [17] or the proof of Lemma 5.2).

(8) Finally, we need to show the uniqueness of $Z(t, x; s, y)$. Suppose that $\widetilde{Z}(t, x; s, y)$ is another kernel that solves (2.14) and has property (2.15). Then for bounded continuous function f on $\mathbb{R}^d$ and $s > 0$, $w(t, x) := \int_{\mathbb{R}^d} \widetilde{Z}(t, x; s, y) f(y)\mathrm{d}y$ is continuous and

$$\partial_t w(t, x) + \mathscr{L}^a_t w(t, x) = 0, \quad \lim_{t \uparrow s} w(t, x) = f(x).$$

It follows from Theorem 6.1, $w(t, x) = P^{(Z)}_{t,s} f(x)$ and so $\widetilde{Z}(t, x; s, y) = Z(t, x; s, y)$. The proof of Theorem 2.3 is now complete. □

References

[1] D. G. Aronson. Non-negative solutions of linear parabolic equations. *Ann. Scuola Norm. Sup. Pisa (3)*, 22:607–694, 1968.

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

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

[4] L. A. Caffarelli, S. Salsa, and L. Silvestre. Regularity estimates for the solution and the free boundary of the obstacle problem for the fractional Laplacian. *Invent. Math.*, 171(2):425–461, 2008.

[5] Z.-Q. Chen. Symmetric jump processes and their heat kernel estimates. *Sci. China Ser. A*, 52(7):1423–1445, 2009.

[6] Z.-Q. Chen and X. Dou. Drift perturbation of subordinate Brownian motions with Gaussian component. *Sci. China Math.*, 59(2):239–260, 2016.

[7] Z.-Q. Chen and E. Hu. Heat kernel estimates for $\Delta + \Delta^{\alpha/2}$ under gradient perturbation. *Stochastic Process. Appl.*, 125(7):2603–2642, 2015.

[8] Z.-Q. Chen, P. Kim, and T. Kumagai. Weighted Poincaré inequality and heat kernel estimates for finite range jump processes. *Math. Ann.*, 342(4):833–883, 2008.

[9] Z.-Q. Chen, P. Kim, and R. Song. Heat kernel estimates for $\Delta + \Delta^{\alpha/2}$ in $C^{1,1}$ open sets. *J. Lond. Math. Soc. (2)*, 84(1):58–80, 2011.

[10] Z.-Q. Chen, P. Kim, and R. Song. Dirichlet heat kernel estimates for fractional Laplacian with gradient perturbation. *Ann. Probab.*, 40(6):2483–2538, 2012.

[11] Z.-Q. Chen and T. Kumagai. Heat kernel estimates for stable-like processes on d-sets. *Stochastic Process. Appl.*, 108(1):27–62, 2003.
[12] Z.-Q. Chen and T. Kumagai. Heat kernel estimates for jump processes of mixed types on metric measure spaces. *Probab. Theory Related Fields*, 140(1-2):277–317, 2008.
[13] Z.-Q. Chen and T. Kumagai. A priori Hölder estimate, parabolic Harnack principle and heat kernel estimates for diffusions with jumps. *Rev. Mat. Iberoam.*, 26(2):551–589, 2010.
[14] Z.-Q. Chen and J.-M. Wang. Perturbation by non-local operators. *ArXiv e-prints*, Dec. 2013.
[15] Z.-Q. Chen and L. Wang. Uniqueness of stable processes with drift. *Proc. Amer. Math. Soc.*, 144(6):2661–2675, 2016.
[16] 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.
[17] E. B. Fabes and D. W. Stroock. A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash. *Arch. Rational Mech. Anal.*, 96(4):327–338, 1986.
[18] A. Friedman. *Partial differential equations of parabolic type*. Prentice-Hall, Englewood Cliffs, NJ [u.a.], 1964.
[19] A. Grigor'yan, J. Hu, and K.-S. Lau. Heat kernels on metric measure spaces. In *Geometry and analysis of fractals*, volume 88 of *Springer Proc. Math. Stat.*, pages 147–207. Springer, Heidelberg, 2014.
[20] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, second edition, 1989.
[21] N. Jacob. *Pseudo differential operators and Markov processes. Vol. III. Markov processes and applications.* North-Holland Mathematical Library. Imperial College Press, 2005.
[22] V. Kolokoltsov. Symmetric stable laws and stable-like jump-diffusions. *Proc. London Math. Soc. (3)*, 80(3):725–768, 2000.
[23] L. Silvestre. Regularity of the obstacle problem for a fractional power of the Laplace operator. *Comm. Pure Appl. Math.*, 60(1):67–112, 2007.
[24] R. Song and Z. Vondraček. Parabolic Harnack inequality for the mixture of Brownian motion and stable process. *Tohoku Math. J. (2)*, 59(1):1–19, 2007.
[25] J.-M. Wang. Laplacian perturbed by non-local operators. *Math. Z.*, 279(1-2):521–556, 2015.
[26] L. Xie and X. Zhang. Heat kernel estimates for critical fractional diffusion operators. *Studia Math.*, 224(3):221–263, 2014.
[27] Q. S. Zhang. Gaussian bounds for the fundamental solutions of $\nabla(A\nabla u) + B\nabla u - u_t = 0$. *Manuscripta Math.*, 93(3):381–390, 1997.

Zhen-Qing Chen: Department of Mathematics, University of Washington, Seattle, WA 98195, USA, Email: zqchen@uw.edu

Eryan Hu: Department of Mathematics, Beijing Institute of Technology, Beijing, 100081, P.R.China, Email: eryanhu@gmail.com

Longjie Xie: School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou, Jiangsu 221000, P.R.China, Email: xlj.98@whu.edu.cn

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