

# Well-posedness of density dependent SDE driven by $\alpha$ -stable process with Hölder drifts<sup>☆</sup>

Mingyan Wu<sup>a</sup>, Zimo Hao<sup>b,c,\*</sup>

<sup>a</sup> School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, Hubei 430072, PR China

<sup>b</sup> School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, PR China

<sup>c</sup> Fakultät für Mathematik, Universität Bielefeld, 33615, Bielefeld, Germany

Received 14 December 2021; received in revised form 1 March 2023; accepted 20 July 2023

Available online 25 July 2023

## Abstract

In this paper, we show the weak and strong well-posedness of density dependent stochastic differential equations driven by  $\alpha$ -stable processes with  $\alpha \in (1, 2)$ . The existence part is based on Euler's approximation as Hao et al. (2021), while, the uniqueness is based on the Schauder estimates in Besov spaces for nonlocal Fokker–Planck equations. For the existence, we only assume the drift being continuous in the density variable. For the weak uniqueness, the drift is assumed to be Lipschitz in the density variable, while for the strong uniqueness, we also need to assume the drift being  $\beta_0$ -order Hölder continuous in the spatial variable, where  $\beta_0 \in (1 - \alpha/2, 1)$ .

© 2023 Elsevier B.V. All rights reserved.

**Keywords:** Lévy process; Density dependent SDE; Heat kernel; Schauder's estimate

## 1. Introduction

Fix  $\alpha \in (1, 2)$ . Let  $(L_t)_{t \geq 0}$  be a  $d$ -dimensional symmetric and rotationally invariant  $\alpha$ -stable process on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . In this paper, we consider the following density

<sup>☆</sup> Mingyan Wu is partially supported by National Natural Science Foundation of China (Grant Nos. 61873320, 12201227). Zimo Hao is grateful to the financial supports of National Natural Science Foundation of China grants of China (Nos. 12131019, 11731009), and the German Research Foundation through the CRC 1283 “Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications”.

\* Corresponding author at: School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, PR China.

E-mail addresses: [mingyanwu@hust.edu.cn](mailto:mingyanwu@hust.edu.cn) (M. Wu), [zimohao@whu.edu.cn](mailto:zimohao@whu.edu.cn) (Z. Hao).

dependent stochastic differential equation (abbreviated as DDSDE):

$$dX_t = b(t, X_t, \rho_t(X_t))dt + dL_t, \quad X_0 \stackrel{(d)}{=} \mu_0, \quad (1.1)$$

where  $b : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}^d$  is a bounded Borel measurable vector field,  $\mu_0$  is a probability measure over  $\mathbb{R}^d$  and for  $t > 0$ ,  $\rho_t(x) = \mathbb{P} \circ X_t^{-1}(dx)/dx$  is the distributional density of  $X_t$  with respect to the Lebesgue measure  $dx$  on  $\mathbb{R}^d$ .

In literature, DDSDE (1.1) is also called McKean–Vlasov SDE of Nemytskii-type which was firstly introduced in [2, Section 2] to give a probabilistic representation for the solutions of nonlinear Fokker–Planck equations. In a series of works [2–5], Barbu and Röckner investigated the following DDSDE driven by Brownian motions:

$$dX_t = b(t, X_t, \rho_t(X_t))dt + \sigma(t, X_t, \rho_t(X_t))dW_t, \quad X_0 \stackrel{(d)}{=} \mu_0, \quad (1.2)$$

where  $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d$  is measurable and  $W$  is a standard  $d$ -dimensional Brownian motion. By Itô's formula, one sees that  $\rho_t$  solves the following nonlinear Fokker–Planck equation (NFPE) in the distributional sense:

$$\partial_t \rho_t - \frac{1}{2} \sum_{i,j=1}^d \partial_i \partial_j \left[ a_{ij}(t, \cdot, \rho_t) \rho_t \right] + \operatorname{div}(b(t, \cdot, \rho_t) \rho_t) = 0, \quad \lim_{t \rightarrow 0} \rho_t(x) dx = \mu_0(dx) \text{ weakly,}$$

where  $\partial_i := \frac{\partial}{\partial x_i}$ ,  $a_{ij} := \sum_{k=1}^d \sigma_{ik} \sigma_{jk}$ , and  $\operatorname{div}$  stands for the divergence. More precisely, for any  $\varphi \in C_c^\infty(\mathbb{R}^d)$ ,

$$\langle \rho_t, \varphi \rangle = \langle \mu_0, \varphi \rangle + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \langle \rho_s, a_{ij}(s, \cdot, \rho_s) \partial_i \partial_j \varphi \rangle ds + \int_0^t \langle \rho_s, b(s, \cdot, \rho_s) \cdot \nabla \varphi \rangle ds,$$

where

$$\langle \rho_t, \varphi \rangle := \int_{\mathbb{R}^d} \varphi(x) \rho_t(x) dx = \mathbb{E} \varphi(X_t).$$

In Barbu and Röckner's works, they obtained the well-posedness for NFPE through analytic methods, and then used the so-called superposition principle to get the well-posedness of DDSDE (2.3). Recently, different from these works, the second named author together with Röckner and Zhang [13] gave a purely probabilistic proof for the existence of the solution to the following DDSDE with additive noises:

$$dX_t = b(t, X_t, \rho_t(X_t))dt + dW_t. \quad (1.3)$$

It is well known that Brownian motion is a continuous Lévy process. Hence, it is natural to consider such density dependent SDEs driven by pure jump Lévy processes. In particular, we consider  $\alpha$ -stable processes which are typical Lévy processes having selfsimilar properties (cf. [26]). Up to now, the study of the well-posedness of SDEs with stable noises has been and remains an important area in stochastic analysis. For the classical case, there are a lot of results about strong solutions, weak solutions, and martingale solutions (see [8, 15, 21, 23, 25] and etc.). We also see that there are many results about McKean–Vlasov SDEs with jumps (see [22] and references therein). Among these results, some applications can be found in financial mathematics (cf. [6]) and neural net-works (cf. [24]). However, under the framework of Lévy noises, there are no results about Nemytskii's type SDEs. Thus, it is natural and interesting to investigate DDSDE (1.1).

On the other hand, McKean–Vlasov SDEs with Lévy noises are related to non-local integral-PDEs. By Itô's formula (cf. [17, Theorem 5.1]) for DDSDE (1.1), we have that for any  $\varphi \in C_c^\infty(\mathbb{R}^d)$ ,

$$\langle \rho_t, \varphi \rangle = \langle \mu_0, \varphi \rangle + \int_0^t \langle \rho_s, b(s, \cdot, \rho_s) \cdot \nabla \varphi \rangle ds + \int_0^t \langle \rho_s, \Delta^{\alpha/2} \varphi \rangle ds, \quad (1.4)$$

where

$$\begin{aligned} \Delta^{\alpha/2} \varphi(x) &:= \int_{\mathbb{R}^d} \left( \varphi(x+z) - \varphi(x) - z \mathbf{1}_{|z| \leq 1} \cdot \nabla \varphi(x) \right) |z|^{-d-\alpha} dz \\ &= \frac{1}{2} \int_{\mathbb{R}^d} \left( \varphi(x+z) + \varphi(x-z) - 2\varphi(x) \right) |z|^{-d-\alpha} dz \end{aligned} \quad (1.5)$$

is the infinitesimal generator of  $(L_t)_{t \geq 0}$  (cf. [26, Theorem 31.5]). Consequently, one sees that  $\rho_t$  solves the following equation in the distributional sense:

$$\partial_t \rho_t - \Delta^{\alpha/2} \rho_t + \operatorname{div}(b(t, \cdot, \rho_t) \rho_t) = 0, \quad \lim_{t \downarrow 0} \rho_t(x) dx = \mu_0(dx) \text{ weakly}, \quad (1.6)$$

where we use the fact that  $\Delta^{\alpha/2}$  is a self-adjoint operator. We point out that the infinitesimal generator of Brownian motion is the Laplacian  $\Delta$ . The fractional Laplacian operator  $\Delta^{\alpha/2}$  is non-local, and is essentially different from the local operator  $\Delta$ . For instance, we can use Leibniz's rule to handle  $\Delta(fg)$  but the non-local case is more difficult. Thus, the Euler type approximation in [13], a purely probabilistic method, is chosen to show the existence of the solutions of DDSDE (1.1) in this paper.

Moreover, when  $b(t, \cdot, u)$  is  $\beta_0$ -order Hölder continuous uniformly in  $t, u$  with  $\beta_0 \in (1 - \alpha/2, 1)$ , we obtain the uniqueness based on some a priori estimates of Besov-type (see Lemma 3.6) for the nonlocal Fokker–Planck Eq. (1.4). This part is not studied in [13]. It is worth noting that the condition  $\beta_0 > 1 - \alpha/2$  is natural. The uniqueness in [13] is obtained based on the well-known pathwise uniqueness for SDE (1.3) with bounded measurable drift  $b(t, x, \rho_t(x))$  (cf. [29]). However, the situation changes when we consider  $\alpha$ -stable noises with  $\alpha \in (0, 2)$ . Let us consider

$$dX_t = b(t, X_t)dt + dL_t,$$

where  $L$  is a  $d$ -dimensional symmetric  $\alpha$ -stable process. When  $d = 1$  and  $\alpha < 1$ , even a bounded and  $\beta_0$ -Hölder continuous  $b$  is not enough to ensure pathwise uniqueness if  $\alpha + \beta_0 < 1$  (see [27] for the counterexample). When  $d \geq 1$  and  $\alpha \in [1, 2)$ , Priola [25] obtained the pathwise uniqueness under  $\beta_0 > 1 - \alpha/2$ . The condition  $\beta_0 > 1 - \alpha/2$  can be found in [11, 15] as well for the supercritical case and the degenerate case respectively.

Before stating the main result, we introduce the classical Hölder spaces in  $\mathbb{R}^d$ . For  $\beta > 0$ , let  $C^\beta(\mathbb{R}^d)$  be the classical  $\beta$ -order Hölder space consisting of all measurable functions  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  with

$$\|f\|_{C^\beta} := \sum_{j=0}^{[\beta]} \|\nabla^j f\|_\infty + [\nabla^{[\beta]} f]_{C^{\beta-[\beta]}} < \infty,$$

where  $[\beta]$  denotes the greatest integer no more than  $\beta$ ,  $\nabla^j$  stands for the  $j$ -order gradient, and

$$\|f\|_\infty := \sup_{x \in \mathbb{R}^d} |f(x)|, \quad [f]_{C^\gamma} := \sup_{h \in \mathbb{R}^d} \frac{\|f(\cdot + h) - f(\cdot)\|_\infty}{|h|^\gamma}, \quad \gamma \in (0, 1).$$

In the sequel, for any  $p \in [1, \infty)$ , we denote by  $L^p$  the space of all  $p$ -order integrable functions on  $\mathbb{R}^d$  with the norm denoted by  $\|\cdot\|_p$ .

As mentioned before, to show the existence of a weak solution, we consider the following Euler scheme to DDSDE (1.1): Let  $T > 0$ ,  $N \in \mathbb{N}$  and  $h := T/N$ . For  $t \in [0, h]$ , define

$$X_t^N := X_0 + L_t,$$

and for  $t \in (kh, (k+1)h]$  with  $k = 1, \dots, N-1$ , we inductively define  $X_t^N$  by

$$X_t^N := X_{kh}^N + \int_{kh}^t b(s, X_{kh}^N, \rho_{kh}^N(X_{kh}^N))ds + (L_t - L_{kh}),$$

where  $\rho_{kh}^N(x)$  is the distributional density of  $X_{kh}^N$ , whose existence is easily seen from the construction.

We give the definition of a weak solution to DDSDE (1.1):

**Definition 1.1** (Weak Solutions). Let  $\mu_0$  be a probability measure on  $\mathbb{R}^d$  and  $\alpha \in (1, 2)$ . We call a filtered probability space  $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t \geq 0})$  together with a pair of  $\mathcal{F}_t$ -adapted processes  $(X_t, L_t)_{t \geq 0}$  defined on it a weak solution of SDE (1.1) with initial distribution  $\mu_0$ , if

- (i)  $\mathbb{P} \circ X_0^{-1} = \mu_0$ , and  $(L_t)_{t \geq 0}$  is a  $d$ -dimensional symmetric and rotationally invariant  $\alpha$ -stable process.
- (ii) for each  $t > 0$ ,  $\rho_t(x) = \mathbb{P} \circ X_t^{-1}(dx)/dx$  and

$$X_t = X_0 + \int_0^t b(s, X_s, \rho_s(X_s))ds + L_t, \quad \mathbb{P} - \text{a.s.}$$

The following existence and uniqueness result is the main theorem of this paper.

**Theorem 1.2.** Assume that  $\alpha \in (1, 2)$ , and  $b$  is bounded measurable and for any  $(t, x, u_0) \in \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}_+$ ,

$$\lim_{u \rightarrow u_0} |b(t, x, u) - b(t, x, u_0)| = 0. \quad (1.7)$$

**(Existence)** For any  $T > 0$  and initial distribution  $\mu_0$ , there are a subsequence  $N_k$  and a weak solution  $(X, L)$  to DDSDE (1.1) in the sense of Definition 1.1. Moreover, for each  $t \in (0, T]$ ,  $X_t$  admits a distributional density  $\rho_t$ , with respect to Lebesgue measure, satisfying the estimate

$$\rho_t(y) \leq c \int_{\mathbb{R}^d} \frac{t}{(t^{1/\alpha} + |x - y|)^{d+\alpha}} \mu_0(dx), \quad (1.8)$$

where the constant  $c > 0$  only depends on  $T, d, \alpha, \|b\|_\infty$ , and the following  $L^1$ -convergences hold:

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} |\rho_t^{N_k}(y) - \rho_t(y)| dy = 0 \quad (1.9)$$

and

$$\lim_{k \rightarrow \infty} \int_0^T \int_{\mathbb{R}^d} |\rho_t^{N_k}(y) - \rho_t(y)| dy dt = 0. \quad (1.10)$$

**(Uniqueness)** Suppose that there is a constant  $c > 0$  such that for all  $(t, x, u_i) \in \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}_+$ ,  $i = 1, 2$ ,

$$|b(t, x, u_1) - b(t, x, u_2)| \leq c|u_1 - u_2|. \quad (1.11)$$

**(i)** If  $\mu_0(dx) = \rho_0(x)dx$  with  $\rho_0 \in L^q(\mathbb{R}^d)$  for some  $q \in (\frac{d}{\alpha-1}, +\infty]$ , then the weak uniqueness holds for DDSDE (1.1).

(ii) If  $\mu_0(dx) = \rho_0(x)dx$  with  $\rho_0 \in \mathbf{C}^{\beta_0}(\mathbb{R}^d)$  for some  $\beta_0 \in (1 - \alpha/2, 1)$  and

$$\sup_{(t,u) \in \mathbb{R}_+^2} \|b(t, \cdot, u)\|_{\mathbf{C}^{\beta_0}} < \infty,$$

then the pathwise uniqueness holds for DDSDE (1.1). Furthermore, there is a unique strong solution to it.

**Remark 1.3.** Although we use the same method as [13] in the existence part, our assumptions on drifts are weaker. Therein, the following local uniform continuity is assumed,

$$\lim_{u \rightarrow u_0} \sup_{|x| < R} |b(t, x, u) - b(t, x, u_0)| = 0, \quad \forall t \geq 0, R > 0. \quad (1.12)$$

For example,  $b(t, x, u) := (|u/x| \wedge 1)\mathbf{1}_{x \neq 0}$ ,  $x \in \mathbb{R}$ ,  $u \in \mathbb{R}_+$ , satisfies the condition (1.7) but does not satisfy (1.12) for any  $R > 0$ . It should be noted that our condition (1.7) is also valid for the Brownian motion case. That is, we improve the result in [13].

**Remark 1.4.** For the uniqueness, the conditions here are natural. Conditions in (i) are the same as [13, Theorem 1.2]; the same condition  $\beta_0 > 1 - \alpha/2$  in (ii) can be found in [11,25] as well.

The paper is organized as follows. In Section 2, we show some estimates of the density to the rotationally invariant and symmetric  $\alpha$ -stable process. In Section 3, we introduce Besov spaces and establish Schauder's estimates for non-local parabolic equations by using Littlewood-Paley's type estimates of heat kernels. In Section 4, we prove some uniform estimates in  $N$  about heat kernels of Euler's scheme  $X_t^N$ . In Section 5, we show the proof of Theorem 1.2.

Throughout this paper, we use the following conventions and notations: As usual, we use  $:=$  as a way of definition. Define  $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$  and  $\mathbb{R}_+ := [0, \infty)$ . The letter  $c = c(\cdot \cdot \cdot)$  denotes an unimportant constant, whose value may change in different places. We use  $A \asymp B$  and  $A \lesssim B$  to denote  $c^{-1}B \leq A \leq cB$  and  $A \leq cB$ , respectively, for some unimportant constant  $c \geq 1$ .

## 2. Preliminaries

### 2.1. $\alpha$ -Stable processes

A càdlàg process  $\{L_t \mid t \geq 0\}$  on  $\mathbb{R}^d$  ( $d \geq 1$ ) is called a Lévy process, if  $L_0 = 0$  almost surely and  $L$  has independent and identically distributed increments. The associated Poisson random measure is defined by

$$N((0, t] \times \Gamma) := \sum_{s \in (0, t]} \mathbf{1}_\Gamma(L_s - L_{s-}), \quad \Gamma \in \mathcal{B}(\mathbb{R}^d \setminus \{0\}), t > 0,$$

and the Lévy measure is given by

$$\nu(\Gamma) := \mathbb{E}N((0, 1] \times \Gamma).$$

Then, the compensated Poisson random measure is defined by

$$\tilde{N}(dr, dz) := N(dr, dz) - \nu(dz)dr.$$

For  $\alpha \in (0, 2)$ , a Lévy process  $L_t$  is called a symmetric and rotationally invariant  $\alpha$ -stable process if the Lévy measure has the form

$$\nu^{(\alpha)}(dz) = c|z|^{-d-\alpha}dz,$$

with some specific constant  $c = c(d, \alpha) > 0$ . In this paper, we only consider the symmetric and rotationally invariant  $\alpha$ -stable process. Without causing confusion, we simply call it the  $\alpha$ -stable process, and assume that  $\nu^{(\alpha)}(dz) = |z|^{-d-\alpha} dz$  here and after. For any  $0 \leq \gamma_1 < \alpha < \gamma_2$ , it is easy to see that

$$\int_{\mathbb{R}^d} (|z|^{\gamma_1} \wedge |z|^{\gamma_2}) \nu^{(\alpha)}(dz) < \infty. \quad (2.1)$$

By Lévy–Itô’s decomposition (cf. [26, Theorem 19.2]), we have

$$L_t = \lim_{\varepsilon \downarrow 0} \int_0^t \int_{\varepsilon < |z| \leq 1} z \tilde{N}(dr, dz) + \int_0^t \int_{|z| > 1} z N(dr, dz). \quad (2.2)$$

By [26, Theorem 31.5], the infinitesimal generator of Lévy process  $(L_t)_{t \geq 0}$  is the fractional Laplacian operator  $\Delta^{\alpha/2}$  defined by (1.5).

Moreover, by Lévy–Khintchine’s formula [26, Theorem 8.1], for all  $|\xi| \geq 1$ , we have

$$\begin{aligned} |\mathbb{E} e^{i\xi \cdot L_t}| &\leq \exp \left( t \int_{\mathbb{R}^d} (\cos(\xi \cdot z) - 1) \nu^{(\alpha)}(dz) \right) \\ &\leq \exp \left( -t |\xi|^\alpha \int_0^\infty \int_{\mathbb{S}^{d-1}} \frac{1 - \cos(\bar{\xi} \cdot r\theta)}{r^{1+\alpha}} \Sigma(d\theta) dr \right) \leq e^{-ct|\xi|^\alpha}, \end{aligned}$$

where  $\bar{\xi} := \xi/|\xi|$ , and  $\Sigma$  is the uniform measure on the sphere  $\mathbb{S}^{d-1}$ , and the constant  $c > 0$  depends only on  $\alpha$  and  $\Sigma(\mathbb{S}^{d-1})$ . Hence, by [26, Proposition 28.1],  $L_t$  admits a smooth density function  $p_\alpha(t, \cdot)$  given by Fourier’s inverse transform

$$p_\alpha(t, x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{-ix \cdot \xi} \mathbb{E} e^{i\xi \cdot L_t} d\xi, \quad \forall t > 0,$$

and the partial derivatives of  $p_\alpha(t, \cdot)$  at any orders tend to 0 as  $|x| \rightarrow \infty$ . Since the  $\alpha$ -stable process  $L_t$  has the scaling property

$$(\lambda^{-1/\alpha} L_{\lambda t})_{t \geq 0} \stackrel{(d)}{=} (L_t)_{t \geq 0}, \quad \forall \lambda > 0,$$

it is easy to see that

$$p_\alpha(t, x) = t^{-d/\alpha} p_\alpha(1, t^{-1/\alpha} x). \quad (2.3)$$

By [7, Theorem 2.1], one knows that there is a constant  $c = c(d, \alpha) > 1$  such that

$$c^{-1} \varrho_\alpha(t, x) \leq p_\alpha(t, x) \leq c \varrho_\alpha(t, x), \quad (2.4)$$

where

$$\varrho_\alpha(t, x) := \frac{t}{(t^{1/\alpha} + |x|)^{d+\alpha}}. \quad (2.5)$$

By [9]\*Lemma 2.2, for any  $j \in \mathbb{N}_0$ , there is a constant  $c = c(d, j, \alpha) > 0$  such that

$$|\nabla^j p_\alpha(t, x)| \leq ct^{-j/\alpha} \varrho_\alpha(t, x). \quad (2.6)$$

Since

$$(t^{1/\alpha} + |x + z|)^{-\gamma} \leq 4^\gamma (t^{1/\alpha} + |x|)^{-\gamma}, \quad \text{for } |z| \leq (2t^{1/\alpha}) \vee (|x|/2),$$

we get that

$$\varrho_\alpha(t, x + z) \leq 4^{d+\alpha} \varrho_\alpha(t, x), \quad \text{for } |z| \leq (2t^{1/\alpha}) \vee (|x|/2). \quad (2.7)$$

Note that  $p_\alpha(t, x)$  is the heat kernel of the operator  $\Delta^{\alpha/2}$ , i.e.,

$$\partial_t p_\alpha(t, x) = \Delta^{\alpha/2} p_\alpha(t, x), \quad \lim_{t \downarrow 0} p_\alpha(t, x) = \delta_0(x), \quad (2.8)$$

where  $\delta_0$  is the Dirac measure at point 0. We also have the following Chapman–Kolmogorov (abbreviated as C-K) equations:

$$(p_\alpha(t) * p_\alpha(s))(x) = \int_{\mathbb{R}^d} p_\alpha(t, x - y) p_\alpha(s, y) dy = p_\alpha(t + s, x), \quad t, s > 0. \quad (2.9)$$

## 2.2. Some estimates of the heat kernel of $\Delta^{\alpha/2}$

Now we give some estimates of the heat kernel of  $\Delta^{\frac{\alpha}{2}}$ . These estimates are straightforward and elementary. Note that [Lemma 2.1](#) and [Corollary 2.2](#) are the same as [9, Lemma 2.2] and [9, Theorem 2.4] respectively when  $j = 0$ .

**Lemma 2.1.** *For any  $j \in \mathbb{N}_0$  and  $\beta \in (0, 1)$ , there is a constant  $c = c(d, \alpha, \beta, j) > 0$  such that for every  $t > 0, x_1, x_2 \in \mathbb{R}^d$ ,*

$$|\nabla^j p_\alpha(t, x_1) - \nabla^j p_\alpha(t, x_2)| \leq c |x_1 - x_2|^\beta t^{-(j+\beta)/\alpha} (p_\alpha(t, x_1) + p_\alpha(t, x_2)). \quad (2.10)$$

**Proof.** If  $|x_1 - x_2| > t^{1/\alpha}$ , then by (2.6) we have

$$\begin{aligned} |\nabla^j p_\alpha(t, x_1) - \nabla^j p_\alpha(t, x_2)| &\lesssim t^{-j/\alpha} (\varrho_\alpha(t, x_1) + \varrho_\alpha(t, x_2)) \\ &\lesssim |x_1 - x_2|^\beta t^{-(j+\beta)/\alpha} (\varrho_\alpha(t, x_1) + \varrho_\alpha(t, x_2)). \end{aligned}$$

If  $|x_1 - x_2| \leq t^{1/\alpha}$ , then by the mean-value formula and (2.6),

$$\begin{aligned} |\nabla^j p_\alpha(t, x_1) - \nabla^j p_\alpha(t, x_2)| &\leq |x_1 - x_2| \int_0^1 |\nabla^{j+1} p_\alpha(t, x_1 + \theta(x_2 - x_1))| d\theta \\ &\lesssim |x_1 - x_2| t^{-(j+1)/\alpha} \int_0^1 \varrho_\alpha(t, x_1 + \theta(x_2 - x_1)) d\theta \\ &\stackrel{(2.7)}{\lesssim} |x_1 - x_2|^\beta t^{-(j+\beta)/\alpha} \varrho_\alpha(t, x_1), \end{aligned}$$

where we have used  $\beta \in (0, 1)$  in the last inequality. Combining the above calculations, we get (2.10) by (2.4).  $\square$

As a consequence of [Lemma 2.1](#), we have the following corollary.

**Corollary 2.2.** *For any  $j \in \mathbb{N}_0$ , there is a constant  $c = c(d, \alpha, j) > 0$  such that for every  $t > 0$  and  $x \in \mathbb{R}^d$ ,*

$$|\Delta^{\frac{\alpha}{2}} \nabla^j p_\alpha(t, x)| \leq c t^{-1-j/\alpha} p_\alpha(t, x). \quad (2.11)$$

**Proof.** First of all, recalling the definition (1.5),

$$\begin{aligned} \Delta^{\frac{\alpha}{2}} \nabla^j p_\alpha(t, x) &= \int_{|z| \leq t^{1/\alpha}} \left( \nabla^j p_\alpha(t, x + z) - \nabla^j p_\alpha(t, x) - z \cdot \nabla^{j+1} p_\alpha(t, x) \right) \frac{dz}{|z|^{d+\alpha}} \\ &\quad + \int_{|z| \geq t^{1/\alpha}} \left( \nabla^j p_\alpha(t, x + z) - \nabla^j p_\alpha(t, x) \right) \frac{dz}{|z|^{d+\alpha}} \\ &:= \mathcal{I}_1 + \mathcal{I}_2. \end{aligned}$$

For  $\mathcal{J}_1$ , by (2.7) and (2.10), we have that for any  $\beta \in (\alpha - 1, 1)$ ,

$$\mathcal{J}_1 \leq \int_{|z| \leq t^{1/\alpha}} |z|^{\beta+1} t^{-\frac{j+1+\beta}{\alpha}} \frac{dz}{|z|^{d+\alpha}} \varrho_\alpha(t, x) \lesssim t^{-1-j/\alpha} \varrho_\alpha(t, x).$$

For  $\mathcal{J}_2$ , by (2.6), one sees that

$$\begin{aligned} \mathcal{J}_2 &\lesssim t^{-j/\alpha} \int_{|z| > t^{1/\alpha}} \left( \varrho_\alpha(t, x+z) + \varrho_\alpha(t, x) \right) \frac{dz}{|z|^{d+\alpha}} \\ &\lesssim t^{-1-j/\alpha} \varrho_\alpha(t, x) + t^{-j/\alpha} \int_{|z| > t^{1/\alpha}} \varrho_\alpha(t, x+z) \frac{dz}{|z|^{d+\alpha}}. \end{aligned}$$

Then, we only need to estimate the second term above denoted by  $\mathcal{J}_3$ . If  $|x| \leq 2t^{1/\alpha}$ , by (2.5) and (2.7), we obtain that

$$\begin{aligned} \mathcal{J}_3 &\lesssim t^{-j/\alpha} \int_{|z| > t^{1/\alpha}} \varrho_\alpha(t, z) \frac{dz}{|z|^{d+\alpha}} \leq t^{-j/\alpha} \frac{t}{t^{(d+\alpha)/\alpha}} \int_{|z| > t^{1/\alpha}} \frac{dz}{|z|^{d+\alpha}} \\ &\lesssim t^{-(d+\alpha+j)/\alpha} = t^{-j/\alpha-1} \frac{t}{t^{(d+\alpha)/\alpha}} \lesssim t^{-j/\alpha-1} \varrho_\alpha(t, x). \end{aligned}$$

If  $|x| > 2t^{1/\alpha}$ , by (2.5) and (2.7), we have

$$\begin{aligned} \mathcal{J}_3 &= t^{-j/\alpha} \left( \int_{\frac{|x|}{2} \geq |z| > t^{1/\alpha}} + \int_{|z| > \frac{|x|}{2}} \right) \varrho_\alpha(t, x+z) \frac{dz}{|z|^{d+\alpha}} \\ &\lesssim t^{-j/\alpha} \varrho_\alpha(t, x) \int_{|z| > t^{1/\alpha}} \frac{dz}{|z|^{d+\alpha}} + t^{-j/\alpha} \frac{1}{|x|^{d+\alpha}} \int_{|z| > \frac{|x|}{2}} \varrho_\alpha(t, x+z) dz \\ &\stackrel{(2.4)}{\lesssim} t^{-j/\alpha-1} \varrho_\alpha(t, x) + t^{-j/\alpha} \frac{1}{|x|^{d+\alpha}} \lesssim t^{-j/\alpha-1} \varrho_\alpha(t, x). \end{aligned}$$

Based on (2.4), the proof is complete.

The following result is also true when we consider Gaussian heat kernels (cf. [13, Lemma 2.1]).

**Lemma 2.3.** For any  $\beta \in (0, \alpha)$  and  $j \in \mathbb{N}_0$ , there is a constant  $c = c(d, \alpha, \beta, j) > 0$  such that for every  $t_1, t_2 > 0$  and  $x \in \mathbb{R}^d$ ,

$$|\nabla^j p_\alpha(t_1, x) - \nabla^j p_\alpha(t_2, x)| \leq c |t_2 - t_1|^{\beta/\alpha} (t_1^{-(j+\beta)/\alpha} p_\alpha(t_1, x) + t_2^{-(j+\beta)/\alpha} p_\alpha(t_2, x)). \quad (2.12)$$

**Proof.** Without loss of generality, we assume that  $t_2 > t_1$ . If  $t_2 - t_1 > t_1$ , then  $t_1 \vee t_2 \leq 2(t_2 - t_1)$  and

$$\begin{aligned} |\nabla^j p_\alpha(t_1, x) - \nabla^j p_\alpha(t_2, x)| &\stackrel{(2.6)}{\lesssim} t_1^{-j/\alpha} \varrho_\alpha(t_1, x) + t_2^{-j/\alpha} \varrho_\alpha(t_2, x) \\ &\lesssim |t_2 - t_1|^{\beta/\alpha} (t_1^{-(j+\beta)/\alpha} \varrho_\alpha(t_1, x) + t_2^{-(j+\beta)/\alpha} \varrho_\alpha(t_2, x)). \end{aligned}$$

For  $t_2 - t_1 \leq t_1$ , notice that by (2.8) and (2.11),

$$|\nabla^j \partial_t p_\alpha(t, x)| = |\nabla^j \Delta^{\alpha/2} p_\alpha(t, x)| = |\Delta^{\alpha/2} \nabla^j p_\alpha(t, x)| \lesssim t^{-1-j/\alpha} \varrho_\alpha(t, x). \quad (2.13)$$

Thus, by the mean-value formula and  $\beta \in (0, \alpha)$ , we have

$$\begin{aligned} |\nabla^j p_\alpha(t_1, x) - \nabla^j p_\alpha(t_2, x)| &\leq |t_2 - t_1| \int_0^1 |\nabla_j \partial_t p_\alpha|(t_1 + \theta(t_2 - t_1), x) d\theta \\ &\lesssim |t_2 - t_1| \int_0^1 (t_1 + \theta(t_2 - t_1))^{-1-j/\alpha} \varrho_\alpha(t_1 + \theta(t_2 - t_1), x) d\theta \\ &\stackrel{(2.5)}{\lesssim} |t_2 - t_1| t_1^{-1-j/\alpha} \varrho_\alpha(t_1, x) \leq |t_2 - t_1|^{\beta/\alpha} t_1^{-(j+\beta)/\alpha} \varrho_\alpha(t_1, x). \end{aligned}$$

By (2.4), the proof is finished.  $\square$

### 3. Besov spaces and Schauder's estimates

In this section, we introduce Besov spaces where we obtain Schauder's estimates for the operator  $\partial_t - \Delta^{\alpha/2}$  (see Lemma 3.6). Let  $\mathcal{S}(\mathbb{R}^d)$  be the Schwartz space of all rapidly decreasing functions on  $\mathbb{R}^d$ , and  $\mathcal{S}'(\mathbb{R}^d)$  the dual space of  $\mathcal{S}(\mathbb{R}^d)$  called Schwartz generalized function (or tempered distribution) space. Given  $f \in \mathcal{S}(\mathbb{R}^d)$ , the Fourier and inverse transforms of  $f$  are defined by

$$\hat{f}(\xi) := \mathcal{F}f(\xi) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{-i\xi \cdot x} f(x) dx, \quad \xi \in \mathbb{R}^d$$

and

$$\check{f}(x) := \mathcal{F}^{-1}f(x) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{i\xi \cdot x} f(\xi) d\xi, \quad x \in \mathbb{R}^d.$$

For any  $f \in \mathcal{S}'(\mathbb{R}^d)$ ,

$$\langle \hat{f}, \varphi \rangle := \langle f, \hat{\varphi} \rangle, \quad \langle \check{f}, \varphi \rangle := \langle f, \check{\varphi} \rangle, \quad \text{for } \forall \varphi \in \mathcal{S}(\mathbb{R}^d).$$

Let  $\chi : \mathbb{R}^d \rightarrow [0, 1]$  be a smooth radial function with

$$\chi(\xi) = \begin{cases} 1, & |\xi| \leq 1, \\ 0, & |\xi| > 3/2. \end{cases}$$

Define  $\psi(\xi) := \chi(\xi) - \chi(2\xi)$  and for  $j \in \mathbb{N}_0$ ,

$$\psi_j(\xi) := \psi(2^{-j}\xi). \quad (3.1)$$

Let  $B_r := \{\xi \in \mathbb{R}^d \mid |\xi| \leq r\}$  for  $r > 0$ . It is easy to see that  $\psi \geq 0$ ,  $\text{supp } \psi \subset B_{3/2}/B_{1/2}$ , and

$$\chi(2\xi) + \sum_{j=0}^k \psi_j(\xi) = \chi(2^{-k}\xi) \rightarrow 1, \quad \text{as } k \rightarrow \infty. \quad (3.2)$$

The block operators  $\mathcal{R}_j$  are defined on  $\mathcal{S}'(\mathbb{R}^d)$  by

$$\mathcal{R}_j f := \begin{cases} \mathcal{F}^{-1}(\chi \mathcal{F}f) = \check{\chi} * f, & j = -1, \\ \mathcal{F}^{-1}(\psi_j \mathcal{F}f) = \check{\psi}_j * f, & j \geq 0. \end{cases}$$

**Remark 3.1.** For  $j \geq -1$ , by definitions, one sees that

$$\mathcal{R}_j = \mathcal{R}_j \tilde{\mathcal{R}}_j, \quad \text{where } \tilde{\mathcal{R}}_j := \sum_{\ell=-1}^1 \mathcal{R}_{j+\ell} \text{ with } \mathcal{R}_{-2} := 0, \quad (3.3)$$

and  $\mathcal{R}_j$  is symmetric in the sense of

$$\int_{\mathbb{R}^d} \mathcal{R}_j f(x) g(x) dx = \int_{\mathbb{R}^d} f(x) \mathcal{R}_j g(x) dx, \quad f \in \mathcal{S}'(\mathbb{R}^d), g \in \mathcal{S}(\mathbb{R}^d). \quad (3.4)$$

Here is the definition of Besov spaces.

**Definition 3.2** (*Besov Spaces*). For any  $\beta \in \mathbb{R}$  and  $p, q \in [1, \infty]$ , the Besov space  $\mathbf{B}_{p,q}^\beta(\mathbb{R}^d)$  is defined by

$$\mathbf{B}_{p,q}^\beta(\mathbb{R}^d) := \left\{ f \in \mathcal{S}'(\mathbb{R}^d) \mid \|f\|_{\mathbf{B}_{p,q}^\beta} := \left[ \sum_{j \geq -1} (2^{\beta j} \|\mathcal{R}_j f\|_p)^q \right]^{1/q} < \infty \right\}.$$

If  $p = q = \infty$ , it is in the sense

$$\mathbf{B}_{\infty,\infty}^\beta(\mathbb{R}^d) := \left\{ f \in \mathcal{S}'(\mathbb{R}^d) \mid \|f\|_{\mathbf{B}_{\infty,\infty}^\beta} := \sup_{j \geq -1} 2^{\beta j} \|\mathcal{R}_j f\|_\infty < \infty \right\}.$$

Recall the following Bernstein's inequality (cf. [1]\*Lemma 2.1).

**Lemma 3.3** (*Bernstein's inequality*). For any  $k \in \mathbb{N}$ , there is a constant  $c = c(d, k) > 0$  such that for all  $j \geq -1$ ,

$$\|\nabla^k \mathcal{R}_j f\|_\infty \leq c 2^{kj} \|\mathcal{R}_j f\|_\infty.$$

In particular, for any  $\alpha \in \mathbb{R}$ ,

$$\|\nabla^k f\|_{\mathbf{B}_{\infty,\infty}^\alpha} \leq c \|f\|_{\mathbf{B}_{\infty,\infty}^{\alpha+k}}. \quad (3.5)$$

**Remark 3.4** (*Equivalence Between Besov Spaces and Hölder Spaces*). If  $\beta > 0$  and  $\beta \notin \mathbb{N}$ , we have the following equivalence between  $\mathbf{B}_{\infty,\infty}^\beta(\mathbb{R}^d)$  and  $\mathbf{C}^\beta(\mathbb{R}^d)$ : (cf. [28])

$$\|f\|_{\mathbf{B}_{\infty,\infty}^\beta} \asymp \|f\|_{\mathbf{C}^\beta}. \quad (3.6)$$

However, for any  $n \in \mathbb{N}_0$ , we only have one side control that is

$$\|f\|_{\mathbf{B}_{\infty,\infty}^n} \lesssim \|f\|_{\mathbf{C}^n}. \quad (3.7)$$

By Bernstein's inequality, we have that for any  $|h| < 1/2$ ,

$$\begin{aligned} |f(x+h) - f(x)| &\stackrel{(3.2)}{\leq} \sum_{j \geq -1} |\mathcal{R}_j f(x+h) - \mathcal{R}_j f(x)| \\ &\lesssim \sum_{j < -\log_2 |h|} \|f\|_{\mathbf{B}_{\infty,\infty}^1} |h| + \sum_{j \geq -\log_2 |h|} 2^{-j} \|f\|_{\mathbf{B}_{\infty,\infty}^1} \\ &\lesssim \|f\|_{\mathbf{B}_{\infty,\infty}^1} |h| (\log_2 |h|^{-1} + 1), \end{aligned}$$

and for any  $|h| \geq 1/2$ ,

$$|f(x+h) - f(x)| \leq 2\|f\|_\infty \leq 4|h| \|f\|_{\mathbf{B}_{\infty,\infty}^1}.$$

Thus, by (3.5), we obtain that

$$\sup_{x \neq y} \frac{|\nabla^k f(x) - \nabla^k f(y)|}{|x - y| (\log_2^+ |x - y|^{-1} + 1)} \lesssim \|f\|_{\mathbf{B}_{\infty,\infty}^{k+1}}, \quad \text{for any } k \in \mathbb{N}_0. \quad (3.8)$$

Now we introduce the estimate of Littlewood-Paley's type for the heat kernel  $p_\alpha(t, x)$ . The same result is proved in [15]\*Lemma 3.1 for  $\alpha = 2$  and [8]\*Lemma 3.3 and [14]\*Lemma 2.12 for  $\alpha \in (0, 2)$ . For reader's convenience, we give a proof here.

**Lemma 3.5.** *Let  $\alpha \in (0, 2)$ . There is a constant  $c = c(\alpha, d) > 1$  such that for all  $j \geq -1$  and  $T > 0$ ,*

$$\int_0^T \int_{\mathbb{R}^d} |\mathcal{R}_j p_\alpha(t, x)| dx dt \leq c(1 + T)2^{-\alpha j}. \quad (3.9)$$

**Proof.** First of all, by the scaling property (2.3), we have that for any  $m \in \mathbb{N}_0$ ,

$$\int_{\mathbb{R}^d} |(\Delta^m p_\alpha)(t, x)| dx = t^{-2m/\alpha} \int_{\mathbb{R}^d} |\Delta^m p_\alpha(1, x)| dx \lesssim t^{-2m/\alpha}. \quad (3.10)$$

For  $j = -1$ , we have

$$\int_0^T \int_{\mathbb{R}^d} |\mathcal{R}_{-1} p_\alpha(t, x)| dx dt \lesssim \int_0^T \|p_\alpha(t, \cdot)\|_1 dt = T2^{-\alpha}2^\alpha \leq T2^\alpha.$$

For  $j \geq 0$ , by (3.1) and the change of variables,

$$\int_{\mathbb{R}^d} |\mathcal{R}_j p_\alpha(t, x)| dx = 2^{-jd} \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} p_\alpha(t, 2^{-j}(x - y)) \check{\psi}(y) dy \right| dx. \quad (3.11)$$

Notice that the support of  $\psi$  is contained in an annulus. By [12, (1.2.1)], we have that  $\Delta^{-m}\check{\psi}$  is a well-defined Schwartz function where

$$\mathcal{F}(\Delta^{-m}\check{\psi})(\xi) := (-|\xi|^2)^{-m} \psi(\xi) \in \mathcal{S}(\mathbb{R}^d), \quad m \in \mathbb{N}_0.$$

Based on this, we have  $\check{\psi} = \Delta^m \Delta^{-m} \check{\psi}$ ,  $m \in \mathbb{N}_0$  and

$$\int_{\mathbb{R}^d} p_\alpha(t, 2^{-j}(x - y)) \check{\psi}(y) dy = \int_{\mathbb{R}^d} \Delta^m p_\alpha(t, 2^{-j}(x - y)) (\Delta^{-m} \check{\psi})(y) dy.$$

Hence,

$$\begin{aligned} \int_{\mathbb{R}^d} |\mathcal{R}_j p_\alpha(t, x)| dx &\stackrel{(3.11)}{\lesssim} 2^{-jd} \int_{\mathbb{R}^d} |\Delta^m p_\alpha(t, 2^{-j}x)| dx \\ &= 2^{-2jm} \int_{\mathbb{R}^d} |(\Delta^m p_\alpha)(t, x)| dx \stackrel{(3.10)}{\lesssim} 2^{-2jm} t^{-2m/\alpha}. \end{aligned}$$

Then, considering the cases  $m = 0$  and  $m = 2$ , one sees that

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} |\mathcal{R}_j p_\alpha(t, x)| dx dt &= \left( \int_0^{2^{-\alpha j}} + \int_{2^{-\alpha j}}^T \right) \int_{\mathbb{R}^d} |\mathcal{R}_j p_\alpha(t, x)| dx dt \\ &\lesssim \int_0^{2^{-\alpha j}} dt + 2^{-4j} \int_{2^{-\alpha j}}^T t^{-4/\alpha} dt \lesssim 2^{-\alpha j}. \end{aligned}$$

The proof is finished.

Following the method used in [14, 15], we give a well-known a priori estimate of Besov-type by (3.9). The result is seen as Schauder's estimate when  $p = q = \infty$  in the literature. In the sequel, for a Banach space  $\mathbb{B}$  and  $T > 0$ ,  $q \in [1, \infty]$ , we denote by

$$\mathbb{L}_T^q(\mathbb{B}) := L^q([0, T]; \mathbb{B}), \quad \mathbb{L}_T^q := L^q([0, T] \times \mathbb{R}^d).$$

**Lemma 3.6.** Let  $\alpha \in (0, 2)$ ,  $\beta \in \mathbb{R}$ . For any  $p \in [1, \infty]$  and  $q \in [1, \infty]$ , there is a constant  $c = c(d, \alpha, \beta, p, q) > 0$  such that for all  $(u, f) \in \mathcal{S}' \times \mathcal{S}'$  with

$$\partial_t u = \Delta^{\alpha/2} u + f, \quad u(0) = u_0,$$

in the following weak sense

$$\langle u(t), \varphi \rangle = \langle u_0, \varphi \rangle + \int_0^t \langle u(s), \Delta^{\alpha/2} \varphi \rangle ds + \int_0^t \langle f(s), \varphi \rangle ds, \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d),$$

and for any  $T > 0$ ,

$$\|u\|_{\mathbb{L}_T^q(\mathbf{B}_{p,q}^{\alpha+\beta})} \leq c \left( T^{1/q} \|u_0\|_{\mathbf{B}_{p,q}^{\alpha+\beta}} + (1+T) \|f\|_{\mathbb{L}_T^q(\mathbf{B}_{p,q}^{\beta})} \right). \quad (3.12)$$

**Proof.** We only give the proof under  $q \in [1, \infty)$ , since the case of  $q = \infty$  is similar and easier. Let  $\{\rho_\varepsilon\}_{\varepsilon>0}$  be a usual mollifier on  $\mathbb{R}^d$ . Then  $u_\varepsilon := u * \rho_\varepsilon$  and  $f_\varepsilon = f * \rho_\varepsilon$  satisfy

$$\partial_t u_\varepsilon(t, x) = \Delta^{\alpha/2} u_\varepsilon(t, x) + f_\varepsilon(t, x), \quad u_\varepsilon(0) = u_0 * \rho_\varepsilon.$$

Thus, without loss of generality, we assume that  $u, f, u_0 \in C^\infty$ . For any  $t \in [0, T]$ , let  $u^t(s) := u(t-s)$  and  $f^t(s) := f(t-s)$  for any  $s \in (0, t)$ . Obviously,

$$\partial_s u^t(s, x) + \Delta^{\alpha/2} u^t(s, x) = -f^t(s, x), \quad u^t(0) = u(t).$$

By Itô's formula (cf. [17, Theorem 5.1]), we have

$$\mathbb{E} u^t(t, x + L_t) = u^t(0, x) - \int_0^t \mathbb{E} f^t(s, x + L_s) ds.$$

Then, we have Duhamel's formula:

$$u(t, x) = \int_{\mathbb{R}^d} p_\alpha(t, x-y) u_0(y) dy + \int_0^t \int_{\mathbb{R}^d} p_\alpha(s, x-y) f(t-s, y) dy ds.$$

Taking  $\mathcal{R}_j$  for both sides, by (3.3) and (3.4), we get

$$\mathcal{R}_j u(t, x) = \int_{\mathbb{R}^d} p_\alpha(t, x-y) \mathcal{R}_j u_0(y) dy + \int_0^t \int_{\mathbb{R}^d} \mathcal{R}_j p_\alpha(s, x-y) \tilde{\mathcal{R}}_j f(t-s, y) dy ds.$$

From this, by Minkowski's inequality and Hölder's inequality, one sees that

$$\begin{aligned} \|\mathcal{R}_j u\|_{\mathbb{L}_T^q(L^p)} &\leq T^{1/q} \|\mathcal{R}_j u_0\|_p + \int_0^T \|\mathcal{R}_j p_\alpha(s)\|_1 \|\tilde{\mathcal{R}}_j f(\cdot - s) 1_{\cdot > s}\|_{\mathbb{L}_T^q(L^p)} ds \\ &\stackrel{(3.9)}{\lesssim} T^{1/q} \|\mathcal{R}_j u_0\|_p + (1+T) 2^{-\alpha j} \|\tilde{\mathcal{R}}_j f\|_{\mathbb{L}_T^q(L^p)}. \end{aligned}$$

By definitions and Fubini's theorem, we have

$$\|u\|_{\mathbb{L}_T^q(\mathbf{B}_{p,q}^{\alpha+\beta})}^q = \sum_{j \geq -1} 2^{(\alpha+\beta)qj} \|\mathcal{R}_j u\|_{\mathbb{L}_T^q(L^p)}^q.$$

Therefore,

$$\begin{aligned} \|u\|_{\mathbb{L}_T^q(\mathbf{B}_{p,q}^{\alpha+\beta})}^q &\lesssim T \sum_{j \geq -1} 2^{(\alpha+\beta)qj} \|\mathcal{R}_j u_0\|_p^q + (1+T)^q \sum_{j \geq -1} 2^{\beta qj} \|\mathcal{R}_j f\|_{\mathbb{L}_T^q(L^p)}^q \\ &\lesssim T \|u_0\|_{\mathbf{B}_{p,q}^{\alpha+\beta}}^q + (1+T)^q \|f\|_{\mathbb{L}_T^q(\mathbf{B}_{p,q}^{\beta})}^q \end{aligned}$$

which implies the desired estimate.

**Remark 3.7** (cf. [15, Section 3]). The above result is true for  $\alpha = 2$ , which means we can get the same result for the Laplacian  $\Delta$  case by considering Brownian motion through the same method.

**Remark 3.8.** Here we compare Schauder estimates in Hölder spaces, Besov spaces and Sobolev spaces. By (3.12) for  $p = q = \infty$  and (3.6), we obtain the classical Schauder's estimate for  $\alpha = 2$ :

$$\|u\|_{\mathbb{L}_T^\infty(C^{2+\beta})} \lesssim \|u_0\|_{C^{2+\beta}} + \|f\|_{\mathbb{L}_T^\infty(C^\beta)}, \quad \beta \in (0, 1).$$

It is well-known that Schauder's estimate is not true for  $\beta = 0$ . But the lemma above tells us that

$$\|u\|_{\mathbb{L}_T^\infty(\mathbf{B}_{\infty,\infty}^2)} \lesssim \|u_0\|_{\mathbf{B}_{\infty,\infty}^2} + \|f\|_{\mathbb{L}_T^\infty(\mathbf{B}_{\infty,\infty}^0)}.$$

Furthermore, by (3.7) and (3.8), we get

$$|\nabla u(t, x) - \nabla u(t, y)| \lesssim |x - y| \left(1 + \log_2^+ |x - y|^{-1}\right) \left(\|u_0\|_{C^2} + \|f\|_{\mathbb{L}_T^\infty}\right).$$

In Sobolev spaces, it holds that

$$\|u\|_{\mathbb{L}_T^q(W^{2,p})} \lesssim \|u_0\|_{\mathbb{L}_T^q(W^{2,p})} + \|f\|_{\mathbb{L}_T^q(L^p)}$$

with  $p, q \neq 1, \infty$  (see [20,30] and references therein). However,  $p, q = 1$  or  $\infty$  are allowed in Besov case.

#### 4. Estimates of heat kernels for Euler–Maruyama scheme

In this section, assume that  $\alpha \in (1, 2)$  and  $b : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a bounded measurable function. Fix  $T > 0$  and  $x \in \mathbb{R}^d$ . Consider the following Euler scheme  $X_t^N(x)$ :  $X_0^N = x$ , and

$$X_t^N = x + \int_0^t b(s, X_{\phi_N(s)}^N) ds + L_t, \quad t \in (0, T], \quad (4.1)$$

where  $N \in \mathbb{N}$ ,  $\phi_N(s) := kh$  for  $s \in [kh, (k+1)h)$  with  $h := T/N$  and  $k = 0, 1, \dots, N-1$ . First of all, we prove the following Duhamel's formula for the Euler scheme.

**Lemma 4.1** (Duhamel's Formula). *Let  $\alpha \in (1, 2)$ . For each  $t \in (0, T]$  and  $x \in \mathbb{R}^d$ ,  $X_t^N(x)$  admits a density  $p_x^N(t, \cdot)$  satisfying the following Duhamel's formula:*

$$p_x^N(t, y) = p_\alpha(t, x - y) + \int_0^t \mathbb{E} \left[ b(s, X_{\phi_N(s)}^N) \cdot \nabla p_\alpha(t - s, X_s^N - y) \right] ds. \quad (4.2)$$

**Proof.** Fix  $t \in (0, T]$  and  $f \in C_c^\infty(\mathbb{R}^d)$ . Letting  $s \in [0, t]$  and

$$u(s, x) := p_\alpha(t - s, \cdot) * f(x) = \int_{\mathbb{R}^d} p_\alpha(t - s, x - y) f(y) dy,$$

by (2.8), it is easy to see that  $u(s, x)$  solves the following equation:

$$(\partial_s + \Delta^{\alpha/2})u = 0, \quad u(t, x) = f(x). \quad (4.3)$$

By Itô's formula (cf. [17, Theorem 5.1]), we have

$$\begin{aligned} u(t, X_t^N) &= u(0, x) + \int_0^t (\partial_s u)(s, X_{s-}^N) ds + \int_0^t b(s, X_{\phi_N(s)}^N) \cdot \nabla u(s, X_{s-}^N) ds \\ &\quad + \int_0^t \int_{|z|>1} \left( u(s, X_{s-}^N + z) - u(s, X_{s-}^N) \right) N(ds, dz) \\ &\quad + \int_0^t \int_{0<|z|\leq 1} \left( u(s, X_{s-}^N + z) - u(s, X_{s-}^N) \right) \tilde{N}(ds, dz) \\ &\quad + \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} \left( u(s, X_s^N + z \mathbf{1}_{|z|\leq 1}) - u(s, X_s^N) - z \mathbf{1}_{|z|\leq 1} \cdot \nabla u(s, X_s^N) \right) \nu^{(\alpha)}(dz) ds. \end{aligned}$$

Observe that a càdlàg function can have at most a countable number of jumps. Taking the expectation for both sides in the above equality, by [17, Section 3], (2.1) and (4.3), we obtain that for any  $f \in C_c^\infty(\mathbb{R}^d)$ ,

$$\mathbb{E} f(X_t^N) = \mathbb{E} u(t, X_t^N) = u(0, x) + \int_0^t \mathbb{E} \left( b(s, X_{\phi_N(s)}^N) \cdot \nabla u(s, X_s^N) \right) ds.$$

Furthermore, since

$$\int_0^t \int_{\mathbb{R}^d} |\nabla p_\alpha(s, y)| dy ds \stackrel{(2.3)}{=} \int_{\mathbb{R}^d} |\nabla p_\alpha(1, x)| dx \int_0^t s^{-1/\alpha} ds < \infty, \quad \text{if } \alpha \in (1, 2),$$

we derive the desired Duhamel's formula.  $\square$

**Remark 4.2.** For any general initial value  $X_0^N = X_0 \in \mathcal{F}_0$ , since  $L$  is independent of  $X_0$ ,  $X_t^N(x)$  defined by (4.1) is independent of  $X_0$ . Consequently, by [19, Lemma 3.11], the Euler scheme  $X_t^N$  with initial value  $X_0$  also has a density  $p_{X_0}^N(t, y)$  given by

$$p_{X_0}^N(t, y) = \int_{\mathbb{R}^d} p_x^N(t, y) \mathbb{P} \circ X_0^{-1}(dx). \quad (4.4)$$

The following uniform estimate for  $p_x^N(t, y)$  was proved by Huang, Suo and Yuan [16] when the coefficient  $b$  takes the form  $b(x)$ . For the convenience of readers, we show it again by the method from [13].

**Theorem 4.3.** *Let  $\alpha \in (1, 2)$ . For any  $T > 0$ , there is a constant  $c = c(d, \alpha, T, \|b\|_\infty) > 0$  such that for any  $N \in \mathbb{N}$ ,  $t \in (0, T]$  and  $x, y \in \mathbb{R}^d$ ,*

$$p_x^N(t, y) \leq c p_\alpha(t, x - y). \quad (4.5)$$

**Proof.** For the simplicity, we use a slight abuse of notation  $\|b\|_\infty := \|b\|_{\mathbb{L}_T^\infty}$  in the following. First of all, by (2.4), (2.6) and (2.7), we know that there is a constant  $c_0 = c_0(d, \alpha) > 2$  such that

$$|\nabla p_\alpha(t, x)| \leq c_0 t^{-1/\alpha} p_\alpha(t, x), \quad (4.6)$$

and

$$p_\alpha(t, x + z) \leq c_0 p_\alpha(t, x), \quad \text{if } |z| \leq 2t^{1/\alpha}. \quad (4.7)$$

Below, we fix this constant  $c_0$  and  $T > 0$ . Let  $\varepsilon > 0$  be small enough such that

$$\ell_\varepsilon := c_0^2 \frac{\alpha}{\alpha-1} \|b\|_\infty \varepsilon^{(\alpha-1)/\alpha} \leq 1/2.$$

Denote

$$h := T/N \quad \text{and} \quad M := [\varepsilon/h] \in \mathbb{N}.$$

Without loss of generality, we assume

$$N \geq (T(\frac{1}{2}\|b\|_\infty)^{\alpha/(\alpha-1)}) \vee (2T/\varepsilon). \quad (4.8)$$

Indeed, if  $N$  is bounded by some positive integer  $N_0$ , an elementary iteration can prove the inequality (4.5). Although the constant  $c$  depends on the number of iterations  $N$ , it is uniformly bounded since  $N$  has an upper bound. Precisely, assuming (4.5) holds for  $p_x^N(kh, y)$  with some constant  $\tilde{c}$ , it suffices to show that (4.5) holds for  $p_x^N((k+1)h, y)$  with another constant. Note that

$$X_{(k+1)h}^N = X_{kh}^N + \int_{kh}^{(k+1)h} b(s, X_{kh}^N) ds + L_{(k+1)h} - L_{kh},$$

where  $L_{(k+1)h} - L_{kh}$  is independent of  $X_{kh}^N$  and has density  $p_\alpha(h, \cdot)$ . Hence, by [19, Lemma 3.11], we have

$$p_x^N((k+1)h, y) = \int_{\mathbb{R}^d} p_x^N(kh, z) p_\alpha\left(h, z + \int_{kh}^{(k+1)h} b(s, z) ds - y\right) dz.$$

Observing that

$$\begin{aligned} h^{\frac{1}{\alpha}} + |y| &\leq h^{\frac{1}{\alpha}} + \left| y + \int_{kh}^{(k+1)h} b(s, z) ds \right| + h\|b\|_\infty \\ &\leq (1 + T^{1-\frac{1}{\alpha}}\|b\|_\infty) \left( h^{\frac{1}{\alpha}} + \left| y + \int_{kh}^{(k+1)h} b(s, z) ds \right| \right), \end{aligned}$$

we obtain, by the definition (2.5), that

$$\begin{aligned} p_\alpha\left(h, z + \int_{kh}^{(k+1)h} b(s, z) ds - y\right) &\stackrel{(2.4)}{\lesssim} \varrho_\alpha\left(h, z - y + \int_{kh}^{(k+1)h} b(s, z) ds\right) \\ &\stackrel{(2.4)}{\lesssim} \varrho_\alpha(h, z - y) \stackrel{(2.4)}{\lesssim} p_\alpha(h, z - y), \end{aligned}$$

where the implicit constants in the above  $\lesssim$  only depend on  $d, \alpha, T, \|b\|_\infty$ . Thus, there is a constant  $\tilde{c}_0 = \tilde{c}_0(d, \alpha, T, \|b\|_\infty)$  such that

$$p_x^N((k+1)h, y) \leq \tilde{c}_0 \int_{\mathbb{R}^d} p_\alpha(kh, x - z) p_\alpha(h, z - y) dz = \tilde{c}_0 p_\alpha((k+1)h, x - y),$$

where we used the C-K Eqs. (2.9) in the last equality. One sees that when the above process is repeated once, the constant  $\tilde{c}_0$  is multiplied once more. Fortunately, due to the  $N_0$  constraint, the iteration can end at  $N$ . Furthermore, by similar calculations or just like (Step 3), we get (4.5) for any  $t \in (kh, (k+1)h)$  with  $k = 0, \dots, N-1$ .

In the sequel, we estimate  $p_{kh}^N$  under the assumption (4.8). Since  $N$  should go to infinity in this case, we have to estimate  $p_{kh}^N$  more carefully. By (4.8), we have  $\|b\|_\infty \leq 2h^{-1+1/\alpha}$  and  $\varepsilon > h$ .

**(Step 1)** In this step, by induction, we prove the following result: for  $k = 1, 2, \dots, M \wedge N$ ,

$$p_x^N(kh, y) \leq c_0 p_\alpha(kh, x - y). \quad (4.9)$$

For  $k = 1$ , noting that  $X_h^N = x + \int_0^h b(s, x)ds + L_h$  with  $\|b\|_\infty \leq 2h^{-(\alpha-1)/\alpha}$ , by (4.7) we get that

$$p_x^N(h, y) = p_\alpha(h, y - x - \int_0^h b(s, x)ds) \leq c_0 p_\alpha(h, x - y).$$

Suppose now that (4.9) holds for  $j = 1, 2, \dots, k-1$ . By Duhamel's formula (4.2), we see that

$$\begin{aligned} p_x^N(kh, y) - p_\alpha(kh, x - y) &= \int_0^{kh} \mathbb{E} \left[ b(s, X_{\phi_N(s)}^N) \cdot \nabla p_\alpha(kh - s, X_s^N - y) \right] ds \\ &= \sum_{j=0}^{k-1} \int_{jh}^{(j+1)h} I_j^N(s) ds, \end{aligned} \quad (4.10)$$

where  $I_j^N(s) := \mathbb{E} \left[ b(s, X_{jh}^N) \cdot \nabla p_\alpha(kh - s, X_s^N - y) \right]$ . Observe that for  $s \in (jh, (j+1)h)$ ,

$$X_s^N = X_{jh}^N + \int_{jh}^s b(r, X_{jh}^N)dr + (L_s - L_{jh}).$$

Since  $L_s - L_{jh}$  is independent of  $X_{jh}^N$  and has density  $p_\alpha(s - jh, \cdot)$ , by [19, Lemma 3.11] and C-K Eqs. (2.9), we have

$$\begin{aligned} I_j^N(s) &= \mathbb{E} \left[ b(s, X_{jh}^N) \cdot \nabla p_\alpha(kh - s) * p_\alpha(s - jh) \left( X_{jh}^N + \int_{jh}^s b(r, X_{jh}^N)dr - y \right) \right] \\ &= \mathbb{E} \left[ b(s, X_{jh}^N) \cdot \nabla p_\alpha \left( kh - jh, X_{jh}^N + \int_{jh}^s b(r, X_{jh}^N)dr - y \right) \right] \\ &\leq \|b\|_\infty \int_{\mathbb{R}^d} |\nabla p_\alpha| \left( kh - jh, z - y + \int_{jh}^s b(r, z)dr \right) p_x^N(jh, z) dz. \end{aligned}$$

Furthermore, by (4.6), (4.7) and induction hypothesis, we obtain that for  $s \in (jh, (j+1)h)$ ,

$$\begin{aligned} I_j^N(s) &\leq \|b\|_\infty (kh - jh)^{-1/\alpha} c_0^2 \int_{\mathbb{R}^d} p_\alpha(kh - jh, z - y) \cdot c_0 p_\alpha(jh, x - z) dz \\ &\leq c_0 \frac{\alpha-1}{\alpha} \ell_\varepsilon \varepsilon^{-(\alpha-1)/\alpha} (kh - s)^{-1/\alpha} p_\alpha(kh, x - y), \end{aligned}$$

where we have used  $h\|b\|_\infty \leq 2h^{1/\alpha}$ . Substituting this into (4.10), we get, since  $kh \leq Mh \leq \varepsilon$  and  $\alpha \in (1, 2)$ , that

$$\begin{aligned} |p_x^N(kh, y) - p_\alpha(kh, x - y)| &\leq c_0 \ell_\varepsilon \varepsilon^{-(\alpha-1)/\alpha} p_\alpha(kh, x - y)^{\frac{\alpha-1}{\alpha}} \int_0^{kh} (kh - s)^{-1/\alpha} ds \\ &= c_0 \ell_\varepsilon \varepsilon^{-(\alpha-1)/\alpha} (kh)^{(\alpha-1)/\alpha} p_\alpha(kh, x - y) \\ &\leq c_0 \ell_\varepsilon p_\alpha(kh, x - y), \end{aligned}$$

which implies that

$$p_x^N(kh, y) \leq (c_0 \ell_\varepsilon + 1) p_\alpha(kh, x - y) \leq c_0 p_\alpha(kh, x - y).$$

**(Step 2)** Next we assume that  $M < N$ . Since  $\phi_N(s + Mh) = \phi_N(s) + Mh$ , we have

$$\begin{aligned} X_{t+Mh}^N &= X_{Mh}^N + \int_{Mh}^{t+Mh} b(s, X_{\phi_N(s)}^N) ds + (L_{t+Mh} - L_{Mh}) \\ &= X_{Mh}^N + \int_0^t b(s + Mh, X_{\phi_N(s)+Mh}^N) ds + (L_{t+Mh} - L_{Mh}). \end{aligned}$$

For  $t \in [0, Mh]$ , letting

$$\tilde{X}_t^N = X_{t+Mh}^N, \quad \tilde{L}_t = L_{t+Mh} - L_{Mh},$$

we have

$$\tilde{X}_t^N = \tilde{X}_0^N + \int_0^t b(s + Mh, \tilde{X}_{\phi_N(s)}^N) ds + \tilde{L}_t.$$

Noting that  $(\tilde{L}_t)_{t \geq 0} \stackrel{d}{=} (L_t)_{t \geq 0}$ , denoting by  $\tilde{p}_z^N(t, \cdot)$  the density of  $\tilde{X}_t^N$  with  $\tilde{X}_0^N = z$ , by Step 1, we have

$$\tilde{p}_z^N(jh, y) \leq c_0 p_\alpha(jh, z - y), \quad j = 1, \dots, M.$$

Hence, for  $j = 1, \dots, M$ , by (4.4), (4.9) and C-K Eqs. (2.9), we obtain that

$$\begin{aligned} p_x^N((j+M)h, y) &= \int_{\mathbb{R}^d} \tilde{p}_z^N(jh, y) p_x^N(Mh, z) dz \\ &\leq c_0^2 \int_{\mathbb{R}^d} p_\alpha(jh, z - y) p_\alpha(Mh, x - z) dz \\ &= c_0^2 p_\alpha((j+M)h, x - y), \end{aligned}$$

that is

$$p_x^N(kh, y) \leq c_0^2 p_\alpha(kh, x - y), \quad k = M+1, \dots, 2M.$$

Repeating the above procedure  $[N/M]$ -times, we get that

$$p_x^N(kh, y) \leq c_0^{[2T/\varepsilon]+1} p_\alpha(kh, x - y), \quad k = 1, \dots, N.$$

We point that the constant  $c_0^{[2T/\varepsilon]+1}$  is independent of  $N$ . **(Step 3)** Observe that for  $t \in (kh, (k+1)h)$ ,

$$X_t^N = X_{kh}^N + \int_{kh}^t b(s, X_{kh}^N) ds + (L_t - L_{kh}),$$

where  $L_t - L_{kh}$  is independent of  $X_{kh}^N$ . Thus, by [19, Lemma 3.11] and (4.7),

$$\begin{aligned} p_x^N(t, y) &= \int_{\mathbb{R}^d} p_x^N(kh, z) p_\alpha(t - kh, z + \int_{kh}^t b(s, z) ds - y) dz \\ &\leq c_0^{[2T/\varepsilon]+2} \int_{\mathbb{R}^d} p_\alpha(kh, x - z) p_\alpha(t - kh, z - y) dz \\ &= c_0^{[2T/\varepsilon]+2} p_\alpha(t, x - y). \end{aligned}$$

Here, we have used  $h\|b\|_\infty \leq 2h^{1/\alpha}$  and C-K Eqs. (2.9).  $\square$

The following corollary is a combination of Theorem 4.3, Lemma 2.1 and Lemma 2.3.

**Corollary 4.4.** Let  $\mu_0(dx) = \mathbb{P} \circ X_0^{-1}(dx)$  be the distribution of  $X_0$  and  $\alpha \in (1, 2)$ .

(i) For any  $T > 0$ , there is a constant  $c = c(d, \alpha, T, \|b\|_\infty) > 0$  such that for all  $N \in \mathbb{N}$ ,  $t \in (0, T]$  and  $y \in \mathbb{R}^d$ ,

$$p_{X_0}^N(t, y) \leq c \int_{\mathbb{R}^d} p_\alpha(t, x - y) \mu_0(dx). \quad (4.11)$$

(ii) For any  $T > 0$  and  $\beta \in (0, \alpha - 1)$ , there is a constant  $c = c(d, \alpha, T, \|b\|_\infty, \beta) > 0$  such that for all  $N \in \mathbb{N}$ ,  $t \in (0, T]$  and  $y_1, y_2 \in \mathbb{R}^d$ ,

$$|p_{X_0}^N(t, y_2) - p_{X_0}^N(t, y_1)| \leq c|y_2 - y_1|^\beta t^{-\beta/\alpha} \sum_{i=1,2} \int_{\mathbb{R}^d} p_\alpha(t, x - y_i) \mu_0(dx). \quad (4.12)$$

(iii) For any  $T > 0$  and  $\beta \in (0, \alpha - 1)$ , there is a constant  $c = c(d, \alpha, T, \|b\|_\infty, \beta) > 0$  such that for all  $N \in \mathbb{N}$ ,  $t_1, t_2 \in (0, T]$  and  $y \in \mathbb{R}^d$ ,

$$|p_{X_0}^N(t_2, y) - p_{X_0}^N(t_1, y)| \leq c|t_2 - t_1|^{\beta/\alpha} \sum_{i=1,2} t_i^{-\beta/\alpha} \int_{\mathbb{R}^d} p_\alpha(t_i, x - y) \mu_0(dx).$$

**Proof.** (i) is a direct consequence of (4.4) and Theorem 4.3.

(ii) By Duhamel's formula (4.2) and (4.4), we have

$$|p_{X_0}^N(t, y_2) - p_{X_0}^N(t, y_1)| \leq \mathcal{I}_1 + \mathcal{I}_2,$$

where

$$\mathcal{I}_1 := \int_{\mathbb{R}^d} |p_\alpha(t, x - y_2) - p_\alpha(t, x - y_1)| \mu_0(dx),$$

and

$$\mathcal{I}_2 := \|b\|_\infty \int_0^t \int_{\mathbb{R}^d} |\nabla p_\alpha(t - s, y_1 - z) - \nabla p_\alpha(t - s, y_2 - z)| p_{X_0}^N(s, z) dz ds,$$

For  $\mathcal{I}_1$ , by (2.10), we have

$$\mathcal{I}_1 \lesssim |y_2 - y_1|^\beta t^{-\beta/\alpha} \sum_{i=1,2} \int_{\mathbb{R}^d} p_\alpha(t, x - y_i) \mu_0(dx).$$

For  $\mathcal{I}_2$ , by (2.10), (i) and C-K Eqs. (2.9), we obtain that

$$\begin{aligned} \mathcal{I}_2 &\lesssim |y_2 - y_1|^\beta \int_0^t (t - s)^{-(1+\beta)/\alpha} \sum_{i=1,2} \left( \int_{\mathbb{R}^d} p_\alpha(t - s, z - y_i) \left[ \int_{\mathbb{R}^d} p_\alpha(s, x - z) \mu_0(dx) \right] dz \right) ds \\ &= |y_2 - y_1|^\beta \int_0^t (t - s)^{-(1+\beta)/\alpha} ds \sum_{i=1,2} \int_{\mathbb{R}^d} p_\alpha(t, x - y_i) \mu_0(dx) \\ &\lesssim |y_2 - y_1|^\beta t^{(\alpha-1-\beta)/\alpha} \sum_{i=1,2} \int_{\mathbb{R}^d} p_\alpha(t, x - y_i) \mu_0(dx), \end{aligned}$$

where we have used  $\beta \in (0, \alpha - 1)$ .

(iii) Suppose that  $t_1 < t_2$ . By Duhamel's formula (4.2) and (4.4), we have

$$|p_{X_0}^N(t_2, y) - p_{X_0}^N(t_1, y)| \leq \mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3,$$

where

$$\mathcal{J}_1 := \int_{\mathbb{R}^d} |p_\alpha(t_2, x - y) - p_\alpha(t_1, x - y)| \mu_0(dx),$$

$$\mathcal{J}_2 := \|b\|_\infty \int_{t_1}^{t_2} \int_{\mathbb{R}^d} |\nabla p_\alpha(t_2 - s, z - y)| p_{X_0}^N(s, z) dz,$$

and

$$\mathcal{J}_3 := \|b\|_\infty \int_0^{t_1} \int_{\mathbb{R}^d} |\nabla p_\alpha(t_2 - s, z - y) - \nabla p_\alpha(t_1 - s, z - y)| p_{X_0}^N(s, z) dz$$

For  $\mathcal{J}_1$ , by (2.12), we have

$$\mathcal{J}_1 \lesssim |t_2 - t_1|^{\beta/\alpha} \sum_{i=1,2} t_i^{-\beta/\alpha} \int_{\mathbb{R}^d} p_\alpha(t_i, x - y) \mu_0(dx).$$

For  $\mathcal{J}_2$ , by (2.6), (i) and C-K Eqs. (2.9), we get

$$\begin{aligned} \mathcal{J}_2 &\lesssim \int_{t_1}^{t_2} (t_2 - s)^{-1/\alpha} \left( \int_{\mathbb{R}^d} p_\alpha(t_2 - s, z - y) \int_{\mathbb{R}^d} p_\alpha(s, x - z) \mu_0(dx) dz \right) ds \\ &= \int_{t_1}^{t_2} (t_2 - s)^{-1/\alpha} ds \int_{\mathbb{R}^d} p_\alpha(t_2, x - y) \mu_0(dx) \\ &\lesssim (t_2 - t_1)^{-1/\alpha+1} \int_{\mathbb{R}^d} p_\alpha(t_2, x - y) \mu_0(dx). \end{aligned}$$

Since  $\beta \in (0, \alpha - 1)$ , we have

$$0 \leq (t_2 - t_1)^{-1/\alpha+1} \leq |t_2 - t_1|^{\beta/\alpha} t_2^{-\beta/\alpha+(\alpha-1)/\alpha} \leq |t_2 - t_1|^{\beta/\alpha} t_2^{-\beta/\alpha} T^{(\alpha-1)/\alpha}.$$

Hence,

$$\mathcal{J}_2 \lesssim |t_2 - t_1|^{\beta/\alpha} t_2^{-\beta/\alpha} \int_{\mathbb{R}^d} p_\alpha(t_2, x - y) \mu_0(dx).$$

For  $\mathcal{J}_3$ , by (2.12), (i) and C-K Eqs. (2.9), we obtain that

$$\begin{aligned} \mathcal{J}_3 &\lesssim |t_2 - t_1|^{\beta/\alpha} \sum_{i=1,2} \int_0^{t_1} (t_i - s)^{-(1+\beta)/\alpha} \left( \int_{\mathbb{R}^d} p_\alpha(t_i - s, z - y) \int_{\mathbb{R}^d} p_\alpha(s, x - z) \mu_0(dx) dz \right) ds \\ &= |t_2 - t_1|^{\beta/\alpha} \sum_{i=1,2} \int_0^{t_1} (t_i - s)^{-(1+\beta)/\alpha} ds \int_{\mathbb{R}^d} p_\alpha(t_i, x - y) \mu_0(dx) \\ &\leq \frac{\alpha}{\alpha-1-\beta} T^{(\alpha-1)/\alpha} |t_2 - t_1|^{\beta/\alpha} \sum_{i=1,2} t_i^{-\beta/\alpha} \int_{\mathbb{R}^d} p_\alpha(t_i, x - y) \mu_0(dx), \end{aligned}$$

where we have used  $\beta \in (0, \alpha - 1)$  and  $0 \leq t_1 < t_2 \leq T$ .

Combining the above calculations, we get the desired estimate.  $\square$

## 5. Proof of Theorem 1.2

Let  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  be a complete filtered probability space,  $L_t$  a  $d$ -dimensional symmetric and rotationally invariant  $\mathcal{F}_t$ -adapted  $\alpha$ -stable process with  $\alpha \in (1, 2)$ ,  $X_0$  an  $\mathcal{F}_0$ -measurable random variable with distribution  $\mu_0$ . Let  $T > 0$ ,  $N \in \mathbb{N}$  and  $h := T/N$ . Let  $X_t^N$  be the Euler approximation of DDSDE (1.1) constructed in the introduction. From the construction, it is easy to see that  $X_t^N$  solves the following SDE:

$$X_t^N = X_0 + \int_0^t b^N(s, X_{\phi_N(s)}^N) ds + L_t, \quad (5.1)$$

where

$$b^N(s, x) = \mathbf{1}_{\{s \geq h\}} b(s, x, \rho_{\phi_N(s)}^N(x))$$

and

$$\phi_N(s) = \sum_{j=0}^{\infty} jh \mathbf{1}_{[jh, (j+1)h)}(s).$$

Trivially,  $s - h \leq \phi_N(s) \leq s$ .

Let  $\mathbb{D}$  be the space of all càdlàg functions from  $[0, T]$  to  $\mathbb{R}^d$ . In the following,  $\mathbb{D}$  is equipped with Skorokhod topology which makes  $\mathbb{D}$  into a Polish space, and use  $d_{\mathbb{D}}$  to denote the associated metric.

**Lemma 5.1.** *The sequence of laws for  $(X_t^N)$  in  $(\mathbb{D}, d_{\mathbb{D}})$  is tight.*

**Proof.** It is trivial that the sequence of distributions for  $(X_0^N, L_0) \equiv (X_0, 0)$  is tight in  $\mathbb{R}^d \times \mathbb{R}^d$ . Taking  $q \in (\alpha/2, \alpha)$ , by Chebyshev's inequality, (5.1) and the fact (cf. [10, Lemma 2.4])

$$\mathbb{E}|L_t - L_s|^q \lesssim |t - s|^{q/\alpha}, \quad q \in (0, \alpha), \quad (5.2)$$

we obtain that for any  $N \in \mathbb{N}$ ,  $R > 0$  and  $0 \leq s < r < t \leq T$ ,

$$\begin{aligned} & \mathbb{P}\left(|X_r^N - X_s^N| \geq R, |X_t^N - X_r^N| \geq R\right) \\ & \leq \mathbb{P}\left(|L_r - L_s| + (r - s)\|b\|_{\infty} \geq R\right) \times \mathbb{P}\left(|L_t - L_r| + (t - r)\|b\|_{\infty} \geq R\right) \\ & \lesssim (r - s)^{q/\alpha} (t - r)^{q/\alpha} R^{-2q} \leq (t - s)^{2q/\alpha} R^{-2q}. \end{aligned} \quad (5.3)$$

Similarly, we have

$$\limsup_{\delta \downarrow 0} \sup_N \mathbb{P}\left(|X_{\delta}^N - X_0^N| \geq \varepsilon\right) = 0, \quad \forall \varepsilon > 0.$$

Hence, combining the above calculations, by [18, Theorem 4.1, p.355], we see that the sequence  $(X_t^N)$  is tight.  $\square$

Let  $p_x^N(t, \cdot)$  be the distributional density of the Euler scheme  $X_t^N(x)$  of SDE (5.1) starting from  $x$  at time 0. Since for each  $x \in \mathbb{R}^d$ ,  $X_t^N(x)$  is independent of  $X_0$ , the distributional density  $\rho_t^N(\cdot)$  of  $X_t^N$  with initial distribution  $\mu_0$  is given by

$$\rho_t^N(y) = \int_{\mathbb{R}^d} p_x^N(t, y) \mu_0(dx). \quad (5.4)$$

Furthermore, by Theorem 4.3, we have that for  $q > 1$ ,

$$\begin{aligned} \left(\int_{\mathbb{R}^d} |\rho_{\phi_N(t)}^N(y)|^q dy\right)^{1/q} &= \left(\int_{\mathbb{R}^d} \left|\int_{\mathbb{R}^d} p_x^N(\phi_N(t), y) \mu_0(dx)\right|^q dy\right)^{1/q} \\ &\lesssim \left(\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |p_{\alpha}(\phi_N(t), x - y)|^q \mu_0(dx) dy\right)^{1/q} \stackrel{(2.3)}{\lesssim} \phi_N(t)^{-\frac{d}{\alpha p}}, \end{aligned} \quad (5.5)$$

where  $1/q + 1/p = 1$ .

**Lemma 5.2.** *For fixed  $T > 0$ , there are a subsequence  $(N_k)_{k \in \mathbb{N}}$  and a continuous function  $\rho \in C((0, T] \times \mathbb{R}^d)$  such that for any  $M \in \mathbb{N}$  with  $M > 1/T$ ,*

$$\lim_{k \rightarrow \infty} \sup_{|y| \leq M} \sup_{1/M \leq t \leq T} |\rho_{t}^{N_k}(y) - \rho_t(y)| = 0. \quad (5.6)$$

**Proof.** By Theorem 4.3 and (2.4), we have that

$$\sup_{|y| \leq M} \sup_{1/M \leq t \leq T} |\rho_t^N(y)| \leq c \int_{\mathbb{R}^d} \sup_{|y| \leq M} \sup_{1/M \leq t \leq T} |p_{\alpha}(t, x - y)| \mu_0(dx) \leq c_M,$$

where  $c_M$  is independent of  $N$ . Moreover, by [Corollary 4.4](#), we have for any  $\beta \in (0, \alpha - 1)$ ,  $t_1, t_2 \in [1/M, T]$  and  $y_1, y_2 \in \mathbb{R}^d$ ,

$$\begin{aligned} |\rho_{t_1}^N(y_1) - \rho_{t_2}^N(y_2)| &\leq |\rho_{t_1}^N(y_1) - \rho_{t_2}^N(y_1)| + |\rho_{t_2}^N(y_1) - \rho_{t_2}^N(y_2)| \\ &\lesssim |t_1 - t_2|^{\beta/\alpha} M^{\beta/\alpha} \sum_{i=1,2} \int_{\mathbb{R}^d} p_\alpha(t_i, x - y_1) \mu_0(dx) \\ &\quad + |y_1 - y_2|^\beta M^{\beta/\alpha} \sum_{i=1,2} \int_{\mathbb{R}^d} p_\alpha(t_2, x - y_i) \mu_0(dx) \\ &\stackrel{(2.4)}{\lesssim} M^{(d+\beta)/\alpha} (|t_1 - t_2|^{\beta/\alpha} + |y_1 - y_2|^\beta), \end{aligned} \quad (5.7)$$

where the implicit constants in the above  $\lesssim$  are independent of  $N$ . Thus, by Ascoli-Arzelà's theorem, we conclude the proof and have [\(5.6\)](#).  $\square$

Now we are in a position to give

**Proof of Theorem 1.2. (Existence)** Fix  $T > 0$ . For the simplicity, we use a slight abuse of notation  $\|\cdot\|_\infty := \|\cdot\|_{\mathbb{L}_T^\infty}$  in some places. Let  $\mathbb{Q}_N$  be the law of  $(X^N, L)$  in  $\mathbb{D} \times \mathbb{D}$ . By [Lemma 5.1](#),  $\mathbb{Q}_N$  is tight. Therefore, by Prokhorov's theorem (cf. [\[19, Theorem 16.3\]](#)), for the subsequence in [Lemma 5.2](#), there are a subsubsequence  $(n_j)_{j \geq 1}$  and a probability measure  $\mathbb{Q}$  on  $\mathbb{D} \times \mathbb{D}$  so that

$$\mathbb{Q}_{n_j} \rightarrow \mathbb{Q} \text{ weakly.}$$

Below, for simplicity of notations, we still denote the above subsequence by  $\mathbb{Q}_N$ ,  $N \in \mathbb{N}$ . Then, by Skorokhod's representation theorem (cf. [\[19, Theorem 4.30\]](#)), there are a probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$  and random variables  $\tilde{X}, \tilde{L}$  thereon such that

$$(\tilde{X}^N, \tilde{L}^N) \rightarrow (\tilde{X}, \tilde{L}), \quad \tilde{\mathbb{P}} - \text{a.s.} \quad (5.8)$$

and

$$\tilde{\mathbb{P}} \circ (\tilde{X}^N, \tilde{L}^N)^{-1} = \mathbb{Q}_N = \mathbb{P} \circ (X^N, L)^{-1}, \quad \tilde{\mathbb{P}} \circ (\tilde{X}, \tilde{L})^{-1} = \mathbb{Q}. \quad (5.9)$$

In particular, the distributional density of  $\tilde{X}_t^N$  is  $\rho_t^N$ . Moreover, by [Lemma 5.2](#) and [\(5.8\)](#), for any  $t \in (0, T)$  and  $\varphi \in C_c^\infty(\mathbb{R}^d)$ ,

$$\tilde{\mathbb{E}}\varphi(\tilde{X}_t) = \lim_{N \rightarrow \infty} \tilde{\mathbb{E}}\varphi(\tilde{X}_t^N) = \lim_{N \rightarrow \infty} \int_{\mathbb{R}^d} \varphi(z) \rho_t^N(z) dz = \int_{\mathbb{R}^d} \varphi(z) \rho_t(z) dz.$$

In other words,  $\rho_t$  is the density of  $\tilde{X}_t$ . Define  $\tilde{\mathcal{F}}_t^N := \sigma\{\tilde{X}_s^N, \tilde{L}_s^N; s \leq t\}$ . Noting that

$$\mathbb{P}[L_t - L_s \in \cdot \mid \mathcal{F}_s] = \mathbb{P}\{L_t - L_s \in \cdot\},$$

we have

$$\tilde{\mathbb{P}}[\tilde{L}_t^N - \tilde{L}_s^N \in \cdot \mid \tilde{\mathcal{F}}_s^N] = \tilde{\mathbb{P}}[\tilde{L}_t^N - \tilde{L}_s^N \in \cdot],$$

which means that  $\tilde{L}_t^N$  is an  $(\tilde{\mathcal{F}}_s^N)$ -adapted  $\alpha$ -stable Lévy process. Thus, by [\(5.1\)](#) and [\(5.9\)](#) we obtain

$$\tilde{X}_t^N = \tilde{X}_0^N + \int_0^t b^N(s, \tilde{X}_{\phi_N(s)}^N) ds + \tilde{L}_t^N, \quad (5.10)$$

where  $b^N(s, \tilde{X}_{\phi_N(s)}^N) = \mathbf{1}_{\{s \geq h\}} b(s, \tilde{X}_{\phi_N(s)}^N, \rho_{\phi_N(s)}^N(\tilde{X}_{\phi_N(s)}^N))$ . We claim that

$$\int_0^t b^N(s, \tilde{X}_{\phi_N(s)}^N) ds \rightarrow \int_0^t b(s, \tilde{X}_s, \rho_s(\tilde{X}_s)) ds, \quad (5.11)$$

in probability as  $N \rightarrow \infty$ . Recalling the results in [18, p.339] and (5.8), one sees that for  $\tilde{P}$ -a.s.  $\omega$ , if  $\Delta \tilde{X}_t(\omega) = \Delta \tilde{L}_t(\omega) = 0$ , then

$$\tilde{X}_t^N(\omega) \rightarrow \tilde{X}_t(\omega).$$

Then, through taking  $N \rightarrow \infty$  in (5.10), it holds that for  $\tilde{\mathbb{P}}$ -a.s.  $\omega$ ,

$$\tilde{X}_t(\omega) = \tilde{X}_0(\omega) + \int_0^t b(s, \tilde{X}_s(\omega), \rho_s(\tilde{X}_s(\omega))) ds + \tilde{L}_t(\omega), \quad t \in D_\omega,$$

where

$$D_\omega := \{t \in \mathbb{R}_+ \mid \Delta \tilde{X}_t(\omega) = \Delta \tilde{L}_t(\omega) = 0\}.$$

Since  $\tilde{X}$  and  $\tilde{L}$  belong to  $\mathbb{D}$ ,  $D_\omega^c$  is a countable set in  $\mathbb{R}_+$  and

$$\tilde{X}_t(\omega) = \tilde{X}_0(\omega) + \int_0^t b(s, \tilde{X}_s(\omega), \rho_s(\tilde{X}_s(\omega))) ds + \tilde{L}_t(\omega), \quad t \in \mathbb{R}_+,$$

which derives the existence.

Let us now prove (5.11). Indeed, observe that

$$\tilde{\mathbb{E}} \left| \int_0^t b^N(s, \tilde{X}_{\phi_N(s)}^N) ds - \int_0^t b(s, \tilde{X}_s, \rho_s(\tilde{X}_s)) ds \right| \leq \mathcal{J}_1^N + \mathcal{J}_2^N + T \|b\|_\infty / N,$$

where

$$\mathcal{J}_1^N := \tilde{\mathbb{E}} \int_h^t \left| b(s, \tilde{X}_{\phi_N(s)}^N, \rho_{\phi_N(s)}^N(\tilde{X}_{\phi_N(s)}^N)) - b(s, \tilde{X}_{\phi_N(s)}^N, \rho_s(\tilde{X}_{\phi_N(s)}^N)) \right| ds$$

and

$$\mathcal{J}_2^N := \tilde{\mathbb{E}} \int_h^t \left| b(s, \tilde{X}_{\phi_N(s)}^N, \rho_s(\tilde{X}_{\phi_N(s)}^N)) - b(s, \tilde{X}_s, \rho_s(\tilde{X}_s)) \right| ds.$$

(1) For  $\mathcal{J}_1^N$ , we have

$$\begin{aligned} \mathcal{J}_1^N &\leq \tilde{\mathbb{E}} \int_h^t \mathbf{1}_{\{|\tilde{X}_{\phi_N(s)}^N| \leq R\}} \left| b(s, \tilde{X}_{\phi_N(s)}^N, \rho_{\phi_N(s)}^N(\tilde{X}_{\phi_N(s)}^N)) - b(s, \tilde{X}_{\phi_N(s)}^N, \rho_s(\tilde{X}_{\phi_N(s)}^N)) \right| ds \\ &\quad + 2 \|b\|_\infty \int_h^t \tilde{\mathbb{P}}(|\tilde{X}_{\phi_N(s)}^N| > R) ds := \mathcal{J}_{11}^N(R) + \mathcal{J}_{12}^N(R). \end{aligned}$$

Since

$$|\rho_{\phi_N(s)}^N(x) - \rho_s(x)| \leq |\rho_{\phi_N(s)}^N(x) - \rho_s^N(x)| + |\rho_s^N(x) - \rho_s(x)|,$$

by (5.6) and (5.7), we see that for each fixed  $(s, x) \in (0, T] \times \mathbb{R}^d$ ,

$$\lim_{N \rightarrow \infty} \mathbf{1}_{\{s \geq h\}} |\rho_{\phi_N(s)}^N(x) - \rho_s(x)| = 0,$$

which implies that for any  $(s, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ , by (1.7) we have

$$\lim_{N \rightarrow \infty} |b(s, x, \rho_{\phi_N(s)}^N(x)) - b(s, x, \rho_s(x))| = 0. \quad (5.12)$$

Moreover, by Hölder's equality and (5.5), we get

$$\begin{aligned}
 \mathcal{J}_{11}^N(R) &= \int_h^t \int_{|x| \leq R} \left| b(s, x, \rho_{\phi_N(s)}^N(x)) - b(s, x, \rho_s(x)) \right| \rho_{\phi_N(s)}^N(x) dx ds \\
 &\leq \left[ \int_h^t \int_{|x| \leq R} \left| b(s, x, \rho_{\phi_N(s)}^N(x)) - b(s, x, \rho_s(x)) \right|^p dx ds \right]^{1/p} \\
 &\quad \times \left[ \int_h^t \int_{|y| \leq R} |\rho_{\phi_N(s)}^N(y)|^q dy ds \right]^{1/q} \\
 &\lesssim \left[ \int_h^t \int_{|x| \leq R} \left| b(s, x, \rho_{\phi_N(s)}^N(x)) - b(s, x, \rho_s(x)) \right|^p dx ds \right]^{1/p} \\
 &\quad \times \left[ \int_h^T (s-h)^{-\frac{d}{\alpha}(q-1)} ds \right]^{1/q}
 \end{aligned} \tag{5.13}$$

provided  $1 < q < 1 + \alpha/d$  and  $1/p + 1/q = 1$ . Note that the implicit constant in the above  $\lesssim$  is independent of  $N, R$ . Thus, for any  $R > 0$ , by dominated convergence theorem and (5.12), we get that

$$\lim_{N \rightarrow \infty} \mathcal{J}_{11}^N(R) = 0. \tag{5.14}$$

For  $\mathcal{J}_{12}^N(R)$ , by (5.1), (5.2) and Chebyshev's inequality, we have

$$\begin{aligned}
 \int_0^t \tilde{\mathbb{P}}(|\tilde{X}_{\phi_N(s)}^N| > R) ds &= \int_0^t \mathbb{P}(|X_{\phi_N(s)}^N| > R) ds \\
 &\lesssim T \mathbb{P}(|X_0| + T \|b\|_\infty > R/2) + \int_0^t \frac{(\phi_N(s))^{1/\alpha}}{R/2} ds \\
 &\leq T \mathbb{P}(|X_0| + T \|b\|_\infty > R/2) + T^{(\alpha+1)/\alpha} (R/2)^{-1}
 \end{aligned} \tag{5.15}$$

which converges to zero uniformly in  $N$  as  $R \rightarrow \infty$ . Consequently, combining (5.14) and (5.15), we obtain that

$$\lim_{N \rightarrow \infty} \mathcal{J}_1^N = 0.$$

(2) For  $\mathcal{J}_2^N$ , let  $K_\varepsilon$  be a family of mollifiers in  $\mathbb{R}^d$  and define

$$B_\varepsilon(t, x) = b(t, \cdot, \rho_t(\cdot)) * K_\varepsilon(x).$$

Notice that  $\|B_\varepsilon\| \leq \|b\|_\infty$  and for any  $R > 0$ ,  $B_R := \{x \in \mathbb{R}^d \mid |x| < R\}$ ,

$$\lim_{\varepsilon \rightarrow 0} \|\mathbf{1}_{B_R}(B_\varepsilon - b)\|_p = 0. \tag{5.16}$$

Then

$$\mathcal{J}_2^N \leq \mathcal{J}_{21}^N(\varepsilon) + \mathcal{J}_{22}^N(\varepsilon) + \mathcal{J}_{23}^N(\varepsilon),$$

where

$$\begin{aligned}
 \mathcal{J}_{21}^N(\varepsilon) &:= \mathbb{E} \int_h^t |B_\varepsilon(s, \tilde{X}_{\phi_N(s)}^N) - B_\varepsilon(s, \tilde{X}_s)| ds, \\
 \mathcal{J}_{22}^N(\varepsilon) &:= \mathbb{E} \int_h^t |b(s, \tilde{X}_{\phi_N(s)}^N, \rho_s(\tilde{X}_{\phi_N(s)}^N)) - B_\varepsilon(s, \tilde{X}_{\phi_N(s)}^N)| ds
 \end{aligned}$$

and

$$\mathcal{J}_{23}^N(\varepsilon) := \mathbb{E} \int_h^t |b(s, \tilde{X}_s, \rho_s(\tilde{X}_s)) - B_\varepsilon(s, \tilde{X}_s)| ds.$$

Thus, by (5.8) and results in [18, p. 339], for any  $s > 0$ ,

$$\tilde{X}_s^N \mathbf{1}_{\Delta \tilde{X}_s=0}(s) \rightarrow \tilde{X}_s \mathbf{1}_{\Delta \tilde{X}_s=0}(s), \quad \text{as } N \rightarrow \infty, \tilde{\mathbb{P}}\text{-a.s.},$$

which, by the dominated convergence theorem, implies that for arbitrary fixed  $\varepsilon > 0$ ,

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E} \int_h^t |B_\varepsilon(s, \tilde{X}_s^N) - B_\varepsilon(s, \tilde{X}_s)| ds &\leq \mathbb{E} \int_0^t \lim_{N \rightarrow \infty} |B_\varepsilon(s, \tilde{X}_s^N) - B_\varepsilon(s, \tilde{X}_s)| \mathbf{1}_{\Delta \tilde{X}_s=0}(s) ds \\ &\quad + 2\|B_\varepsilon\|_\infty \mathbb{E} \int_0^t \mathbf{1}_{\Delta \tilde{X}_s>0}(s) ds = 0, \end{aligned}$$

where we use the fact that for Lebesgue a.e.  $s \in [0, t]$ ,  $\Delta \tilde{X}_s = 0$  since  $\tilde{X} \in \mathbb{D}$ . On the other hand, by (5.2), we have

$$\begin{aligned} \mathbb{E} \int_h^t |B_\varepsilon(s, \tilde{X}_{\phi_N(s)}^N) - B_\varepsilon(s, \tilde{X}_s^N)| ds &\leq \|\nabla B_\varepsilon\|_\infty \int_h^t \mathbb{E} |\tilde{X}_{\phi_N(s)}^N - \tilde{X}_s^N| ds \\ &\lesssim \|\nabla B_\varepsilon\|_\infty (|h| \|b\|_\infty + |h|^{1/\alpha}), \end{aligned}$$

where  $h = T/N$ . Consequently, for fixed  $\varepsilon > 0$ ,

$$\lim_{N \rightarrow \infty} \mathcal{J}_{21}^N(\varepsilon) = 0.$$

For  $\mathcal{J}_{22}^N(\varepsilon)$ , we have

$$\begin{aligned} \mathcal{J}_{22}^N(\varepsilon) &\leq \mathbb{E} \int_h^t \mathbf{1}_{\{|\tilde{X}_{\phi_N(s)}^N| \leq R\}} |b(s, \tilde{X}_{\phi_N(s)}^N, \rho_s(\tilde{X}_{\phi_N(s)}^N)) - B_\varepsilon(s, \tilde{X}_{\phi_N(s)}^N)| ds \\ &\quad + 2\|b\|_\infty \int_h^t \tilde{\mathbb{P}}(|\tilde{X}_{\phi_N(s)}^N| > R) ds := I_R^N(\varepsilon) + J_R^N. \end{aligned}$$

Samely as (5.13), by Hölder's inequality with  $1 < q < \alpha/d + 1$  and  $q = \frac{p}{p-1}$ , we see that

$$I_R^N(\varepsilon) \lesssim \left[ \int_0^T \int_{|y| \leq R} |b(s, y, \rho_s(y)) - B_\varepsilon(s, y)|^p dy ds \right]^{1/p} \left[ \int_h^T (s-h)^{-\frac{dq}{\alpha p}} ds \right]^{1/q},$$

where the implicit constant in the above  $\lesssim$  is independent of  $N, R$  and  $\varepsilon$ . Hence, for each  $R > 0$ , by the dominated convergence theorem and (5.16), we obtain

$$\limsup_{\varepsilon \rightarrow 0} \sup_N I_R^N(\varepsilon) = 0.$$

By (5.15), we have  $\lim_{R \rightarrow \infty} \sup_N J_R^N = 0$ . For  $\mathcal{J}_{23}^N(\varepsilon)$ , it is similar to  $\mathcal{J}_{22}^N(\varepsilon)$ .

Combining the above calculations, we get (5.11). The proof of the existence is finished.

**(Uniqueness)** For  $i = 1, 2$ , let  $\rho_t^{(i)}$  be two densities of two weak solutions  $X^{(i)}$  for DDSDE (1.1) respectively:

$$X_t^{(i)} = X_0 + \int_0^t b(s, X_s^{(i)}, \rho_s^{(i)}(X_s^{(i)})) ds + L_t^{(i)}.$$

For each  $i$ , by the well-known result (see [23] for example), for any  $x \in \mathbb{R}^d$ , there is a unique weak solution  $(\tilde{X}_t^{(i)}(x), \tilde{L}_t^{(i)})$  for the following classical SDE with bounded drift  $b(t, x, \rho_t^{(i)}(x))$

and  $\alpha \in (1, 2)$ :

$$\tilde{X}_t^{(i)}(x) = x + \int_0^t b(s, \tilde{X}_s^{(i)}(x), \rho_s^{(i)}(\tilde{X}_s^{(i)}(x))) ds + \tilde{L}_t^{(i)}.$$

Denote by  $p_x^{(i)}(t, \cdot)$  the density of  $\tilde{X}_t^{(i)}(x)$ . Then, we have

$$\rho_t^{(i)}(y) = \int_{\mathbb{R}^d} p_x^{(i)}(t, y) \rho_0(x) dx. \quad (5.17)$$

Moreover, by (4.5) and (5.6),

$$p_x^{(i)}(t, y) \leq c p_\alpha(t, x - y), \quad \forall (t, x, y) \in (0, T] \times \mathbb{R}^d \times \mathbb{R}^d. \quad (5.18)$$

In the following, we prove (i) and (ii) respectively.

(i) Define  $u_t := \rho_t^{(1)} - \rho_t^{(2)}$ . Hence, by (1.6),

$$\partial_t u_t = \Delta^{\alpha/2} u_t + \operatorname{div}(\mathcal{B}(t, x) u_t), \quad u_0 = 0, \quad (5.19)$$

in weak sense, where

$$\mathcal{B}(t, x) := b(t, x, \rho_t^{(1)}(x)) + \rho_t^{(2)}(x) \frac{b(t, x, \rho_t^{(1)}(x)) - b(t, x, \rho_t^{(2)}(x))}{u_t(x)},$$

and use the convention  $\frac{0}{0} = 0$ . In this case, we only need to prove that  $u_t = 0$ . By (5.17) and (5.18), we have

$$\|\rho_t^{(i)}\|_\infty \leq \|p_\alpha(t, \cdot)\|_p \|\rho_0\|_q \stackrel{(2.3)}{\lesssim} t^{-d/(\alpha q)} \in L^1([0, T]), \quad i = 1, 2,$$

where  $1/p + 1/q = 1$ . Notice that, by the assumption (1.11),

$$\|\mathcal{B}(t, \cdot)\|_\infty \lesssim \|b\|_\infty + \|\rho_t^{(2)}\|_\infty \lesssim 1 + t^{-d/(\alpha q)}. \quad (5.20)$$

By Duhamel's formula and (5.19), we have

$$u_t(x) = \int_0^t \left( p_\alpha(t-s) * \operatorname{div}(\mathcal{B}(s) u_s) \right)(x) ds.$$

Hence, for  $q > d/(\alpha - 1)$ , by (5.20) we have

$$\begin{aligned} \|u_t\|_\infty &\leq \int_0^t \|\nabla p_\alpha(t-s)\|_1 \|\mathcal{B}(s) u_s\|_\infty ds \\ &\stackrel{(2.3)}{\lesssim} \int_0^t (t-s)^{-1/\alpha} (1 + s^{-d/(\alpha q)}) \|u_s\|_\infty ds, \end{aligned}$$

which completes the proof by Gronwall's inequality (cf. [31, Example 2.4]).

(ii) Based on the weak uniqueness result (i) with  $q = \infty$ , we denote by

$$\rho_t := \rho_t^1 = \rho_t^2.$$

By the well-known result (see [11]\*Theorem 1.1 for example), if we have

$$\mathcal{A}(t, x) := b(t, x, \rho_t(x)) \in L^\infty([0, T]; \mathbf{C}^{\beta_0}) \quad (5.21)$$

for  $\beta_0 \in (1 - \alpha/2, 1)$  and any  $T > 0$ , then the strong uniqueness holds. Thus, it is enough to show that  $\rho_t(x) \in L^\infty([0, T]; \mathbf{C}^{\beta_0})$  for  $\beta_0 > 1 - \alpha/2$ . Unfortunately, we cannot obtain

it directly from  $\rho_0 \in \mathbf{C}^{\beta_0}$  and (5.17), since  $p_t(x, y) \neq p_t(x - y)$ . Firstly, by (5.17) and (5.18), we have

$$\sup_{t \in [0, T]} \|\rho_t\|_\infty \lesssim \|\rho_0\|_\infty \sup_{t \in [0, T]} \int_{\mathbb{R}^d} p_\alpha(t, y) dy = \|\rho_0\|_\infty,$$

which implies that

$$\bar{b}(t, x) := b(t, x, \rho_t(x))\rho_t(x) \in L^\infty([0, T] \times \mathbb{R}^d).$$

Hence, by (1.6), (3.12), (3.5) and (3.6), for any  $T > 0$ , there is a constant  $c_T$  such that for all  $t \in [0, T]$ ,

$$\|\rho_t\|_{\mathbf{C}^{(\alpha-1) \wedge \beta_0}} \leq c_T (\|\rho_0\|_{\mathbf{C}^{\beta_0}} + \|\bar{b}\|_{L_T^\infty}). \quad (5.22)$$

If  $\alpha - 1 > \beta_0$ , (5.21) is straightforward. Otherwise, by (1.11) and (5.22), we have

$$\|\bar{b}\|_{L_T^\infty(\mathbf{C}^{\alpha-1})} < \infty.$$

Thus, by (1.6), (3.12), (3.5) and (3.6) again, there is a constant  $c_T^{(2)}$  such that for all  $t \in [0, T]$ ,

$$\|\rho_t\|_{\mathbf{C}^{(2(\alpha-1)) \wedge \beta_0}} \leq c_T^{(2)}.$$

By induction, there are a  $N \in \mathbb{N}$  with

$$(\alpha - 1)N > \beta_0$$

and a constant  $c_T^{(N)}$  such that

$$\|\rho_t\|_{\mathbf{C}^{\beta_0}} \leq c_T^{(N)}, \quad \forall t \in [0, T].$$

Then, we have (5.21) and complete the proof.  $\square$

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

The authors are deeply grateful to Prof. Xicheng Zhang and Prof. Fuke Wu for their valuable suggestions and for correcting errors. Moreover, the authors would like to thank the reviewers for their reviews and helpful comments on an early draft of the paper.

## References

- [1] H. Bahouri, J.-Y. Chemin, R. Danchin, Fourier analysis and nonlinear partial differential equations, in: *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Vol. 343, Springer, Heidelberg, 2011.
- [2] V. Barbu, M. Röckner, Probabilistic representation for solutions to nonlinear Fokker-Planck equations, *SIAM J. Math. Anal.* 50 (4) (2018) 4246–4260.
- [3] V. Barbu, M. Röckner, From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE, *Ann. Probab.* 48 (4) (2020) 1902–1920.
- [4] V. Barbu, M. Röckner, Solutions for nonlinear Fokker-Planck equations with measures as initial data and McKean-Vlasov equations, *J. Funct. Anal.* 280 (7) (2021) Paper No. 108926, 35.

- [5] V. Barbu, M. Röckner, Uniqueness for nonlinear Fokker-Planck equations and weak uniqueness for McKean-Vlasov SDEs, *Stoch. Partial Differ. Equ. Anal. Comput.* 9 (3) (2021) 702–713.
- [6] C. Benazzoli, L. Campi, L.D. Persio, Mean field games with controlled jump-diffusion dynamics: existence results and an illiquid interbank market model, *Stochastic Process. Appl.* 130 (2020) 6927–6964.
- [7] R.M. Blumenthal, R.K. Gettoor, Some theorems on stable processes, *Trans. Amer. Math. Soc.* 95 (1960) 263–273.
- [8] Z.-Q. Chen, Z. Hao, X. Zhang, Hölder regularity and gradient estimates for SDEs driven by cylindrical  $\alpha$ -stable processes, *Electron. J. Probab.* 25 (2020) Paper No. 137, 23.
- [9] Z.-Q. Chen, X. Zhang, Heat kernels and analyticity of non-symmetric jump diffusion semigroups, *Probab. Theory Related Fields* 165 (1–2) (2016) 267–312.
- [10] Z.-Q. Chen, X. Zhang,  $L^p$ -maximal hypoelliptic regularity of nonlocal kinetic Fokker-Planck operators, *J. Math. Pures Appl.* (9) 116 (2018) 52–87.
- [11] Z.-Q. Chen, X. Zhang, G. Zhao, Supercritical SDEs driven by multiplicative stable-like Lévy processes, *Trans. Amer. Math. Soc.* 374 (11) (2021) 7621–7655.
- [12] L. Grafakos, *Modern Fourier Analysis*, third ed., in: *Graduate Texts in Mathematics*, Vol. 250, Springer, New York, 2014.
- [13] Z. Hao, M. Röckner, X. Zhang, Euler scheme for density dependent stochastic differential equations, *J. Differential Equations* 274 (2021) 996–1014.
- [14] Z. Hao, Z. Wang, M. Wu, Schauder’s estimates for nonlocal equations with singular Lévy measures, available at [arXiv:2002.09887](https://arxiv.org/abs/2002.09887).
- [15] Z. Hao, M. Wu, X. Zhang, Schauder estimates for nonlocal kinetic equations and applications, *J. Math. Pures Appl.* (9) 140 (2020) 139–184.
- [16] X. Huang, Y. Suo, C. Yuan, Estimate of heat kernel for Euler-Maruyama scheme of SDEs driven by  $\alpha$ -stable noise and applications, Available at [arXiv:2103.01323](https://arxiv.org/abs/2103.01323).
- [17] N. Ikeda, S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, second ed., in: *North-Holland Mathematical Library*, Vol. 24, North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989.
- [18] J. Jacod, A.N. Shiryaev, *Limit Theorems for Stochastic Processes*, second ed., in: *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Vol. 288, Springer-Verlag, Berlin, 2003.
- [19] O. Kallenberg, *Foundations of Modern Probability*, second ed., in: *Probability and its Applications* (New York), Springer-Verlag, New York, 2002.
- [20] N. Krylov, The heat equation in  $L^q((0, T), L^p)$ -spaces with weights, *SIAM J. Math. Anal.* 32 (2001) 1117–1141.
- [21] F. Kühn, R.L. Schilling, Strong convergence of the Euler-Maruyama approximation for a class of Lévy-driven SDEs, *Stochastic Process. Appl.* 129 (8) (2019) 2654–2680.
- [22] M. Liang, M.B. Majka, J. Wang, Exponential ergodicity for SDEs and McKean-Vlasov processes with Lévy noise, *Ann. Inst. Henri Poincaré Probab. Stat.* 57 (3) (2021) 1665–1701.
- [23] C. Ling, G. Zhao, Nonlocal elliptic equation in Hölder space and the martingale problem, *J. Differential Equations* 314 (2022) 653–699.
- [24] H. Masuda, Ergodicity and exponential  $\beta$ -mixing bounds for multidimensional diffusions with jumps, *Stochastic Process. Appl.* 117 (2007) 35–56.
- [25] E. Priola, Pathwise uniqueness for singular SDEs driven by stable processes, *Osaka J. Math.* 49 (2) (2012) 421–447.
- [26] K.-i. Sato, *Lévy Processes and Infinitely Divisible Distributions*, in: *Cambridge Studies in Advanced Mathematics*, Vol. 68, Cambridge University Press, Cambridge, 1999, Translated from the 1990 Japanese original, Revised by the author.
- [27] H. Tanaka, M. Tsuchiya, S. Watanabe, Perturbation of drift-type for Lévy processes, *J. Math. Kyoto Univ.* 14 (1974) 73–92.
- [28] H. Triebel, *Theory of Function Spaces. II*, in: *Monographs in Mathematics*, Vol. 84, Birkhäuser Verlag, Basel, 1992.
- [29] A.Y. Veretennikov, On the strong solutions of stochastic differential equations, *Theory Probab. Appl.* 24 (4) (1979) 354–366.
- [30] P. Xia, L. Xie, G. Zhao, X. Zhang,  $L^q(L^p)$ -theory of stochastic differential equations, *Stochastic Process. Appl.* 130 (2020) 5188–5211.
- [31] X. Zhang, Stochastic Volterra equations in Banach spaces and stochastic partial differential equation, *J. Funct. Anal.* 258 (4) (2010) 1361–1425.