Estimate of Transition Kernel for Euler-Maruyama Scheme for SDEs Driven by α-Stable Noise and Applications

Xing Huang[a)], **Yongqiang Suo**[*b)], **Chenggui Yuan**[c)]

[a)]Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

[b)]School of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

[c)]Department of Mathematics, Swansea University, Bay Campus, Swansea, SA1 8EN, UK

xinghuang@tju.edu.cn, suoyongqiang@nuaa.edu.cn, c.yuan@swansea.ac.uk

December 31, 2023

Abstract

In this paper, the discrete parametrix method is adopted to investigate the estimation of transition kernel for Euler-Maruyama scheme SDEs driven by α-stable noise, which implies Krylov's estimate and Khasminskii's estimate. As an application, the convergence rate of Euler-Maruyama scheme for a class of multidimensional SDEs with singular drift (by the use of Zvonkin's transformation) is obtained.

AMS subject Classification: 60H10, 34K26, 39B72.

Keywords: Zvonkin's transformation, Euler-Maruyama scheme, Transition kernel, Krylov's estimate.

1 Introduction

We consider the following $\mathbb{R}^d$-valued stochastic differential equation (SDE for short)

$$X_t = x + \int_0^t b(X_s)\mathrm{d}s + \int_0^t f(X_{s-})\mathrm{d}L_s, \tag{1.1}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d, f : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ are measurable functions, and $(L_t)_{t\geq 0}$ is a general $\mathbb{R}^d$-valued Lévy process defined on a complete filtration probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$.

If the coefficients b and f are Lipschitz continuous, the existence and uniqueness of strong solutions to (1.1) are established by the Picard iteration. Moreover, the SDE (1.1) can also be numerically solved with the Euler-Maruyama (EM for short) scheme, see [17] and references therein.

When the coefficients b and f are irregular, there has been a great interest in investigating pathwise uniqueness for SDE (1.1) in the past decades. A useful method in this direction is Zvonkin's transformation which was introduced in [34]. This method has been applied to various SDEs, see e.g. [7, 9, 13, 18, 29, 30, 31, 32] and references therein. Furthermore, Zvonkin's transformation has also been applied to investigate the convergence rate of EM scheme for SDEs, we refer to [1, 6, 8, 10, 21, 22, 23] for more details. There are also other transformation methods applied to EM scheme, see [19, 21, 22, 23].

In the continuous case, $L_s = as + \sigma W_s$, where $(W_s)_{\{s\geq 0\}}$ is a Brownian motion. [21] establishes the existence and uniqueness result and convergence of a numerical scheme for (1.1) in the one-dimensional case, the drift therein is piecewise Lipschitz. Their proof is based on a transformation, which globally transforms the piecewise Lipschitz drifts into Lipschitz ones. [22, 23] present a transformation for the multidimensional case which allows to prove an existence and uniqueness result for d-dimensional SDEs with discontinuous drift and degenerate diffusion coefficients. Compared with the Zvonkin's transformation, the transformation in [22, 23] does not have to solve a system of parabolic partial differential equations in each step. Recently, the strong convergence of EM scheme for SDEs with integrable drift has been obtained by the first author with his co-authors in [2], and the proof is based on the "parametrix method", which was introduced to obtain existence and estimates on the fundamental solutions of PDEs, see [15, 20].

For a general Lévy process, [11] shows that the EM scheme converges strongly with convergence rate $\frac{1}{2}$ for (1.1) with additive noise and a one-sided Lipschitz continuous drift b. [25] establishes the convergence rate for SDEs with Hölder continuous coefficients driven by Brownian motion and by truncated α-stable processes with index $\alpha > 1$. Moreover, [19] studies the strong convergence of the EM scheme for a large class of SDEs driven by Lévy processes such as isotropic α-stable, relative stable, layered stable processes, whose proofs rely on the so-called Itô-Tanaka trick which relates the time average $\int_0^t b(X_s)\mathrm{d}s$ of the solution $(X_t)_{t\geq 0}$ to (1.1). Under the assumption that the drift is Hölder continuous, the EM scheme for stochastic functional differential equations with α-stable noise is shown to converge in [12]. It is worth noting that the drifts of SDEs in these literature are assumed to be Hölder(-Dini) continuous or piecewise Lipschitz continuous. However, essential difficulty comes up when the drifts only belong to some Sobolev space. More precisely, it is not easy to obtain an estimate like

$$\mathbb{E}\left|\int_0^T b(X_t^{(\delta)}) - b(X_{t_\delta}^{(\delta)})\mathrm{d}t\right|^q \leq \Phi(\delta)$$

for some function $\Phi : [0,\infty) \to [0,\infty)$ with $\lim_{\delta\to 0}\Phi(\delta) = 0$. Here, $X_t^{(\delta)}$ stands for the numerical solution to the SDE and b is the singular drift, $t_\delta := \lfloor t/\delta \rfloor \delta$, and $\lfloor t/\delta \rfloor$ denotes the integer part of t/δ.

In this work, we consider the case where $(L_t)_{t\geq 0}$ is a d-dimensional Brownian motion subordinated by a subordinator. More precisely, let $W := (W_t)_{t\geq 0}$ and $S := (S_t)_{t\geq 0}$ be independent stochastic processes, where W is the Brownian motion on $\mathbb{R}^d$ with $W_0 = 0$ and S is an $\alpha/2$-stable (with $\alpha \in (1,2)$) subordinator independent of W, i.e., S is a one dimensional non-negative increasing Lévy process with $S_0 = 0$ and with Laplace transformation $\mathbb{E}\mathrm{e}^{-\gamma S_t} = \mathrm{e}^{-t\gamma^{\frac{\alpha}{2}}}, \gamma, t \geq 0$. By the scaling property, the process S_t has the same law as $t^{\frac{2}{\alpha}}S_1$. Then

$(W_{S_t})_{t\geq 0}$ is a Lévy process. We consider the following SDE:

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}W_{S_t}, \quad t \geq 0, \quad X_0 = x. \tag{1.2}$$

Herein, $b : \mathbb{R}^d \to \mathbb{R}^d$, $(W_{S_t})_{t\geq 0}$ is a rotationally invariant d-dimensional α-stable process, with the Lévy measure $\nu(\mathrm{d}z) = \frac{c_\alpha}{|z|^{d+\alpha}}\mathrm{d}z$ for some constant $c_\alpha > 0$.

Under assumptions **(A1)-(A2)** below, we will adopt the discrete parametrix method used in [15] to obtain explicit upper bounds on the transition kernel of the discrete-time EM scheme of (1.2) . As an application, we investigate the strong convergence rate of the EM scheme.

To end this section, we outline the structure of the remaining contents as follows: In Section 2 we state our assumptions and main results. Section 3 is devoted to the notation and some preliminaries. In Section 4, we investigate the estimate of transition kernel of (2.3). The convergence rate of (2.3) is discussed in Section 5.

2 Assumptions and Main results

To state our main result, we first introduce some notation and facts about Sobolev spaces, which can be found in [26].

For $(p, \gamma) \in [1, \infty] \times [0, 2]$, let $H^\gamma_p := (I - \Delta)^{-\frac{\gamma}{2}}(L^p(\mathbb{R}^d))$ be the usual Bessel potential space with the norm

$$\|f\|_{\gamma,p} := \|(I - \Delta)^{\frac{\gamma}{2}} f\|_p \asymp \|f\|_p + \|(-\Delta)^{\frac{\gamma}{2}} f\|_p,$$

where $\|\cdot\|_p$ is the usual L^p-norm in $\mathbb{R}^d$, and $(I - \Delta)^{\frac{\gamma}{2}} f$ and $(-\Delta)^{\frac{\gamma}{2}} f$ are defined by the Fourier transformation

$$(I - \Delta)^{\frac{\gamma}{2}} f := \mathcal{F}^{-1}((1 + |\cdot|^2)^{\frac{\gamma}{2}}\mathcal{F}f), \quad (-\Delta)^{\frac{\gamma}{2}} f := \mathcal{F}^{-1}(|\cdot|^\gamma \mathcal{F}f).$$

For $p = \infty, \gamma = 1$, we define H^1_∞ as the space of Lipschitz functions with norm

$$\|f\|_{1,\infty} := \|f\|_\infty + \|\nabla f\|_\infty.$$

In the sequel, we introduce the Sobolev embedding and an important inequality on the norm of elements in H^γ_p, which will be used to construct the Zvonkin transformation and obtain a priori estimate of solution to the associated elliptic equation.

For $p \in [1, \infty]$ and $\gamma \in [0, 2]$,

$$\begin{cases} H^\gamma_p \subset L^q, \ \ q \in \left[p, \dfrac{dp}{d-\gamma p}\right], & \gamma p < d; \\ H^\gamma_p \subset H^{\gamma-\frac{d}{p}}_\infty \subset C_b^{\gamma-\frac{d}{p}}, & \gamma p > d, \end{cases} \tag{2.1}$$

where C_b^β is the usual Hölder space. Moreover, for $\gamma \in [0, 1]$ and $p \in (1, \infty]$, there exists a constant c such that for all $f \in H^\gamma_p$,

$$\|f(\cdot + z) - f(\cdot)\|_p \leq c(|z|^\gamma \wedge 1)\|f\|_{\gamma,p}. \tag{2.2}$$

Throughout the paper, we impose the following assumptions on the drift b.

(A1) $\|b\|_\infty := \sup_{x\in\mathbb{R}^d} |b(x)| < \infty$.

(A2) There exist constants $\beta \in (1 - \frac{\alpha}{2}, 1)$ and $p > (\frac{2d}{\alpha} \vee 2)$ such that $b \in H^\beta_p$.

Under (**A2**), (1.2) has a unique strong solution $(X_t)_{t\geq 0}$, (see, for instance, [29, Theorem 2.4]).

The EM scheme corresponding to (1.2) is defined as follows: for any $\delta \in (0, 1)$,

$$\mathrm{d}X^{(\delta)}_t = b(X^{(\delta)}_{t_\delta})\mathrm{d}t + \mathrm{d}W_{S_t}, \quad t \geq 0, \quad X^{(\delta)}_0 = X_0 \tag{2.3}$$

with $t_\delta := \lfloor t/\delta \rfloor \delta$, where $\lfloor t/\delta \rfloor$ denotes the integer part of t/δ. We emphasize that $(X^{(\delta)}_{k\delta})_{k\geq 0}$ is a homogeneous Markov process. For $t > s$ and $x \in \mathbb{R}^d$, $p^{(\delta)}(s, x; t, \cdot)$ denotes the transition density of $X^{(\delta)}_t$ with the starting point $X^{(\delta)}_s = x$.

Let $p_\alpha(t, x)$ be the density of W_{S_t}. Our first main result gives an explicit upper bound of the transition kernel $p^{(\delta)}$.

Theorem 2.1. *Under* (**A1**), *there exists a constant* $C > 0$ *such that*

$$p^{(\delta)}(j\delta, x; t, y) \leq C p_\alpha(t - j\delta, y - x), \quad x, y \in \mathbb{R}^d, \quad t > j\delta, \quad \delta \in (0, 1). \tag{2.4}$$

Remark 2.2. *[14, Theorem 2.1] proved that the parametrix construction is feasible under assumptions that the indensity coefficient* $a(x) = \sigma\sigma^*(x)$ *is strictly positive, bounded from above and below and Hölder continuous, and the drift term* $b \in C_b^\gamma(\mathbb{R}^d)$*, the Hölder index* γ *satisfies three conditions. We also would like to point out that the authors in [5] established the two sided estimate of the transition density function* p*. Moreover, they only assume that* b *belongs to some Kato's class when* $\alpha \in (1, 2)$*, which is a sharp result. However, in order to establish the strong convergence for the EM scheme of SDEs with irregular drift, our method is based on the Zvonkin transformation, which cannot deal with the singular noise term. So we only consider the additive noise case in this work, and leave the multiplicative noise case as our future work.*

As an application of Theorem 2.1, the rate of strong convergence for the EM Scheme (2.3) can be obtained as follows.

Theorem 2.3. *Assume* **(A1)**-**(A2)**. *Then, for* $\eta \in (0, 2)$*, there exist constants* $C_1, C_2 > 0$ *independent of* $\delta \in (0, 1)$*, such that for any* $\epsilon \in (0, 1)$

$$\mathbb{E}\Big(\sup_{0\le t\le T} |X_t - X_t^{(\delta)}|^\eta\Big) \le C_1 2^{C_2(1+\|b\|_{\beta,p}^{\frac{2\alpha p}{\alpha p-2d}})} \Big(\delta^{\frac{\eta\beta}{\alpha}} 1_{\{2\beta<\alpha\}} + \mathbb{E}(S_1^{\frac{\alpha}{2}\epsilon})\delta^{\frac{\eta}{2}\epsilon} 1_{\{2\beta\ge\alpha\}}\Big). \tag{2.5}$$

Remark 2.4. *(1) Under a certain balance condition, and when bis bounded and Hölder continuous with respect to both the space, time variable, the rate of strong convergence of EM scheme* (2.3) *(inhomogenous case) was established in [19, Corollary 2.6]. Since Theorem 2.3 is also available for* $p = \infty$*, a close inspection of the Sobolev embedding* (2.1) *reveals that the result in [19, Corollary 2.6] is only the special case of our setting for* $p = \infty$*. Moreover, the optimal balance condition* $\beta > \frac{2}{\alpha} - 1$ *for* $\alpha \in (1, 2)$ *in [19] is stronger than* $\beta \in (1 - \frac{\alpha}{2}, 1)$ *in* **(A2)**.

(2) Compared with the result in [19, Corollary 2.6], the moment index η *in Theorem 2.3 is allowed to be greater than* α*, which is reasonable since* $X_t - X_t^{(\delta)}$ *is a bounded process.*

3 Notation and Preliminaries

Throughout the paper we use the following notation:

We write $f(x) \asymp g(x)$ to mean that there exist positive constants C_1, C_2 such that $C_1 g(x) \le f(x) \le C_2 g(x)$, $x \vee y = \max\{x, y\}$, and $x \wedge y = \min\{x, y\}$.

Let η_t be the density of S_t for $t > 0$. It follows from [3, Lemma 2.1] that the density of W_{S_t} has the following expression

$$p_\alpha(t,x) = \int_0^\infty (2\pi s)^{-\frac{d}{2}} \mathrm{e}^{-\frac{|x|^2}{2s}} \eta_t(s)\mathrm{d}s \asymp t(t^{1/\alpha} + |x|)^{-d-\alpha}, \ (t,x) \in (0,\infty) \times \mathbb{R}^d. \tag{3.1}$$

Moreover, according to [4, Lemma 2.2], it is clear that

$$|\nabla^k p_\alpha(t,x)| \le Ct(t^{1/\alpha} + |x|)^{-d-\alpha-k} \le Ct^{-k/\alpha} p_\alpha(t,x), \quad k \in \mathbb{N}, \tag{3.2}$$

where ∇^k stands for the kth-order gradient with respect to the spatial variable x.

According to the Markov property of W_{S_t}, we have

$$\int_{\mathbb{R}^d} p_\alpha(t-r, x'-y) p_\alpha(r-s, y-x)\mathrm{d}y = p_\alpha(t-s, x'-x). \tag{3.3}$$

In addition, for any $p \ge 1$, (3.1) implies that there exists a constant $C > 0$ such that

$$\begin{aligned} \|p_\alpha(t,\cdot)\|_p &\le \left(\int_{\mathbb{R}^d} Ct^p (t^{1/\alpha} + |x|)^{-pd-p\alpha} \mathrm{d}x \right)^{\frac{1}{p}} \\ &= \left(\int_{\mathbb{R}^d} Ct^p t^{-pd/\alpha - p\alpha/\alpha} (1 + |x|/t^{1/\alpha})^{-pd-p\alpha} \mathrm{d}x \right)^{\frac{1}{p}} \\ &= \left(\int_{\mathbb{R}^d} Ct^p t^{-pd/\alpha - p\alpha/\alpha} t^{d/\alpha} (1 + |y|)^{-pd-p\alpha} \mathrm{d}y \right)^{\frac{1}{p}} \le Ct^{-d/\alpha + d/(\alpha p)}. \end{aligned} \tag{3.4}$$

Recall that η_t is the density of S_t. Let

$$\Theta(r) = \frac{\mathbb{E}\left((2\pi S_1)^{-\frac{d}{2}} \mathrm{e}^{\frac{r}{2S_1}} \right)}{\mathbb{E}(2\pi S_1)^{-\frac{d}{2}}} = \frac{\int_0^\infty (2\pi s)^{-\frac{d}{2}} \mathrm{e}^{\frac{r}{2s}} \eta_1(s)\mathrm{d}s}{\int_0^\infty (2\pi s)^{-\frac{d}{2}} \eta_1(s)\mathrm{d}s}, \quad r \ge 0,$$

which is well defined due to

$$\mathbb{E}\mathrm{e}^{\gamma S_t^{-1}} \le \exp\left[\frac{c\gamma}{t^{\frac{2}{\alpha}}} + \frac{c\gamma^{\frac{\alpha}{2(\alpha-1)}}}{t^{\frac{1}{\alpha-1}}} \right] < \infty, \quad \gamma, t > 0,$$

for some constant $c>0$, see [28, Proof of Corollary 2.2]. It is not difficult to see that the function Θ is increasing and continuous on $[0,\infty)$ with $\Theta(0)=1$.

Next, we state a lemma which will be used frequently in the forthcoming sections.

Lemma 3.1. *There exists a constant $C>0$ such that for any $x, M\in\mathbb{R}^d$ and $r>0$,*

$$p_\alpha(r,x+M)\le C4^{d+\alpha}p_\alpha(r,x)\Theta(|M|^2r^{-\frac{2}{\alpha}}). \tag{3.5}$$

Proof. Note that the function $\mathrm{e}^{-\frac{|x|^2}{4s}}$ is increasing w.r.t $s\in[0,\infty)$ and the function $\mathrm{e}^{\frac{|M|^2}{2s}}$ is decreasing w.r.t $s\in[0,\infty)$. Applying the FKG inequality to the probability measure $\mu_r(\mathrm{d}s):=\frac{(2\pi s)^{-\frac{d}{2}}\eta_r(s)\mathrm{d}s}{\int_0^\infty(2\pi u)^{-\frac{d}{2}}\eta_r(u)\mathrm{d}u}$, functions $\mathrm{e}^{-\frac{|x|^2}{4s}}$ and $\mathrm{e}^{\frac{|M|^2}{2s}}$, it yields that

$$\int_0^\infty(2\pi s)^{-\frac{d}{2}}\mathrm{e}^{-\frac{|x|^2}{4s}}\mathrm{e}^{\frac{|M|^2}{2s}}\eta_r(s)\mathrm{d}s\le\frac{\int_0^\infty(2\pi s)^{-\frac{d}{2}}\mathrm{e}^{-\frac{|x|^2}{4s}}\eta_r(s)\mathrm{d}s\int_0^\infty(2\pi s)^{-\frac{d}{2}}\mathrm{e}^{\frac{|M|^2}{2s}}\eta_r(s)\mathrm{d}s}{\int_0^\infty(2\pi s)^{-\frac{d}{2}}\eta_r(s)\mathrm{d}s}.$$

This, together with the elementary inequality $|a-\overline{a}|^2\ge\frac{1}{2}|a|^2-|\overline{a}|^2$, $a,\overline{a}\in\mathbb{R}^d$ and the scaling property of S_t, yields that

$$\begin{aligned}p_\alpha(r,x+M)&=\int_0^\infty(2\pi s)^{-\frac{d}{2}}\mathrm{e}^{-\frac{|x+M|^2}{2s}}\eta_r(s)\mathrm{d}s\\&\le\int_0^\infty(2\pi s)^{-\frac{d}{2}}\mathrm{e}^{-\frac{|x|^2}{4s}}\mathrm{e}^{\frac{|M|^2}{2s}}\eta_r(s)\mathrm{d}s\\&\le\frac{\int_0^\infty(2\pi s)^{-\frac{d}{2}}\mathrm{e}^{-\frac{|x|^2}{4s}}\eta_r(s)\mathrm{d}s\int_0^\infty(2\pi s)^{-\frac{d}{2}}\mathrm{e}^{\frac{|M|^2}{2s}}\eta_r(s)\mathrm{d}s}{\int_0^\infty(2\pi s)^{-\frac{d}{2}}\eta_r(s)\mathrm{d}s}\\&\le p_\alpha\left(r,\frac{x}{\sqrt{2}}\right)\frac{\int_0^\infty(2\pi s)^{-\frac{d}{2}}\mathrm{e}^{\frac{|M|^2}{2s}}\eta_r(s)\mathrm{d}s}{\int_0^\infty(2\pi s)^{-\frac{d}{2}}\eta_r(s)\mathrm{d}s}\\&=p_\alpha\left(r,\frac{x}{\sqrt{2}}\right)\frac{\int_0^\infty(2\pi s)^{-\frac{d}{2}}\mathrm{e}^{\frac{|M|^2r^{-\frac{2}{\alpha}}}{2s}}\eta_1(s)\mathrm{d}s}{\int_0^\infty(2\pi s)^{-\frac{d}{2}}\eta_1(s)\mathrm{d}s}\\&=p_\alpha\left(r,\frac{x}{\sqrt{2}}\right)\Theta(|M|^2r^{-\frac{2}{\alpha}}).\end{aligned} \tag{3.6}$$

Noting that for all $a>1, |z|\le\frac{1}{a}(t^{1/\alpha}\vee|x|)$, we have

$$(t^{1/\alpha}+|x+z|)\ge(t^{1/\alpha}+|x|-|z|)\ge\frac{a-1}{a}(t^{1/\alpha}+|x|),$$

which yields that

$$(t^{1/\alpha}+|x+z|)^{-\gamma}\le(\frac{a}{a-1})^{\gamma}(t^{1/\alpha}+|x|)^{-\gamma},\quad \gamma>0.$$

This together with (3.1) and (3.6) for $z=\frac{1-\sqrt{2}}{\sqrt{2}}x$, $a=\frac{4}{3}$ and $\gamma=d+\alpha$ yields (3.5). □

4 Transition Kernel of EM scheme (2.3)

In this section, we first express the transition kernel of the approximate solution given by the discrete-time EM scheme in terms of a sum of convolutional terms with iterated kernels $H^{(\delta),(k)}$ and the density of the frozen homogeneous scheme defined in (4.1) below, and reveal its explicit upper bounds. Following this result, we then finish the proof of Theorem 2.1.

For $x\in\mathbb{R}^d$ and $j\ge 0$, let us begin with the "frozen" homogeneous scheme $(\widetilde{X}_{i\delta}^{(\delta),j,x,x'})_{i\ge j}$, which is defined by

$$\widetilde{X}_{(i+1)\delta}^{(\delta),j,x,x'}=\widetilde{X}_{i\delta}^{(\delta),j,x,x'}+b(x')\delta+(W_{S_{(i+1)\delta}}-W_{S_{i\delta}}),\quad i\ge j,\quad \widetilde{X}_{j\delta}^{(\delta),j,x,x'}=x. \tag{4.1}$$

Note that, the drift b is frozen at x' in the above definition. In what follows, $p^{(\delta)}(j\delta,x;j'\delta,\cdot)$ and $\widetilde{p}^{(\delta),x'}(j\delta,x;j'\delta,\cdot)$ denote the transition densities between times $j\delta$ and $j'\delta$ of the discretization scheme (2.3) and the above "frozen" scheme, respectively.

To derive the transition kernel of the EM scheme (2.3), we introduce discrete and homogeneous infinitesimal generators as follows:

For $\psi\in C^2(\mathbb{R}^d;\mathbb{R})$ and $j\ge 0$, we define the family of operators $\mathscr{L}_{j\delta}^{(\delta)}$ and $\hat{\mathscr{L}}_{j\delta}^{(\delta)}$ by

$$\begin{aligned}(\mathscr{L}_{j\delta}^{(\delta)}\psi)(x)&:=\delta^{-1}\left\{\mathbb{E}(\psi(X_{(j+1)\delta}^{(\delta)})|X_{j\delta}^{(\delta)}=x)-\psi(x)\right\},\\(\hat{\mathscr{L}}_{j\delta}^{(\delta)}\psi)(x)&:=\delta^{-1}\left\{\mathbb{E}\psi(\widetilde{X}_{(j+1)\delta}^{(\delta),j,x,x'})-\psi(x)\right\},\end{aligned}$$

and the discrete kernel $H^{(\delta)}$ as:

$$H^{(\delta)}(j\delta,x;j'\delta,x'):=(\mathscr{L}_{j\delta}^{(\delta)}-\hat{\mathscr{L}}_{j\delta}^{(\delta)})\widetilde{p}^{(\delta),x'}((j+1)\delta,\cdot;j'\delta,x')(x),\ \ j'\ge j+1, \tag{4.2}$$

here we use the convention $\widetilde{p}^{(\delta),x'}((j+1)\delta,\cdot;(j+1)\delta,x')=\delta_{\{x'\}}(\cdot)$, where $\delta_{\{x'\}}(\cdot)$ is the Delta function at the point x'.

Similar to to the parametrix method in [20, Proposition 4.1], we obtain the expansion of the transition kernel of the EM scheme (2.3) in terms of the transition kernel of the frozen scheme (4.1) and (4.2), it is stated as the following lemma.

Lemma 4.1. *For $0\le j<j'\le\lfloor T/\delta\rfloor$, the following expansion holds.*

$$p^{(\delta)}(j\delta,x;j'\delta,x')=\sum_{k=0}^{j'-j}(\widetilde{p}^{(\delta),x'}\otimes_\delta H^{(\delta),(k)})(j\delta,x;j'\delta,x'), \tag{4.3}$$

where $\widetilde{p}^{(\delta),x'}\otimes_\delta H^{(\delta),(0)}=\widetilde{p}^{(\delta),x'}$, $\widetilde{p}^{(\delta),x'}\otimes_\delta H^{(\delta),(k)}=(\widetilde{p}^{(\delta),x'}\otimes_\delta H^{(\delta),(k-1)})\otimes_\delta H^{(\delta)}$ with $\otimes_\delta$ being the convolution type binary operation defined by

$$(f\otimes_\delta g)(j\delta,x;j'\delta,x')=\delta\sum_{k=j}^{j'-1}\int_{\mathbb{R}^d}f(j\delta,x;k\delta,z)g(k\delta,z;j'\delta,x')\mathrm{d}z.$$

Proof. By the definition of (4.2), it yields that

$$H^{(\delta)}(j\delta,x;j'\delta,x')=\delta^{-1}\int_{\mathbb{R}^d}(p^{(\delta)}-\widetilde{p}^{\delta,x'})(j\delta,x;(j+1)\delta,z)\widetilde{p}^{\delta,x'}((j+1)\delta,z;j'\delta,x')\mathrm{d}z.$$

This, together with the Markov property, yields the following identity:

$$\begin{aligned}
&p^{(\delta)}(j\delta,x;j'\delta,x')-\widetilde{p}^{(\delta),x'}(j\delta,x;j'\delta,x')\\
&=\sum_{k=j}^{j'-1}\int_{\mathbb{R}^d}p^{(\delta)}(j\delta,x;(k+1)\delta,z')\widetilde{p}^{(\delta,x')}((k+1)\delta,z';j'\delta,y)\mathrm{d}z'\\
&\qquad\qquad-\int_{\mathbb{R}^d}p^{(\delta)}(j\delta,x;k\delta,z)\widetilde{p}^{(\delta,x')}(k\delta,z;j'\delta,y)\mathrm{d}z\\
&=\sum_{k=j}^{j'-1}\delta\int_{\mathbb{R}^d}p^{(\delta)}(j\delta,x;k\delta,z)\mathrm{d}z\\
&\qquad\qquad\cdot\int_{\mathbb{R}^d}\frac{(p^{(\delta)}-\widetilde{p}^{(\delta,x')})(k\delta,z;(k+1)\delta,z')\widetilde{p}^{(\delta,x')}((k+1)\delta,z';j'\delta,x')}{\delta}\mathrm{d}z'\\
&=\sum_{k=j}^{j'-1}\delta\int_{\mathbb{R}^d}p^{(\delta)}(j\delta,x;k\delta,z)H^{(\delta)}(k\delta,z;j'\delta,x')\mathrm{d}z
\end{aligned}$$

$$= p^{(\delta)} \otimes_\delta H^{(\delta)}(j\delta, x; j'\delta, x').$$

Then, the assertion (4.3) follows by by iterative application of this identity. □

The following lemma gives the smoothing properties of the discrete convolution kernel and the estimate of $p^{(\delta)}(j\delta, x; j'\delta, x')$.

Lemma 4.2. *Assume* **(A1)**. *Then there exists a constant $\hat{C}_T > 0$ independent of δ such that for any $0 \le j < j' \le \lfloor T/\delta \rfloor$,*

$$\begin{aligned} &|(\widetilde{p}^{(\delta),x'} \otimes_\delta H^{(\delta),(m)})|(j\delta, x; j'\delta, x') \\ &\le \hat{C}_T^m \|b\|_\infty^m ((j'-j)\delta)^{m(1-\frac{1}{\alpha})} \frac{\Gamma(1-\frac{1}{\alpha})^m}{\Gamma(1+m(1-\frac{1}{\alpha}))} p_\alpha((j'-j)\delta, x'-x), \qquad m \ge 0, \end{aligned} \tag{4.4}$$

where $\Gamma(x) := \int_0^\infty t^{x-1}\mathrm{e}^{-t}\mathrm{d}t$ is the gamma function. Consequently, it holds

$$p^{(\delta)}(j\delta, x; j'\delta, x') \le \sum_{m=0}^{j'-j} \frac{[\hat{C}_T \|b\|_\infty T^{(1-\frac{1}{\alpha})}\Gamma(1-\frac{1}{\alpha})]^m}{\Gamma(1+m(1-\frac{1}{\alpha}))} p_\alpha((j'-j)\delta, x'-x). \tag{4.5}$$

Proof. We divide the proof into two steps.

Step 1. We claim that

$$|H^{(\delta)}|(j\delta, x; j'\delta, x') \le \hat{C}_T \|b\|_\infty ((j'-j)\delta)^{\frac{-1}{\alpha}} p_\alpha((j'-j)\delta, x'-x), \quad j' > j. \tag{4.6}$$

Firstly, we prove (4.6) for $j' = j+1$. It follows from (3.1), (3.2), (3.5) and (4.2) that

$$
\begin{aligned}
&|H^{(\delta)}|(j\delta, x; (j+1)\delta, x') \\
&= \left|(\mathscr{L}_{j\delta}^{(\delta)} - \hat{\mathscr{L}}_{j\delta}^{(\delta)})\widetilde{p}^{(\delta),x'}((j+1)\delta, \cdot; (j+1)\delta, x')(x)\right| \\
&= \left|\frac{1}{\delta}\Big\{\mathbb{E}\left(\delta_{\{x'\}}(X_{(j+1)\delta}^{(\delta)})|X_{j\delta}^{(\delta)} = x\right) - \mathbb{E}\left(\delta_{\{x'\}}(\widetilde{X}_{(j+1)\delta}^{(\delta),j,x,x'})|\widetilde{X}_{j\delta}^{(\delta),j,x,x'} = x\right)\Big\}\right| \\
&= \frac{1}{\delta}|p^{(\delta)} - \widetilde{p}^{(\delta),x'}|(j\delta, x; (j+1)\delta, x') \\
&= \frac{1}{\delta}\left|p_\alpha(\delta, x' - x - b(x)\delta) - p_\alpha(\delta, x' - x - b(x')\delta)\right| \\
&\le 2\|b\|_\infty \left|\int_0^1 \nabla p_\alpha(\delta, x' - x - b(x')\delta + \theta(b(x') - b(x))\delta)\mathrm{d}\theta\right| \\
&\le C\|b\|_\infty \delta^{-\frac{1}{\alpha}} \sup_{\theta\in[0,1]} p_\alpha(\delta, x' - x - b(x')\delta + \theta(b(x') - b(x))\delta) \\
&\le C\|b\|_\infty \delta^{-\frac{1}{\alpha}} p_\alpha(\delta, x' - x)\Theta(9\delta^{2-\frac{2}{\alpha}}\|b\|_\infty^2) \\
&=: \hat{C}_T \delta^{-\frac{1}{\alpha}} p_\alpha(\delta, x' - x).
\end{aligned}
$$

Thus, (4.6) holds for $j' = j+1$.

Next, we are going to show that (4.6) holds for $j' > j+1$. According to (3.1), (3.2), (3.3), (3.5) and (4.2), it holds that

$$
\begin{aligned}
&|H^{(\delta)}|(j\delta, x; j'\delta, x') \\
&= \frac{1}{\delta}\Big|\Big\{\int_{\mathbb{R}^d} p_\alpha(\delta, z)\widetilde{p}^{(\delta),x'}((j+1)\delta, x + b(x)\delta + z; j'\delta, x')\mathrm{d}z - \widetilde{p}^{(\delta),x'}((j+1)\delta, x; j'\delta, x')\Big\} \\
&\quad - \Big\{\int_{\mathbb{R}^d} p_\alpha(\delta, z)\widetilde{p}^{(\delta),x'}((j+1)\delta, x + b(x')\delta + z; j'\delta, x')\mathrm{d}z - \widetilde{p}^{(\delta),x'}((j+1)\delta, x; j'\delta, x')\Big\}\Big| \\
&= \frac{1}{\delta}\Big|\int_{\mathbb{R}^d} p_\alpha(\delta, z)\Big\{\widetilde{p}^{(\delta),x'}((j+1)\delta, x + b(x)\delta + z; j'\delta, x') \\
&\quad - \widetilde{p}^{(\delta),x'}((j+1)\delta, x + b(x')\delta + z; j'\delta, x')\Big\}\mathrm{d}z\Big| \\
&= \frac{1}{\delta}\Big|\int_{\mathbb{R}^d} p_\alpha(\delta, z)\Big\{p_\alpha\left((j' - (j+1))\delta, x' - x - b(x)\delta - z - b(x')(j' - (j+1))\delta\right) \\
&\quad - p_\alpha\left((j' - (j+1))\delta, x' - x - b(x')\delta - z - b(x')(j' - (j+1))\delta\right)\Big\}\mathrm{d}z\Big| \\
&= \frac{1}{\delta}\Big|p_\alpha((j' - j)\delta, x' - x - b(x)\delta - b(x')(j' - (j+1))\delta) \\
&\quad - p_\alpha((j' - j)\delta, x' - x - b(x')\delta - b(x')(j' - (j+1))\delta)\Big|
\end{aligned}
$$

$$
\begin{aligned}
&\le 2\|b\|_\infty \sup_{\theta\in[0,1]} |\nabla p_\alpha((j'-j)\delta, x'-x-b(x')\delta - b(x')(j'-(j+1))\delta + \theta(b(x')-b(x))\delta)| \\
&\le C\|b\|_\infty((j'-j)\delta)^{-\frac{1}{\alpha}} \sup_{\theta\in[0,1]} p_\alpha((j'-j)\delta, x'-x-b(x')(j'-j-1)\delta + \theta(b(x')-b(x))\delta) \\
&\le C\|b\|_\infty((j'-j)\delta)^{-\frac{1}{\alpha}} p_\alpha((j'-j)\delta, x'-x)\Theta(9[(j'-j-1)\delta]^{2-\frac{2}{\alpha}}\|b\|_\infty^2) \\
&= \hat{C}_T\|b\|_\infty((j'-j)\delta)^{-\frac{1}{\alpha}} p_\alpha((j'-j)\delta, x'-x).
\end{aligned}
$$

Step 2. We are going to prove (4.4).

Due to (3.5) and the increasing property of function Θ, it is not difficult to see that

$$
\begin{aligned}
\widetilde{p}^{(\delta),x'}(j\delta, x; j'\delta, x') &= p_\alpha((j'-j)\delta, x'-x-b(x')(j'-j)\delta) \\
&\le C4^{d+\alpha}\Theta(\|b\|_\infty^2((j'-j)\delta)^{2-\frac{2}{\alpha}})p_\alpha((j'-j)\delta, x'-x) \\
&\le C4^{d+\alpha}\Theta(\|b\|_\infty^2 T^{2-\frac{2}{\alpha}})p_\alpha((j'-j)\delta, x'-x) \\
&=: \hat{C}_T p_\alpha((j'-j)\delta, x'-x).
\end{aligned}
\tag{4.7}
$$

Combining this with (4.6), we obtain from the definition of operator $\otimes_\delta$ that

$$
\begin{aligned}
&|\widetilde{p}^{(\delta),x'} \otimes_\delta H^{(\delta)}|(j\delta, x; j'\delta, x') \\
&= \delta \sum_{k=j}^{j'-1} \int_{\mathbb{R}^d} \widetilde{p}^{(\delta),x'}(j\delta, x; k\delta, z) H^{(\delta)}(k\delta, z; j'\delta, x')\mathrm{d}z \\
&\le \delta \sum_{k=j}^{j'-1} \int_{\mathbb{R}^d} \hat{C}_T p_\alpha((k-j)\delta, z-x)\hat{C}_T\|b\|_\infty((j'-k)\delta)^{\frac{-1}{\alpha}} p_\alpha((j'-k)\delta, x'-z)\mathrm{d}z \\
&= \delta\hat{C}_T\|b\|_\infty p_\alpha((j'-j)\delta, x'-x)\sum_{k=j}^{j'-1}((j'-k)\delta)^{\frac{-1}{\alpha}} \\
&\le \hat{C}_T\|b\|_\infty p_\alpha((j'-j)\delta, x'-x)\int_{j\delta}^{j'\delta}(j'\delta - v)^{\frac{-1}{\alpha}}\mathrm{d}v \\
&= \hat{C}_T\|b\|_\infty((j'-j)\delta)^{1-\frac{1}{\alpha}}\beta(1, 1-\frac{1}{\alpha})p_\alpha((j'-j)\delta, x'-x).
\end{aligned}
$$

In the above equation $\beta(m,n) := \int_0^t s^{m-1}(1-s)^{n-1}\mathrm{d}s$ stands for the *beta* function.

Using this and (4.6), we get

$$
|\widetilde{p}^{(\delta),x'} \otimes_\delta H^{(\delta),(2)}|(j\delta, x; j'\delta, x')
$$

$$
\begin{aligned}
&\le \delta \sum_{k=j}^{j'-1} \int_{\mathbb{R}^d} |\widetilde{p}^{(\delta),x'} \otimes_\delta H^{(\delta)}(j\delta, x; k\delta, z)||H^{(\delta)}(k\delta, z; j'\delta, x')|\mathrm{d}z \\
&\le \delta \sum_{k=j}^{j'-1} \int_{\mathbb{R}^d} \hat{C}_T^2 \|b\|_\infty^2 ((k-j)\delta)^{1-\frac{1}{\alpha}} \beta(1, 1-\frac{1}{\alpha}) p_\alpha((k-j)\delta, z-x) \\
&\qquad\qquad \times ((j'-k)\delta)^{\frac{-1}{\alpha}} p_\alpha((j'-k)\delta, x'-z)\mathrm{d}z \\
&= \delta \hat{C}_T^2 \|b\|_\infty^2 \beta(1, 1-\frac{1}{\alpha}) \sum_{k=j}^{j'-1} ((k-j)\delta)^{1-\frac{1}{\alpha}} ((j'-k)\delta)^{\frac{-1}{\alpha}} p_\alpha((j'-j)\delta, x'-x) \\
&\le \hat{C}_T^2 \|b\|_\infty^2 \beta(1, 1-\frac{1}{\alpha}) \int_{j\delta}^{j'\delta} (v-j\delta)^{1-\frac{1}{\alpha}} (j'\delta - v)^{\frac{-1}{\alpha}} \mathrm{d}v p_\alpha((j'-j)\delta, x'-x) \\
&= \hat{C}_T^2 \|b\|_\infty^2 ((j'-j)\delta)^{2(1-\frac{1}{\alpha})} \beta(1, 1-\frac{1}{\alpha}) \beta(2-\frac{1}{\alpha}, 1-\frac{1}{\alpha}) p_\alpha((j'-j)\delta, x'-x).
\end{aligned}
$$

By an induction argument, one has

$$
\begin{aligned}
&|\widetilde{p}^{(\delta),x'} \otimes_\delta H^{(\delta),(m)}|(j\delta, j'\delta, x, x') \\
&\le \hat{C}_T^m \|b\|_\infty^m ((j'-j)\delta)^{m(1-\frac{1}{\alpha})} p_\alpha((j'-j)\delta, x'-x) \prod_{i=1}^m \beta(i - \frac{i-1}{\alpha}, 1-\frac{1}{\alpha}) \\
&= \hat{C}_T^m \|b\|_\infty^m ((j'-j)\delta)^{m(1-\frac{1}{\alpha})} p_\alpha((j'-j)\delta, x'-x) \prod_{i=1}^m \frac{\Gamma(i-\frac{i-1}{\alpha})\Gamma(1-\frac{1}{\alpha})}{\Gamma(i+1-\frac{i}{\alpha})} \\
&= \hat{C}_T^m \|b\|_\infty^m ((j'-j)\delta)^{m(1-\frac{1}{\alpha})} \frac{\Gamma^m(1-\frac{1}{\alpha})}{\Gamma(1+m(1-\frac{1}{\alpha}))} p_\alpha((j'-j)\delta, x'-x).
\end{aligned}
$$

Therefore, (4.4) is proved, and (4.5) follows from (4.3) and (4.4). □

We are now in the position to prove Theorem 2.1.

Proof of Theorem 2.1. For fixed $t > 0$, there is an integer $k \ge 0$ such that $t \in [k\delta, (k+1)\delta)$. It follows from Lemma 3.1 that

$$
\begin{aligned}
p^{(\delta)}(k\delta, x; t, y) &= p_\alpha(t-k\delta, y-x-b(x)(t-k\delta)) \\
&\le C4^{d+\alpha}\Theta((t-k\delta)^{2-\frac{2}{\alpha}}\|b\|_\infty^2) p_\alpha(t-k\delta, y-x) \qquad (4.8) \\
&= C_1 p_\alpha(t-k\delta, y-x).
\end{aligned}
$$

Note that Lemma 4.2 implies

$$p^{(\delta)}(j\delta, x; j'\delta, x') \leq \sum_{m=0}^{j'-j} \frac{[\hat{C}_T \|b\|_\infty T^{(1-\frac{1}{\alpha})}\Gamma(1-\frac{1}{\alpha})]^m}{\Gamma(1+m(1-\frac{1}{\alpha}))} p_\alpha((j'-j)\delta, x'-x)$$
$$\leq C_2 p_\alpha((j'-j)\delta, x'-x), \quad j' > j,\ x, x' \in \mathbb{R}^d. \tag{4.9}$$

Combining this with (4.8) and the Chapman-Kolmogrov equation, we obtain

$$\begin{aligned} p^{(\delta)}(j\delta, x; t, y) &= \int_{\mathbb{R}^d} p^{(\delta)}(j\delta, x; \lfloor t/\delta \rfloor \delta, z) p^{(\delta)}(\lfloor t/\delta \rfloor \delta, z; t, y) \mathrm{d}z \\ &\leq C_1 C_2 \int_{\mathbb{R}^d} p_\alpha(t - \lfloor t/\delta \rfloor \delta, y - z) p_\alpha((\lfloor t/\delta \rfloor - j)\delta, z - x) \mathrm{d}z \\ &= C p_\alpha(t - j\delta, y - x). \end{aligned}$$

The proof is therefore completed. □

5 Proof of Theorem 2.3

Before finishing the proof of Theorem 2.3, we prepare some auxiliary lemmas.

Lemma 5.1. *Assume* **(A1)**-**(A2)**, *and let $T > 0$ be fixed. Then there exists a constant $C_T > 0$ independent of $\delta \in (0,1)$, such that for any $\epsilon \in (0,1)$,*

$$\int_0^T \mathbb{E}|b(X_t^{(\delta)}) - b(X_{t_\delta}^{(\delta)})|^2 \mathrm{d}t \leq C_T \left(\delta^{\frac{2\beta}{\alpha}} 1_{\{2\beta < \alpha\}} + \mathbb{E}(S_1^{\frac{\alpha}{2}\epsilon}) \delta^\epsilon 1_{\{2\beta \geq \alpha\}} \right). \tag{5.1}$$

Proof. Observe that

$$\begin{aligned} \int_0^T \mathbb{E}|b(X_t^{(\delta)}) - b(X_{t_\delta}^{(\delta)})|^2 \mathrm{d}t = &\int_0^\delta \mathbb{E}|b(X_t^{(\delta)}) - b(X_0^{(\delta)})|^2 \mathrm{d}t \\ &+ \sum_{k=1}^{\lfloor T/\delta \rfloor} \int_{k\delta}^{T \wedge (k+1)\delta} \mathbb{E}|b(X_t^{(\delta)}) - b(X_{k\delta}^{(\delta)})|^2 \mathrm{d}t. \end{aligned}$$

It follows from **(A1)** that

$$\int_0^\delta \mathbb{E}|b(X_t^{(\delta)}) - b(X_0^{(\delta)})|^2 \mathrm{d}t \leq 4\|b\|_\infty^2 \delta. \tag{5.2}$$

For $t\in[k\delta,(k+1)\delta)$, noting the independence between $X^{(\delta)}_{k\delta}$ and $W_{S_t}-W_{S_{k\delta}}$. Similar to the Gaussian bounds for the EM scheme in [24, Lemma 3.5], we use the explicit upper bound of the transition kernel $p^{(\delta)}$, i.e. Theorem 2.1 and derive that

$$\begin{aligned}
&\mathbb{E}|b(X^{(\delta)}_t)-b(X^{(\delta)}_{k\delta})|^2\\
&=\mathbb{E}|b(X^{(\delta)}_{k\delta}+b(X^{(\delta)}_{k\delta})(t-k\delta)+(W_{S_t}-W_{S_{k\delta}}))-b(X^{(\delta)}_{k\delta})|^2\\
&=\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|b(y+z)-b(y)|^2p^{(\delta)}(0,x;k\delta,y)p^{(\delta)}(k\delta,y;t,z+y)\mathrm{d}y\mathrm{d}z\\
&\le C\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|b(y+z)-b(y)|^2p_\alpha(k\delta,y-x)p_\alpha(t-k\delta,z)\mathrm{d}y\mathrm{d}z.
\end{aligned}\tag{5.3}$$

By **(A1)**, **(A2)**, Hölder's inequality, (3.4) and (2.2), we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^d}|b(y+z)-b(y)|^2p_\alpha(k\delta,y-x)\mathrm{d}y\\
&\le\Big\{\Big(\int_{\mathbb{R}^d}|b(y+z)-b(y)|^p\mathrm{d}y\Big)^{\frac{2}{p}}\Big(\int_{\mathbb{R}^d}p_\alpha(k\delta,y-x)^{\frac{p}{p-2}}\mathrm{d}y\Big)^{\frac{p-2}{p}}\Big\}\\
&\le C(k\delta)^{-2d/(\alpha p)}\|b\|^2_{\beta,p}(|z|^{2\beta}\wedge 1).
\end{aligned}$$

We therefore infer from (5.3) that for any $\epsilon\in(0,1)$ and $t\in[k\delta,(k+1)\delta)$,

$$\begin{aligned}
&\mathbb{E}|b(X^{(\delta)}_t)-b(X^{(\delta)}_{k\delta})|^2\\
&\le C(k\delta)^{-2d/(\alpha p)}\|b\|^2_{\beta,p}\int_{\mathbb{R}^d}\{|z|^{2\beta}\wedge 1\}p_\alpha(t-k\delta,z)\{1_{\{2\beta\ge\alpha\}}+1_{\{2\beta<\alpha\}}\}\mathrm{d}z\\
&\le C(k\delta)^{-2d/(\alpha p)}\|b\|^2_{\beta,p}\Big(\mathbb{E}(|W_{S_\delta}|^{2\beta}\wedge 1)1_{\{2\beta\ge\alpha\}}+\delta^{\frac{2\beta}{\alpha}}1_{\{2\beta<\alpha\}}\Big)\\
&\le C(k\delta)^{-2d/(\alpha p)}\|b\|^2_{\beta,p}\Big(\mathbb{E}(|W_{S_\delta}|^{\alpha\epsilon})1_{\{2\beta\ge\alpha\}}+\delta^{\frac{2\beta}{\alpha}}1_{\{2\beta<\alpha\}}\Big)\\
&=C(k\delta)^{-2d/(\alpha p)}\|b\|^2_{\beta,p}\Big(\mathbb{E}|W_1|^{\alpha\epsilon}\mathbb{E}(S_1^{\frac{\alpha}{2}\epsilon})\delta^{\epsilon}1_{\{2\beta\ge\alpha\}}+\delta^{\frac{2\beta}{\alpha}}1_{\{2\beta<\alpha\}}\Big),
\end{aligned}\tag{5.4}$$

where the second inequality is due to the fact that for $2\beta<\alpha$,

$$\begin{aligned}
\int_{\mathbb{R}^d}\{|z|^{2\beta}\wedge 1\}p_\alpha(t,z)\mathrm{d}z&\asymp\int_{\mathbb{R}^d}\{|z|^{2\beta}\wedge 1\}t(t^{\frac{1}{\alpha}}+|z|)^{-d-\alpha}\mathrm{d}z\\
&\le C\int_0^\infty t\frac{r^{2\beta+d-1}}{(t^{\frac{1}{\alpha}}+r)^{d+\alpha}}\mathrm{d}r\\
&=\Big(\int_0^{t^{\frac{1}{\alpha}}}+\int_{t^{\frac{1}{\alpha}}}^\infty\Big)\frac{tr^{2\beta+d-1}}{(t^{\frac{1}{\alpha}}+r)^{d+\alpha}}\mathrm{d}r\le Ct^{\frac{2\beta}{\alpha}}.
\end{aligned}$$

Noting that $\int_0^T r^{-2d/(\alpha p)}\mathrm{d}r < \infty$ due to **(A2)**, we arrive at

$$\sum_{k=1}^{\lfloor T/\delta\rfloor} \int_{k\delta}^{T\wedge(k+1)\delta} \mathbb{E}|b(X_t^{(\delta)}) - b(X_{k\delta}^{(\delta)})|^2 \mathrm{d}t \le C_T \left(\delta^{\frac{2\beta}{\alpha}} 1_{\{2\beta<\alpha\}} + \mathbb{E}|W_1|^{\alpha\epsilon}\mathbb{E}(S_1^{\frac{\alpha}{2}\epsilon})\delta^{\epsilon} 1_{\{2\beta\ge\alpha\}} \right).$$

This combined with (5.2) implies (5.1). □

Next, we use Theorem 2.1 to derive the Krylov estimate and the Khasminskii estimate of $(X_t^{(\delta)})_{t\ge0}$, see [9, 18, 29, 30, 32] for more results about Krylov's estimate and Khasminskii's estimate.

Lemma 5.2. *Assume* **(A1)**. *Then, for any $q > (d/\alpha \vee 1)$, there exist constants $C, c > 0$ such that Krylov's estimate*

$$\mathbb{E}\Big(\int_s^t |f(X_r^{(\delta)})|\mathrm{d}r \Big| \mathscr{F}_s \Big) \le C \|f\|_q (t-s)^{1-d/(\alpha q)}, \quad f \in L^q(\mathbb{R}^d), 0 \le s \le t \le T, \tag{5.5}$$

holds, which implies the Khasminskii estimate

$$\mathbb{E}\exp\left(\lambda \int_0^T |f(X_t^{(\delta)})|\mathrm{d}t \right) \le 2^{1+T(c\lambda\|f\|_q)^{\frac{1}{1-d/(\alpha q)}}}, \quad f \in L^q(\mathbb{R}^d), \lambda > 0. \tag{5.6}$$

Proof. For $0 \le s \le t \le T$, note that

$$\begin{aligned} \mathbb{E}\Big(\int_s^t |f(X_r^{(\delta)})|\mathrm{d}r \Big| \mathscr{F}_s \Big) &= \mathbb{E}\Big(\int_s^{t\wedge(s_\delta+\delta)} |f(X_r^{(\delta)})|\mathrm{d}r \Big| \mathscr{F}_s \Big) + \mathbb{E}\Big(\int_{t\wedge(s_\delta+\delta)}^t |f(X_r^{(\delta)})|\mathrm{d}r \Big| \mathscr{F}_s \Big) \\ &=: I_1(s,t) + I_2(s,t). \end{aligned}$$

For $t \in [s, s_\delta + \delta]$,

$$X_r^{(\delta)} = X_{s_\delta}^{(\delta)} + b(X_{s_\delta}^{(\delta)})(r - s_\delta) + (W_{S_s} - W_{S_{s_\delta}}) + (W_{S_r} - W_{S_s}), \quad r \in [s, s_\delta + \delta),$$

in view of the independence between $W_{S_r} - W_{S_s}$ and $\mathscr{F}_s$, we derive from (3.4) and Hölder's inequality that for any $q > d/\alpha$,

$$\begin{aligned} I_1(s,t) &= \int_s^{t\wedge(s_\delta+\delta)} \int_{\mathbb{R}^d} f(x + b(x)(r - s_\delta) + w + z) p_\alpha(r - s, z)\mathrm{d}z \Big|_{x = X_{s_\delta}^{(\delta)}}^{w = W_s - W_{s_\delta}} \mathrm{d}r \\ &\le \|f\|_q \int_s^{t\wedge(s_\delta+\delta)} (r-s)^{-d/\alpha + d(q-1)/(\alpha q)}\mathrm{d}r \le \frac{(t-s)^{1-d/(\alpha q)}}{1 - d/(\alpha q)} \|f\|_q. \end{aligned} \tag{5.7}$$

Using the Markov property and the tower property of conditional expectations, one has that

$$
\begin{aligned}
I_2(s,t) &\le \int_{s_\delta+\delta}^t \mathbb{E}\Big(|f(X_r^{(\delta)})|\Big|\mathscr{F}_s\Big)\mathrm{d}r = \int_{s_\delta+\delta}^t \mathbb{E}\Big(\mathbb{E}\Big(|f(X_r^{(\delta)})|\Big|\mathscr{F}_{s_\delta+\delta}\Big)\big|\mathscr{F}_s\Big)\mathrm{d}r \\
&= \int_{s_\delta+\delta}^t \mathbb{E}\Big(\mathbb{E}|f(X_r^{(\delta)})|\Big|X_{s_\delta+\delta}^{(\delta)}\Big|\mathscr{F}_s\Big)\mathrm{d}r.
\end{aligned}
$$

Applying Theorem 2.1, Hölder's inequality and (3.4), we conclude that

$$
\begin{aligned}
\mathbb{E}|(f(X_r^{(\delta)})|X_{s_\delta+\delta}^{(\delta)})| &= \int_{\mathbb{R}^d} |f(y)| p^{(\delta)}(s_\delta+\delta, X_{s_\delta+\delta}^{(\delta)}; r, y)\mathrm{d}y \\
&\le C\int_{\mathbb{R}^d} |f(y)| p_\alpha(r-s_\delta-\delta, y-X_{s_\delta+\delta}^{(\delta)})\mathrm{d}y \le C(r-s_\delta-\delta)^{-\frac{d}{\alpha q}}\|f\|_q.
\end{aligned}
$$

Therefore, it holds that

$$
I_2(s,t) \le C(t-s)^{1-d/(\alpha q)}\|f\|_q, \tag{5.8}
$$

which, together with (5.7), implies (5.5).

The remaining procedure of deriving (5.6) from (5.5) is standard. For readers' convenience, we sketch it here. For each $k \ge 1$, applying inductively (5.5) gives

$$
\begin{aligned}
\mathbb{E}\bigg(\Big(\int_s^t |f(X_r^{(\delta)})|\mathrm{d}r\Big)^k\bigg|\mathscr{F}_s\bigg) &= k!\mathbb{E}\Big(\int_{\triangle_{k-1}(s,t)} |f(X_{r_1}^{(\delta)})|\cdots|f(X_{r_{k-1}}^{(\delta)})|\mathrm{d}r_1 \\
&\qquad \cdots\mathrm{d}r_{k-1} \times \mathbb{E}\Big(\int_{r_{k-1}}^t |f(X_{r_k}^{(\delta)})|\mathrm{d}r_k\Big|\mathscr{F}_{r_{k-1}}\Big)\Big|\mathscr{F}_s\Big) \\
&\le k!(C(t-s)^{1-d/(\alpha q)}\|f\|_q)^k, \quad 0 \le s \le t \le T,
\end{aligned} \tag{5.9}
$$

where

$$
\triangle_k(s,t) := \{(r_1,\cdots,r_k) \in \mathbb{R}^k : s \le r_1 \le \cdots \le r_k \le t\}.
$$

Taking $\delta_0 = (2C\lambda\|f\|_q)^{-\frac{1}{1-d/(\alpha q)}}$, and combining this with (5.9), we derive

$$
\mathbb{E}\Big(\exp\Big(\lambda\int_{(i-1)\delta_0}^{i\delta_0\wedge T} |f(X_t^{(\delta)})|\mathrm{d}t\Big)\Big|\mathscr{F}_{(i-1)\delta_0}\Big) \le \sum_{k=0}^\infty \frac{1}{2^k} = 2, \quad i \ge 1, \tag{5.10}
$$

which further implies that

$$
\begin{aligned}
\mathbb{E}\exp\Big(\lambda\int_0^T|f(X_t^{(\delta)})|\mathrm{d}t\Big) &= \mathbb{E}\Big\{\mathbb{E}\Big(\exp\Big(\int_0^T|f(X_t^{(\delta)})|\mathrm{d}t\Big)\Big|\mathscr{F}_{\lfloor T/\delta_0\rfloor\delta_0}\Big\}\\
&= \mathbb{E}\Big(\exp\Big(\lambda\sum_{i=1}^{\lfloor T/\delta_0\rfloor}\int_{(i-1)\delta_0}^{i\delta_0}|f(X_t^{(\delta)})|\mathrm{d}t\Big)\\
&\quad\times\mathbb{E}\Big(\exp\Big(\lambda\int_{\lfloor T/\delta_0\rfloor\delta_0}^T|f(X_t^{(\delta)})|\mathrm{d}t\Big)\Big|\mathscr{F}_{\lfloor T/\delta_0\rfloor\delta_0}\Big)\Big)\\
&\le 2\,\mathbb{E}\exp\Big(\lambda\sum_{i=1}^{\lfloor T/\delta_0\rfloor}\int_{(i-1)\delta_0}^{i\delta_0}|f(X_t^{(\delta)})|\mathrm{d}t\Big)\\
&\le\cdots\le 2^{1+T/\delta_0}.
\end{aligned}
\tag{5.11}
$$

Therefore, (5.6) holds. □

The following lemma is concerned with Krylov's and Khasminskii's estimates for the solution process $(X_t)_{t\ge0}$ to (1.2), which is more or less standard; see, for instance, [9, 18, 29, 30, 32]. Whereas, we herein state them and provide a sketch of its proof by using explicit upper bound of heat kernel.

Lemma 5.3. *Assume* **(A1)**. *Then for any* $q>(d/\alpha\vee 1)$,

$$
\mathbb{E}\Big(\int_s^t|f(X_r)|\mathrm{d}r\Big|\mathscr{F}_s\Big)\le C\,\|f\|_q(t-s)^{1-d/(\alpha q)},\quad f\in L^q(\mathbb{R}^d),0\le s\le t\le T,
\tag{5.12}
$$

and

$$
\mathbb{E}\exp\Big(\lambda\int_0^T|f(X_t)|\mathrm{d}t\Big)\le 2^{1+T(\lambda c\|f\|_q)^{\frac{1}{1-d/(\alpha q)}}},\quad f\in L^q(\mathbb{R}^d),\lambda>0
\tag{5.13}
$$

hold for some constants $C,c>0$.

Proof. By [5, Theorem 1.5 and Remark 1.6], we conclude that the solution X_t to SDE (1.2) at time t has a probability density under condition **(A1)**, and we denote it as $p(0,x;t,y)$. Moreover, it satisfies the upper bound as follows:

$$
p(0,x;t,y)\le Cp_\alpha(t,y-x),\quad 0<t\le T,x,y\in\mathbb{R}^d.
\tag{5.14}
$$

This, together with Hölder's inequality, Markov property and (3.4), yields that

$$
\begin{aligned}
\mathbb{E}\Big(\int_s^t |f(X_r)|\mathrm{d}r\Big|\mathscr{F}_s\Big) &= \int_s^t \mathbb{E}(|f(X_r^{s,x})|)\Big|_{x=X_s}\mathrm{d}r \\
&\le C\int_s^t\int_{\mathbb{R}^d}|f(y)|p_\alpha(r-s,y-x)\mathrm{d}y\Big|_{x=X_s}\mathrm{d}r \\
&\le C(t-s)^{1-d/(\alpha q)}\|f\|_q,
\end{aligned}
\tag{5.15}
$$

By repeating the same procedure as in the proof of (5.6), we can derive (5.11). □

For a locally integrable function $h:\mathbb{R}^d\to\mathbb{R}$, the Hardy-Littlewood maximal operator $\mathscr{M}h$ is defined as below

$$
(\mathscr{M}h)(x)=\sup_{r>0}\frac{1}{|B_r(x)|}\int_{B_r(x)}h(y)\mathrm{d}y,\qquad x\in\mathbb{R}^d,
$$

where $B_r(x)$ is the ball with radius r centered at the point x and $|B_r(x)|$ denotes the volume of $B_r(x)$. Then, we obtain the following estimates.

Lemma 5.4. *There exists a constant $C>0$ such that for any continuous and weakly differentiable function $f:\mathbb{R}^d\to\mathbb{R}$,*

$$
|f(x)-f(y)|\le C|x-y|\{(\mathscr{M}|\nabla f|)(x)+(\mathscr{M}|\nabla f|)(y)\},\qquad a.e.\ x,y\in\mathbb{R}^d. \tag{5.16}
$$

Moreover, there exists a constant $C_q>0$ such that for any $q>1$ and $f\in L^q(\mathbb{R}^d)$,

$$
\|\mathscr{M}f\|_q\le C_q\|f\|_q. \tag{5.17}
$$

The proof of (5.16) can be found in [33, Lemma 3.5]. (5.17) is a well-known inequality as "Hardy-Littlewood maximal strong type estimate", please refer to [27, Theorem 1] for more proof details, we omit it here.

We cite the following lemma [31, Lemma 2.3] for future use.

Lemma 5.5. *Let $q>1$ and $\gamma\in[1,2]$. There exists a constant $C=C(q,\gamma,d)$ such that for any $f\in H_q^\gamma$,*

$$
\|f(\cdot+z)-f(\cdot)\|_{1,q}\le|z|^{\gamma-1}\|f\|_{\gamma,q}.
$$

To overcome the difficulty caused by the possible discontinuity of the drift b, we give some results on Zvonkin's transformation. More precisely, for any $\lambda > 0$, consider the following elliptic equation for $u^\lambda : \mathbb{R}^d \to \mathbb{R}^d$

$$\mathscr{L} u^\lambda + b + \nabla_b u^\lambda = \lambda u^\lambda, \tag{5.18}$$

where b is given in (1.2), $\mathscr{L}$ is defined as

$$\mathscr{L} f(x) = \int_{\mathbb{R}^d - \{\mathbf{0}\}} \{f(x+z) - f(x) - \langle \nabla f(x), z\rangle 1_{\{|z|\le 1\}}\}\nu(\mathrm{d}z). \tag{5.19}$$

According to [29, Theorem 4.11] and Sobolev's embedding (2.1), we have the following lemma.

Lemma 5.6. *Assume* **(A2)**. *Then, for any* $\gamma \in ((1+\alpha/2-\beta)\vee 1 \vee (d/p-\beta+1), \alpha)$, *there exists a constant* $\lambda_0 > 0$ *such that for any* $\lambda \ge \lambda_0$ (5.18) *has a unique solution* $u^\lambda \in H_p^{\gamma+\beta}$ *satisfying*

$$\|\nabla u^\lambda\|_\infty \le \frac{1}{2}, \quad \|u^\lambda\|_{\gamma+\beta,p} \le C_1 \|b\|_{\beta,p}, \tag{5.20}$$

for some constant $C_1 > 0$.

Now we are in position to complete the proof of Theorem 2.3.

Proof of Theorem 2.3. Firstly, recall that the Lévy-Itô decomposition of W_{S_t} is

$$W_{S_t} = \int_0^t \int_{\{|z|\le 1\}} z\widetilde{N}(\mathrm{d}s, \mathrm{d}z) + \int_0^t \int_{\{|z|>1\}} zN(\mathrm{d}s,\mathrm{d}z), \quad t \ge 0, \tag{5.21}$$

here, N is the Poisson random measure with compensator $\nu(\mathrm{d}z)\mathrm{d}t$. Set $\theta^\lambda(x) := x + u^\lambda(x), x \in \mathbb{R}^d$, and $Z_t^{(\delta)} := X_t - X_t^{(\delta)}$. According to Itô's formula in [29, Lemma 6.4], we obtain from (5.18) that

$$\begin{aligned}
\mathrm{d}\theta^\lambda(X_t) &= \lambda u^\lambda(X_t)\mathrm{d}t + \mathrm{d}W_{S_t} + \int_{\mathbb{R}^d-\{\mathbf{0}\}} [u^\lambda(X_{t-}+z) - u^\lambda(X_{t-})]\widetilde{N}(\mathrm{d}t,\mathrm{d}z) \\
\mathrm{d}\theta^\lambda(X_t^{(\delta)}) &= \lambda u^\lambda(X_t)\mathrm{d}t + \nabla\theta^\lambda(X_t^{(\delta)})(b(X_{t_\delta}^{(\delta)}) - b(X_t^{(\delta)}))\mathrm{d}t \\
&\quad + \mathrm{d}W_{S_t} + \int_{\mathbb{R}^d-\{\mathbf{0}\}} [u^\lambda(X_{t-}^{(\delta)}+z) - u^\lambda(X_{t-}^{(\delta)})]\widetilde{N}(\mathrm{d}t,\mathrm{d}z).
\end{aligned}$$

Set $\overline{\theta^\lambda}(X_t, X_t^{(\delta)}) := \theta^\lambda(X_t) - \theta^\lambda(X_t^{(\delta)})$ and $g(x,z) := u^\lambda(x+z) - u^\lambda(x)$. Then, it follows that

$$\begin{aligned}\mathrm{d}\overline{\theta^\lambda}(X_t, X_t^{(\delta)}) &= \lambda(u^\lambda(X_t) - u^\lambda(X_t^{(\delta)}))\mathrm{d}t + \nabla\theta^\lambda(X_t^{(\delta)})(b(X_t^{(\delta)}) - b(X_{t_\delta}^{(\delta)}))\mathrm{d}t \\ &\quad + \int_{|z|>1} \Big[g(X_{t-}, z) - g(X_{t-}^{(\delta)}, z)\Big]\, \widetilde{N}(\mathrm{d}t, \mathrm{d}z) \\ &\quad + \int_{0<|z|\le 1} \Big[g(X_{t-}, z) - g(X_{t-}^{(\delta)}, z)\Big]\, \widetilde{N}(\mathrm{d}t, \mathrm{d}z).\end{aligned} \tag{5.22}$$

We obtain from (5.20) that

$$\frac{1}{4}|Z_t^{(\delta)}|^2 \le |\overline{\theta^\lambda}(X_t, X_t^{(\delta)})|^2 \le \frac{9}{4}|Z_t^{(\delta)}|^2. \tag{5.23}$$

By Itô's formula and (5.23), we arrive at

$$\begin{aligned}|Z_t^{(\delta)}|^2 &\le 4|\overline{\theta^\lambda}(X_t, X_t^{(\delta)})|^2 \\ &\le 8\lambda \int_0^t \langle \overline{\theta^\lambda}(X_s, X_s^{(\delta)}), u^\lambda(X_s) - u^\lambda(X_s^{(\delta)})\rangle \mathrm{d}s \\ &\quad + 8\int_0^t \langle \overline{\theta^\lambda}(X_s, X_s^{(\delta)}), \nabla\theta^\lambda(X_s^{(\delta)})(b(X_s^{(\delta)}) - b(X_{s_\delta}^{(\delta)}))\rangle \mathrm{d}s \\ &\quad + 4\int_0^t \int_{|z|\ge 1} \Big|g(X_{s-}, z) - g(X_{s-}^{(\delta)}, z)\Big|^2 \nu(\mathrm{d}z)\mathrm{d}s \\ &\quad + 4\int_0^t \int_{0<|z|\le 1} \Big|g(X_{s-}, z) - g(X_{s-}^{(\delta)}, z)\Big|^2 \nu(\mathrm{d}z)\mathrm{d}s \\ &\quad + 4\Big\{\int_0^t \int_{|z|\ge 1} \big\{|\hat{\theta^\lambda}(X_s, X_s^{(\delta)}, g)|^2 - |\overline{\theta^\lambda}(X_s, X_s^{(\delta)})|^2\big\}\widetilde{N}(\mathrm{d}s, \mathrm{d}z) \\ &\qquad + \int_0^t \int_{0<|z|\le 1} \big\{|\hat{\theta^\lambda}(X_s, X_s^{(\delta)}, g)|^2 - |\overline{\theta^\lambda}(X_s, X_s^{(\delta)})|^2\big\}\widetilde{N}(\mathrm{d}s, \mathrm{d}z)\Big\} \\ &=: \sum_{i=1}^4 I_i^{(\delta)}(t) + M(t),\end{aligned} \tag{5.24}$$

where $\hat{\theta^\lambda}(X_t, X_t^{(\delta)}, g) = \overline{\theta^\lambda}((X_t, X_t^{(\delta)}) + g(X_{t-}, z) - g(X_{t-}^{(\delta)}, z))$. By means of (5.20) and (5.23), we obtain

$$I_1^{(\delta)}(t) \le 6\lambda \int_0^t |Z_s^{(\delta)}|^2 \mathrm{d}s. \tag{5.25}$$

Similarly, by virtue of (5.20), (5.23) and Young's inequality, we arrive at

$$I_2^{(\delta)}(t) \le C\Big\{\int_0^t |Z_s^{(\delta)}|^2 \mathrm{d}s + \int_0^t |b(X_s^{(\delta)}) - b(X_{s_\delta}^{(\delta)})|^2 \mathrm{d}s\Big\}. \tag{5.26}$$

Thanks to [12, (2.9), (2.12)], there exists a constant $C(t,\nu)>0$ such that

$$I_3^{(\delta)}(t) \le C\int_0^t \int_{|z|>1} \Big|g(X_{u-},z)-g(X_{u-}^{(\delta)},z)\Big|^2 \mathrm{d}u\nu(\mathrm{d}z) \le C(t,\nu)\int_0^t |Z_u^{(\delta)}|^2 \mathrm{d}u. \tag{5.27}$$

Let $\gamma\in((1+\alpha/2-\beta)\vee 1\vee(d/p-\beta+1),\alpha)$. Define

$$U(x,z) := |z|^{1-\beta-\gamma}|\nabla u^\lambda(x+z)-\nabla u^\lambda(x)|.$$

It follows that

$$|\nabla g(\cdot,z)(x)| = U(x,z)|z|^{\beta+\gamma-1}. \tag{5.28}$$

Noting that $(\mathscr{M}(f))^2(x)\le \mathscr{M}(f^2)(x)$ due to Jensen's inequality. By (5.16), (5.28) and [29, (3.3)] for $\mathbb{B}=L^2(\{0<|z|\le 1\},\nu)$, we get

$$\begin{aligned}
&I_4^{(\delta)}(t)\\
&= \int_0^t \int_{0<|z|\le 1} \Big|g(X_{s-},z)-g(X_{s-}^{(\delta)},z)\Big|^2 \mathrm{d}s\nu(\mathrm{d}z)\\
&\le 2C_2\int_0^t |Z_s^{(\delta)}|^2\Big\{\mathscr{M}\Big(\int_{0<|z|\le1}|\nabla g(\cdot,z)|^2\nu(\mathrm{d}z)\Big)(X_{s-})\\
&\qquad\qquad + \mathscr{M}\Big(\int_{0<|z|\le1}|\nabla g(\cdot,z)|^2\nu(\mathrm{d}z)\Big)(X_{s-}^{(\delta)})\Big\}\mathrm{d}s\\
&= 2C_2\int_0^t |Z_s^{(\delta)}|^2\Big\{\mathscr{M}\Big(\int_{0<|z|\le1}|U(\cdot,z)|^2|z|^{2(\beta+\gamma-1)}\nu(\mathrm{d}z)\Big)(X_{s-})\\
&\qquad\qquad + \mathscr{M}\Big(\int_{0<|z|\le1}|U(\cdot,z)|^2|z|^{2(\beta+\gamma-1)}\nu(\mathrm{d}z)\Big)(X_{s-}^{(\delta)})\Big\}\mathrm{d}s.
\end{aligned}$$

As a result, plugging this with (5.25)-(5.27) into (5.24) gives that

$$|Z_t^{(\delta)}|^2 \le C_3\int_0^t \sup_{r\in[0,s]}|Z_r^{(\delta)}|^2(\mathrm{d}s+\mathrm{d}A_s) + \int_0^t C_3|b(X_s^{(\delta)})-b(X_{s_\delta}^{(\delta)})|^2\mathrm{d}s + M_t,$$

where M_t is a local martingale, and

$$\begin{aligned}
A_t = \int_0^t &\Big\{\mathscr{M}\Big(\int_{0<|z|\le1}|U(\cdot,z)|^2|z|^{2(\beta+\gamma-1)}\nu(\mathrm{d}z)\Big)(X_{s-})\\
&\qquad + \mathscr{M}\Big(\int_{0<|z|\le1}|U(\cdot,z)|^2|z|^{2(\beta+\gamma-1)}\nu(\mathrm{d}z)\Big)(X_{s-}^{(\delta)})\Big\}\mathrm{d}s.
\end{aligned}$$

By (5.17) and Minkowski's inequality, we have

$$\Big\|\mathscr{M}\Big(\int_{0<|z|\le 1}|U(\cdot,z)|^2|z|^{2(\beta+\gamma-1)}\nu(\mathrm{d}z)\Big)\Big\|_{\frac{p}{2}}\le C\int_{0<|z|\le 1}\|U(\cdot,z)\|_p^2\,|z|^{2(\beta+\gamma-1)}\nu(\mathrm{d}z).$$

This together with Hölder's inequality and the fact $\int_{0<|z|\le 1}|z|^{2(\beta+\gamma-1)}\nu(\mathrm{d}z)<\infty$ due to $2(\beta+\gamma-1)>\alpha$, we derive that for any $\zeta>0$,

$$\begin{aligned}\mathbb{E}\exp\{\zeta A_t\}&\le\bigg(\mathbb{E}\exp\Big\{2\zeta\int_0^t\mathscr{M}\Big(\int_{0<|z|\le 1}|U(\cdot,z)|^2|z|^{2(\beta+\gamma-1)}\nu(\mathrm{d}z)\Big)(X_{s-})\mathrm{d}s\Big\}\bigg)^{1/2}\\&\quad\times\bigg(\mathbb{E}\exp\Big\{2\zeta\int_0^t\mathscr{M}\Big(\int_{0<|z|\le 1}|U(\cdot,z)|^2|z|^{2(\beta+\gamma-1)}\nu(\mathrm{d}z)\Big)(X_{s-}^{(\delta)})\mathrm{d}s\Big\}\bigg)^{1/2}\\&\le 2^{\left\{1+t\left(2\zeta c\int_{0<|z|\le 1}\|U(\cdot,z)\|_p^2|z|^{2(\beta+\gamma-1)}\nu(\mathrm{d}z)\right)^{\frac{\alpha p}{\alpha p-2d}}\right\}}\\&\le 2^{\left\{1+t\left(2\zeta C\|b\|_{\beta,p}^2\right)^{\frac{\alpha p}{\alpha p-2d}}\right\}},\end{aligned}$$

where the second inequality is due to (5.6), (5.13) for taking parameters $\lambda=2\zeta$ and $q=\frac{p}{2}$, and in the last display we used the fact that

$$\begin{aligned}\|U(\cdot,z)\|_p&=|z|^{1-\beta-\gamma}\Big(\int_{\mathbb{R}^d}|\nabla u^\lambda(x+z)-\nabla u^\lambda(x)|^p\mathrm{d}x\Big)^{1/p}\\&\le|z|^{1-\beta-\gamma}|z|^{\beta+\gamma-1}\|\nabla u\|_{\beta+\gamma-1,p}\le\|u\|_{\beta+\gamma,p}\le C\|b\|_{\beta,p},\end{aligned}$$

which is due to Lemma 5.6 and Lemma 5.5.

Consequently, we deduce by stochastic Gronwall's inequality (see e.g. [29, Lemma 3.8]) that, for $0<\kappa'<\kappa<1$,

$$\begin{aligned}&\Big(\mathbb{E}\Big(\sup_{0\le s\le t}|Z_s^{(\delta)}|^{2\kappa'}\Big)\Big)^{1/\kappa'}\\&\le\Big(\frac{\kappa}{\kappa-\kappa'}\Big)^{1/\kappa'}\Big(\mathbb{E}\mathrm{e}^{\kappa A_t/(1-\kappa)}\Big)^{(1-\kappa)/\kappa}\times\int_0^t\Big\{C_3\mathbb{E}|b(X_s^{(\delta)})-b(X_{s_\delta}^{(\delta)})|^2\Big\}\mathrm{d}s.\end{aligned}$$

Taking $\kappa'=\frac{\eta}{2}$ and combining with Lemma 5.1 implies that (2.5) holds. □

References

[1] Bao, J., Huang, X., Yuan, C., Convergence rate of Euler–Maruyama Scheme for SDEs with Hölder-Dini continuous drifts, *J. Theor. Probab.*, **32** (2019), 848-871.

[2] Bao, J., Huang, X., Zhang, S., Convergence rate of EM algorithm for SDEs under integrability condition, *J. Appl. Probab.*, (2022), 1-24. doi:10.1017/jpr.2021.56

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

[4] Chen, Z.-Q., Zhang, X., Heat kernels and analyticity of non-symmetric jump diffusion semigroups, *Probab. Theory Relat. Fields*, **165** (2016), 267-312.

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

[6] Dareiotis, K., Kumar, C., Sabanis, S., On tamed Euler approximations of SDEs driven by Lévy noise with applications to delay equations, *SIAM J. Numer. Anal.*, **54** (2016), 1840-1872.

[7] Flandoli, M., Gubinelli, M., Priola, E., Flow of diffeomorphisms for SDEs with unbounded Hölder continuous drift, *Bull. Sci. Math.*, **134** (2010), 405-422.

[8] Gottlich, S., Lux, K., Neuenkirch, A., The Euler scheme for stochastic differential equations with discontinuous drift coefficient: A numerical study of the convergence rate, *Adv. Differ. Equ.*, **429** (2019), 1-21.

[9] Gyöngy, I., Martinez, T., On stochastic differential equations with locally unbounded drift, *Czechoslovak Math.J.*, **51** (2001), 763-783.

[10] Halidias, N., Kloeden, P. E., A note on the Euler-Maruyama scheme for stochastic differential equations with a discontinuous monotone drift coefficient, *BIT Numer. Math.*, **48** (2008), 51-59.

[11] Higham,D. J, Kloeden, P. E., Strong convergence rates for backward Euler on a class of nonlinear jump diffusion problems, *J. Comput. Appl. Math.*, **205** (2007), 949-956.

[12] Huang, X., Liao, Z., The Euler-Maruyama method for S(F)DEs with Hölder drift and α-stable noise, *Stoch. Anal. Appl.*, **36** (2018), 28-39.

[13] Huang, X., Wang, F.-Y., Degenerate SDEs with singular drift and applications to Heisenberg groups, *J. Differ. Equ.*, **265** (2018), 2745-2777.

[14] Knopova, V., Kulik, A., Parametrix construction of the transition probability density of the solution to an SDE driven by α-stable noise, *Ann. Inst. H. Poincaré Probab. Statist.* , **54** (2018), 100-140.

[15] Konakov, V., Mammen, E., Local limits theorems for transition densities of Markov chains converging to diffusions, *Probab. Theory Relat. Fields.*, **117** (2000), 551-587.

[16] Konakov, V., Menozzi, S., Weak error for stable driven stochastic differential equations: Expansion of densities, *J. Theor. Probab.*, **24** (2011), 454-478.

[17] Kloden, P.E., Platen, E., Numerical solutions of stochastic differential equations, *Springer, Berlin-Heidelberg.* 1992.

[18] Krylov, N. V., Röckner, M., Strong solutions of stochastic equations with singular time dependent drift, *Probab. Theory Relat. Fields.*, **131** (2005), 154-196.

[19] Kühn F., Schilling R.-L., Strong convergence of the Euler-Maruyama approximation for a class of Lévy-driven SDEs. *Stoch. Process Their Appl.*, **129** (2019), 2654-2680.

[20] Lemaire, V., Menozzi, S., On some non asymptotic bounds for the Euler scheme, *Electron. J. Probab.*, **15** (2010), 1645-1681.

[21] Leobacher, G., Szölgyenyi, M., A numerical method for SDEs with discontinuous drift, *BIT Numer. Math.*, **56** (2016), 151-162.

[22] Leobacher, G., Szölgyenyi, M., A strong order 1/2 method for multidimensional SDEs with discontinuous drift, *Ann. Appl. Probab.*, **27** (2017), 2383-2418.

[23] Leobacher, G., Szölgyenyi, M. , Convergence of the Euler-Maruyama method for multidimensional SDEs with discontinuous drift and degenerate diffusion coefficient, *Numer. Math.*, **138** (2018), 219-239.

[24] Ngo, H-L., Taguchi, D., Strong rate of convergence for the Euler-Maruya approximation of stochatic differential equations with irregular coefficients, *Math. Comp.* **85** (2016), 1793-1819.

[25] Pamen, O. M., Taguchi, D., Strong rate of convergence for the Euler-Maruyama approximation of SDE with Hölder continuous drift coefficient, *Stoch. Process Their Appl.*, **127** (2017), 2542-2559.

[26] Triebel, H., Interpolation theory, Functional Spaces, Differential Operators, *North-Holland Publishing company.* 1978.

[27] Stein, E. M. Singular integrals and differentiability properties of functions, Princeton University Press. 1970.

[28] Wang,. F.-Y., Wang, J., Harnack inequalities for stochastic equations driven by Lévy noise, *J. Math. Anal. Appl.* **410** (2014), 513-523.

[29] Xie, L., Zhang, X., Ergodicity of stochastic differential equations with jumps and singular coefficients, *Ann. Inst. H. Poincaré Probab. Statist.*, **56** (2020), 175-229.

[30] Zhang, X., Strong solutions of SDEs with singural drift and Sobolev diffusion coefficients, *Stoch. Process Their Appl.*, **115** (2005), 1805-1818.

[31] Zhang, X., Stochastic differential equations with Sobolev drifts and driven by α-stable processes, *Ann. Inst. H. Poincaré Probab. Statist.* , **49** (2013), 1057-1079.

[32] Zhang, X., Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients, *Electron. J. Probab.*, **16** (2011), 1096-1116.

[33] Zhang, X., Well-posedness and large deviation for degenerate SDEs with Sobolev coefficients, *Rev. Mat. Iberoam.*, **29** (2013), 25-52.

[34] Zvonkin, A. K., A transformation of the phase space of a diffusion process that removes the drift, *Math. USSR Sb.*, **93** (1974), 129-149.