Rev. Mat. Iberoam. **34** (2018), no. 4, 1755–1788
DOI 10.4171/RMI/1042
© European Mathematical Society

Stochastic flows for Lévy processes with Hölder drifts

Zhen-Qing Chen, Renming Song and Xicheng Zhang

Abstract. In this paper, we study the following stochastic differential equation (SDE) in $\mathbb{R}^d$:

$$\mathrm{d}X_t = \mathrm{d}Z_t + b(t, X_t)\,\mathrm{d}t, \quad X_0 = x,$$

where Z is a Lévy process. We show that for a large class of Lévy processes Z and Hölder continuous drifts b, the SDE above has a unique strong solution for every starting point $x \in \mathbb{R}^d$. Moreover, these strong solutions form a C^1-stochastic flow. As a consequence, we show that, when Z is an α-stable-type Lévy process with $\alpha \in (0,2)$ and b is a bounded β-Hölder continuous function with $\beta \in (1-\alpha/2, 1)$, the SDE above has a unique strong solution. When $\alpha \in (0,1)$, this in particular partially solves an open problem from Priola. Moreover, we obtain a Bismut type derivative formula for $\nabla \mathbb{E}_x f(X_t)$ when Z is a subordinate Brownian motion. To study the SDE above, we first study the following nonlocal parabolic equation with Hölder continuous b and f:

$$\partial_t u + \mathscr{L} u + b \cdot \nabla u + f = 0, \quad u(1, \cdot) = 0,$$

where $\mathscr{L}$ is the generator of the Lévy process Z.

1. Introduction

Consider the following stochastic differential equation (SDE) in $\mathbb{R}^d$:

$$\mathrm{d}X_t = \mathrm{d}Z_t + b(t, X_t)\,\mathrm{d}t, \quad X_0 = x, \tag{1.1}$$

where $b(t,x)\colon [0,1]\times\mathbb{R}^d \to \mathbb{R}^d$ is a bounded Borel function and Z is a Lévy process in $\mathbb{R}^d$. When Z is a Brownian motion, one can use the Girsanov transform to show

Mathematics Subject Classification (2010): Primary 60H10, 35K05; Secondary 60H30, 47G20.
Keywords: SDE, supercritical, subcritical, stable process, subordinate Brownian motion, gradient estimate, strong existence, pathwise uniqueness, Bismut formula, stochastic flow, C^1-diffeomorphism.

that SDE (1.1) has a unique weak solution for a large class of $b(t,x)$, for example, for bounded measurable Borel $b(t,x)$. However, when Z is a discontinuous Lévy process without Gaussian component, the problem becomes much harder for one can not use the Girsanov transform to solve (1.1). When $d=1$, Kurenok [14] showed that (1.1) has a weak solution for a class of one-dimensional Lévy processes and bounded time-dependent drift b, based on Krylov type estimates for Lévy processes with time-dependent drift. When $d=1$, Z is an isotropic α-stable process in $\mathbb{R}^d$ with $\alpha\in(1,2)$ and the drift $b=b(x)$ is L^p-integrable for $p>d/(\alpha-1)$, Portenko [17] used a perturbation approach to construct a weak solution to the SDE (1.1); it is extended to $d\geqslant 2$ in [16]. Recently, Chen and Wang [5], using heat kernel estimates and the martingale problem approach, showed that (1.1) has a unique weak solution when Z is an isotropic α-stable process with $\alpha>1$ and $b=b(x)$ is in some Kato class that includes $L^p(\mathbb{R}^d)$ with $p>d/(\alpha-1)$ and bounded Borel functions. In this paper, we will concentrate on the existence and uniqueness of strong solutions of (1.1) for non-Lipschitz continuous drift b. We refer the reader to [10] for the definitions of and the relations between weak solution, uniqueness of weak solution, strong solution, pathwise uniqueness. In non-technical terms, a weak solution to (1.1) means that we can find a pair (X,Z) on some probability space so that Z has the same distribution as the pre-given Lévy process and (1.1) holds. A strong solution to (1.1) means that given a Lévy process Z, there is a solution X to (1.1) on the same probability space on which Z is defined and is adapted to the filtration generated by Z.

When $d=1$, Z is a Brownian motion and b is a bounded Borel function on $\mathbb{R}$, Zvonkin [30] used a transformation (one-to-one map) to remove the drift from (1.1) and show (1.1) has a unique strong solution for every starting point x. When $b(t,x)$ depends on x only, this transformation is just a scale function for X. Zvonkin's approach was extended to the multi-dimensional case by Veretennikov [24]. Since then, many people have made contributions to the pathwise uniqueness problem for SDEs driven by Brownian motion (see [12], [9], [8], [26] and references therein). However, when Z is a pure jump Lévy process, strong existence and pathwise uniqueness of SDE (1.1) become quite involved for drift b which is not Lipschitz continuous. When $d=1$, $b(t,x)=b(x)$ and Z is a symmetric α-stable process in $\mathbb{R}$ with $\alpha\in(0,1)$, Tanaka, Tsuchiya and Watanabe [23] proved that pathwise uniqueness fails for (1.1) even for bounded $b\in C_b^\beta(\mathbb{R})$. On the other hand, when $d=1$ and Z is a symmetric α-stable process in $\mathbb{R}$ with $\alpha\in[1,2)$, it is shown in [23] that pathwise uniqueness holds for (1.1) for any bounded continuous $b(t,x)=b(x)$. For $d\geqslant 2$, using Zvonkin's transform, Priola [18] obtained pathwise uniqueness for SDE (1.1) when Z is a non-degenerate symmetric (but possibly non-isotropic) α-stable process in $\mathbb{R}^d$ with $\alpha\in[1,2)$ and time-independent $b(t,x)=b(x)\in C_b^\beta(\mathbb{R}^d)$ with $\beta\in(1-\alpha/2,1)$. Note that in this case, the infinitesimal generator corresponding to the solution X of (1.1) is $\mathscr{L}^{(\alpha)}+b\cdot\nabla$. Here $\mathscr{L}^{(\alpha)}$ is the infinitesimal generator of the Lévy process Z, which is a nonlocal operator of order α. When $\alpha>1$, $\mathscr{L}^{(\alpha)}$ is the dominant term, which is called the subcritical case. When $\alpha\in(0,1)$, the gradient ∇ is of higher order than the nonlocal operator $\mathscr{L}^{(\alpha)}$ so the corresponding SDE (1.1) is called supercritical. The critical case corresponds to $\alpha = 1$. Priola's result was extended to drifts b in some fractional Sobolev spaces in the subcritical case in Zhang [28] and to more general Lévy processes in the subcritical and critical cases in Priola [19]. However, when $d \geqslant 2$, $\alpha \in (0,1)$ and Z is a symmetric non-isotropic α-stable process in $\mathbb{R}^d$, even for time-independent Hölder continuous drift $b = b(x)$, pathwise uniqueness for SDE (1.1) was an open question until now; see Remark 5.5 in [19]. When Z is an isotropic α-stable process, SDE (1.1) is connected with the following nonlocal PDE:

$$\partial_t u + \Delta^{\alpha/2} u + b \cdot \nabla u + f = 0,$$

where $\Delta^{\alpha/2} := -(-\Delta)^{\alpha/2}$ is the usual fractional Laplacian. In order to solve SDE (1.1) driven by an isotropic stable process Z, one needs to understand the above PDE better. In this direction, Silvestre [21] obtained the following a priori interior estimate:

$$\|u\|_{L^\infty([0,1];C^{\alpha+\beta}(B_1))} \leqslant C \left(\|u\|_{L^\infty([0,2]\times B_2)} + \|f\|_{L^\infty([0,2];C^\beta(B_2))} \right), \tag{1.2}$$

where, for any $r > 0$, B_r stands for the open ball of radius r centered at the origin, provided $b \in L^\infty([0,2]; C^\beta(B_2))$ and $\alpha + \beta > 1$. This estimate, as pointed out in Remark 5.5 of [19], can be combined with the argument of [18] to show that the SDE (1.1) has a pathwise unique strong solution when Z is an isotropic α-stable process with $\alpha \in (0,1)$ and $b \in C_b^\beta(\mathbb{R}^d)$ with $\beta \in (1-\alpha/2, 1)$. However, the approach of [21] to establish (1.2) strongly depends on realizing the fractional Laplacian in $\mathbb{R}^d$ as the boundary trace of an elliptic operator in the upper half space of $\mathbb{R}^{d+1}$. Extending this approach to other nonlocal operators, such as α-stable-type operators, would be very hard if not impossible.

The goal of this paper is to establish strong existence and pathwise uniqueness for SDE (1.1) with, possibly time-dependent, Hölder continuous drift b for a large class of Lévy processes that have no Gaussian component, including stable-type Lévy processes. Our approach also uses Zvonkin's transform. One of the main contributions of this paper is a new approach of establishing estimates analogous to (1.2) for a large class of Lévy processes and for time-dependent drift $b(t,x)$; see Theorem 2.3. Probabilistic consideration played a key role in our approach. With this new approach, we not only extend the main result of [19] in the subcritical case to more general Lévy processes and time-dependent drifts, but also establish strong existence and pathwise uniqueness result in the supercritical case for a large class of Lévy processes where the drift b can be time-dependent. We emphasize that the Lévy process Z in this paper can be non-symmetric and may also have drift. Throughout this paper, we assume the Lévy process Z has no Gaussian component. If Z has a non-degenerate Gaussian component, then the Gaussian part will play the dominant role and one can obtain results similar to the case of Brownian motion. The case where Z has a degenerate Gaussian component will be different and we will not deal with this case in the present paper.

One of the main results of this paper (see Corollary 1.4 (i) below) in particular partially solves an open problem raised in Remark 5.5 of [19], where Z is a symmetric α-stable process with $\alpha \in (0,1)$; see (i) and (iii) of Corollary 1.4 below. Our approach is mainly probabilistic.

In this paper, we use ":=" as a way of definition. For $a, b \in \mathbb{R}$, $a \wedge b := \min\{a, b\}$, $a \vee b := \max\{a, b\}$, and $a^+ := a \vee 0$. Let $\mathscr{L}_{\nu,\eta}$ be the infinitesimal generator of the Lévy process Z, that is,

$$\mathscr{L}_{\nu,\eta} f(x) = \int_{\mathbb{R}^d} \big(f(x+z) - f(x) - \mathbf{1}_{\{|z|\leqslant 1\}} z \cdot \nabla f(x)\big)\, \nu(\mathrm{d}z) + \eta \cdot \nabla f(x),$$

where ν is the Lévy measure of Z and η is a vector in $\mathbb{R}^d$. For any $\eta \in \mathbb{R}^d$ and any Lévy measure ν, i.e., a measure on $\mathbb{R}^d \setminus \{0\}$ with $\int (1 \wedge |z|^2)\nu(\mathrm{d}z) < \infty$, we will use $\{T_t^{\nu,\eta}; t \geqslant 0\}$ to denote the transition semigroup of the Lévy process Z with infinitesimal generator $\mathscr{L}_{\nu,\eta}$, i.e.,

$$T_t^{\nu,\eta} f(x) := \mathbb{E} f(x + Z_t). \tag{1.3}$$

For any $r \in (0, 1)$, the operator $\mathscr{L}_{\nu,\eta}$ can be rewritten as

$$\mathscr{L}_{\nu,\eta} f(x) = \int_{\mathbb{R}^d} \big(f(x+z) - f(x) - \mathbf{1}_{\{|z|\leqslant r\}} z \cdot \nabla f(x)\big)\, \nu(\mathrm{d}z) + \eta_r \cdot \nabla f(x)$$

with

$$\eta_r = \eta - \int_{r<|z|\leqslant 1} z\, \nu(\mathrm{d}z).$$

Let $N(\mathrm{d}t, \mathrm{d}z)$ be the Poisson random measure associated with Z, i.e.,

$$N((0,t] \times \Gamma) := \sum_{0<s\leqslant t} 1_\Gamma(Z_s - Z_{s-}), \quad t > 0,\ \Gamma \in \mathcal{B}(\mathbb{R}^d \setminus \{0\}).$$

Let $\tilde{N}(\mathrm{d}t, \mathrm{d}z) := N(\mathrm{d}t, \mathrm{d}z) - \mathrm{d}t\nu(\mathrm{d}z)$ be the compensated Poisson random measure. By the Lévy–Itô decomposition, we can write for each $r > 0$,

$$Z_t = \int_0^t \int_{|z|<r} z\, \tilde{N}(\mathrm{d}s, \mathrm{d}z) + \int_0^t \int_{|z|\geqslant r} z\, N(\mathrm{d}s, \mathrm{d}z) + \eta_r\, t.$$

Before we present the main results, we give the main idea of this paper and a rough description of Zvonkin's transform. Consider the following backward parabolic system:

$$\partial_t u_t + (\mathscr{L}_{\nu,\eta} - \lambda) u_t + b_t \cdot \nabla u_t + b_t = 0, \qquad u_1 = 0, \tag{1.4}$$

where $\lambda \geqslant 0$ is a parameter to be chosen later. Suppose one could prove that the above system has a unique solution and further show that $\|\nabla u_t\|_\infty \leqslant c(1 \vee \lambda)^{-\theta} \|b\|_{\infty,\beta}$, where $\|\cdot\|_{\infty,\beta}$ is the Hölder norm of order β, see the beginning of Section 2 for a definition. Then one takes a fixed large λ so that $\|\nabla u_t\|_\infty \leqslant 1/2$. Define

$$\Phi_t(x) = x + u_t(x).$$

Since $\|\nabla u_t\|_\infty \leqslant 1/2$, $x \mapsto \Phi_t(x)$ is a flow of diffeomorphisms for which we have good control. Direct computations show that

$$\partial_t \Phi_t + \int_{|z|<r} [\Phi_s(x+z) - \Phi_s(x) - z \cdot \nabla \Phi_s(x)] \nu(\mathrm{d}z) + (b_t + \eta_r) \cdot \nabla \Phi_t = \bar{a}_t$$

with

$$\bar{a}_t(x) = \lambda u_t(x) + \eta_r - \int_{|z|\geqslant r} \big[u_t(x+z) - u_t(x)\big]\, \nu(\mathrm{d}z).$$

So, if X is a solution of (1.1), then using Itô's formula (see, e.g., [10]),

$$\begin{aligned} Y_t := \Phi_t(X_t) = \Phi_0(x) &+ \int_0^t \int_{|z|<r} (\Phi_s(X_{s-}+z) - \Phi_s(X_{s-}))\tilde{N}(\mathrm{d}s, \mathrm{d}z) \\ &+ \int_0^t \int_{|z|\geqslant r} (\Phi_s(X_{s-}+z) - \Phi_s(X_{s-}))N(\mathrm{d}s, \mathrm{d}z) + \int_0^t \bar{a}_s(X_s)\, \mathrm{d}s. \end{aligned}$$

Since $X_s = \Phi_s^{-1}(Y_s)$, we get

$$\begin{aligned} Y_t = \Phi_0(x) &+ \int_0^t \int_{|z|<r} (\Phi_s(\Phi_s^{-1}(Y_{s-}) + z) - Y_{s-})\tilde{N}(\mathrm{d}s, \mathrm{d}z) \\ &+ \int_0^t \int_{|z|\geqslant r} (\Phi_s(\Phi_s^{-1}(Y_{s-}) + z) - Y_{s-})N(\mathrm{d}s, \mathrm{d}z) + \int_0^t \bar{a}_s(\Phi_s^{-1}(Y_s))\, \mathrm{d}s. \end{aligned} \tag{1.5}$$

In the last equation, b no longer appears and the regularity of the coefficients depends only on the regularity of Φ which is the same as that of the solution u of (1.4). Suppose that we have solved (1.4) and established enough regularity on the solution u. We can then show (1.5) has a strong solution Y. Clearly, $X_t = \Phi_t^{-1}(Y_t)$ will be a strong solution of (1.1). Uniqueness of solutions for (1.1) follows from the uniqueness for (1.5). We call the transform $Y_t = \Phi_t(X_t)$, which transforms (1.1) to (1.5), Zvonkin's transform.

Now solving (1.1) reduces to studying (1.4). We seek minimal conditions on the Lévy process and the drift $b(t,x)$ to guarantee a sufficiently regular solution of (1.4). We will assume that b is Hölder continuous of suitable order $\beta \in (0,1)$ and that the semigroup of Z has some regularization effect which will be spelled out precisely below. The regularization effect of the semigroup has to be strong enough to compensate for the lack of regularity of b. The interplay of the regularization effect of the semigroup and the Hölder continuity of b is the key to the argument of this paper, which will be realized by freezing the coefficient b at point $x_0 \in \mathbb{R}^d$ along the characterizing equation

$$\dot{y}_t = -b(y_t), \quad y_0 = x_0,$$

and using the pointwise estimate (1.7) below (for more details, see the proof of Lemma 2.6 (i) below).

We now describe the setup and the main results of this paper. Suppose that ν can be decomposed as

$$\nu = \nu_0 + \nu_1 + \nu_2, \tag{1.6}$$

where ν_1, ν_2 are two Lévy measures, and ν_0 is a *finite signed* measure supported on the set $\{z \in \mathbb{R}^d : |z| > 1\}$ so that

$$\nu_0 + \nu_1 \quad \text{is still a Lévy measure.}$$

The reason for this seemingly opaque decomposition is that it not only allows us to easily verify the condition of our main theorems, but also give us more freedom to include a larger class of processes in our framework as our main assumption will be only on ν_1 through the transition semigroup $\{T_t^{\nu,0}; t \geqslant 0\}$. In many circumstances, we do not know if this assumption holds directly on $T^{\nu,0}$ or not. This is the case, for example, when Z is a truncated α-stable-like Lévy process where $\nu(\mathrm{d}z) = c(z)|z|^{-(d+\alpha)}\mathbf{1}_{\{|z|\leqslant 1\}}\mathrm{d}z$ with $0 < c_1 \leqslant c(z) \leqslant c_2 < \infty$. However we can take $\nu_0(\mathrm{d}z) = -c|z|^{-(d+\alpha)}\mathbf{1}_{\{|z|>1\}}\mathrm{d}z$, $\nu_1 = c_1|z|^{-(d+\alpha)}\mathrm{d}z$ and $\nu_2(\mathrm{d}z) = (c(z) - c_1)|z|^{-(d+\alpha)}\mathbf{1}_{\{|z|\leqslant 1\}}\mathrm{d}z$ so that (1.6) holds. Since ν_1 is the Lévy measure for the rotationally symmetric α-stable process, which has nice scaling property, one can easily verify that its transition semigroup $\{T_t^{\nu,0}; t \geqslant 0\}$ satisfies the condition of our main theorems. See Example 4.2 below for another example of such a decomposition where we treat general α-stable type Lévy measures $\nu(\mathrm{d}z) = \kappa(z)\,\mathrm{d}z$ with $c_1|z|^{-d-\alpha} \leqslant \kappa(z) \leqslant c_2|z|^{-d-\alpha}$ for $|z| \leqslant 1$. The idea behind the decomposition of (1.6) is that the Lévy process $Z^{(1)}$ corresponding to $\nu_0 + \nu_1$ should share many properties with the Lévy process having Lévy measure ν_1, as it can be obtained from it by adding or removing jumps of size larger than 1, while the original Lévy process Z has the same distribution as the sum of $Z^{(1)}$ and a Lévy process $Z^{(2)}$ having Lévy measure ν_2 that is independent of $Z^{(1)}$, so many properties obtained for $Z^{(1)}$ can be transferred to Z. See the paragraph before the statement of Lemma 2.6 below for further motivation behind the decomposition (1.6) and its utility.

We now make the following assumption about $T_t^{\nu_1,0}$. There exist $\alpha \in (0,2)$, $\bar\alpha, \delta \in (0,1]$ and $K_0 > 0$ so that the following gradient estimates for the semigroup $\{T_t^{\nu_1,0}; t \geqslant 0\}$ hold:

$\left(\mathbf{H}^{\alpha,\bar\alpha,\delta}_{\nu_1,K_0}\right)$ If $\alpha \in (0,1]$, then for any $x \in \mathbb{R}^d$, $\beta \in [0,\bar\alpha]$ and bounded Borel function f satisfying

$$|f(x+y) - f(x)| \leqslant \Lambda\,|y|^\beta \quad \text{for all } y \in \mathbb{R}^d,$$

with some $\Lambda > 0$, it holds that

$$|\nabla T_t^{\nu_1,0} f(x)| \leqslant K_0\,\Lambda\, t^{(\delta\beta-1)/\alpha} \quad \text{for all } t \in (0,1). \tag{1.7}$$

$\left(\mathbf{H}^{\alpha}_{\nu_1,K_0}\right)$ If $\alpha \in (1,2)$, then for any bounded Borel function f, it holds that

$$\|\nabla T_t^{\nu_1,0} f\|_\infty \leqslant K_0\,\|f\|_\infty\, t^{-1/\alpha} \quad \text{for all } t \in (0,1). \tag{1.8}$$

Remark 1.1. (i) Estimates (1.7) and (1.8) allow us to borrow the Hölder regularity of the drift to compensate the time singularity.

(ii) Condition (1.8) in the subcritical case is the same as Hypothesis 1 of Priola [19] for Lévy process with Lévy measure $\nu = \nu_1$. In the subcritical case, under condition (1.8) for $\nu = \nu_1$ and condition (1.9) below for some $\gamma > \alpha/2$ (which is Hypothesis 2 in [19]), Priola [19] derived Hölder estimate for solutions of (1.4) which enabled him to show that SDE (1.1) has a unique strong solution.

(iii) The pointwise estimate (1.7) is crucial for the well-posedness of SDEs with Hölder drifts in the supercritical case. The reason for the complicated formulation of $(\mathbf{H}^{\alpha,\bar\alpha,\delta}_{\nu_1,K_0})$ is that it allows us to cover a larger class of processes. The parameters $\bar\alpha$ and δ are mainly designed to treat the case when Z is a cylindrical stable processes with possibly different stable indices for which previous approaches fail to work; see Example 4.3 below. In many other cases, for example, in Examples 4.1 and 4.2 below, $\bar\alpha$ and δ can all be chosen to be 1.

The first main result of this paper is the following.

Theorem 1.2. (i) (Supercritical and critical case) *Suppose that* $(\mathbf{H}^{\alpha,\bar\alpha,\delta}_{\nu_1,K_0})$ *holds for some* $\alpha \in (0,1]$, $\bar\alpha, \delta \in (0,1]$ *and* $K_0 > 0$. *Assume that there is* $\gamma > 0$ *with* $\gamma + (1-\alpha)/\delta < \bar\alpha$ *such that*

$$\int_{|z|\leqslant 1} |z|^{2\gamma}\nu(\mathrm{d}z) < \infty, \tag{1.9}$$

and there is some $\beta \in (\gamma + (1-\alpha)/\delta, 1]$ *so that*

$$\sup_{t\in[0,1]} \|b(t,\cdot)\|_\infty + \sup_{t\in[0,1]} \sup_{x\neq y\in\mathbb{R}^d} \frac{|b(t,x)-b(t,y)|}{|x-y|^\beta} < \infty. \tag{1.10}$$

Then for every $x \in \mathbb{R}^d$, *there is a unique strong solution* $\{X_t(x); t\in[0,1]\}$ *to the SDE* (1.1). *Moreover,* $\{X_t(x), t\in[0,1], x\in\mathbb{R}^d\}$ *forms a* C^1*-stochastic diffeomorphism flow, and for each* $x\in\mathbb{R}^d$, $t\mapsto \nabla X_t(x)$ *is continuous, and*

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}\Big[\sup_{t\in[0,1]} |\nabla X_t(x)|^p\Big] \leqslant C_p < \infty \quad \text{for every } p \geqslant 1, \tag{1.11}$$

where C_p *only depends on* $p, d, \alpha, \beta, \gamma, \nu, K_0, \bar\alpha, \delta$ *and the Hölder norm of* b.

(ii) (Subcritical case) *Suppose that* $(\mathbf{H}^{\alpha}_{\nu_1,K_0})$ *holds for some* $\alpha \in (1,2)$ *and* $K_0 > 0$, *and there is some* $\gamma \in (0,1)$ *so that* (1.9) *holds. Assume that* (1.10) *holds for some* $\beta \in ((\gamma+1-\alpha)^+, 1]$. *Then for every* $x\in\mathbb{R}^d$, *there is a unique strong solution* $\{X_t(x); t\in[0,1]\}$ *to SDE* (1.1). *Moreover,* $\{X_t(x), t\in[0,1], x\in\mathbb{R}^d\}$ *forms a* C^1*-stochastic diffeomorphism flow, and for each* $x\in\mathbb{R}^d$, $t\mapsto\nabla X_t(x)$ *is continuous, and* (1.11) *holds with constant* C_p *only depending on* $p, d, \alpha, \beta, \gamma, \nu, K_0$ *and the Hölder norm of* b.

Remark 1.3. (i) By a suitable localization argument (cf. [28]), for the local uniqueness of SDE (1.1), the global condition (1.10) can be replaced with a local condition. Moreover, although $t\mapsto X_t(x)$ is not continuous, since we are considering an additive noise, the conclusion that $t\mapsto\nabla X_t(x)$ is continuous is not surprising.

(ii) For the subcritical case, we only assume (1.9) to hold for some $\gamma\in(0,1)$ rather than for $\gamma > \alpha/2$ as assumed in [19]. So even in the subcritical case, Theorem 1.2 (ii) yields new result; see Remark 1.5.

Various examples of Lévy processes satisfying the conditions $\big(\mathbf{H}^{\alpha,\bar{\alpha},\delta}_{\nu_1,K_0}\big)$ with $\alpha \in (0,1]$, $\big(\mathbf{H}^{\alpha}_{\nu_1,K_0}\big)$ with $\alpha \in (1,2)$, and (1.9) (and hence the conclusion of Theorem 1.2 holds for these Lévy processes) are given in Section 4. To illustrate the scope and applicability of Theorem 1.2, here we only give the following corollary, which is a direct consequence of these examples.

Corollary 1.4. (i) (Stable-type Lévy process) *Let Z be a Lévy process in $\mathbb{R}^d$ whose Lévy measure ν has a density $\kappa(z)$. Assume that for some $0<\alpha_1 \leqslant \alpha_2 < 2$,*

$$c_1\,|z|^{-d-\alpha_1} \leqslant \kappa(z) \leqslant c_2\,|z|^{-d-\alpha_2} \quad \text{for} \;\; 0<|z|\leqslant 1.$$

Assume that $\alpha_2 < 2\alpha_1$, and $b(t,x)$ is bounded and β-Hölder continuous in x uniformly in $t\in[0,1]$, for some $\beta \in (1+\alpha_2/2-\alpha_1, 1]$. Then SDE (1.1) has a unique strong solution for every $x\in\mathbb{R}^d$ and (1.11) holds.

(ii) (Subordinate Brownian motion) *Let Z be a subordinate Brownian motion in $\mathbb{R}^d$ with characteristic function $\Phi(z)$. Suppose that there are $0<\alpha_1\leqslant\alpha_2<2$ such that*

$$C_1\,|z|^{\alpha_1} \leqslant \Phi(z) \leqslant C_2\,|z|^{\alpha_2} \quad \text{for } |z|\geqslant 1.$$

Assume that $\alpha_2 < 2\alpha_1$, and $b(t,x)$ is bounded and β-Hölder continuous in x uniformly in $t\in[0,1]$, for some $\beta \in (1+\alpha_2/2-\alpha_1, 1]$. Then SDE (1.1) has a unique strong solution for every $x\in\mathbb{R}^d$ and (1.11) holds.

(iii) (Cylindrical stable process) *Let $Z=(Z^1,\dots,Z^k)$, where Z^j, $1\leqslant j\leqslant k$, are independent d_j-dimensional rotationally symmetric α_j-stable processes, respectively, with $\alpha_j\in(0,2)$ and $d_j\geqslant 1$. Let $\alpha := \min_{1\leqslant j\leqslant k}\alpha_j$ and $\alpha_{\max} := \max_{1\leqslant j\leqslant k}\alpha_j$. Suppose that*

$$\text{(1.12)} \qquad \text{either} \quad \alpha>1 \quad \text{or} \quad \alpha\in(0,1] \;\; \text{and} \;\; \alpha_{\max} < 2\alpha^2/(2-\alpha),$$

and that $b(t,x)$ is bounded and β-Hölder continuous in x uniformly in $t\in[0,1]$, for some

$$\text{(1.13)} \qquad \beta\in(\beta_0,1] \;\; \text{with } \beta_0 := \alpha_{\max}/2 + (\alpha_{\max}/\alpha\mathbf{1}_{\{\alpha\leqslant 1\}} + \mathbf{1}_{\{\alpha>1\}})(1-\alpha).$$

Then SDE (1.1) has a unique strong solution for every $x\in\mathbb{R}^d$, where $d:=\sum_{j=1}^k d_j$, and (1.11) holds.

Note that condition (1.12) implies that $\alpha < 2\alpha^2/(2-\alpha)$. The latter is equivalent to $\alpha > 2/3$. If in Corollary 1.4 (iii), $\alpha_j=\alpha$ for every $1\leqslant j\leqslant k$, then conditions (1.12) and (1.13) become

$$\alpha > 2/3 \quad \text{and} \quad \beta\in(1-\alpha/2, 1], \quad \text{respectively.}$$

An interesting open question is whether the constraint $\alpha>2/3$ can be dropped.[1]

[1]Note added in proof: this question has been answered affirmatively by Z.-Q. Chen, X. Zhang and G. Zhao in a recent work [7].

Remark 1.5. Corollary 1.4 (iii) in particular covers some cases of α-stable processes with $\alpha \in (1,2)$ for which the results from Priola [19] are not applicable. Let ν be the Lévy measure of the cylindrical stable process Z in Corollary 1.4 (iii). We will in fact show in Example 4.3 that, when $\alpha = \min_{1\leqslant j\leqslant k} \alpha_j \in (1,2)$, condition $\big(\mathbf{H}^{\alpha}_{\nu,K_0}\big)$ holds for some $K_0 > 0$ but condition $\big(\mathbf{H}^{\alpha^*}_{\nu,K_0}\big)$ fails for any $\alpha^* > \alpha$. So Hypothesis 1 of [19] holds with this α for the cylindrical stable process Z. On the other hand, condition (1.9) holds if and only if $2\gamma > \alpha_{\max}$. Hence in the case $\alpha \in (1,2)$, Hypothesis 2 of [19] fails when α_j's are not identical (i.e., when $\alpha_{\max} > \alpha$), and so the main results of [19] are not applicable to these Lévy processes.

The second main result of this paper is the following derivative formula of $\mathbb{E} f(X_t(x))$.

Theorem 1.6. *Under the assumptions of Theorem 1.2, if $Z_t = W_{S_t}$ is a subordinate Brownian motion as described in Example 4.1 below, then we have the following derivative formula*:

$$\nabla \mathbb{E} f(X_t(x)) = \mathbb{E}\Big[\frac{f(X_t(x))}{S_t}\int_0^t \nabla X_s(x)\,\mathrm{d}W_{S_s}\Big], \quad f \in C_b^1(\mathbb{R}^d). \tag{1.14}$$

In particular, for any $p > 1$, there is a constant $C_p > 0$ such that for any $f \in C_b^1(\mathbb{R}^d)$ and $(t,x) \in (0,1) \times \mathbb{R}^d$,

$$|\nabla \mathbb{E} f(X_t(x))| \leqslant C_p\, t^{-1/\alpha} \left(\mathbb{E}|f(X_t(x))|^p\right)^{1/p}. \tag{1.15}$$

This paper is organized as follows. In Section 2, we solve a nonlocal advection equation which is slightly more general than (1.4) and obtain estimates on the gradient of the solutions. In particular, we derive a priori uniform $C^{1+\gamma}$ estimate on the solution of the nonlocal advection equation. This is crucial for applying Zvonkin's transform. Even when Z is a rotationally symmetric stable process, our approach to the a priori estimate is simpler and more elementary than that of [21]. In Section 3, we shall prove our main results by using Zvonkin's transform. In Section 3, we give three examples to illustrate the main results of this paper, from which Corollary 1.4 follows. In the Appendix, we prove a continuous dependence result about the SDEs with jumps with respect to the coefficients and the initial values.

2. Differentiability of solutions of nonlocal advection equations

In this paper we use the following conventions. The letter C with or without subscripts will denote a positive constant, whose value is not important and may change from one appearance to another. We write $f(x) \preceq g(x)$ to mean that there exists a constant $C_0 > 0$ such that $f(x) \leqslant C_0 g(x)$; and $f(x) \asymp g(x)$ to mean that there exist $C_1, C_2 > 0$ such that $C_1 g(x) \leqslant f(x) \leqslant C_2 g(x)$.

For a function $u(t,x)$ defined on $[0,1]\times\mathbb{R}^d$, sometimes we use $u_t(x)$ for $u(t,x)$. Denote by $C_c^\infty(\mathbb{R}^d)$ the space of smooth functions with compact support on $\mathbb{R}^d$. For $\beta\in(0,1]$ and a function f on $\mathbb{R}^d$,

$$[f]_\beta := \sup_{x\neq y}\frac{|f(x)-f(y)|}{|x-y|^\beta},\quad \|f\|_\beta := \|f\|_\infty + [f]_\beta,$$

and for a function $f\colon [0,1]\times\mathbb{R}^d\to\mathbb{R}$,

$$[f]_{\infty,\beta} := \sup_{s\in[0,1]}[f_s]_\beta,\quad \|f\|_{\infty,\beta} := \sup_{s\in[0,1]}\|f_s\|_\beta.$$

Recall the following characterization for a Hölder continuous function f. Let $P_\theta f$ be the Poisson integral of f defined by

$$P_\theta f(x) := \int_{\mathbb{R}^d} f(y)\,p_\theta(x-y)\,\mathrm{d}y,\quad \theta>0,$$

where $p_\theta(x)$ is the density of a Cauchy process Z_θ given by

$$p_\theta(x) := c_d\,\theta(\theta^2+|x|^2)^{-(d+1)/2}\asymp \theta(\theta+|x|)^{-d-1}.$$

It is well known (cf. [22], Proposition 7 on p. 142) that $\|f\|_\beta<\infty$ if and only if f is bounded and

$$\|\partial_\theta P_\theta f\|_\infty \leqslant C\,\theta^{\beta-1}\quad \text{for every } \theta>0$$

and

$$\|f\|_\beta \asymp \|f\|_\infty + \sup_{\theta>0}\|\theta^{1-\beta}\partial_\theta P_\theta f\|_\infty. \tag{2.1}$$

The following commutator estimate result plays an important role in our proof of the Hölder regularity of the gradient in the case of $\alpha\in(0,1]$.

Lemma 2.1. *For any $\beta,\gamma\in(0,1)$ with $\gamma\leqslant\beta$, there is a positive constant $C=C(\beta,\gamma,d)$ such that for any Borel functions f,g on $\mathbb{R}^d$,*

$$[\partial_\theta P_\theta(fg)-f\partial_\theta P_\theta g]_{\beta-\gamma}\leqslant C\,[f]_\beta\,\|g\|_\infty\,\theta^{\gamma-1},\quad \theta>0,$$

provided that $[f]_\beta$ and $\|g\|_\infty$ are finite. In particular, if $g\equiv 1$, then

$$[\partial_\theta P_\theta f]_{\beta-\gamma}\leqslant C\,[f]_\beta\,\theta^{\gamma-1},\quad \theta>0.$$

Proof. It suffices to prove that

$$\begin{aligned}|\partial_\theta P_\theta(fg)(x) - f\partial_\theta P_\theta g(x) - \partial_\theta P_\theta\,(fg)(x') + f\partial_\theta P_\theta g(x')|\\ \leqslant C\,[f]_\beta\,\|g\|_\infty\,\theta^{\gamma-1}\,|x-x'|^{\beta-\gamma}.\end{aligned}\tag{2.2}$$

By definition, we have

$$\partial_\theta P_\theta(fg)(x) - f\partial_\theta P_\theta g(x) = \int_{\mathbb{R}^d}(f(y)-f(x))\,g(y)\,\partial_\theta p_\theta(x-y)\,\mathrm{d}y. \tag{2.3}$$

Notice the following easy estimates:

$$|\partial_\theta p_\theta(x)| \preceq (\theta+|x|)^{-d-1}, \quad |\nabla\partial_\theta p_\theta(x)| \preceq (\theta+|x|)^{-d-2} \tag{2.4}$$

and

$$\int_{\mathbb{R}^d} |x|^\beta (\theta+|x|)^{-d-k}\,\mathrm{d}x \preceq \theta^{\beta-k}, \quad k\in\mathbb{N}. \tag{2.5}$$

If $|x-x'|\geqslant \theta/2$, then (2.2) follows from

$$\begin{aligned}\|\partial_\theta P_\theta(fg) - f\partial_\theta P_\theta g\|_\infty &\overset{(2.3)}{\leqslant} [f]_\beta\,\|g\|_\infty \int_{\mathbb{R}^d} |y|^\beta |\partial_\theta p_\theta(y)|\mathrm{d}y \\ &\overset{(2.4)}{\preceq} [f]_\beta\,\|g\|_\infty \int_{\mathbb{R}^d} |y|^\beta (\theta+|y|)^{-d-1}\mathrm{d}y \\ &\overset{(2.5)}{\preceq} [f]_\beta\,\|g\|_\infty \theta^{\beta-1} \preceq [f]_\beta\,\|g\|_\infty\,\theta^{\gamma-1}|x-x'|^{\beta-\gamma}.\end{aligned}$$

Next, we assume

$$|x-x'|\leqslant \theta/2. \tag{2.6}$$

Notice that

$$\begin{aligned}\partial_\theta P_\theta(fg)(x) &- f\partial_\theta P_\theta g(x) - (\partial_\theta P_\theta(fg)(x') - f\partial_\theta P_\theta g(x')) \\ &= \int_{\mathbb{R}^d} (f(y)-f(x))\,g(y)\,(\partial_\theta p_\theta(x-y) - \partial_\theta p_\theta(x'-y))\,\mathrm{d}y \\ &\quad + \int_{\mathbb{R}^d} (f(x')-f(x))\,g(y)\,\partial_\theta p_\theta(x'-y)\,\mathrm{d}y =: I_1 + I_2.\end{aligned}$$

For I_1, we have

$$\begin{aligned}|I_1| &\leqslant [f]_\beta\,\|g\|_\infty \int_{\mathbb{R}^d} |x-y|^\beta |x-x'| \Big(\int_0^1 |\nabla\partial_\theta p_\theta(x-y+r(x'-x))|\mathrm{d}r\Big)\,\mathrm{d}y \\ &\overset{(2.4)}{\preceq} [f]_\beta\,\|g\|_\infty\,|x-x'| \int_{\mathbb{R}^d} |x-y|^\beta \Big(\int_0^1 (\theta+|x-y+r(x'-x)|)^{-d-2}\mathrm{d}r\Big)\,\mathrm{d}y \\ &\overset{(2.6)}{\preceq} [f]_\beta\,\|g\|_\infty\,|x-x'| \int_{\mathbb{R}^d} |x-y|^\beta (\theta+|x-y|)^{-d-2}\mathrm{d}y \\ &\overset{(2.5)}{\preceq} [f]_\beta\,\|g\|_\infty\,|x-x'|\theta^{\beta-2} \overset{(2.6)}{\preceq} [f]_\beta\,\|g\|_\infty\,|x-x'|^{\beta-\gamma}\theta^{\gamma-1}.\end{aligned}$$

For I_2, we similarly have

$$|I_2| \preceq |x-x'|^\beta\,[f]_\beta\,\|g\|_\infty \int_{\mathbb{R}^d} (\theta+|y|)^{-d-1}\mathrm{d}y \preceq [f]_\beta\,\|g\|_\infty\,|x-x'|^{\beta-\gamma}\,\theta^{\gamma-1}.$$

Combining the above estimates, we obtain (2.2). □

We also need the following lemma for treating the case of $\alpha \in (1,2)$.

Lemma 2.2. *If* $\big(\mathbf{H}^{\alpha}_{\nu_1,K_0}\big)$ *holds for some* $\alpha \in (1,2)$ *and* $K_0 > 0$, *then for any* $\beta, \gamma \in [0,1]$, *there is a constant* $K_1 > 0$ *such that*

$$\|\nabla T^{\nu_1,0}_t f\|_\gamma \leqslant K_1\, t^{(\beta-1-\gamma)/\alpha}\, \|f\|_\beta \quad \text{for all } t \in (0,1).$$

Proof. Note that $\|\nabla T^{\nu_1,0}_t f\|_\infty \leqslant \|\nabla f\|_\infty$. By (1.8) and Proposition 1.2.6 of [15] (with $X_1 = C^0_b(\mathbb{R}^d)$, $Y_1 = C^1_b(\mathbb{R}^d)$ and $X_2 = Y_2 = L^\infty(\mathbb{R}^d)$), we get

$$\|\nabla T^{\nu_1,0}_t f\|_\infty \preceq t^{(\beta-1)/\alpha}\|f\|_\beta.$$

On the other hand, by $\big(\mathbf{H}^{\alpha}_{\nu_1,K_0}\big)$ we have

$$\|\nabla^2 T^{\nu_1,0}_t f\|_\infty = \|\nabla T^{\nu_1,0}_{t/2}\nabla T^{\nu_1,0}_{t/2} f\|_\infty \preceq (t/2)^{-1/\alpha}\|\nabla T^{\nu_1,0}_{t/2} f\|_\infty \preceq (t/2)^{(\beta-2)/\alpha}\|f\|_\beta.$$

Hence,

$$[\nabla T^{\nu_1,0}_t f]_\gamma \leqslant 2\,\|\nabla^2 T^{\nu_1,0}_t f\|^\gamma_\infty\, \|\nabla T^{\nu_1,0}_t f\|^{1-\gamma}_\infty \preceq t^{(\beta-1-\gamma)/\alpha}\, \|f\|_\beta. \qquad \square$$

For $\lambda \geqslant 0$, consider the following linear backward nonlocal parabolic system:

$$\partial_t u_t + (\mathscr{L}_{\nu,\eta} - \lambda)u_t + b_t \cdot \nabla u_t + f_t = 0, \quad u_1 = 0, \tag{2.7}$$

where $\mathscr{L}_{\nu,\eta}$ is the infinitesimal generator of the Lévy process Z, and $b, f\colon [0,1]\times \mathbb{R}^d \to \mathbb{R}^d$ are bounded Borel functions.

The following theorem is the main result of this section and it is crucial in our proof of Theorem 1.2 in the next section.

Theorem 2.3. (i) (Supercritical and critical case) *Suppose* $\alpha \in (0,1]$ *and* $\big(\mathbf{H}^{\alpha,\bar\alpha,\delta}_{\nu_1,K_0}\big)$ *holds for some* $\bar\alpha, \delta \in (0,1]$ *and* $K_0 > 0$ *with* $1-\alpha < \delta\bar\alpha$. *If*

$$\|b\|_{\infty,\beta} < \infty, \quad \|f\|_{\infty,\beta} < \infty \tag{2.8}$$

for some $\beta \in ((1-\alpha)/\delta, \bar\alpha]$, *then for any* $\gamma \in (0, \beta - (1-\alpha)/\delta)$, *there exists a continuous function* $u\colon [0,1]\times\mathbb{R}^d \to \mathbb{R}^d$ *such that for all* $t \in [0,1]$ *and* $\varphi \in C^\infty_c(\mathbb{R}^d)$,

$$\langle u_t, \varphi\rangle = \int_t^1 \langle u_s, (\mathscr{L}^*_{\nu,\eta} - \lambda)\varphi\rangle\, \mathrm{d}s + \int_t^1 \langle b_s \cdot \nabla u_s, \varphi\rangle \mathrm{d}s + \int_t^1 \langle f_s, \varphi\rangle\, \mathrm{d}s \tag{2.9}$$

with

$$\sup_{t\in[0,1]} \|u_t(\cdot)\|_\infty \leqslant \sup_{t\in[0,1]} \|f_t(\cdot)\|_\infty, \tag{2.10}$$

and for some $\theta_0 > 0$ *and all* $\lambda \geqslant 0$,

$$\|\nabla u\|_{\infty,\gamma} \leqslant C(1\vee\lambda)^{-\theta_0}\|f\|_{\infty,\beta}. \tag{2.11}$$

Here $C = C(d,\alpha,\beta,K_0,\bar\alpha,\delta,\|b\|_{\infty,\beta},\gamma,|\nu_0|(\mathbb{R}^d))$, $\langle u,\varphi\rangle := \int u\varphi \mathrm{d}x$ *and* $\mathscr{L}^*_{\nu,\eta}$ *is the adjoint operator of* $\mathscr{L}_{\nu,\eta}$.

(ii) (Subcritical case) *Suppose* $\alpha \in (1,2)$ *and* $\big(\mathbf{H}^{\alpha}_{\nu_1,K_0}\big)$ *holds for some* $K_0 > 0$. *If* (2.8) *holds for some* $\beta \in [0,1]$, *then for any* $\gamma \in (0, (\beta+\alpha-1)\wedge 1)$, *there exists a continuous function* $u\colon [0,1]\times\mathbb{R}^d \to \mathbb{R}^d$ *such that for all* $t \in [0,1]$ *and* $\varphi \in C^\infty_c(\mathbb{R}^d)$, (2.9)–(2.11) *hold with* $C = C(d,\alpha,\beta,K_0,\|b\|_{\infty,\beta},\gamma,|\nu_0|(\mathbb{R}^d))$.

Remark 2.4. At this stage one can not show the uniqueness of weak solutions for (2.7) with regularities (2.10) and (2.11) since u may be not in the domain of $\mathscr{L}_{\nu,\eta}$. In Corollary 2.10 below, under an additional assumption (2.28), we will show the existence of classical solutions, which automatically yields the well-posedness in the class of classical solutions.

We will first prove several lemmas before we present the proof of the theorem above.

Lemma 2.5. *If $b, f \in L^\infty([0,1]; C_b^\infty(\mathbb{R}^d;\mathbb{R}^d))$, then there exists a unique solution $u_t(x)$ in the space $C([0,1]; C_b^\infty(\mathbb{R}^d;\mathbb{R}^d))$ to equation (2.7) with the following probabilistic representation:*

$$u_t(x) = \int_t^1 \mathrm{e}^{\lambda(t-s)}\, \mathbb{E} f_s(X_{t,s}(x))\, \mathrm{d}s, \tag{2.12}$$

where $X_{t,s}(x) = X_{t,s}$ is the unique solution to the following SDE:

$$X_{t,s} = x + \int_t^s b_r(X_{t,r})\, \mathrm{d}r + Z_s - Z_t, \quad s \geqslant t. \tag{2.13}$$

Moreover, we have the following a priori estimate:

$$\sup_{t\in[0,1]} \|u_t\|_\infty \leqslant \sup_{t\in[0,1]} \|f_t\|_\infty. \tag{2.14}$$

Proof. The existence and uniqueness of $u_t(x)$ and the representation (2.12) follow from Theorem 4.4 in [27]. The estimate (2.14) immediately follows from (2.12). □

Recalling decomposition (1.6), we can write

$$\mathscr{L}_{\nu,\eta} = \mathscr{L}_{\nu_0,0} + \mathscr{L}_{\nu_1,0} + \mathscr{L}_{\nu_2,\eta}, \tag{2.15}$$

where $\mathscr{L}_{\nu_0,0}$ is given by

$$\mathscr{L}_{\nu_0,0} f(x) = \int_{|z|>1} (f(x+z) - f(x))\, \nu_0(\mathrm{d}z).$$

Let $Z^{(1)}$ and $Z^{(2)}$ be two independent Lévy processes with generators $\mathscr{L}_{\nu_0+\nu_1,0}$ and $\mathscr{L}_{\nu_2,\eta}$. Clearly,

$$Z_t \overset{(d)}{=} Z_t^{(1)} + Z_t^{(2)}. \tag{2.16}$$

The following is a key lemma on the gradient estimate for solutions u of (2.7) in the supercritical case. One can relatively easily obtain such a gradient estimate if the semigroup $T_t^{\nu,0}$ has the property (1.7). But as we mentioned earlier, for many interesting cases of Lévy processes, such as truncated stable processes and general stable-type Lévy processes in Corollary 1.4 (i), we do not know if (1.7) holds directly for them (that is, with $\nu = \nu_1$) or not. Our idea is as follows. Decompose the Lévy measure ν as in (1.6). Since $\mathscr{L}_{\nu_0+\nu_1,0}$ is a lower order perturbation of $\mathscr{L}_{\nu_1,0}$ by a finite measure ν_0 on $\{z \in \mathbb{R}^d : |z| > 1\}$, under condition (1.7), one can obtain the desired gradient estimate for Lévy process $Z^{(1)}$. Since $Z_t \overset{(d)}{=} Z_t^{(1)} + Z_t^{(2)}$, intuitively, the solution X_t to the original SDE (1.1) should have the same distribution as $Y_t + Z_t^{(2)}$, where conditional on $Z^{(2)}$, Y_t is a weak solution of

$$Y_t = X_0 + Z_t^{(1)} + \int_0^t b(t, Y_s + Z_s^{(2)})\, dt.$$

With this intuition in mind, using the probabilistic representation of u and by conditioning on $Z^{(2)}$, we can reduce the solution u of (2.7) for $\mathscr{L}_{\nu,\eta}$ to a solution of (2.7) for $\mathscr{L}_{\nu_0+\nu_1,\eta}$ with $\tilde{b}(t,z) := b(t, x + Z_t^{(2)})$ in place of $b(t,x)$ and then get the desired gradient estimate. See the proof of part (ii) of the following lemma.

Lemma 2.6. *Suppose that $(\mathbf{H}^{\alpha,\bar{\alpha},\delta}_{\nu_1,K_0})$ holds for some $K_0 > 0$ and $\alpha \in (0,1]$, $\bar{\alpha}, \delta \in (0,1]$ with $1-\alpha < \delta\bar{\alpha}$, and that $b, f \in L^\infty([0,1]; C_b^\infty(\mathbb{R}^d;\mathbb{R}^d))$. Let u be the solution of (2.7). Then for any $\beta_1, \beta_2 \in ((1-\alpha)/\delta, \bar{\alpha}]$, there is a constant $C > 0$ depending only on $K_0, \alpha, \delta, \beta_1, \beta_2$, $[b]_{\infty,\beta_1}$ and $|\nu_0|(\mathbb{R}^d)$ such that for all $\lambda \geqslant 0$,*

$$\sup_{t\in[0,1]} \|\nabla u_t\|_\infty \leqslant C\,(1 \vee \lambda)^{(1-\alpha-\delta\beta_2)/\alpha}\,[f]_{\infty,\beta_2}. \tag{2.17}$$

Proof. (i) We first assume that $\eta = 0$ and that $\nu_2 = 0$ in decomposition (1.6). Fix $x_0 \in \mathbb{R}^d$ and let y_t satisfy the following ODE:

$$\dot{y}_t = b_t(x_0 + y_t) \quad \text{with } y_0 = 0.$$

Define

$$\tilde{u}_t(x) := u_t(x + x_0 + y_t), \quad \tilde{f}_t(x) := f_t(x + x_0 + y_t) \tag{2.18}$$

and

$$\tilde{b}_t(x) := b_t(x + x_0 + y_t) - b_t(x_0 + y_t).$$

Clearly, by (2.7) and (2.15), $\tilde{u}$ satisfies

$$\partial_t \tilde{u}_t + (\mathscr{L}_{\nu_1,0} - \lambda)\tilde{u}_t + \tilde{b}_t \cdot \nabla\tilde{u}_t + \mathscr{L}_{\nu_0,0}\tilde{u}_t + \tilde{f}_t = 0, \quad \tilde{u}_1 = 0.$$

We have by the representation (2.12) (with $b = 0$ and f_s replaced by $g_s = \tilde{b}_s \cdot \nabla\tilde{u}_s + \mathscr{L}_{\nu_0,0}\tilde{u}_s + \tilde{f}_s$ there)

$$\tilde{u}_t(x) = \int_t^1 \mathrm{e}^{\lambda(t-s)}\, T^{\nu_1,0}_{s-t}\big(\tilde{b}_s \cdot \nabla\tilde{u}_s + \mathscr{L}_{\nu_0,0}\tilde{u}_s + \tilde{f}_s\big)(x)\, \mathrm{d}s.$$

Fix $\beta_1, \beta_2 \in ((1-\alpha)/\delta, \bar{\alpha}]$. Note that by the definition of $\tilde{b}_s$,

$$|\tilde{b}_s(y) \cdot \nabla\tilde{u}_s(y)| \leqslant \|\nabla\tilde{u}_s(\cdot)\|_\infty\, [b_s(\cdot)]_{\beta_1}\, |y|^{\beta_1} \quad \text{for all } y \in \mathbb{R}^d,$$

and that by $\mathscr{L}_{\nu_0,0}f(x)=\int_{\{|z|>1\}}(f(x+z)-f(x))\nu_0(\mathrm{d}z)$,

$$|\nabla T^{\nu_1,0}_{s-t}(\mathscr{L}_{\nu_0,0}\tilde u_s)(x)|\leqslant\|\nabla(\mathscr{L}_{\nu_0,0}\tilde u_s)\|_\infty\leqslant 2\,|\nu_0|(\mathbb{R}^d)\,\|\nabla\tilde u_s(\cdot)\|_\infty.$$

We have by (1.7) that for $t\in[0,1]$,

$$\begin{aligned}|\nabla\tilde u_t(0)|\leqslant&\int_t^1\mathrm{e}^{\lambda(t-s)}\big(K_0[b_s(\cdot)]_{\beta_1}(s-t)^{(\delta\beta_1-1)/\alpha}+2|\nu_0|(\mathbb{R}^d)\big)\|\nabla\tilde u_s(\cdot)\|_\infty\,\mathrm{d}s\\&+K_0\int_t^1\mathrm{e}^{\lambda(t-s)}(s-t)^{(\delta\beta_2-1)/\alpha}\,[\tilde f_s(\cdot)]_{\beta_2}\,\mathrm{d}s.\end{aligned}$$

By (2.18) and the arbitrariness of x_0, one in fact has

(2.19)
$$\|\nabla u_t(\cdot)\|_\infty\leqslant C\int_t^1(s-t)^{(\delta\beta_1-1)/\alpha}\|\nabla u_s(\cdot)\|_\infty\mathrm{d}s+C\,[f]_{\infty,\beta_2}(1\vee\lambda)^{(1-\alpha-\delta\beta_2)/\alpha}.$$

By Gronwall's inequality, we obtain (2.17).

(ii) Next we consider the general case. Fix $t_0\in[0,1)$ and a càdlàg function $\ell\colon[0,1]\to\mathbb{R}^d$, and define

$$b^\ell_r(x):=b_r(x-\ell_{t_0}+\ell_r),\quad f^\ell_r(x):=f_r(x-\ell_{t_0}+\ell_r).$$

Let $Y^\ell_{t,s}(x):=Y^\ell_{t,s}$ be the solution to the following SDE:

$$Y^\ell_{t,s}=x+\int_t^s b^\ell_r(Y^\ell_{t,r})\,\mathrm{d}r+Z^{(1)}_s-Z^{(1)}_t,\quad s\geqslant t.$$

Since $Z^{(1)}$ and $Z^{(2)}$ are independent, by (2.16) and the uniqueness in law of the solution to SDE (2.13), we have

$$X_{t_0,\cdot}(x)\overset{(d)}{=}Y^{Z^{(2)}}_{t_0,\cdot}(x)-Z^{(2)}_{t_0}+Z^{(2)}_\cdot,$$

and so by (2.12),

$$u_{t_0}(x)=\mathbb{E}\Big(\int_{t_0}^1\mathrm{e}^{\lambda(t_0-s)}\,\mathbb{E}\big[f^\ell_s(Y^\ell_{t_0,s}(x))\big]\,\mathrm{d}s\Big|_{\ell=Z^{(2)}}\Big).$$

Now we define

$$u^\ell_t(x):=\int_t^1\mathrm{e}^{\lambda(t-s)}\,\mathbb{E}\big[f^\ell_s(Y^\ell_{t,s}(x))\big]\,\mathrm{d}s.$$

Then by Lemma 2.5, $u^\ell_t(x)$ is a solution to the following equation:

$$\partial_t u^\ell+(\mathscr{L}_{\nu_0+\nu_1,0}-\lambda)u^\ell+b^\ell\cdot\nabla u^\ell+f^\ell=0,\quad u^\ell_1=0.$$

In view of

$$[b^\ell]_{\infty,\beta_1}=[b]_{\infty,\beta_1},\quad[f^\ell]_{\infty,\beta_2}=[f]_{\infty,\beta_2},$$

by what has been proved in (i), we have for any càdlàg function ℓ,

$$\|\nabla u^\ell\|_\infty \leqslant C_2\,(1\vee\lambda)^{(1-\alpha-\delta\beta_2)/\alpha}\,[f]_{\infty,\beta_2},$$

which in turn gives (2.17) by noting that $\nabla u_{t_0}(x) = \mathbb{E}\big[\nabla u^\ell_{t_0}(x)|_{\ell=Z^{(2)}}\big]$ and t_0 is arbitrary. □

Lemma 2.7. *Suppose that $(\mathbf{H}^{\alpha,\bar\alpha,\delta}_{\nu_1,K_0})$ holds for some $K_0>0$ and $\alpha\in(0,1]$, $\bar\alpha,\delta\in(0,1]$ with $1-\alpha<\delta\bar\alpha$, and that $b,f\in L^\infty([0,1];C^\infty_b(\mathbb{R}^d;\mathbb{R}^d))$. Let u be the solution of (2.7). Then for any $\beta\in((1-\alpha)/\delta,\bar\alpha]$ and $\gamma\in(0,\beta-(1-\alpha)/\delta)$, there exists a constant $C>0$ depending only on $d,\alpha,\bar\alpha,\delta,K_0,\beta,\gamma$, $[b]_{\infty,\beta}$ and $|\nu_0|(\mathbb{R}^d)$ such that for all $\lambda\geqslant 0$,*

$$[\nabla u]_{\infty,\gamma}\leqslant C\,(1\vee\lambda)^{(1-\alpha-\delta(\beta-\gamma))/\alpha}\,[f]_{\infty,\beta}. \tag{2.20}$$

Proof. Fix $\gamma\in(0,\beta-(1-\alpha)/\delta)$. For $\theta>0$, define

$$w^\theta_t(x) := \partial_\theta P_\theta u_t(x)$$

and

$$g^\theta_t(x) := \partial_\theta P_\theta(b_t\cdot\nabla u_t)(x) - b_t(x)\cdot\nabla\partial_\theta P_\theta u_t(x) + \partial_\theta P_\theta f_t(x),$$

then

$$\partial_t w^\theta_t + (\mathscr{L}_{\nu,\eta}-\lambda)w^\theta_t + b_t\cdot\nabla w^\theta_t + g^\theta_t = 0,\qquad w^\theta_1 = 0.$$

Since $\beta-\gamma>(1-\alpha)/\delta$, by (2.17) with $\beta_1=\beta$ and $\beta_2=\beta-\gamma$, we have

$$\sup_{t\in[0,1]}\|\nabla w^\theta_t(\cdot)\|_\infty \leqslant C(1\vee\lambda)^{(1-\alpha-\delta(\beta-\gamma))/\alpha}[g^\theta]_{\infty,\beta-\gamma},$$

and by Lemma 2.1,

$$[g^\theta]_{\infty,\beta-\gamma}\preceq [b]_{\infty,\beta}\sup_{t\in[0,1]}\|\nabla u_t\|_\infty\,\theta^{\gamma-1} + [f]_{\infty,\beta}\,\theta^{\gamma-1} \overset{(2.17)}{\preceq} [f]_{\infty,\beta}\,\theta^{\gamma-1}.$$

Hence,

$$\sup_{t\in[0,1]}\|\partial_\theta P_\theta\nabla u_t\|_\infty \leqslant C(1\vee\lambda)^{(1-\alpha-\delta(\beta-\gamma))/\alpha}\,[f]_{\infty,\beta}\,\theta^{\gamma-1}\quad \text{for every } \theta>0,$$

which yields (2.20) by (2.1) and (2.17). □

Lemma 2.8. *Suppose that $(\mathbf{H}^{\alpha}_{\nu_1,K_0})$ holds for some $\alpha\in(1,2)$ and $K_0>0$, and that $b,f\in L^\infty([0,1];C^\infty_b(\mathbb{R}^d;\mathbb{R}^d))$. Let u be the solution of (2.7). Then for any $\beta\in[0,1]$ and $\gamma\in[0,(\beta+\alpha-1)\wedge 1)$, there exists a constant $C>0$ depending only on $d,\alpha,K_0,\gamma,\beta$, $\|b\|_{\infty,\beta}$ and $|\nu_0|(\mathbb{R}^d)$ such that for all $\lambda\geqslant 0$,*

$$\|\nabla u\|_{\infty,\gamma}\leqslant C\,(1\vee\lambda)^{(1-\alpha-\beta+\gamma)/\alpha}\,\|f\|_{\infty,\beta}. \tag{2.21}$$

Proof. As in the proof of Lemma 2.6, we first assume that $\eta = 0$ and that $\nu_2 = 0$ in decomposition (1.6). By the representation (2.12) (with $b = 0$ there), we have

$$u_t(x) = \int_t^1 \mathrm{e}^{\lambda(t-s)}\, T^{\nu_1,0}_{s-t}\left(b_s \cdot \nabla u_s + \mathscr{L}_{\nu_0,0} u_s + f_s\right)(x)\,\mathrm{d}s.$$

Without loss of generality, we assume $\gamma \in [\beta, (\beta+\alpha-1)\wedge 1)$. By Lemma 2.2, we have

$$\begin{aligned}
&\|\nabla u_t\|_\gamma \\
&\preceq \int_t^1 \mathrm{e}^{\lambda(t-s)}(s-t)^{\frac{\beta-1-\gamma}{\alpha}}\left(\|b_s\cdot\nabla u_s\|_\beta + \|f_s\|_\beta\right)\mathrm{d}s + \int_t^1 \mathrm{e}^{\lambda(t-s)}\|\nabla T^{\nu_1,0}_{s-t}\mathscr{L}_{\nu_0,0}u_s\|_\gamma\,\mathrm{d}s \\
&\preceq \int_t^1 \mathrm{e}^{\lambda(t-s)}(s-t)^{(\beta-1-\gamma)/\alpha}\left(\|b_s\|_\beta\|\nabla u_s\|_\beta + \|f_s\|_\beta\right)\mathrm{d}s + \int_t^1 \mathrm{e}^{\lambda(t-s)}\|\nabla u_s\|_\gamma\,\mathrm{d}s \\
&\preceq \int_t^1 \left((s-t)^{(\beta-1-\gamma)/\alpha}+1\right)\|\nabla u_s\|_\gamma\,\mathrm{d}s + \|f\|_{\infty,\beta}\int_t^1 \mathrm{e}^{\lambda(t-s)}(s-t)^{(\beta-1-\gamma)/\alpha}\,\mathrm{d}s,
\end{aligned}$$

which yields (2.21) by Gronwall's inequality. For the general case, we can follow the same argument as in (ii) of Lemma 2.6 to derive (2.21). □

We also need the following simple lemma.

Lemma 2.9. *Let $\mathscr{U}$ be a family of uniformly bounded continuous functions. If*

$$\lim_{|y|\to 0}\sup_{f\in\mathscr{U}}\|f(\cdot+y)-f(\cdot)\|_\infty = 0,$$

then

$$\lim_{t\to 0}\sup_{f\in\mathscr{U}}\|T^{\nu,0}_t f - f\|_\infty = 0.$$

Proof. Notice that for any $\varepsilon > 0$,

$$|T^{\nu,0}_t f(x) - f(x)| \leqslant 2\,\|f\|_\infty\,\mathbb{P}(|Z_t|\geqslant\varepsilon) + \mathbb{E}(|f(x+Z_t)-f(x)|\cdot 1_{|Z_t|<\varepsilon}).$$

The desired limit follows by the assumption and the uniform boundedness assumption and the fact $\lim_{t\to 0}\mathbb{P}(|Z_t|\geqslant\varepsilon) = 0$. □

Now we are ready to give:

Proof of Theorem 2.3. Suppose that b and f satisfy (2.8). Let ϱ be a non-negative smooth function with compact support in $\mathbb{R}^d$ satisfying $\int_{\mathbb{R}^d}\varrho(x)\mathrm{d}x = 1$. For $n\in\mathbb{N}$, define $\varrho_n(x) := n^d\varrho(nx)$ and

$$b^n_t := \varrho_n * b_t, \quad f^n_t := \varrho_n * f_t. \tag{2.22}$$

Clearly, $b^n, f^n \in L^\infty([0,1]; C^\infty_b(\mathbb{R}^d,\mathbb{R}^d))$ and

$$\|b^n\|_{\infty,\beta} \leqslant \|b\|_{\infty,\beta}, \quad \|f^n\|_{\infty,\beta}\leqslant\|f\|_{\infty,\beta}.$$

Let u_t^n be the solution to the following equation:

$$\partial_t u_t^n + (\mathscr{L}_{\nu,\eta} - \lambda)u_t^n + b_t^n \cdot \nabla u_t^n + f_t^n = 0, \quad u_1^n = 0. \tag{2.23}$$

By (2.17), (2.20) and (2.21), there is a $\theta_0 > 0$ such that for all $\lambda \geqslant 0$,

$$\sup_n \|\nabla u^n\|_{\infty,\gamma} \leqslant C\,(1 \vee \lambda)^{-\theta_0}\,\|f\|_{\infty,\beta}, \tag{2.24}$$

and by (2.14),

$$\sup_{t\in[0,1]} \|u_t^n\|_\infty \leqslant \sup_{t\in[0,1]} \|f_t^n\|_\infty \leqslant \sup_{t\in[0,1]} \|f_t\|_\infty. \tag{2.25}$$

Moreover, by the representation (2.12) (with $b = 0$ there), we can write

$$u_t^n(x) = \int_t^1 \mathrm{e}^{\lambda(t-s)}\, T_{s-t}^{\nu,0}((\eta + b_s^n)\cdot \nabla u_s^n + f_s^n)(x)\,\mathrm{d}s.$$

Using this representation, (2.24), (2.25) and Lemma 2.9, one can easily show that

$$\lim_{|t-t'|\to 0} \sup_n \|u_t^n - u_{t'}^n\|_\infty = 0.$$

Hence, by the Ascoli–Arzelà lemma, there is a subsequence (still denoted by u^n) and a function u with

$$\|\nabla u\|_{\infty,\gamma} \leqslant C(1\vee\lambda)^{-\theta_0}\|f\|_{\infty,\beta}, \qquad \sup_{t\in[0,1]}\|u_t\|_\infty \leqslant \sup_{t\in[0,1]}\|f_t\|_\infty$$

such that

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \sup_{|x|\leqslant R} |u_t^n(x) - u_t(x)| = 0, \quad \text{for all } R > 0. \tag{2.26}$$

On the other hand, noticing the following interpolation inequality (cf. Theorem 3.2.1 in [11]):

$$\|\nabla\phi\|_\infty \leqslant C\|\nabla\phi\|_\gamma^{1/(1+\gamma)}\|\phi\|_\infty^{\gamma/(1+\gamma)},$$

by (2.24) and (2.26), we further have

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \sup_{|x|\leqslant R} |\nabla u_t^n(x) - \nabla u_t(x)| = 0 \quad \text{for every } R > 0. \tag{2.27}$$

Thus by (2.23), (2.26) and (2.27), it is easy to see that u satisfies (2.9). □

Corollary 2.10. *Under the assumption of Theorem 2.3, if we further assume that for some $\gamma_0 \in (0, \beta - (1-\alpha)/\delta)$ in the case of $\alpha \in (0,1]$ and $\gamma_0 \in (0, (\beta+\alpha-1)\wedge 1)$ in the case of $\alpha \in (1,2)$,*

$$\int_{|z|\leqslant 1} |z|^{1+\gamma_0}\,\nu(\mathrm{d}z) < \infty, \tag{2.28}$$

then the solution u of equation (2.7) satisfying (2.10) and (2.11) for some $\gamma > \gamma_0$ is a classical solution; that is, $\mathscr{L}_{\nu,\eta} u_s(x)$ and $\nabla u_s(x)$ exist pointwise and are continuous in x, and for all $t \in [0,1]$ and $x \in \mathbb{R}^d$,

$$u_t(x) = \int_t^1 (\mathscr{L}_{\nu,\eta} - \lambda) u_s(x)\, \mathrm{d}s + \int_t^1 b_s(x) \cdot \nabla u_s(x)\, \mathrm{d}s + \int_t^1 f_s(x)\, \mathrm{d}s. \tag{2.29}$$

Proof. Since $\|\nabla u\|_{\infty,\gamma} < \infty$ for some $\gamma \in (\gamma_0, \beta-(1-\alpha)/\delta)$ in the case of $\alpha \in (0,1]$ and $\gamma \in (\gamma_0, (\beta+\alpha-1)\wedge 1)$ in the case of $\alpha \in (1,2)$, the function $x \mapsto \nabla u_s(x)$ is continuous. Now using (2.28), it is easy to check that

$$x \mapsto \mathscr{L}_{\nu,\eta} u_s(x) \ \text{ is continuous.}$$

Hence, by (2.9), equation (2.29) is satisfied for all $t \in [0,1]$ and $x \in \mathbb{R}^d$. □

3. Stochastic flow and Bismut formula

Throughout this section, we assume that either $(\mathbf{H}^{\alpha,\bar\alpha,\delta}_{\nu_1,K_0})$ holds for some $\alpha \in (0,1]$, $\bar\alpha, \delta \in (0,1]$ and $K_0 > 0$ or $(\mathbf{H}^{\alpha}_{\nu_1,K_0})$ holds for some $\alpha \in (1,2)$ and $K_0 > 0$. Suppose also that (1.9) and (1.10) hold for some

$\gamma \in (0,1)$ with $\gamma + (1-\alpha)/\delta < \bar\alpha$ and $\beta \in (\gamma + (1-\alpha)/\delta, \bar\alpha]$ in the case of $\alpha \in (0,1]$

and

$$\gamma \in (0,1) \text{ and } \beta \in ((\gamma+1-\alpha)^+, 1] \text{ in the case of } \alpha \in (1,2).$$

Notice that (1.9) implies (2.28) with $\gamma_0 = \gamma$. Hence, for $\lambda \geqslant 0$, by Corollary 2.10, the following nonlocal equation has a classical solution u:

$$\partial_t u_t + (\mathscr{L}_{\nu,\eta} - \lambda) u_t + b_t \cdot \nabla u_t + b_t = 0, \quad u_1(x) = 0.$$

Similarly, let b^n be defined by (2.22) and let u^n be the solution to the following equation:

$$\partial_t u^n_t + (\mathscr{L}_{\nu,\eta} - \lambda) u^n_t + b^n_t \cdot \nabla u^n_t + b^n_t = 0, \quad u^n_1(x) = 0. \tag{3.1}$$

Using the same argument leading to (2.26) and (2.27), we see that there is a subsequence, still denoted by u^n, such that

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \sup_{|x|\leqslant R} |\nabla^j u^n_t(x) - \nabla^j u_t(x)| = 0 \quad \text{for every } R > 0 \text{ and } j = 0, 1. \tag{3.2}$$

For simplicity, we use the following convention:

$$u^\infty := u, \qquad b^\infty := b, \qquad \mathbb{N}_\infty := \mathbb{N} \cup \{\infty\}.$$

By (2.11), one can choose λ sufficiently large, independent of $n \in \mathbb{N}_\infty$, such that

$$\|\nabla u^n_t(\cdot)\|_\infty + \sup_{x\neq x'} \frac{|\nabla u^n_t(x) - \nabla u^n_t(x')|}{|x-x'|^\gamma} \leqslant \frac{1}{2} \quad \text{for every } n \in \mathbb{N}_\infty \text{ and } t \in [0,1]. \tag{3.3}$$

From now on we will fix such a λ. Define

$$\Phi_t^n(x) = x + u_t^n(x), \quad n \in \mathbb{N}_\infty. \tag{3.4}$$

Since for each $t \in [0,1]$,

$$\frac{1}{2}|x-y| \leqslant |\Phi_t^n(x) - \Phi_t^n(y)| \leqslant \frac{3}{2}|x-y|,$$

$x \mapsto \Phi_t^n(x)$ is a diffeomorphism with

$$1/2 \leqslant |\nabla\Phi_t^n(x)| \leqslant 3/2 \quad \text{and} \quad |\nabla(\Phi_t^n)^{-1}(x)| \leqslant 2, \tag{3.5}$$

where $(\Phi_t^n)^{-1}$ denotes the inverse function of $x \mapsto \Phi_t^n(x)$.

Lemma 3.1. *Under the assumptions above, there is a constant $C = C(d,\gamma) > 0$ such that for all $t \in [0,1]$ and $n \in \mathbb{N}_\infty$,*

$$\|\nabla\Phi_t^n\|_\gamma + \|\nabla(\Phi_t^n)^{-1}\|_\gamma \leqslant C. \tag{3.6}$$

Moreover, for each $t \in [0,1]$, $R > 0$ and $j = 0,1$, we have

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \sup_{|x|\leqslant R} \left|\nabla^j\Phi_t^n(x) - \nabla^j\Phi_t^\infty(x)\right| = 0 \tag{3.7}$$

and

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \sup_{|x|\leqslant R} \left|\nabla^j(\Phi_t^n)^{-1}(x) - \nabla^j(\Phi_t^\infty)^{-1}(x)\right| = 0. \tag{3.8}$$

Proof. (i) For notational simplicity, we drop the superscript "n". Clearly,

$$\sup_{t\in[0,1]} \|\nabla\Phi_t(\cdot)\|_\gamma < d+1.$$

In view of

$$(\nabla\Phi_s)^{-1}(x) - (\nabla\Phi_s)^{-1}(x') = (\nabla\Phi_s)^{-1}(x)\,(\nabla\Phi_s(x') - \nabla\Phi_s(x))\,(\nabla\Phi_s)^{-1}(x'),$$

we have by (3.3) and (3.5),

$$[(\nabla\Phi_s)^{-1}]_\gamma \leqslant \|(\nabla\Phi_s)^{-1}\|_\infty^2[\nabla\Phi_s]_\gamma = \|(\nabla\Phi_s)^{-1}\|_\infty^2[\nabla u_s]_\gamma \leqslant 2 \quad \text{for all } s \in [0,1].$$

Hence by (3.5) again, for all $s \in [0,1]$,

$$\|\nabla\Phi_s^{-1}\|_\gamma = \|(\nabla\Phi_s)^{-1}(\Phi_s^{-1})\|_\gamma \leqslant \|(\nabla\Phi_s)^{-1}\|_\infty + \|\nabla\Phi_s^{-1}\|_\infty^\gamma[(\nabla\Phi_s)^{-1}]_\gamma \leqslant 2 + 2^{\gamma+1}.$$

(ii) Properties (3.7) and (3.8) follow from the definitions of Φ_t and Φ_t^{-1}, (3.2) and (3.3). □

For any given $n \in \mathbb{N}_\infty$, define

$$g_s^n(y,z) := \Phi_s^n\left((\Phi_s^n)^{-1}(y) + z\right) - y. \tag{3.9}$$

Lemma 3.2. *Under the assumption stated at the beginning of this section, for $\gamma_1, \gamma_2 \geqslant 0$ with $\gamma_1 + \gamma_2 = \gamma \in (0,1)$, there is a positive constant $C_1 = C_1(d, \gamma_1, \gamma_2)$ such that for all $n \in \mathbb{N}_\infty$, $t \in [0,1]$ and $y, z \in \mathbb{R}^d$,*

$$\|\nabla g^n(\cdot, z)\|_{\infty,\gamma_1} \leqslant C_1(1 \wedge |z|^{\gamma_2}) \quad \text{and} \quad |g^n_s(y,z)| \leqslant 3|z|/2. \tag{3.10}$$

Moreover, for each $t \in [0,1]$, $y, z \in \mathbb{R}^d$ and $j = 0, 1$, we have

$$\lim_{n\to\infty} \nabla^j_y g^n_t(y,z) = \nabla^j_y g^\infty_t(y,z). \tag{3.11}$$

Proof. For notational simplicity, we drop the superscript "n" in this proof. Since

$$\nabla_y g_s(y,z) = \nabla\Phi_s\left(\Phi_s^{-1}(y) + z\right) \cdot \nabla\Phi_s^{-1}(y) - \mathbb{I},$$

where $\mathbb{I}$ denotes the identity $d \times d$ matrix, we have

$$\|\nabla g_s(\cdot, z)\|_\infty \leqslant 2\|\nabla\Phi_s\|_\gamma (1 \wedge |z|^\gamma)\|\nabla\Phi_s^{-1}\|_\infty \overset{(3.6)}{\leqslant} C(1 \wedge |z|^\gamma)$$

and

$$\begin{aligned}
[\nabla g_s(\cdot, z)]_\gamma &\leqslant [\nabla\Phi_s(\Phi_s^{-1}(\cdot) + z)]_\gamma \|\nabla\Phi_s^{-1}\|_\infty + \|\nabla\Phi_s\|_\infty [\nabla\Phi_s^{-1}]_\gamma \\
&\leqslant [\nabla\Phi_s]_\gamma \|\nabla\Phi_s^{-1}\|_\infty^{1+\gamma} + \|\nabla\Phi_s\|_\infty [\nabla\Phi_s^{-1}]_\gamma \overset{(3.6)}{\leqslant} C.
\end{aligned}$$

Thus, by definition, for $\gamma_1 + \gamma_2 = \gamma$, we have

$$[\nabla g_s(\cdot, z)]_{\gamma_1} \leqslant (2\|\nabla g_s(\cdot, z)\|_\infty)^{\gamma_2/\gamma} [\nabla g_s(\cdot, z)]_\gamma^{\gamma_1/\gamma} \leqslant C(1 \wedge |z|^{\gamma_2}),$$

which in turn gives the first estimate in (3.10). The second inequality in (3.10) follows from (3.5) and the definition of g^n. Property (3.11) follows from (3.7), (3.8), and the definition of g^n. □

Taking $\gamma_1 = 0$ in Lemma 3.2 yields that there is a $C_0 = C_0(d, \gamma) > 0$ so that

$$\|\nabla g^n(\cdot, z)\|_\infty \leqslant C_0(1 \wedge |z|^\gamma) \quad \text{and} \quad |g^n_s(y,z)| \leqslant 3|z|/2. \tag{3.12}$$

Choose $r_0 \in (0,1)$ so that

$$C_0 r_0^\gamma + 3\,r_0/2 < 1. \tag{3.13}$$

Such a choice of r_0 will be used below to establish the C^1-stochastic diffeomorphic property of the unique solution Y^n of SDE (3.18) below. For any given $n \in \mathbb{N}_\infty$, define

(3.14)

$$a^n_s(y) := \lambda u^n_s\big((\Phi^n_s)^{-1}(y)\big) + \eta_{r_0} - \int_{|z|\geqslant r_0} \big(u^n_s\big((\Phi^n_s)^{-1}(y) + z\big) - u^n_s\big((\Phi^n_s)^{-1}(y)\big)\big)\,\nu(\mathrm{d}z).$$

We have:

Lemma 3.3. *Under the assumptions stated at the beginning of this section, there is a positive constant* $C_2 = C_2\,(d,\lambda,\gamma,r_0,\nu(|z|\geqslant r_0))$ *such that for all* $n\in\mathbb{N}_\infty$, $t\in[0,1]$ *and* $y\in\mathbb{R}^d$,

$$\|\nabla a^n\|_{\infty,\gamma}\leqslant C_2 \quad \textit{and} \quad |a^n_s(y)|\leqslant C_2(1+\|b\|_\infty). \tag{3.15}$$

Moreover, for each $t\in[0,1]$, $y\in\mathbb{R}^d$ *and* $j=0,1$, *we have*

$$\lim_{n\to\infty}\nabla^j_y a^n_t(y)=\nabla^j_y a^\infty_t(y). \tag{3.16}$$

Proof. For notational simplicity, we drop the superscript "n". Since

$$\nabla(u_s(\Phi_s^{-1}))=(\nabla u_s)(\Phi_s^{-1})\cdot\nabla\Phi_s^{-1},$$

we have by (3.6) that for all $s\in[0,1]$,

$$\begin{aligned}\|\nabla(u_s(\Phi_s^{-1}))\|_\gamma &\leqslant \|\nabla u_s(\Phi_s^{-1})\|_\gamma\|\nabla\Phi_s^{-1}\|_\infty+\|\nabla u_s\|_\infty\|\nabla\Phi_s^{-1}\|_\gamma\\ &\leqslant \|\nabla u_s\|_\gamma\|\nabla\Phi_s^{-1}\|_\infty(1+\|\Phi_s^{-1}\|_\infty^\gamma)+\|\nabla u_s\|_\infty\|\nabla\Phi_s^{-1}\|_\gamma\leqslant C,\end{aligned}$$

where the constant C only depends on d and γ. Similarly, we have

$$\|\nabla(u_s(\Phi_s^{-1}(\cdot)+z))\|_\gamma\leqslant C.$$

Hence,

$$\left\|\nabla\int_{|z|\geqslant r_0}\big(u_s\big(\Phi_s^{-1}(\cdot)+z\big)-u_s(\Phi_s^{-1}(\cdot))\big)\,\nu(\mathrm{d}z)\right\|_\gamma\leqslant C\cdot\nu(|z|\geqslant r_0).$$

Therefore $\|\nabla a\|_{\infty,\gamma}\leqslant C_2$ by (3.14). The second inequality in (3.15) follows from the definition of $a_s(y)$ and the fact that u^n is uniformly bounded due to (2.10). Property (3.16) follows from (3.2), (3.3), (3.7), (3.8), and the definition of a^n. □

Now recalling the definitions of random measures N and $\tilde N$ associated with Z in the introduction, we can present the following Zvonkin's transformation by Itô's formula.

Lemma 3.4. *Suppose that the assumptions stated at the beginning of this section hold. Let* $\Phi^n_t(x)$ *be defined as in* (3.4). *For* $n\in\mathbb{N}_\infty$, X^n_t *satisfies*

$$X^n_t=x+\int_0^t b^n_s(X^n_s)\,\mathrm{d}s+Z_t,\quad t\in[0,1] \tag{3.17}$$

if and only if $Y^n_t=\Phi^n_t(X^n_t)$ *solves the following SDE for* $t\in[0,1]$:

$$\begin{aligned}Y^n_t=\Phi^n_0(x)&+\int_0^t a^n_s(Y^n_s)\,\mathrm{d}s+\int_0^t\int_{|z|<r_0}g^n_s(Y^n_{s-},z)\tilde N(\mathrm{d}s,\mathrm{d}z)\\ &+\int_0^t\int_{|z|\geqslant r_0}g^n_s(Y^n_{s-},z)N(\mathrm{d}s,\mathrm{d}z),\end{aligned} \tag{3.18}$$

where a^n *and* g^n *are defined by* (3.14) *and* (3.9).

Proof. For $n \in \mathbb{N}$, since $x \mapsto \Phi_t^n(x)$ and $x \mapsto (\Phi_t^n)^{-1}(x)$ are smooth, the assertion of this lemma follows from Itô's formula as calculated in the introduction (see (1.5)). For $n = \infty$, since we only have $\|\nabla\Phi^\infty\|_{\infty,\gamma} < \infty$, one needs suitable mollifying technique. This is standard and can be found in [18] and [28]. We omit the details. □

Lemma 3.5. *Suppose that the assumptions stated at the beginning of this section hold. For $n \in \mathbb{N}_\infty$, let $Y_t^n(x)$ be the solution of (3.18) with initial value $\Phi_0^n(x)$. We have*

$$\lim_{n\to\infty} \mathbb{E}\Big[\sup_{t\in[0,1]} |Y_t^n(x) - Y_t^\infty(x)| \wedge 1\Big] = 0. \tag{3.19}$$

Moreover, for any $p > 1$, we have

$$\sup_{n\in\mathbb{N}_\infty} \sup_{x\in\mathbb{R}^d} \mathbb{E}\Big[\sup_{t\in[0,1]} |\nabla Y_t^n(x)|^p\Big] < \infty, \tag{3.20}$$

and for each $x \in \mathbb{R}^d$,

$$\lim_{n\to\infty} \mathbb{E}\Big[\sup_{t\in[0,1]} |\nabla Y_t^n(x) - \nabla Y_t^\infty(x)|^p\Big] = 0. \tag{3.21}$$

Proof. (3.19) follows from Lemmas 3.2, 3.3 and Proposition 5.1 below. In this proof, we shall drop the superscript "∞". Notice that

$$\begin{aligned}\nabla Y_t^n = \nabla\Phi_0^n(x) + \int_0^t \nabla a_s^n(Y_s^n)\nabla Y_s^n \mathrm{d}s + \int_0^t \int_{|z|<r_0} \nabla_y g_s^n(Y_{s-}^n, z)\nabla Y_{s-}^n \tilde{N}(\mathrm{d}s, \mathrm{d}z)\\ + \int_0^t \int_{|z|\geqslant r_0} \nabla_y g_s^n(Y_{s-}^n, z)\nabla Y_{s-}^n N(\mathrm{d}s, \mathrm{d}z).\end{aligned} \tag{3.22}$$

By the Burkholder–Davis–Gundy inequality (Theorem 2.11 in [13]), and (3.10) and (3.15), we have for $p \geqslant 2$,

$$\begin{aligned}\mathbb{E}\Big[\sup_{s\in[0,t]} |\nabla Y_s^n|^p\Big] &\preceq |\nabla\Phi_0^n(x)|^p + \int_0^t \mathbb{E}|\nabla a_s^n(Y_s^n)\nabla Y_s^n|^p \,\mathrm{d}s\\ &\quad + \mathbb{E}\Big[\int_0^t \int_{|z|<r_0} |\nabla_y g_s^n(Y_s^n, z)\nabla Y_s^n|^2 \,\nu(\mathrm{d}z)\,\mathrm{d}s\Big]^{p/2}\\ &\quad + \mathbb{E}\Big[\int_0^t \int_{|z|\geqslant r_0} |\nabla_y g_s^n(Y_s^n, z)\nabla Y_s^n| \,\nu(\mathrm{d}z)\,\mathrm{d}s\Big]^{p}\\ &\quad + \mathbb{E}\Big[\int_0^t \int_{\mathbb{R}^d} |\nabla_y g_s^n(Y_s^n, z)\nabla Y_s^n|^p \,\nu(\mathrm{d}z)\,\mathrm{d}s\Big]\\ &\preceq 1 + \Big(1 + \Big(\int_{|z|<r_0} |z|^{2\gamma}\,\nu(\mathrm{d}z)\Big)^{p/2}\Big) \int_0^t \mathbb{E}|\nabla Y_s^n|^p \,\mathrm{d}s,\end{aligned}$$

which gives (3.20) by Gronwall's inequality.

Next, set $U_t^n := \nabla Y_t^n - \nabla Y_t$. By equations (3.22), (3.10), (3.15) and Theorem 2.11 in [13], we have

$$\mathbb{E}\Big[\sup_{s\in[0,t]}|U_s^n|^p\Big] \preceq h_n + \int_0^t \mathbb{E}|U_s^n|^p \mathrm{d}s,$$

where

$$\begin{aligned} h_n &:= |\nabla\Phi_0^n(x) - \nabla\Phi_0(x)|^p + \int_0^1 (\mathbb{E}|\nabla a_s^n(Y_s^n) - \nabla a_s(Y_s)|^{2p})^{1/2}\mathrm{d}s \\ &+ \bigg(\mathbb{E}\bigg[\int_0^1\int_{|z|<r_0} |\nabla_y g_s^n(Y_s^n,z) - \nabla_y g_s(Y_s,z)|^2\,\nu(\mathrm{d}z)\,\mathrm{d}s\bigg]^p\bigg)^{1/2} \\ &+ \bigg(\mathbb{E}\bigg[\int_0^1\int_{|z|\geqslant r_0} |\nabla_y g_s^n(Y_s^n,z) - \nabla_y g_s(Y_s,z)|\,\nu(\mathrm{d}z)\,\mathrm{d}s\bigg]^{2p}\bigg)^{1/2} \\ &+ \bigg(\mathbb{E}\bigg[\int_0^1\int_{\mathbb{R}^d} |\nabla_y g_s^n(Y_s^n,z) - \nabla_y g_s(Y_s,z)|^p\,\nu(\mathrm{d}z)\,\mathrm{d}s\bigg]^2\bigg)^{1/2}. \end{aligned}$$

By Gronwall's inequality, (3.10), (3.11), (3.15), (3.16) and (3.19), it is easy to see that

$$\lim_{n\to 0}\mathbb{E}\Big[\sup_{t\in[0,1]}|U_t^n|^p\Big] \preceq \lim_{n\to 0} h_n = 0.$$

The proof is complete. □

We are now in a position to give a:

Proof of Theorem 1.2. Let $a = a^\infty$ and $g = g^\infty$ be defined by (3.14) and (3.9), respectively. By Lemmas 3.2 and 3.3, we have

$$|a_s(y) - a_s(y')| \leqslant C_1|y - y'|$$

and

$$\int_{|z|\leqslant r_0} |g_s(y,z) - g_s(y',z)|^2\nu(\mathrm{d}z) \leqslant C_2^2|y-y'|^2\int_{|z|\leqslant r_0}|z|^{2\gamma}\nu(\mathrm{d}z) \overset{(1.9)}{\leqslant} C|y-y'|^2.$$

Hence, by Theorem IV.9.1 in [10], (3.18) has a unique strong solution. (1.11) follows from $X_t(x) = \Phi_t^{-1}(Y_t(\Phi_0(x)))$ and (3.20). Moreover, let $Y_t(y)$ be the solution of SDE (3.18) with starting point y. By (3.12) and the choice of r_0 in (3.13), $\{Y_t(y), t\in[0,1], y\in\mathbb{R}^d\}$ defines a C^1-stochastic diffeomorphism flow (cf. [18], p. 442–445), so does $\{X_t(x), t\in[0,1], x\in\mathbb{R}^d\}$. Next we show that $t\mapsto\nabla X_t(x)$ is continuous. Let $X_t^n(x)$ satisfy (3.17). Clearly, $t\mapsto\nabla X_t^n(x)$ is continuous for each $n\in\mathbb{N}$. On the other hand, by Lemma 3.5 and (3.8), we also have

$$\lim_{n\to\infty}\mathbb{E}\Big[\sup_{t\in[0,1]}|\nabla X_t^n(x) - \nabla X_t(x)|^p\Big] = 0. \tag{3.23}$$

From this, we immediately obtain the desired continuity. □

Proof of Theorem 1.6. First of all, we show that the right-hand side of (1.14) is no bigger than the right-hand side of (1.15). By Hölder's inequality, it suffices to show that for any $p > 1$,

$$I(t) := \mathbb{E}\Big[\frac{1}{S_t^p}\Big|\int_0^t \nabla X_s(x)\,\mathrm{d}W_{S_s}\Big|^p\Big] \leqslant C\,t^{-p/\alpha}.$$

By [29, (2.11)], one has

$$I(t) \preceq \mathbb{E}\Big[\frac{1}{S_t^p}\Big(\int_0^t |\nabla X_s(x)|^2 \mathrm{d}S_s\Big)^{p/2}\Big] \leqslant \mathbb{E}\Big[\frac{1}{S_t^{p/2}}\sup_{s\in[0,1]}|\nabla X_s(x)|^p\Big]$$
$$\overset{(1.11)}{\preceq} \left(\mathbb{E}\left[S_t^{-p}\right]\right)^{1/2} \overset{(4.3)}{\preceq} t^{-p/\alpha}.$$

Let b^n be defined as in (2.22) and X^n be the unique solution to SDE (3.17). For $f \in C_b^1(\mathbb{R}^d)$, by Theorem 1.1 in [29] or Theorem 1.1 in [25], we have

$$\nabla \mathbb{E} f(X_t^n(x)) = \mathbb{E}\Big[\frac{f(X_t^n(x))}{S_t}\int_0^t \nabla X_s^n(x)\,\mathrm{d}W_{S_s}\Big], \quad n \in \mathbb{N}.$$

Thus, in order to prove (1.14), it suffices to show the following two relations:

$$\begin{aligned}\lim_{n\to\infty} \nabla \mathbb{E} f(X_t^n(x)) &= \lim_{n\to\infty} \mathbb{E}\big[(\nabla f)(X_t^n(x))\nabla X_t^n(x)\big]\\ &= \mathbb{E}\Big[(\nabla f)(X_t(x))\nabla X_t(x)\Big] = \nabla \mathbb{E} f(X_t(x))\end{aligned} \tag{3.24}$$

and

$$\lim_{n\to\infty} \mathbb{E}\Big[\frac{f(X_t^n(x))}{S_t}\int_0^t \nabla X_s^n(x)\,\mathrm{d}W_{S_s}\Big] = \mathbb{E}\Big[\frac{f(X_t(x))}{S_t}\int_0^t \nabla X_s(x)\,\mathrm{d}W_{S_s}\Big]. \tag{3.25}$$

Notice that by (3.19) and (3.8),

$$\lim_{n\to\infty} \mathbb{E}\left[|X_t^n(x) - X_t(x)| \wedge 1\right] = 0. \tag{3.26}$$

The relations (3.24) and (3.25) follow by (3.23), (3.26) and the dominated convergence theorem. □

4. Examples

Now we give some examples for which the assumptions of Theorem 1.2 are satisfied.

Example 4.1 (Subordinate Brownian motions)**.** Let $Z_t := W_{S_t}$, where W is a Brownian motion in $\mathbb{R}^d$ with infinitesimal generator $\Delta/2$ and S is a subordinator (i.e., an increasing real-valued Lévy process starting from 0), which is independent of W_t. The process Z defined above is called a subordinate Brownian motion, for some basic properties of subordinate Brownian motion one refers to, for example, Chapter 5 in [2]. Let $\phi(\lambda)$ be the Laplace exponent of S, i.e., $\mathbb{E}\mathrm{e}^{-\lambda S_t} = \mathrm{e}^{-t\phi(\lambda)}$. If for some $\alpha \in (0,2)$,

$$\phi(\lambda) \geqslant C\lambda^{\alpha/2}, \quad \lambda \geqslant 1, \tag{4.1}$$

then $(\mathbf{H}^{\alpha,1,1}_{\nu,K_0})$ holds for some $K_0 > 0$. Indeed, using the independence of S and W, one can easily check that for any bounded Borel function f on $\mathbb{R}^d$,

$$\nabla T^{\nu,0}_t f(x) = \mathbb{E}\Big[f(x + W_{S_t})\frac{W_{S_t}}{S_t}\Big].$$

Thus, if, for some $\beta \in (0,1)$, $\Lambda_x := \sup_{y\in\mathbb{R}^d} |f(x+y) - f(x)|/|y|^\beta < \infty$, then

$$\begin{aligned} |\nabla T^{\nu,0}_t f(x)| &= \Big|\mathbb{E}\Big[\Big(f(x + W_{S_t}) - f(x)\Big)\frac{W_{S_t}}{S_t}\Big]\Big| \\ &\leqslant \Lambda_x\,\mathbb{E}\Big[\frac{|W_{S_t}|^{1+\beta}}{S_t}\Big] \leqslant C\,\Lambda_x\,\mathbb{E}\big[S_t^{-(1-\beta)/2}\big] \leqslant K_0\,\Lambda_x\,t^{(\beta-1)/\alpha}, \end{aligned} \tag{4.2}$$

where the last step is due to the fact that, for any $p \in (0,1)$,

$$\begin{aligned} \mathbb{E}S_t^{-p} &= \frac{1}{\Gamma(p)}\,\mathbb{E}\int_0^\infty \lambda^{p-1}\,\mathrm{e}^{-\lambda S_t}\,\mathrm{d}\lambda = \frac{1}{\Gamma(p)}\int_0^\infty \lambda^{p-1}\,\mathrm{e}^{-t\phi(\lambda)}\,\mathrm{d}\lambda \\ &\overset{(4.1)}{\leqslant} \frac{1}{\Gamma(p)}\left(\frac{1}{p} + \int_1^\infty \lambda^{p-1}\,\mathrm{e}^{-C\,t\lambda^{\alpha/2}}\,\mathrm{d}\lambda\right) \leqslant C\,t^{-2p/\alpha}, \quad t \in (0,1]. \end{aligned} \tag{4.3}$$

The constant C can be chosen to be independent of $p \in (0,1)$ so that the constant K_0 in (4.2) is independent of $\beta \in (0,1)$. Moreover, it follows from (15) in [3] that

$$\nu(\mathrm{d}z) \leqslant \frac{c_0\,\phi(|z|^{-2})}{|z|^d}\,\mathrm{d}z.$$

Thus if there exists $\tilde\alpha \in (0,2)$ such that

$$\phi(\lambda) \leqslant C\,\lambda^{\tilde\alpha/2} \quad \text{for } \lambda \geqslant 1, \tag{4.4}$$

then (1.9) is satisfied for any $\gamma \in (\tilde\alpha/2, 1]$. This implies that we need to take $\beta \in (\tilde\alpha/2 + 1 - \alpha, 1]$ in Theorem 1.2.

There are many examples of subordinate Brownian motions satisfying (4.1) and (4.4). One important example is the symmetric relativistic α-stable process in $\mathbb{R}^d$. (For some basic information on symmetric relativistic α-stable processes, see, for instance, Chapter 5 in [2] or [4].) In this case, $\phi(\lambda) = (\lambda + m^{2/\alpha})^{\alpha/2} - m$ for some $m > 0$, (4.1) holds and (4.4) is satisfied with $\tilde\alpha = \alpha$. This implies that in this case we can take any $\beta \in (1 - \alpha/2, 1]$ in Theorem 1.2.

Example 4.2 (Stable-type Lévy processes)**.** Let Z be a Lévy process in $\mathbb{R}^d$ whose Lévy measure $\nu(\mathrm{d}z) = \kappa(z)\,\mathrm{d}z$. Assume that for some $0 < \alpha_1 \leqslant \alpha_2 < 2$,

$$c_1\,|z|^{-d-\alpha_1} \leqslant \kappa(z) \leqslant c_2\,|z|^{-d-\alpha_2} \quad \text{for } |z| \leqslant 1. \tag{4.5}$$

We say that a Lévy process satisfying the above condition is of stable-type. In this case, we can make the following decomposition for ν:

$$\nu = \nu_0 + \nu_1 + \nu_2$$

with $\nu_0(\mathrm{d}z) := -c_1|z|^{-d-\alpha_1}\mathbf{1}_{\{|z|>1\}}\mathrm{d}z$ and

$$\nu_1(\mathrm{d}z) := c_1|z|^{-d-\alpha_1}\mathrm{d}z, \quad \nu_2(\mathrm{d}z) := (\kappa(z)-c_1|z|^{-d-\alpha_1})\mathbf{1}_{\{|z|\leqslant 1\}}\mathrm{d}z+\kappa(z)\mathbf{1}_{\{|z|>1\}}\mathrm{d}z.$$

By Example 4.1, $(\mathbf{H}^{\alpha_1,1,1}_{\nu_1,K_0})$ holds for some $K_0 > 0$. Condition (1.9) holds for any $\gamma \in (\alpha_2/2, 1]$. This implies that in this case we need to take $\beta \in (\alpha_2/2+1-\alpha_1, 1]$ in Theorem 1.2. One particular example is the case when $\alpha_1 = \alpha_2 = \alpha$ and the relation in (4.5) is satisfied for all $z \in \mathbb{R}^d$. The corresponding Lévy process is called an α-stable-like Lévy process. Another particular example is the case when $\kappa(z) = 0$ for $|z| > 1$ and $\alpha_1 = \alpha_2 = \alpha$. The corresponding Lévy process is called a truncated α-stable-like Lévy process. Observe that the relativistic α-stable process satisfies condition (4.5) with $\alpha_1 = \alpha_2 = \alpha$. The third particular example is the case where $\kappa(z)$ is comparable to the Lévy kernel of the relativistic α-stable process. The corresponding Lévy process can be called relativistic α-stable-like.

Example 4.3 (Cylindrical stable processes)**.** In this example we consider a cylindrical stable process $Z = (Z^1, \ldots, Z^k)$ in $\mathbb{R}^d$, where Z^j, $1 \leqslant j \leqslant k$, are independent d_j-dimensional rotationally symmetric α_j-stable processes with $\alpha_j \in (0, 2)$ and $\sum_{j=1}^k d_j = d$. We can realize Z as follows:

$$Z_t = W_{S_t} := \big(W^1_{S^1_t}, \ldots, W^k_{S^k_t}\big),$$

where W^j, $1 \leqslant j \leqslant k$, are independent d_j-dimensional standard Brownian motions with infinitesimal generator $\Delta/2$ in $\mathbb{R}^{d_j}$ and S^j, $1 \leqslant j \leqslant k$, are independent $\alpha_j/2$-stable subordinators with $\alpha_j \in (0, 2)$ for $1 \leqslant j \leqslant k$, that are also independent of Brownian motions $\{W^1, \ldots, W^k\}$. Define

$$\alpha := \min_{1\leqslant j\leqslant k} \alpha_j \quad \text{and} \quad \alpha_{\max} := \max_{1\leqslant j\leqslant k} \alpha_j.$$

We claim that if $\alpha \in (0, 1]$, then $(\mathbf{H}^{\alpha,\alpha,\delta}_{\nu,K_0})$ holds with some $K_0 > 0$ and $\delta := \alpha/\alpha_{\max}$; and if $\alpha \in (1, 2)$, then $(\mathbf{H}^{\alpha}_{\nu,K_0})$ holds for some $K_0 > 0$.

Indeed, for $1 \leqslant i \leqslant k$, let $\nabla_i = (\partial_{x_{j_i+1}}, \ldots, \partial_{x_{j_i+d_i}})$, where $j_i := d_0 + \cdots + d_{i-1}$ with $d_0 := 0$. As in Example 4.1, we also have the following derivative formula for any bounded Borel function f on $\mathbb{R}^d$:

$$\nabla_i T^{\nu,0}_t f(x) = \mathbb{E}\big[(S^i_t)^{-1} W^i_{S^i_t} f(x + W_{S_t})\big].$$

Suppose $\alpha \in (0, 1]$. For $\beta \in [0, \alpha]$ and $x \in \mathbb{R}$, if

$$\Lambda_x := \sup_{y\in\mathbb{R}^d} \frac{|f(x+y) - f(x)|}{|y|^\beta} < \infty,$$

then we have by (4.3) that for $t \in (0,1]$,

$$\begin{aligned}|\nabla_i T_t^{\nu,0} f(x)| &= \big|\mathbb{E}\big[(S_t^i)^{-1} W^i_{S_t^i}\big(f(x+W_{S_t}) - f(x)\big)\big]\big| \leqslant \Lambda_x\, \mathbb{E}\big[(S_t^i)^{-1}|W^i_{S_t^i}|\,|W_{S_t}|^\beta\big] \\ &\leqslant \Lambda_x \Big(\mathbb{E}\big[(S_t^i)^{-1}|W^i_{S_t^i}|^{1+\beta}\big] + \mathbb{E}\big[(S_t^i)^{-1}|W^i_{S_t^i}|\big] \sum_{j\neq i} \mathbb{E}\big[|W^j_{S_t^j}|^\beta\big]\Big) \\ &\leqslant C\,\Lambda_x \Big(t^{(\beta-1)/\alpha_i} + t^{-1/\alpha_i} \sum_{j\neq i} t^{\beta/\alpha_j}\, \mathbb{E}\big[|W^j_{S_1^j}|^\beta\big]\Big) \\ &\leqslant K_0\,\Lambda_x \big(t^{(\beta-1)/\alpha} + t^{\beta/\alpha_{\max}-1/\alpha}\big) \qquad (\text{since } \beta < \alpha) \\ &\leqslant K_0\,\Lambda_x\, t^{(\alpha\beta/\alpha_{\max}-1)/\alpha} = K_0\,\Lambda_x\, t^{(\delta\beta-1)/\alpha};\end{aligned}$$

that is, $(\mathbf{H}^{\alpha,\alpha,\delta}_{\nu,K_0})$ holds. If $\alpha \in (1,2)$, then we have by (4.3) that for $t \in (0,1]$,

$$\begin{aligned}|\nabla_i T_t^{\nu,0} f(x)| &\leqslant \|f\|_\infty\, \mathbb{E}[|W^i_{S_t^i}|/(S_t^i)] \\ &\preceq \|f\|_\infty\, \mathbb{E}[(S_t^i)^{-1/2}] \preceq \|f\|_\infty t^{-1/\alpha_i} \leqslant K_0\, \|f\|_\infty t^{-1/\alpha}.\end{aligned}$$

Thus in this case, $(\mathbf{H}^{\alpha}_{\nu,K_0})$ holds. The claim is now verified.

It is not difficult to see by using the property of the rotationally symmetric α_j-stable process $W^j_{S_j}$ that the parameter α in the now verified property $(\mathbf{H}^{\alpha,\alpha,\delta}_{\nu,K_0})$ and $(\mathbf{H}^{\alpha}_{\nu,K_0})$ is best possible. For example, it can be shown that when $\alpha \in (1,2)$, property $(\mathbf{H}^{\alpha^*}_{\nu,K_0})$ fails for any $\alpha^* > \alpha$.

Note that (1.9) holds for any $\gamma > \alpha_{\max}/2$. For Theorem 1.2 to be valid, the following constraint needs to be satisfied:

$$1 \geqslant \beta > \alpha_{\max}/2 + \alpha_{\max}(1-\alpha)/\alpha \;\text{ if } \alpha \leqslant 1, \quad \text{and} \quad \alpha_{\max} < 2\alpha \;\text{ if } \alpha > 1.$$

Clearly, when $\alpha > 1$, the condition $\alpha_{\max} < 2\alpha$ is automatically satisfied. Consequently, in this case for Theorem 1.2 to be applicable, we need α_i's to satisfy

$$\text{either} \quad \alpha > 1 \quad \text{or} \quad \alpha \in (0,1] \;\text{ and }\; \alpha_{\max} < 2\alpha^2/(2-\alpha), \tag{4.6}$$

and take

$$\beta \in (\beta_0, 1] \;\text{ with } \beta_0 := \alpha_{\max}/2 + (\alpha_{\max}/\alpha \mathbf{1}_{\{\alpha\leqslant 1\}} + \mathbf{1}_{\{\alpha>1\}})(1-\alpha).$$

Condition (4.6) implies that $\alpha > 2/3$. An open question is whether constraint (4.6) can be dropped. It boils down to the question whether $(\mathbf{H}^{\alpha,1,1}_{\nu,K_0})$ holds for any cylindrical stable process.

This example can be extended in two directions. First, as in Example 4.1, we can consider more general subordinators $\{S^1,\dots,S^k\}$. Second, as in Example 4.2, we can consider more general Lévy process, whose Lévy measure is bounded by the Lévy measure of the cylindrical α-stable process W_S (or, more generally, the cylindrical subordinate Brownian motion) from below.

Proof of Corollary 1.4. It follows from Examples 4.1, 4.2 and 4.3. □

5. Appendix

In this appendix, we prove the continuous dependence of solutions to SDEs with jumps with respect to the initial values and coefficients, which is used in the proof of Lemma 3.5.

Proposition 5.1. *Fix $r > 0$. Let $a^n, g^n, n \in \mathbb{N}_\infty$ be two families of uniformly Lipschitz continuous functions in the sense that for some $C > 0$, and all $n \in \mathbb{N}_\infty$ and $t \in [0,1], x, y, z \in \mathbb{R}^d$,*

$$|a^n_t(x) - a^n_t(y)| \leqslant C\,|x-y|, \quad |g^n_t(x,z) - g^n_t(y,z)| \leqslant C\,|x-y|\,h(z), \tag{5.1}$$

where $\int_{|z|\leqslant r} |h(z)|^2 \nu(\mathrm{d}z) < \infty$. Suppose that for each $t \in [0,1]$ and $x, z \in \mathbb{R}^d$,

$$\lim_{n\to\infty} a^n_t(x) = a^\infty_t(x), \quad \lim_{n\to\infty} g^n_t(x,z) = g^\infty_t(x,z) \tag{5.2}$$

and

$$\sup_{n\in\mathbb{N}_\infty} \sup_{t\in[0,1]} \sup_{x\in\mathbb{R}^d} \left(\frac{|a^n_t(x)|}{1+|x|} + \sup_{0<|z|\leqslant r} \frac{|g^n_t(x,z)|}{|z|} \right) < \infty. \tag{5.3}$$

For $n \in \mathbb{N}_\infty$, let Y^n_t be the solution to the following SDE:

$$Y^n_t = \xi_n + \int_0^t a^n_s(Y^n_s)\,\mathrm{d}s + \int_0^t \int_{|z|\leqslant r} g^n_s(Y^n_{s-}, z)\,\tilde{N}(\mathrm{d}s, \mathrm{d}z) + \int_0^t \int_{|z|>r} g^n_s(Y^n_{s-}, z)\,N(\mathrm{d}s, \mathrm{d}z).$$

If ξ_n converges to ξ_∞ in probability as $n \to \infty$, then

$$\lim_{n\to\infty} \mathbb{E}\Big(\sup_{t\in[0,1]} |Y^n_t - Y^\infty_t| \wedge 1 \Big) = 0, \tag{5.4}$$

which implies that Y^n_t converges to Y^∞_t in probability.

We begin with the following lemma.

Lemma 5.2. *There is a nonnegative smooth function f on $\mathbb{R}^d$ with the following properties:*

$$\begin{aligned} &f(x) = |x|^2 \ \ \text{if } |x| \leqslant 1, \quad f(x) = 2 \ \ \text{if } |x| \geqslant 2, \\ &\text{and} \quad |\nabla f| + |\nabla^2 f| \leqslant C_1 \mathbf{1}_{\{|x|\leqslant 2\}}, \end{aligned} \tag{5.5}$$

for some constant $C_1 > 0$, and that for any constant $C_2 > 0$, there exists a constant $C_3 > 0$ such that for all $\delta > 0$, $r \in [0,1]$ and $|y| \leqslant C_2((|x|+\delta) \wedge 1)$,

$$|y|\,|\nabla f(x+ry)| \leqslant C_3(f(x)+\delta), \quad |y|^2\,|\nabla^2 f(x+ry)| \leqslant C_3(f(x)+\delta^2). \tag{5.6}$$

Proof. Let ϕ be an increasing smooth function on $(0,\infty)$ with $\phi(r)=r$ for $r\leqslant 1$ and $\phi(r)=2$ for $r\geqslant 4$. Let $f(x):=\phi(|x|^2)$. It is easy to check that f has the desired properties. □

We also need the following key lemma.

Lemma 5.3. *Let τ_1 and τ_2 be two stopping times with $0\leqslant\tau_1\leqslant\tau_2\leqslant 1$. In the setup of Proposition 5.1, let Y^n solve the following SDE on $[\tau_1,\tau_2]$:*

$$Y^n_t=Y^n_{\tau_1}+\int_{\tau_1}^t a^n_s(Y^n_s)\,\mathrm{d}s+\int_{\tau_1}^t\int_{|z|\leqslant r} g^n_s(Y^n_{s-},z)\,\tilde N(\mathrm{d}s,\mathrm{d}z).$$

If $Y^n_{\tau_1}$ converges to $Y^\infty_{\tau_1}$ in probability, then

$$\lim_{n\to\infty}\mathbb{E}\Big[\sup_{t\in[\tau_1,\tau_2]}|Y^n_t-Y^\infty_t|\wedge 1\Big]=0.$$

Proof. In this proof we will drop the superscript "∞" and write

$$U^n_s:=Y^n_s-Y_s,\quad A^n_s:=a^n_s(Y^n_s)-a_s(Y_s),\quad \Gamma^n_s(z):=g^n_s(Y^n_{s-},z)-g_s(Y_{s-},z).$$

Let f be as in Lemma 5.2. By Itô's formula, we have

$$\begin{aligned}
&f(U^n_t)\\
&=f(U^n_{\tau_1})+\int_{\tau_1}^t\langle A^n_s,\nabla f(U^n_s)\rangle\,\mathrm{d}s+\int_{\tau_1}^t\int_{|z|\leqslant r}[f(U^n_{s-}+\Gamma^n_s(z))-f(U^n_{s-})]\tilde N(\mathrm{d}s,\mathrm{d}z)\\
&\quad+\int_{\tau_1}^t\int_{|z|\leqslant r}[f(U^n_{s-}+\Gamma^n_s(z))-f(U^n_{s-})-\Gamma^n_s(z)\cdot\nabla f(U^n_{s-})]\,\nu(\mathrm{d}z)\,\mathrm{d}s.
\end{aligned}$$

For $R>0$, define a stopping time

$$\tau_R:=\inf\{t\geqslant\tau_1:|Y_s|>R\}\wedge\tau_2.$$

For any $T\in[0,1]$, by the Burkholder–Davis–Gundy inequality ([13], Theorem 2.11), we have

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{t\in[\tau_1,T\wedge\tau_R]}|f(U^n_t)|^2\Big)\\
&\preceq\mathbb{E}|f(U^n_{\tau_1})|^2+\mathbb{E}\int_{\tau_1}^{T\wedge\tau_R}|\langle A^n_s,\nabla f(U^n_s)\rangle|^2\,\mathrm{d}s\\
&\quad+\mathbb{E}\int_{\tau_1}^{T\wedge\tau_R}\int_{|z|\leqslant r}|f(U^n_{s-}+\Gamma^n_s(z))-f(U^n_{s-})|^2\,\nu(\mathrm{d}z)\,\mathrm{d}s\\
&\quad+\mathbb{E}\int_{\tau_1}^{T\wedge\tau_R}\Big|\int_{|z|\leqslant r}[f(U^n_{s-}+\Gamma^n_s(z))-f(U^n_{s-})-\Gamma^n_s(z)\cdot\nabla f(U^n_{s-})]\nu(\mathrm{d}z)\Big|^2\,\mathrm{d}s\\
&=:\mathbb{E}|f(U^n_{\tau_1})|^2+I^n_1+I^n_2+I^n_3.
\end{aligned}$$

For I_1^n, by (5.1) and (5.5), we have

$$I_1^n \preceq \mathbb{E}\int_{\tau_1}^{T\wedge\tau_R} |a_s^n(Y_s)-a_s(Y_s)|^2\,\mathrm{d}s + \mathbb{E}\int_{\tau_1}^{T\wedge\tau_R} |f(U_s^n)|^2\,\mathrm{d}s.$$

For I_2^n and I_3^n, we note that by (5.1) and (5.3), we have

$$|\Gamma_s^n(z)| \leqslant C\,((|U_{s-}^n| + |g_s^n(Y_{s-},z)-g_s(Y_{s-},z)|)\wedge 1), \quad |z|\leqslant r,$$

Thus by (5.6), we have

$$\begin{aligned} I_2^n &\preceq \mathbb{E}\int_{\tau_1}^{T\wedge\tau_R}\int_{|z|\leqslant r} |\Gamma_s^n(z)|^2\Big(\int_0^1 |\nabla f(U_{s-}^n + r\Gamma_s^n(z))|^2\mathrm{d}r\Big)\nu(\mathrm{d}z)\,\mathrm{d}s \\ &\preceq \mathbb{E}\int_{\tau_1}^{T\wedge\tau_R}\int_{|z|\leqslant r} |g_s^n(Y_s,z)-g_s(Y_s,z)|^2\,\nu(\mathrm{d}z)\,\mathrm{d}s + \mathbb{E}\int_{\tau_1}^{T\wedge\tau_R} |f(U_s^n)|^2\mathrm{d}s \end{aligned}$$

and

$$\begin{aligned} I_3^n &\preceq \mathbb{E}\int_{\tau_1}^{T\wedge\tau_R}\Big(\int_{|z|\leqslant r} |\Gamma_s^n(z)|^2\Big(\int_0^1\int_0^1 |\nabla^2 f(U_{s-}^n + rr'\Gamma_s^n(z))|\mathrm{d}r\mathrm{d}r'\Big)\nu(\mathrm{d}z)\Big)^2\mathrm{d}s \\ &\preceq \mathbb{E}\int_{\tau_1}^{T\wedge\tau_R}\Big(\int_{|z|\leqslant r} |g_s^n(Y_s,z)-g_s(Y_s,z)|^2\nu(\mathrm{d}z)\Big)^2\mathrm{d}s + \mathbb{E}\int_{\tau_1}^{T\wedge\tau_R} |f(U_s^n)|^2\mathrm{d}s. \end{aligned}$$

Combining the above calculations, we obtain

$$\mathbb{E}\Big[\sup_{t\in[\tau_1,T\wedge\tau_R]} |f(U_t^n)|^2\Big] \preceq h_n + \mathbb{E}\int_{\tau_1}^{T\wedge\tau_R} |f(U_s^n)|^2\,\mathrm{d}s,$$

where

$$\begin{aligned} h_n := \mathbb{E}\left[|f(U_{\tau_1}^n)|^2\right] &+ \mathbb{E}\int_{\tau_1}^{T\wedge\tau_R} |a_s^n(Y_s)-a_s(Y_s)|^2\,\mathrm{d}s \\ &+ \mathbb{E}\int_{\tau_1}^{T\wedge\tau_R}\int_{|z|\leqslant r} |g_s^n(Y_s,z)-g_s(Y_s,z)|^2\nu(\mathrm{d}z)\,\mathrm{d}s \\ &+ \mathbb{E}\int_{\tau_1}^{T\wedge\tau_R}\Big(\int_{|z|\leqslant r} |g_s^n(Y_s,z)-g_s(Y_s,z)|^2\nu(\mathrm{d}z)\Big)^2\mathrm{d}s. \end{aligned}$$

By Gronwall's inequality, (5.2), (5.3) and the dominated convergence theorem, we have for each $R>0$,

$$\lim_{n\to 0}\mathbb{E}\Big[\sup_{t\in[\tau_1,\tau_R]} |f(U_t^n)|^2\Big] \leqslant \lim_{n\to 0} C\,h_n = 0.$$

In particular,

$$\lim_{n\to\infty}\mathbb{E}\Big[\sup_{t\in[\tau_1,\tau_R]} |Y_t^n - Y_t^\infty|^4\wedge 1\Big] = 0,$$

which together with $\lim_{R\to\infty}\mathbb{P}(\tau_R<\tau_2)=0$ gives the desired limit. □

Now we give:

Proof of Proposition 5.1. Let $\tau_1 := 0$ and for $m \in \mathbb{N}$, define recursively

$$\tau_{m+1} := \inf\{t > \tau_m : |Z_s - Z_{s-}| > r\} \wedge 1.$$

Since Z only has finite many jumps greater than r before time 1, we have $\lim_{m\to\infty} \tau_m = 1$. Clearly, for $t \in [\tau_m, \tau_{m+1})$, Y_t^n satisfies

$$Y_t^n = Y_{\tau_m}^n + \int_{\tau_m}^t a_s^n(Y_s^n)\,\mathrm{d}s + \int_{\tau_m}^t \int_{|z|\leqslant r} g_s^n(Y_{s-}^n, z)\,\tilde{N}(\mathrm{d}s, \mathrm{d}z),$$

where

$$Y_{\tau_m}^n := Y_{\tau_m-}^n + g_{\tau_m}^n(Y_{\tau_m-}^n, Z_{\tau_m} - Z_{\tau_m-}). \tag{5.7}$$

Since $\xi_n \to \xi_\infty$ in probability as $n \to \infty$, by Lemma 5.3 and induction, we have for each $m \in \mathbb{N}$,

$$\lim_{n\to\infty} \mathbb{E}\Big[\sup_{t\in[\tau_m,\tau_{m+1})} |Y_t^n - Y_t^\infty| \wedge 1\Big] = 0.$$

Condition (5.1) and (5.7) with $m+1$ in place of m there imply that the above property extends to the right endpoint τ_{m+1} of the random interval; that is,

$$\lim_{n\to\infty} \mathbb{E}\Big[\sup_{t\in[\tau_m,\tau_{m+1}]} |Y_t^n - Y_t^\infty| \wedge 1\Big] = 0.$$

This gives the desired result as for any $m_0 \in \mathbb{N}$,

$$\mathbb{E}\Big[\sup_{t\in[0,1]} |Y_t^n - Y_t^\infty| \wedge 1\Big] \leqslant \sum_{m=1}^{m_0} \mathbb{E}\Big[\sup_{t\in[\tau_m,\tau_{m+1}]} |Y_t^n - Y_t^\infty| \wedge 1\Big] + \mathbb{P}(\tau_{m_0+1} < 1),$$

and $\lim_{m_0\to\infty} \mathbb{P}(\tau_{m_0+1} < 1) = 0$. □

Acknowledgement. We thank Zenghu Li, Feng-Yu Wang and Jian Wang for helpful discussions, and Enrico Priola for communications about his Remark 5.5 of [19]. We also thank the anonymous referee for helpful comments.

References

[1] APPLEBAUM, D.: *Lévy processes and stochastic calculus.* Cambridge Studies in Advance Mathematics 93, Cambridge University Press, Cambridge, 2004.

[2] BOGDAN, K., BYCZKOWSKI, T., KULCZYCKI, T., RYZNAR, M., SONG, R. AND VONDRACEK, Z.: *Potential analysis of stable processes and its extensions.* Lecture Notes in Mathematics 1980, Springer, Berlin, 2009.

[3] BOGDAN, K., GRZYWNY, T. AND RYZNAR, M.: Density and tails of unimodal convolution semigroups. *J. Funct. Anal.* **266** (2014), no. 6, 3543–3571.

[4] Chen, Z.-Q. and Song, R.: Drift transforms and Green function estimates for discontinuous processes. *J. Funct. Anal.* **201** (2003), no. 1, 262–281.

[5] Chen, Z.-Q. and Wang, L.: Uniqueness of stable processes with drift. *Proc. Amer. Math. Soc.* **144** (2016), no. 6, 2661–2675.

[6] Chen, Z.-Q. and Zhang, X.: Heat kernels and analyticity of non-symmetric jump diffusion semigroups. *Probab. Theory Relat. Fields* **165** (2016), no. 1-2, 267–312.

[7] Chen, Z.-Q., Zhang, X. and Zhao, G.: Well-posedness of supercritical SDE driven by Lévy processes with irregular drifts. Preprint, arXiv: 1709.04632, 2017.

[8] Fedrizzi, E. and Flandoli F.: Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics* **83** (2011), no. 3, 241–257.

[9] Flandoli, F., Gubinelli, M. and Priola, E.: Well-posedness of the transport equation by stochastic perturbation. *Invent. Math.* **180** (2010), no. 1, 1–53.

[10] Ikeda, N. and Watanabe, S.: *Stochastic differential equations and diffusion processes.* Second edition. North-Holland Mathematical Library 24, North-Holland/Kodansha, Tokyo, 1989.

[11] Krylov, N. V.: *Lectures on elliptic and parabolic equations in Hölder spaces.* Graduate Studies in Mathematics 12, Amer. Math. Soc., Providence, RI, 1996.

[12] Krylov, N. V. and Röckner, M.: Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields* **131** (2005), no. 2, 154–196.

[13] Kunita, H.: Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In *Real and stochastic analysis,* 305–373. Trends Math., Birkhäuser, Boston, MA, 2004.

[14] Kurenok, V. P.: Stochastic equations with time-dependent drift driven by Lévy processes. *J. Theoret. Probab.* **20** (2007), no. 4, 859–869.

[15] Lunardi, A.: *Analytic semigroups and optimal regularity in parabolic problems.* Progress in Nonlinear Differential Equations and their Applications 16, Birkhäuser, Basel, 1995.

[16] Podolynny, S. I. and Portenko, N. I.: On multidimensional stable processes with locally unbounded drift. *Random Oper. Stochastic Equations* **3** (1995), no. 2, 113–124.

[17] Portenko, N. I.: Some perturbations of drift-type for symmetric stable processes. *Random Oper. Stochastic Equations* **2** (1994), no. 3, 211–224.

[18] Priola, E.: Pathwise uniqueness for singular SDEs driven by stable processes. *Osaka J. Math.* **49** (2012), no. 2, 421–447.

[19] Priola, E.: Stochastic flow for SDEs with jumps and irregular drift term. In *Stochastic analysis,* 193–210. Banach Center Publ. 105, Polish Acad. Sci. Inst. Math., Warsaw, 2015.

[20] Protter, P.: *Stochastic integration and differential equations.* Second edition. Applications of Mathematics 21, Stochastic Modelling and Applied Probability, Springer, Berlin, 2004.

[21] Silvestre, L.: On the differentiability of the solution to an equation with drift and fractional diffusion. *Indiana Univ. Math. J.* **61** (2012), no. 2, 557–584.

[22] Stein, E. M.: *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series 30, Princeton University Press, Princeton, NJ, 1970.

[23] Tanaka, H., Tsuchiya, M. and Watanabe, S.: Perturbation of drift-type for Lévy processes. *J. Math. Kyoto Univ.* **14** (1974), 73–92.

[24] Veretennikov, A. Ju.: On the strong solutions of stochastic differential equations. *Theory Probability Appl.* **24** (1979), no. 2, 354-366.

[25] Wang, F.-Y., Xu, L. and Zhang, X.: Gradient estimates for SDEs driven by multiplicative Lévy noise. *J. Funct. Anal.* **269** (2015), no. 10, 3195–3219.

[26] Zhang, X.: Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.* **16** (2011), no. 38, 1096–1116.

[27] Zhang, X.: Stochastic functional differential equations driven by Lévy processes and quasi-linear partial integro-differential equations. *Ann. Appl. Probab.* **22** (2012), no. 6, 2505–2538.

[28] Zhang, X.: Stochastic differential equations with Sobolev drifts and driven by α-stable processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** (2013), no. 4, 1057–1079.

[29] Zhang, X.: Derivative formula and gradient estimate for SDEs driven by α-stable processes. *Stochastic Process. Appl.* **123** (2013), no. 4, 1213–1228.

[30] Zvonkin, A. K.: A transformation of the phase space of a diffusion process that removes the drift. *Mat. Sb. (N.S.)* **93 (135)** (1974), 129–149, 152.

Received August 8, 2016. Published online December 3, 2018.

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

Renming Song: Department of Mathematics, University of Illinois, Urbana, IL 61801, USA.
E-mail: rsong@illinois.edu

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

The research of Z.-Q. Chen is partially supported by NSF grant DMS-1206276. The research of R. Song is partially supported by a grant from the Simons Foundation (208236). The research of X. Zhang is partially supported by NNSFC grant of China (Nos. 11271294, 11325105, 11731009).