

# 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  $\mathbb{R}^d$ :

$$dX_t = dZ_t + b(t, X_t)dt, \quad X_0 = x,$$

where  $Z$  is a Lévy process. We show that for a large class of Lévy processes  $Z$  and Hölder continuous drift  $b$ , 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  $\alpha$ -stable-type Lévy process with  $\alpha \in (0, 2)$  and  $b$  is bounded and  $\beta$ -Hölder continuous with  $\beta \in (1 - \alpha/2, 1)$ , the SDE above has a unique strong solution. When  $\alpha \in (0, 1)$ , this in particular solves an open problem from Priola [11]. 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  $f$ :

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

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

**AMS 2010 Mathematics Subject Classification:** Primary 60H10, 35K05; Secondary 60H30, 47G20

**Keywords and Phrases:** SDE, supercritical, subcritical, stable process, subordinate Brownian motion, gradient estimate, strong existence, pathwise uniqueness, Bismut formula, stochastic flow,  $C^1$ -diffeomorphism

## 1. INTRODUCTION

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

$$dX_t = dZ_t + b(t, X_t)dt, \quad X_0 = x, \tag{1.1}$$

where  $b(t, x) : [0, 1] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a bounded Borel function and  $Z$  is a Lévy process in  $\mathbb{R}^d$ . When  $d = 1$ ,  $Z$  is a Brownian motion and  $b$  is a bounded Borel function on  $\mathbb{R}$ , Zvonkin [22] proved that the above SDE admits a unique strong solution for every starting point  $x$ . Zvonkin's result was extended to the multi-dimensional case by Veretennikov [16]. Since then, many people have made contributions to this problem (see [8, 5, 4, 18] 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  $\alpha$ -stable process in  $\mathbb{R}$  with  $\alpha \in (0, 1)$ , Tanaka, Tsuchiya and Watanabe [15] 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  $\alpha$ -stable process in  $\mathbb{R}$  with  $\alpha \in [1, 2)$ , it is shown in [15] that pathwise uniqueness holds for (1.1) for any bounded continuous  $b(t, x) = b(x)$ . For  $d \geq 2$ , using Zvonkin's transform, Priola [10] obtained pathwise uniqueness for SDE (1.1) when  $Z$  is a non-degenerate symmetric (but possibly non-isotropic)  $\alpha$ -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  $\mathcal{L}^{(\alpha)} + b \cdot \nabla$ .

---

The research of ZC is partially supported by NSF grant DMS-1206276. The research of RS is partially supported by a grant from the Simons Foundation (208236). The research of XZ is partially supported by NNSFC grant of China (Nos. 11271294, 11325105).

Here  $\mathcal{L}^{(\alpha)}$  is the infinitesimal generator of the Lévy process  $Z$ , which is a nonlocal operator of order  $\alpha$ . When  $\alpha > 1$ ,  $\mathcal{L}^{(\alpha)}$  is the dominant term, which is called the subcritical case. When  $\alpha \in (0, 1)$ , the gradient  $\nabla$  is of higher order than the nonlocal operator  $\mathcal{L}^{(\alpha)}$  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 [20] and to more general Lévy processes in the subcritical case in [11]. However, when  $d \geq 2$ ,  $\alpha \in (0, 1)$  and  $Z$  is a symmetric  $\alpha$ -stable process in  $\mathbb{R}^d$ , even for time-independent Hölder continuous drift  $b$ , pathwise uniqueness for SDE (1.1) was an open question until now; see [11, Remark 5.5]. When  $Z$  is a rotationally symmetric  $\alpha$ -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 a rotationally symmetric stable process  $Z$ , one needs to understand the above PDE better. In this direction, Silvestre [13] obtained the following a priori interior estimate:

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

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  $\alpha + \beta > 1$ . Such an estimate suggests that one could solve the supercritical SDE (1.1) uniquely when  $Z$  is a rotationally symmetric  $\alpha$ -stable process with  $\alpha \in (0, 1)$  and  $b \in C_b^\beta(\mathbb{R}^d)$  with  $\beta \in (1 - \alpha/2, 1)$ . However, this is not an easy task since one needs additional asymptotic estimates in the time variable. Furthermore, the approach of [13] strongly depends on realizing the fractional Laplacian in  $\mathbb{R}^d$  as the boundary trace of an elliptic operator in upper half space of  $\mathbb{R}^{d+1}$ . Extending this approach to other nonlocal operators, such as  $\alpha$ -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 including stable-type Lévy processes. We not only extend the main result of [11] 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. One of the main results of this paper in particular solves the open problem raised in [11, Remark 5.5] where  $Z$  is a symmetric  $\alpha$ -stable process with  $\alpha \in (0, 1)$ . 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  $a^+ := a \vee 0$ . We now describe the setup and the main results of this paper. Let  $\mathcal{L}_{\nu,\eta}$  be the infinitesimal generator of the Lévy process  $Z$ , that is,

$$\mathcal{L}_{\nu,\eta} f(x) = \int_{\mathbb{R}^d} (f(x+z) - f(x) - \mathbf{1}_{\{|z|\leq 1\}} z \cdot \nabla f(x)) \nu(dz) + \eta \cdot \nabla f(x),$$

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

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

Suppose that  $\nu$  can be decomposed as

$$\nu = \nu_0 + \nu_1 + \nu_2, \tag{1.2}$$

where  $\nu_1, \nu_2$  are two Lévy measures, and  $\nu_0$  is a *finite signed* measure so that

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

We 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  $\{T_t^{\nu_1,0}; t \geq 0\}$  hold.

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

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

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

$$|\nabla T_t^{\nu_1,0} f(x)| \leq K_0 \Lambda t^{(\delta\beta-1)/\alpha} \quad \text{for all } t \in (0, 1). \quad (1.3)$$

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

$$\|\nabla T_t^{\nu_1,0} f\|_\infty \leq K_0 \|f\|_\infty t^{-1/\alpha} \quad \text{for all } t \in (0, 1). \quad (1.4)$$

**Remark 1.1.** The pointwise estimate (1.3) allows us to borrow the Hölder regularity of the drift to compensate the time singularity, which is crucial for the well-posedness of SDEs with Hölder drifts in the supercritical case. Condition (1.4) in the subcritical case is the same as Hypothesis 1 of Priola [11]. Moreover, the parameters  $\bar{\alpha}$  and  $\delta$  are mainly designed for Example 4.3 below, and in Examples 4.1 and 4.2 below,  $\bar{\alpha}$  and  $\delta$  can all be chosen to be 1.

The first main result of this paper is the following

**Theorem 1.2.** *Suppose that either  $(\mathbf{H}_{\nu_1, K_0}^{\alpha, \bar{\alpha}, \delta})$  holds for some  $\alpha \in (0, 1]$ ,  $\bar{\alpha}, \delta \in (0, 1]$  and  $K_0 > 0$ , or  $(\mathbf{H}_{\nu_1, K_0}^\alpha)$  holds for some  $\alpha \in (1, 2)$  and  $K_0 > 0$ . Let  $\gamma \in (0, 1)$  be such that*

$$\int_{|z| \leq 1} |z|^{2\gamma} \nu(dz) < \infty. \quad (1.5)$$

*Assume further that  $\gamma + (1 - \alpha)/\delta < \bar{\alpha}$  in the case  $\alpha \in (0, 1]$ . If for some  $\beta \in (\gamma + (1 - \alpha)/\delta, 1]$  in the case  $\alpha \in (0, 1]$ , and for some  $\beta \in ((\gamma + 1 - \alpha)^+, 1]$  in the case  $\alpha \in (1, 2)$ , it holds that*

$$\sup_{t \in [0, 1]} \|b(t, \cdot)\|_\infty + \sup_{t \in [0, 1]} \sup_{x \neq y \in \mathbb{R}^d} \frac{|b(t, x) - b(t, y)|}{|x - y|^\beta} < \infty, \quad (1.6)$$

*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*

$$\sup_{x \in \mathbb{R}^d} \mathbb{E} \left[ \sup_{t \in [0, 1]} |\nabla X_t(x)|^p \right] \leq C_p < \infty \quad \text{for every } p \geq 1, \quad (1.7)$$

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

**Remark 1.3.** By a suitable localization argument (cf. [20]), for the local uniqueness of SDE (1.1), the global condition (1.6) 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.

Various examples of Lévy processes satisfying the conditions  $(\mathbf{H}_{\nu_1, K_0}^{\alpha, \bar{\alpha}, \delta})$  with  $\alpha \in (0, 1]$ ,  $(\mathbf{H}_{\nu_1, K_0}^\alpha)$  with  $\alpha \in (1, 2)$ , and (1.5) (and hence the conclusion of Theorem 1.2 holds for these Lévy processes) are given in Section 4. To illustrate 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  $\nu$  has a density  $\kappa(z)$ . Assume that for some  $0 < \alpha_1 \leq \alpha_2 < 2$ ,*

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

Assume that  $\alpha_2 < 2\alpha_1$ , and  $b(t, x)$  is bounded and  $\beta$ -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 (1.7) 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  $0 < \alpha_1 \leq \alpha_2 < 2$  such that

$$C_1|z|^{\alpha_1} \leq \Phi(z) \leq C_2|z|^{\alpha_2} \quad \text{for } |z| \geq 1.$$

Assume that  $\alpha_2 < 2\alpha_1$ , and  $b(t, x)$  is bounded and  $\beta$ -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 (1.7) holds.

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

$$\text{either } \alpha > 1 \quad \text{or} \quad \alpha \in (0, 1] \quad \text{and} \quad \alpha_{\max} < 2\alpha^2/(2 - \alpha), \quad (1.8)$$

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

$$\beta \in (\beta_0, 1] \quad \text{with} \quad \beta_0 := \alpha_{\max}/2 + (\alpha_{\max}/\alpha \mathbf{1}_{\{\alpha \leq 1\}} + \mathbf{1}_{\{\alpha > 1\}})(1 - \alpha). \quad (1.9)$$

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 (1.7) holds.

Note that condition (1.8) 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  $1 \leq j \leq k$ , then conditions (1.8) and (1.9) 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.

For Corollary 1.4 (iii), let  $\nu$  be the Lévy measure of the cylindrical stable process  $Z$ . We will in fact show in Example 4.3 that, when  $\alpha = \min_{1 \leq j \leq k} \alpha_j \in (1, 2)$ , condition  $(\mathbf{H}_{\nu, K_0}^\alpha)$  holds for some  $K_0 > 0$  but condition  $(\mathbf{H}_{\nu, K_0}^{\alpha^*})$  fails for any  $\alpha^* > \alpha$ . So Hypothesis 1 of [11] holds with this  $\alpha$  for the cylindrical stable process  $Z$ . On the other hand, condition (1.5) holds if and only if  $2\gamma > \alpha_{\max}$ . Hence in the case  $\alpha \in (1, 2)$ , Hypothesis 2 of [11] fails when  $\alpha_j$ 's are not identical (i.e., when  $\alpha_{\max} > \alpha$ ), and so the main results of [11] are not applicable.

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

**Theorem 1.5.** *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:*

$$\nabla \mathbb{E}f(X_t(x)) = \mathbb{E} \left[ \frac{f(X_t(x))}{S_t} \int_0^t \nabla X_s(x) dW_{S_s} \right], \quad f \in C_b^1(\mathbb{R}^d). \quad (1.10)$$

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  $(t, x) \in (0, 1) \times \mathbb{R}^d$ ,

$$|\nabla \mathbb{E}f(X_t(x))| \leq C_p t^{-1/\alpha} (\mathbb{E}|f(X_t(x))|^p)^{1/p}. \quad (1.11)$$

This paper is organized as follows: In Section 2, we solve a nonlocal advection equation 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. Even when  $Z$  is a rotationally symmetric stable process, our approach to the a priori estimate is simpler and more elementary than that of [13]. In Section 3, we shall prove our main results by using Zvonkin's transform. In Section 4, we give three examples to illustrate the main results of this paper, from which

Corollary 1.4 follows. In 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  $f(x) \leq g(x)$  to mean that there exists a constant  $C_0 > 0$  such that  $f(x) \leq C_0 g(x)$ ; and  $f(x) \asymp g(x)$  to mean that there exist  $C_1, C_2 > 0$  such that  $C_1 g(x) \leq f(x) \leq C_2 g(x)$ .

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  $f : [0, 1] \times \mathbb{R}^d \rightarrow \mathbb{R}$ ,

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

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

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

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

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

It is well-known (cf. [14, Proposition 7 on p.142]) that  $\|f\|_\beta < \infty$  if and only if  $f$  is bounded and

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

and

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

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  $\gamma \leq \beta$ , there is a positive constant  $C = C(\beta, \gamma, d)$  such that for any Borel functions  $f, g$  on  $\mathbb{R}^d$ ,*

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

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

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

*Proof.* It suffices to prove that

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

By definition, we have

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

Notice the following easy estimates:

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

and

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

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

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

Next, we assume

$$|x - x'| \leq \theta/2. \quad (2.6)$$

Notice that

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

For  $I_1$ , we have

$$\begin{aligned} |I_1| &\leq [f]_\beta \|g\|_\infty \int_{\mathbb{R}^d} |x-y|^\beta |x-x'| \left( \int_0^1 |\nabla \partial_\theta p_\theta(x-y+r(x'-x))| dr \right) dy \\ &\stackrel{(2.4)}{\leq} [f]_\beta \|g\|_\infty |x-x'| \int_{\mathbb{R}^d} |x-y|^\beta \left( \int_0^1 (\theta + |x-y+r(x'-x)|)^{-d-2} dr \right) dy \\ &\stackrel{(2.6)}{\leq} [f]_\beta \|g\|_\infty |x-x'| \int_{\mathbb{R}^d} |x-y|^\beta (\theta + |x-y|)^{-d-2} dy \\ &\stackrel{(2.5)}{\leq} [f]_\beta \|g\|_\infty |x-x'| \theta^{\beta-2} \stackrel{(2.6)}{\leq} [f]_\beta \|g\|_\infty |x-x'|^{\beta-\gamma} \theta^{\gamma-1}. \end{aligned}$$

For  $I_2$ , we similarly have

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

Combining the above estimates, we obtain (2.2).  $\square$

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

**Lemma 2.2.** *Suppose that  $(\mathbf{H}_{v_1, K_0}^\alpha)$  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*

$$\|\nabla T_t^{v_1, 0} f\|_\gamma \leq K_1 t^{(\beta-1-\gamma)/\alpha} \|f\|_\beta \quad \text{for all } t \in (0, 1).$$

*Proof.* Note that  $\|\nabla T_t^{\nu_1,0} f\|_\infty \leq \|\nabla f\|_\infty$ . By (1.4) and the interpolation theorem, we have

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

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

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

Hence,

$$[\nabla T_t^{\nu_1,0} f]_\gamma \leq 2 \|\nabla^2 T_t^{\nu_1,0} f\|_\infty^\gamma \|\nabla T_t^{\nu_1,0} f\|_\infty^{1-\gamma} \leq t^{(\beta-1-\gamma)/\alpha} \|f\|_\beta.$$

□

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

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

where  $\mathcal{L}_{\nu, \eta}$  is the infinitesimal generator of the Lévy process  $Z$ , and  $b, f : [0, 1] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  are bounded Borel functions. Recalling decomposition (1.2), we can write

$$\mathcal{L}_{\nu, \eta} = \mathcal{L}_{\nu_0, 0} + \mathcal{L}_{\nu_1, 0} + \mathcal{L}_{\nu_2, \eta}. \quad (2.8)$$

The following theorem is the main result of this section.

**Theorem 2.3.** *Suppose that either  $(\mathbf{H}_{\nu_1, K_0}^{\alpha, \bar{\alpha}, \delta})$  holds for some  $\alpha \in (0, 1]$ ,  $\bar{\alpha}, \delta \in (0, 1]$  and  $K_0 > 0$  or  $(\mathbf{H}_{\nu_1, K_0}^\alpha)$  holds for some  $\alpha \in (1, 2)$  and  $K_0 > 0$ . Suppose further that  $1 - \alpha < \delta \bar{\alpha}$  in the case  $\alpha \in (0, 1]$ . If for some  $\beta \in ((1 - \alpha)/\delta, \bar{\alpha}]$  in the case  $\alpha \in (0, 1]$  and  $\beta \in [0, 1]$  in the case  $\alpha \in (1, 2)$ , it holds that*

$$\|b\|_{\infty, \beta} < \infty, \quad \|f\|_{\infty, \beta} < \infty, \quad (2.9)$$

then for any  $\gamma \in (0, \beta - (1 - \alpha)/\delta)$  when  $\alpha \in (0, 1]$  and  $\gamma \in (0, (\beta + \alpha - 1) \wedge 1)$  when  $\alpha \in (1, 2)$ , there exists a continuous function  $u : [0, 1] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that for all  $t \in [0, 1]$  and  $\varphi \in C_c^\infty(\mathbb{R}^d)$ ,

$$\langle u_t, \varphi \rangle = \int_0^t \langle u_s, (\mathcal{L}_{\nu, \eta}^* - \lambda) \varphi \rangle ds + \int_0^t \langle b_s \cdot \nabla u_s, \varphi \rangle ds + \int_0^t \langle f_s, \varphi \rangle ds \quad (2.10)$$

with

$$\sup_{t \in [0, 1]} \|u_t(\cdot)\|_\infty \leq \sup_{t \in [0, 1]} \|f_t(\cdot)\|_\infty, \quad (2.11)$$

and for some  $\theta_0 > 0$  and all  $\lambda \geq 0$ ,

$$\|\nabla u\|_{\infty, \gamma} \leq C(1 \vee \lambda)^{-\theta_0} \|f\|_{\infty, \beta}. \quad (2.12)$$

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

We will first prove several lemmas before we present the proof of the theorem above. Let  $Z^{(1)}$  and  $Z^{(2)}$  be two independent Lévy processes with generators  $\mathcal{L}_{\nu_0 + \nu_1, 0}$  and  $\mathcal{L}_{\nu_2, \eta}$ . Clearly,

$$Z_t \stackrel{(d)}{=} Z_t^{(1)} + Z_t^{(2)}. \quad (2.13)$$

**Lemma 2.4.** *Assume that  $b, f \in L^\infty([0, 1]; C_b^\infty(\mathbb{R}^d; \mathbb{R}^d))$ . There exists a unique solution  $u_t(x) \in C([0, 1]; C_b^\infty(\mathbb{R}^d; \mathbb{R}^d))$  to equation (2.7) with the following probabilistic representation:*

$$u_t(x) = \int_t^1 e^{\lambda(t-s)} \mathbb{E} f_s(X_{t,s}(x)) ds, \quad (2.14)$$

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

$$X_{t,s} = x + \int_t^s b_r(X_{t,r}) dr + Z_s - Z_t, \quad s \geq t. \quad (2.15)$$

Moreover, we have the following a priori estimate

$$\sup_{t \in [0,1]} \|u_t\|_\infty \leq \sup_{t \in [0,1]} \|f_t\|_\infty. \quad (2.16)$$

*Proof.* The existence and uniqueness of  $u_t(x)$  and the representation (2.14) follow from [19, Theorem 4.4]. The estimate (2.16) immediately follows from (2.14).  $\square$

**Lemma 2.5.** *Suppose that  $(\mathbf{H}_{\nu_1, K_0}^{\alpha, \bar{\alpha}, \delta})$  holds for some  $\alpha \in (0, 1]$ ,  $\bar{\alpha}, \delta \in (0, 1]$  and  $K_0 > 0$ , and that  $b, f \in L^\infty([0, 1]; C_b^\infty(\mathbb{R}^d; \mathbb{R}^d))$ . Suppose further that  $1 - \alpha < \delta \bar{\alpha}$  in the case  $\alpha \in (0, 1]$ . 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  $\lambda \geq 0$ ,*

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

*Proof.* (i) We first assume that  $\eta = 0$  and that  $\nu_2 = 0$  in decomposition (1.2). 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

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

and

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

Clearly, by (2.7) and (2.8),  $\tilde{u}$  satisfies

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

We have by the representation (2.14) (with  $b = 0$  there)

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

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

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

and that

$$|\nabla T_{s-t}^{\nu_1, 0}(\mathcal{L}_{\nu_0, 0} \tilde{u}_s)(x)| \leq \|\nabla(\mathcal{L}_{\nu_0, 0} \tilde{u}_s)\|_\infty \leq 2|\nu_0|(\mathbb{R}^d) \|\nabla \tilde{u}_s(\cdot)\|_\infty.$$

We have by (1.3) that for  $t \in [0, 1]$ ,

$$\begin{aligned} |\nabla \tilde{u}_t(0)| &\leq \int_t^1 e^{\lambda(t-s)} \left( K_0 [b_s(\cdot)]_{\beta_1} (s-t)^{(\delta\beta_1-1)/\alpha} + 2|\nu_0|(\mathbb{R}^d) \right) \|\nabla \tilde{u}_s(\cdot)\|_\infty ds \\ &\quad + K_0 \int_t^1 e^{\lambda(t-s)} (s-t)^{(\delta\beta_2-1)/\alpha} [\tilde{f}_s(\cdot)]_{\beta_2} ds. \end{aligned}$$

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

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

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  $\ell : [0, 1] \rightarrow \mathbb{R}^d$ , and define

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



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

$$Y_{t,s}^\ell = x + \int_t^s b_r^\ell(Y_{t,r}^\ell) dr + Z_s^{(1)} - Z_t^{(1)}, \quad s \geq t.$$

Since  $Z^{(1)}$  and  $Z^{(2)}$  are independent, by (2.13) and the uniqueness in law of the solution to SDE (2.15), we have

$$X_{t_0,\cdot}(x) \stackrel{(d)}{=} Y_{t_0,\cdot}^{Z^{(2)}}(x) - Z_{t_0}^{(2)} + Z^{(2)},$$

and so by (2.14),

$$u_{t_0}(x) = \mathbb{E} \left( \int_{t_0}^1 e^{\lambda(t_0-s)} \mathbb{E} [f_s^\ell(Y_{t_0,s}^\ell(x))] ds \Big|_{\ell=Z^{(2)}} \right).$$

Now we define

$$u_t^\ell(x) := \int_t^1 e^{\lambda(t-s)} \mathbb{E} [f_s^\ell(Y_{t,s}^\ell(x))] ds.$$

Then by Lemma 2.4,  $u_t^\ell(x)$  is a solution to the following equation:

$$\partial_t u^\ell + (\mathcal{L}_{\nu_0+\nu_1,0} - \lambda)u^\ell + b^\ell \cdot \nabla u^\ell + f^\ell = 0, \quad u_1^\ell = 0.$$

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  $\ell$ ,

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

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

**Lemma 2.6.** *Suppose that  $(\mathbf{H}_{\nu_1, K_0}^{\alpha, \bar{\alpha}, \delta})$  holds for some  $\alpha \in (0, 1]$ ,  $\bar{\alpha}, \delta \in (0, 1]$  and  $K_0 > 0$ , and that  $b, f \in L^\infty([0, 1]; C_b^\infty(\mathbb{R}^d; \mathbb{R}^d))$ . Suppose further that  $1 - \alpha < \delta \bar{\alpha}$  in the case  $\alpha \in (0, 1]$ . 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  $\lambda \geq 0$ ,*

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

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

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

and

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

then

$$\partial_t w_t^\theta + (\mathcal{L}_{\nu,\eta} - \lambda)w_t^\theta + b_t \cdot \nabla w_t^\theta + g_t^\theta = 0, \quad w_1^\theta = 0.$$

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

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

and by Lemma 2.1,

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

Hence,

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

which yields (2.20) by (2.1) and (2.17).  $\square$

**Lemma 2.7.** *Suppose that  $(\mathbf{H}_{\nu_1, K_0}^\alpha)$  holds for some  $\alpha \in (1, 2)$  and  $K_0 > 0$ , 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  $\beta \in [0, 1]$  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  $\lambda \geq 0$ ,*

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

*Proof.* As in the proof of Lemma 2.5, we first assume that  $\eta = 0$  and that  $\nu_2 = 0$  in decomposition (1.2). By the representation (2.14) (with  $b = 0$  there), we have

$$u_t(x) = \int_t^1 e^{\lambda(t-s)} T_{s-t}^{\nu_1, 0} (b_s \cdot \nabla u_s + \mathcal{L}_{\nu_0, 0} u_s + f_s)(x) ds.$$

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

$$\begin{aligned} \|\nabla u_t\|_\gamma &\leq \int_t^1 e^{\lambda(t-s)} (s-t)^{\frac{\beta-1-\gamma}{\alpha}} (\|b_s \cdot \nabla u_s\|_\beta + \|f_s\|_\beta) ds + \int_t^1 e^{\lambda(t-s)} \|\nabla T_{s-t}^{\nu_1, 0} \mathcal{L}_{\nu_0, 0} u_s\|_\gamma ds \\ &\leq \int_t^1 e^{\lambda(t-s)} (s-t)^{(\beta-1-\gamma)/\alpha} (\|b_s\|_\beta \|\nabla u_s\|_\beta + \|f_s\|_\beta) ds + \int_t^1 e^{\lambda(t-s)} \|\nabla u_s\|_\gamma ds \\ &\leq \int_t^1 ((s-t)^{(\beta-1-\gamma)/\alpha} + 1) \|\nabla u_s\|_\gamma ds + \|f\|_{\infty, \beta} \int_t^1 e^{\lambda(t-s)} (s-t)^{(\beta-1-\gamma)/\alpha} ds, \end{aligned}$$

which yields (2.21) by Gronwall's inequality. For the general case, we can follow the same argument in (ii) of Lemma 2.5 to derive (2.21).  $\square$

Now we are ready to give

*Proof of Theorem 2.3.* Suppose that  $b$  and  $f$  satisfy (2.9). Let  $\varrho$  be a non-negative smooth function with compact support in  $\mathbb{R}^d$  having  $\int_{\mathbb{R}^d} \varrho(x) dx = 1$ . For  $n \in \mathbb{N}$ , define  $\varrho_n(x) := n^d \varrho(nx)$  and

$$b_t^n := \varrho_n * b_t, \quad f_t^n := \varrho_n * f_t. \quad (2.22)$$

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

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

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

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

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

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

and by (2.16),

$$\sup_{t \in [0, 1]} \|u_t^n\|_\infty \leq \sup_{t \in [0, 1]} \|f_t^n\|_\infty \leq \sup_{t \in [0, 1]} \|f_t\|_\infty. \quad (2.25)$$

Moreover, by the representation (2.14) (with  $b = 0$  there), we can write

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

Using (2.24) and (2.25), one can easily show that for any  $R > 0$ ,

$$\lim_{|t-t'| \rightarrow 0} \sup_n \sup_{|x| \leq R} |u_t^n(x) - u_{t'}^n(x)| = 0.$$

Hence, by Ascoli-Arzela's lemma, there is a subsequence (still denoted by  $u^n$ ) and a function  $u$  with

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

such that

$$\lim_{n \rightarrow \infty} \sup_{t \in [0, 1]} \sup_{|x| \leq R} |u_t^n(x) - u_t(x)| = 0, \quad \forall R > 0. \quad (2.26)$$

On the other hand, noticing the following interpolation inequality (cf. [7, Theorem 3.2.1])

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

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

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

Thus by (2.23), (2.26) and (2.27), it is easy to see that  $u$  satisfies (2.10).  $\square$

**Corollary 2.8.** *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  $\alpha \in (1, 2)$ ,*

$$\int_{|z| \leq 1} |z|^{1+\gamma_0} \nu(dz) < \infty, \quad (2.28)$$

*then the solution  $u$  of equation (2.7) satisfying (2.11) and (2.12) for some  $\gamma > \gamma_0$  is a classical solution; that is,  $\mathcal{L}_{v, \eta} u_s(x)$  and  $\nabla u_s(x)$  exists pointwise and is continuous in  $x$ , and for all  $t \in [0, 1]$  and  $x \in \mathbb{R}^d$ ,*

$$u_t(x) + \int_t^1 (\mathcal{L}_{v, \eta} - \lambda) u_s(x) ds + \int_t^1 b_s(x) \cdot \nabla u_s(x) ds + \int_t^1 f_s(x) ds = 0. \quad (2.29)$$

*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  $\alpha \in (1, 2)$ , by (2.28), it is easy to check that

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

Hence, by (2.10), equation (2.29) is satisfied for all  $t \in [0, 1]$  and  $x \in \mathbb{R}^d$ .  $\square$

### 3. STOCHASTIC FLOW AND BISMUT FORMULA

Suppose that either  $(\mathbf{H}_{v_1, K_0}^{\alpha, \bar{\alpha}, \delta})$  holds for some  $\alpha \in (0, 1]$ ,  $\bar{\alpha}, \delta \in (0, 1]$  and  $K_0 > 0$  or  $(\mathbf{H}_{v_1, K_0}^{\alpha})$  holds for some  $\alpha \in (1, 2)$  and  $K_0 > 0$ . Suppose also that (1.5) and (1.6) hold for some

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

and

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

Notice that (1.5) implies (2.28) with  $\gamma_0 = \gamma$ . Hence, for  $\lambda \geq 0$ , by Corollary 2.8, the following nonlocal equation has a classical solution  $u$ :

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

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

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

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

$$\lim_{n \rightarrow \infty} \sup_{t \in [0, 1]} \sup_{|x| \leq R} |\nabla^j u_t^n(x) - \nabla^j u_t(x)| = 0 \quad \text{for every } R > 0 \text{ and } j = 0, 1. \quad (3.1)$$

For simplicity, we use the following convention:

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

By (2.12), one can choose  $\lambda$  sufficiently large, independent of  $n \in \mathbb{N}_\infty$ , such that

$$\|\nabla u_t^n(\cdot)\|_\infty + \sup_{x \neq x'} \frac{|\nabla u_t^n(x) - \nabla u_t^n(x')|}{|x - x'|^\gamma} \leq \frac{1}{2} \quad \text{for every } n \in \mathbb{N}_\infty \text{ and } t \in [0, 1]. \quad (3.2)$$

Define

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

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

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

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

$$1/2 \leq |\nabla \Phi_t^n(x)| \leq 3/2 \quad \text{and} \quad |\nabla(\Phi_t^n)^{-1}(x)| \leq 2, \quad (3.3)$$

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

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

$$\|\nabla \Phi_t^n\|_\gamma + \|\nabla(\Phi_t^n)^{-1}\|_\gamma \leq C. \quad (3.4)$$

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

$$\lim_{n \rightarrow \infty} \sup_{t \in [0, 1]} \sup_{|x| \leq R} |\nabla^j \Phi_t^n(x) - \nabla^j \Phi_t^\infty(x)| = 0, \quad (3.5)$$

and

$$\lim_{n \rightarrow \infty} \sup_{t \in [0, 1]} \sup_{|x| \leq R} |\nabla^j(\Phi_t^n)^{-1}(x) - \nabla^j(\Phi_t^\infty)^{-1}(x)| = 0. \quad (3.6)$$

*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 (3.2) and (3.3),

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

Hence by (3.3) again, for all  $s \in [0, 1]$ ,

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

(ii) Properties (3.5) and (3.6) follow from the definitions of  $\Phi_t$  and  $\Phi_t^{-1}$ , (3.1) and (3.2).  $\square$

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

$$g_s^n(y, z) := \Phi_s^n \left( (\Phi_s^n)^{-1}(y) + z \right) - y. \quad (3.7)$$

**Lemma 3.2.** *For  $\gamma_1, \gamma_2 \geq 0$  with  $\gamma_1 + \gamma_2 = \gamma \in (0, 1)$ , there is a positive constant  $C_1 = C_1(d, \lambda, \gamma_1, \gamma_2)$  such that for all  $n \in \mathbb{N}_\infty$ ,  $t \in [0, 1]$  and  $y, z \in \mathbb{R}^d$ ,*

$$\|\nabla g^n(\cdot, z)\|_{\infty, \gamma_1} \leq C_1(1 \wedge |z|^{\gamma_2}) \quad \text{and} \quad |g_s^n(y, z)| \leq 3|z|/2. \quad (3.8)$$

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

$$\lim_{n \rightarrow \infty} \nabla_y^j g_t^n(y, z) = \nabla_y^j g_t^\infty(y, z). \quad (3.9)$$

*Proof.* For notational simplicity, we drop the superscript “ $n$ ”. Since

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

we have

$$\|\nabla g_s(\cdot, z)\|_\infty \leq 2\|\nabla \Phi_s\|_\gamma (1 \wedge |z|^\gamma) \|\nabla \Phi_s^{-1}\|_\infty \stackrel{(3.4)}{\leq} C(1 \wedge |z|^\gamma)$$

and

$$\begin{aligned} [\nabla g_s(\cdot, z)]_\gamma &\leq [\nabla \Phi_s(\Phi_s^{-1}(\cdot) + z)]_\gamma \|\nabla \Phi_s^{-1}\|_\infty + \|\nabla \Phi_s\|_\infty [\nabla \Phi_s^{-1}]_\gamma \\ &\leq [\nabla \Phi_s]_\gamma \|\nabla \Phi_s^{-1}\|_\infty^{1+\gamma} + \|\nabla \Phi_s\|_\infty [\nabla \Phi_s^{-1}]_\gamma \stackrel{(3.4)}{\leq} C. \end{aligned}$$

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

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

which in turn gives the first estimate in (3.8). The second inequality in (3.8) follows from (3.3) and the definition of  $g^n$ . Property (3.9) follows from (3.5), (3.6), and the definition of  $g^n$ .  $\square$

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

$$\|\nabla g^n(\cdot, z)\|_\infty \leq C_0(1 \wedge |z|^\gamma) \quad \text{and} \quad |g_s^n(y, z)| \leq 3|z|/2. \quad (3.10)$$

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

$$C_0 r_0^\gamma + 3r_0/2 < 1. \quad (3.11)$$

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 (3.16).

Let  $N(dt, dz)$  be the Poisson random measure associated with  $Z$ , i.e.,

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

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

$$Z_t = \int_0^t \int_{|z| < r} z \tilde{N}(ds, dz) + \int_0^t \int_{|z| \geq r} z N(ds, dz) + \eta_r t,$$

where  $\eta_r \in \mathbb{R}^d$  is a constant vector depending on  $r$ .

Recall that  $r_0 \in (0, 1)$  is the constant in (3.11). For any given  $n \in \mathbb{N}_\infty$ , define

$$a_s^n(y) := \eta_{r_0} + \lambda u_s^n((\Phi_s^n)^{-1}(y)) - \int_{|z| \geq r_0} \left( u_s^n((\Phi_s^n)^{-1}(y) + z) - u_s^n((\Phi_s^n)^{-1}(y)) \right) \nu(dz). \quad (3.12)$$

We have

**Lemma 3.3.** *There is a positive constant  $C_2 = C_2(d, \lambda, \gamma, r_0)$  such that for all  $n \in \mathbb{N}_\infty$ ,  $t \in [0, 1]$  and  $y \in \mathbb{R}^d$ ,*

$$\|\nabla a^n\|_{\infty, \gamma} \leq C_2 \quad \text{and} \quad |a_s^n(y)| \leq C_2(1 + \|b\|_\infty). \quad (3.13)$$

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

$$\lim_{n \rightarrow \infty} \nabla_y^j a_t^n(y) = \nabla_y^j a_t^\infty(y). \quad (3.14)$$

*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 (3.4) that for all  $s \in [0, 1]$ ,

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

Hence  $\|\nabla a\|_{\infty, \gamma} \leq C_1$  by (3.12). The second inequality in (3.13) follows from the definition of  $a_s(y)$  and the fact that  $u^n$  is uniformly bounded due to (2.11) of Theorem 2.3. Property (3.14) follows from (3.1), (3.2), (3.5), (3.6), and the definition of  $a^n$ .  $\square$

The following lemma is a direct application of Itô’s formula.

**Lemma 3.4.** *Let  $\Phi_t^n(x)$  be defined as above. For  $n \in \mathbb{N}_\infty$ ,  $X_t^n$  satisfies*

$$X_t^n = x + \int_0^t b_s^n(X_s^n) ds + Z_t, \quad t \in [0, 1] \quad (3.15)$$

if and only if  $Y_t^n = \Phi_t^n(X_t^n)$  solves the following SDE for  $t \in [0, 1]$ :

$$Y_t^n = \Phi_0^n(x) + \int_0^t a_s^n(Y_s^n) ds + \int_0^t \int_{|z| < r_0} g_s^n(Y_{s-}^n, z) \tilde{N}(ds, dz) + \int_0^t \int_{|z| \geq r_0} g_s^n(Y_{s-}^n, z) N(ds, dz), \quad (3.16)$$

where  $a^n$  and  $g^n$  are defined by (3.12) and (3.7).

*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. 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 [10] and [20]. We omit the details.  $\square$

**Lemma 3.5.** *For  $n \in \mathbb{N}_\infty$ , let  $Y_t^n(x)$  be the solution of (3.16) with initial value  $\Phi_0^n(x)$ . We have*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \sup_{t \in [0, 1]} |Y_t^n(x) - Y_t^\infty(x)| \wedge 1 \right] = 0. \quad (3.17)$$

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

$$\sup_{n \in \mathbb{N}_\infty} \sup_{x \in \mathbb{R}^d} \mathbb{E} \left[ \sup_{t \in [0, 1]} |\nabla Y_t^n(x)|^p \right] < \infty, \quad (3.18)$$

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

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \sup_{t \in [0, 1]} |\nabla Y_t^n(x) - \nabla Y_t^\infty(x)|^p \right] = 0. \quad (3.19)$$

*Proof.* (3.17) follows from Lemmas 3.2, 3.3 and Proposition 5.1 below. In this proof, we shall drop the superscript “ $\infty$ ”. Notice that

$$\begin{aligned} \nabla Y_t^n &= \nabla \Phi_0^n(x) + \int_0^t \nabla a_s^n(Y_s^n) \nabla Y_s^n ds + \int_0^t \int_{|z| < r_0} \nabla_y g_s^n(Y_{s-}^n, z) \nabla Y_{s-}^n \tilde{N}(ds, dz) \\ &\quad + \int_0^t \int_{|z| \geq r_0} \nabla_y g_s^n(Y_{s-}^n, z) \nabla Y_{s-}^n N(ds, dz). \end{aligned} \quad (3.20)$$

By the Burkholder-Davis-Gundy inequality [9, Theorem 2.11] and (3.8), (3.13), we have for  $p \geq 2$ ,

$$\mathbb{E} \left[ \sup_{s \in [0, t]} |\nabla Y_s^n|^p \right] \leq |\nabla \Phi_0^n(x)|^p + \int_0^t \mathbb{E} |\nabla a_s^n(Y_s^n) \nabla Y_s^n|^p ds + \mathbb{E} \left[ \int_0^t \int_{|z| < r_0} |\nabla_y g_s^n(Y_s^n, z) \nabla Y_s^n|^2 \nu(dz) ds \right]^{p/2}$$

$$\begin{aligned}
& + \mathbb{E} \left[ \int_0^t \int_{|z| \geq r_0} |\nabla_y g_s^n(Y_s^n, z) \nabla Y_s^n|^p \nu(dz) ds \right]^p + \mathbb{E} \left[ \int_0^t \int_{\mathbb{R}^d} |\nabla_y g_s^n(Y_s^n, z) \nabla Y_s^n|^p \nu(dz) ds \right] \\
& \leq 1 + \left( 1 + \left( \int_{|z| < r_0} |z|^{2\gamma} \nu(dz) \right)^{p/2} \right) \int_0^t \mathbb{E} |\nabla Y_s^n|^p ds,
\end{aligned}$$

which gives (3.18) by Gronwall's inequality.

Next, set  $U_t^n := \nabla Y_t^n - \nabla Y_t$ . By equation (3.20), (3.8), (3.13) and [9, Theorem 2.11],

$$\mathbb{E} \left[ \sup_{s \in [0, t]} |U_s^n|^p \right] \leq h_n + \int_0^t \mathbb{E} |U_s^n|^p ds,$$

where

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

By Gronwall's inequality, (3.8), (3.9), (3.13), (3.14) and (3.17), it is easy to see that

$$\lim_{n \rightarrow 0} \mathbb{E} \left[ \sup_{t \in [0, 1]} |U_t^n|^p \right] \leq \lim_{n \rightarrow 0} h_n = 0.$$

The proof is complete.  $\square$

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 (3.12) and (3.7), respectively. By Lemmas 3.2 and 3.3, we have

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

and

$$\int_{|z| \leq r_0} |g_s(y, z) - g_s(y', z)|^2 \nu(dz) \leq C_2^2 |y - y'|^2 \int_{|z| \leq r_0} |z|^{2\gamma} \nu(dz) \stackrel{(1.5)}{\leq} C |y - y'|^2.$$

Hence, (3.16) has a unique strong solution by the classical result (cf. [6, Theorem IV.9.1]). (1.7) follows from  $X_t(x) = \Phi_t^{-1}(Y_t(\Phi_0(x)))$  and (3.18). Moreover, let  $Y_t(y)$  be the solution of SDE (3.16) with starting point  $y$ . By (3.10) and the choice of  $r_0$  in (3.11),  $\{Y_t(y), t \in [0, 1], y \in \mathbb{R}^d\}$  defines a  $C^1$ -stochastic diffeomorphism flow (cf. [10, 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 (3.15). 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 (3.6), we also have

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

From this, we immediately obtain the desired continuity.  $\square$

*Proof of Theorem 1.5.* First of all, we show that the right hand side of (1.10) is no bigger than the right hand side of (1.11). By Hölder's inequality, it suffices to show that for any  $p > 1$ ,

$$I(t) := \mathbb{E} \left[ \frac{1}{S_t^p} \left| \int_0^t \nabla X_s(x) dW_{S_s} \right|^p \right] \leq C t^{-p/\alpha}.$$

By [21, (2.11)], one has

$$I(t) \leq \mathbb{E} \left[ \frac{1}{S_t^p} \left( \int_0^t |\nabla X_s(x)|^2 dS_s \right)^{p/2} \right] \leq \mathbb{E} \left[ \frac{1}{S_t^{p/2}} \sup_{s \in [0,1]} |\nabla X_s(x)|^p \right] \stackrel{(1.7)}{\leq} \left( \mathbb{E} [S_t^{-p}] \right)^{1/2} \stackrel{(4.3)}{\leq} t^{-p/\alpha}.$$

Let  $b^n$  be defined as in (2.22) and  $X^n$  be the unique solution to SDE (3.15). For  $f \in C_b^1(\mathbb{R}^d)$ , by [21, Theorem 1.1] or [17, Theorem 1.1], we have

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

Thus, in order to show formula (1.10), it suffices to show the following two relations:

$$\lim_{n \rightarrow \infty} \nabla \mathbb{E} f(X_t^n(x)) = \lim_{n \rightarrow \infty} \mathbb{E} \left[ (\nabla f)(X_t^n(x)) \nabla X_t^n(x) \right] = \mathbb{E} \left[ (\nabla f)(X_t(x)) \nabla X_t(x) \right] = \nabla \mathbb{E} f(X_t(x)) \quad (3.22)$$

and

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{f(X_t^n(x))}{S_t} \int_0^t \nabla X_s^n(x) dW_{S_s} \right] = \mathbb{E} \left[ \frac{f(X_t(x))}{S_t} \int_0^t \nabla X_s(x) dW_{S_s} \right]. \quad (3.23)$$

Notice that by (3.17) and (3.6),

$$\lim_{n \rightarrow \infty} \mathbb{E} [|X_t^n(x) - X_t(x)| \wedge 1] = 0. \quad (3.24)$$

(3.22) and (3.23) follow by (3.21), (3.24) and the dominated convergence theorem.  $\square$

#### 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  $d$ -dimensional Brownian motion with infinitesimal generator  $\Delta/2$  and  $S$  is a one-dimensional subordinator, which is independent of  $W_t$ . Let  $\phi(\lambda)$  be the Laplace exponent of  $S$ , i.e.,  $\mathbb{E} e^{-\lambda S_t} = e^{-t\phi(\lambda)}$ . If for some  $\alpha \in (0, 2)$ ,

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

then  $(\mathbf{H}_{\nu, K_0}^{\alpha, 1, 1})$  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  $\mathbb{R}^d$ ,

$$\nabla T_t^{\nu, 0} f(x) = \mathbb{E} \left[ f(x + W_{S_t}) \frac{W_{S_t}}{S_t} \right].$$

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

$$\begin{aligned} |\nabla T_t^{\nu, 0} f(x)| &= \left| \mathbb{E} \left[ \left( f(x + W_{S_t}) - f(x) \right) \frac{W_{S_t}}{S_t} \right] \right| \\ &\leq \Lambda_x \mathbb{E} \left[ \frac{|W_{S_t}|^{1+\beta}}{S_t} \right] \leq C \Lambda_x \mathbb{E} \left[ S_t^{-\frac{1-\beta}{2}} \right] \leq K_0 \Lambda_x t^{(\beta-1)/\alpha}, \end{aligned} \quad (4.2)$$

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

$$\mathbb{E} S_t^{-p} = \frac{1}{\Gamma(p)} \mathbb{E} \int_0^\infty \lambda^{p-1} e^{-\lambda S_t} d\lambda = \frac{1}{\Gamma(p)} \int_0^\infty \lambda^{p-1} e^{-t\phi(\lambda)} d\lambda$$



$$(4.1) \quad \frac{1}{\Gamma(p)} \left( \frac{1}{p} + \int_1^\infty \lambda^{p-1} e^{-Ct\lambda^{\alpha/2}} d\lambda \right) \leq Ct^{-\frac{2p}{\alpha}}, \quad t \in (0, 1]. \quad (4.3)$$

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 [2, (15)] that

$$\nu(dz) \leq \frac{c_0 \phi(|z|^{-2})}{|z|^d} dz.$$

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

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

then (1.5) 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  $\alpha$ -stable process in  $\mathbb{R}^d$ . 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  $\nu(dz) = \kappa(z)dz$ . Assume that for some  $0 < \alpha_1 \leq \alpha_2 < 2$ ,

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

We call a Lévy process satisfying the above condition of stable-type. In this case, we can make the following decomposition for  $\nu$ :

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

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

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

By Example 4.1,  $(\mathbf{H}_{\nu_1, K_0}^{\alpha_1, 1, 1})$  holds for some  $K_0 > 0$ . Condition (1.5) 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  $\alpha$ -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  $\alpha$ -stable-like Lévy process. Observe that relativistic  $\alpha$ -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 relativistic  $\alpha$ -stable process. The corresponding Lévy process can be called relativistic  $\alpha$ -stable-like.

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

$$Z_t = W_{S_t} := (W_{S_t^1}^1, \dots, W_{S_t^k}^k),$$

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

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

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

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

$$\nabla_i T_t^{\nu,0} f(x) = \mathbb{E} \left[ (S_t^i)^{-1} W_{S_t^i}^i f(x + W_{S_t^i}) \right].$$

Suppose  $\alpha \in (0, 1]$ . For  $\beta \in [0, \alpha]$  and  $x \in \mathbb{R}$ , if  $\Lambda_x := \sup_{y \in \mathbb{R}^d} |f(x+y) - f(x)|/|y|^\beta < \infty$ , then we have by (4.3) that for  $t \in (0, 1]$ ,

$$\begin{aligned} |\nabla_i T_t^{\nu,0} f(x)| &= \left| \mathbb{E} \left[ (S_t^i)^{-1} W_{S_t^i}^i (f(x + W_{S_t^i}) - f(x)) \right] \right| \\ &\leq \Lambda_x \mathbb{E} \left[ (S_t^i)^{-1} |W_{S_t^i}^i| |W_{S_t^i}|^\beta \right] \\ &\leq \Lambda_x \left( \mathbb{E} \left[ (S_t^i)^{-1} |W_{S_t^i}^i|^{1+\beta} \right] + \mathbb{E} \left[ (S_t^i)^{-1} |W_{S_t^i}^i| \right] \sum_{j \neq i} \mathbb{E} \left[ |W_{S_t^j}^j|^\beta \right] \right) \\ &\leq C \Lambda_x \left( t^{(\beta-1)/\alpha_i} + t^{-1/\alpha_i} \sum_{j \neq i} t^{\beta/\alpha_j} \mathbb{E} \left[ |W_{S_t^j}^j|^\beta \right] \right) \\ &\leq K_0 \Lambda_x \left( t^{(\beta-1)/\alpha} + t^{\beta/\alpha_{\max}-1/\alpha} \right) \quad (\text{since } \beta < \alpha) \\ &\leq K_0 \Lambda_x t^{(\alpha\beta/\alpha_{\max}-1)/\alpha} = K_0 \Lambda_x t^{(\delta\beta-1)/\alpha}, \end{aligned}$$

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

$$|\nabla_i T_t^{\nu,0} f(x)| \leq \|f\|_\infty \mathbb{E} \left[ |W_{S_t^i}^i| / (S_t^i) \right] \leq \|f\|_\infty \mathbb{E} \left[ (S_t^i)^{-1/2} \right] \leq \|f\|_\infty t^{-1/\alpha_i} \leq K_0 \|f\|_\infty t^{-1/\alpha}.$$

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

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

Note that (1.5) holds for any  $\gamma > \alpha_{\max}/2$ . For Theorem 1.2 to be valid, the following constraint needs to be satisfied:

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

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  $\alpha_i$ 's to satisfy

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

and take

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

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  $(\mathbf{H}_{\nu, K_0}^{\alpha, 1, 1})$  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  $\alpha$ -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.

**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  $t \in [0, 1]$ ,  $x, y, z \in \mathbb{R}^d$ ,*

$$|a_t^n(x) - a_t^n(y)| \leq C|x - y|, \quad |g_t^n(x, z) - g_t^n(y, z)| \leq C|x - y|h(z), \quad (5.1)$$

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

$$\lim_{n \rightarrow \infty} a_t^n(x) = a_t^\infty(x), \quad \lim_{n \rightarrow \infty} g_t^n(x, z) = g_t^\infty(x, z) \quad (5.2)$$

and

$$\sup_{n \in \mathbb{N}_\infty} \sup_{t \in [0, 1]} \sup_{x \in \mathbb{R}^d} \left( \frac{|a_t^n(x)|}{1 + |x|} + \sup_{0 < |z| \leq r} \frac{|g_t^n(x, z)|}{|z|} \right) < \infty. \quad (5.3)$$

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

$$Y_t^n = \xi_n + \int_0^t a_s^n(Y_s^n) ds + \int_0^t \int_{|z| \leq r} g_s^n(Y_{s-}^n, z) \tilde{N}(ds, dz) + \int_0^t \int_{|z| > r} g_s^n(Y_{s-}^n, z) N(ds, dz).$$

Assume that  $\xi_n$  converges to  $\xi_\infty$  in probability as  $n \rightarrow \infty$ . Then we have

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \sup_{t \in [0, 1]} |Y_t^n - Y_t^\infty| \wedge 1 \right) = 0, \quad (5.4)$$

which implies that  $Y_t^n$  converges to  $Y_t^\infty$  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:*

$$f(x) = |x|^2 \text{ if } |x| \leq 1, \quad f(x) = 2 \text{ if } |x| \geq 2, \quad \text{and} \quad |\nabla f| + |\nabla^2 f| \leq C_1 \mathbf{1}_{\{|x| \leq 2\}}, \quad (5.5)$$

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  $|y| \leq C_2(|x| + \delta) \wedge 1$ ,

$$|y| |\nabla f(x + ry)| \leq C_3(f(x) + \delta), \quad |y|^2 |\nabla^2 f(x + ry)| \leq C_3(f(x) + \delta^2). \quad (5.6)$$

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

We also need the following key lemma.

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

$$Y_t^n = Y_{\tau_1}^n + \int_{\tau_1}^t a_s^n(Y_s^n) ds + \int_{\tau_1}^t \int_{|z| \leq r} g_s^n(Y_{s-}^n, z) \tilde{N}(ds, dz).$$

Assume that  $Y_{\tau_1}^n$  converges to  $Y_{\tau_1}^\infty$  in probability, then we have

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \sup_{t \in [\tau_1, \tau_2]} |Y_t^n - Y_t^\infty| \wedge 1 \right] = 0.$$

*Proof.* In this proof we will drop the superscript “ $\infty$ ” and write

$$U_s^n := Y_s^n - Y_s, \quad A_s^n := a_s^n(Y_s^n) - a_s(Y_s), \quad \Gamma_s^n(z) := g_s^n(Y_{s-}^n, z) - g_s(Y_{s-}, z).$$

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

$$\begin{aligned} f(U_t^n) &= f(U_{\tau_1}^n) + \int_{\tau_1}^t \langle A_s^n, \nabla f(U_s^n) \rangle ds + \int_{\tau_1}^t \int_{|z| \leq r} [f(U_{s-}^n + \Gamma_s^n(z)) - f(U_{s-}^n)] \tilde{N}(ds, dz) \\ &\quad + \int_{\tau_1}^t \int_{|z| \leq r} [f(U_{s-}^n + \Gamma_s^n(z)) - f(U_{s-}^n) - \Gamma_s^n(z) \cdot \nabla f(U_{s-}^n)] \nu(dz) ds. \end{aligned}$$

For  $R > 0$ , define a stopping time

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

For any  $T \in [0, 1]$ , by the Burkholder-Davis-Gundy inequality [9, Theorem 2.11], we have

$$\begin{aligned} &\mathbb{E} \left( \sup_{t \in [\tau_1, T \wedge \tau_R]} |f(U_t^n)|^2 \right) \\ &\leq \mathbb{E} |f(U_{\tau_1}^n)|^2 + \mathbb{E} \int_{\tau_1}^{T \wedge \tau_R} |\langle A_s^n, \nabla f(U_s^n) \rangle|^2 ds \\ &\quad + \mathbb{E} \int_{\tau_1}^{T \wedge \tau_R} \int_{|z| \leq r} |f(U_{s-}^n + \Gamma_s^n(z)) - f(U_{s-}^n)|^2 \nu(dz) ds \\ &\quad + \mathbb{E} \int_{\tau_1}^{T \wedge \tau_R} \left| \int_{|z| \leq r} [f(U_{s-}^n + \Gamma_s^n(z)) - f(U_{s-}^n) - \Gamma_s^n(z) \cdot \nabla f(U_{s-}^n)] \nu(dz) \right|^2 ds \\ &=: \mathbb{E} |f(U_{\tau_1}^n)|^2 + I_1^n + I_2^n + I_3^n. \end{aligned}$$

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

$$I_1^n \leq \mathbb{E} \int_{\tau_1}^{T \wedge \tau_R} |a_s^n(Y_s) - a_s(Y_s)|^2 ds + \mathbb{E} \int_{\tau_1}^{T \wedge \tau_R} |f(U_s^n)|^2 ds.$$

For  $I_2^n$  and  $I_3^n$ , noticing that by (5.1) and (5.3),

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

by (5.6), we have

$$\begin{aligned} I_2^n &\leq \mathbb{E} \int_{\tau_1}^{T \wedge \tau_R} \int_{|z| \leq r} |\Gamma_s^n(z)|^2 \left( \int_0^1 |\nabla f(U_{s-}^n + r\Gamma_s^n(z))|^2 dr \right) \nu(dz) ds \\ &\leq \mathbb{E} \int_{\tau_1}^{T \wedge \tau_R} \int_{|z| \leq r} |g_s^n(Y_s, z) - g_s(Y_s, z)|^2 \nu(dz) ds + \mathbb{E} \int_{\tau_1}^{T \wedge \tau_R} |f(U_s^n)|^2 ds, \end{aligned}$$

and

$$\begin{aligned} I_3^n &\leq \mathbb{E} \int_{\tau_1}^{T \wedge \tau_R} \left( \int_{|z| \leq r} |\Gamma_s^n(z)|^2 \left( \int_0^1 \int_0^1 |\nabla^2 f(U_{s-}^n + rr'\Gamma_s^n(z))| dr dr' \right) \nu(dz) \right)^2 ds \\ &\leq \mathbb{E} \int_{\tau_1}^{T \wedge \tau_R} \left( \int_{|z| \leq r} |g_s^n(Y_s, z) - g_s(Y_s, z)|^2 \nu(dz) \right)^2 ds + \mathbb{E} \int_{\tau_1}^{T \wedge \tau_R} |f(U_s^n)|^2 ds. \end{aligned}$$

Combining the above calculations, we obtain

$$\mathbb{E} \left[ \sup_{t \in [\tau_1, T \wedge \tau_R]} |f(U_t^n)|^2 \right] \leq h_n + \mathbb{E} \int_{\tau_1}^{T \wedge \tau_R} |f(U_s^n)|^2 ds,$$

where

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

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

$$\lim_{n \rightarrow 0} \mathbb{E} \left[ \sup_{t \in [\tau_1, \tau_R]} |f(U_t^n)|^2 \right] \leq \lim_{n \rightarrow 0} Ch_n = 0.$$

In particular,

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

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

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  $\lim_{m \rightarrow \infty} \tau_m = 1$ . Clearly, for  $t \in (\tau_m, \tau_{m+1}]$ ,  $Y_t^n$  satisfies

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

where  $Y_{\tau_m}^n := Y_{\tau_m-}^n + g_{\tau_m}^n(Y_{\tau_m-}^n, Z_{\tau_m} - Z_{\tau_m-})$ . Since  $\xi_n \rightarrow \xi_\infty$  in probability as  $n \rightarrow \infty$ , by Lemma 5.3 and induction, we have for each  $m \in \mathbb{N}$ ,

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

which gives the desired limit by noticing that for any  $m_0 \in \mathbb{N}$ ,

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

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

**Acknowledgement.** We thank Zenghu Li, Feng-Yu Wang and Jian Wang for helpful discussions.

## REFERENCES

- [1] Applebaum, D.: *Lévy Processes and Stochastic Calculus*. Cambridge Studies in Advance Mathematics 93, Cambridge University Press, 2004.
- [2] Bogdan, K., Grzywny, T. and Ryznar, M.: Density and tails of unimodal convolution semigroups. *J. Funct. Anal.* **266** (2014), 3543–3571.
- [3] Chen, Z.-Q. and Zhang, X.: Heat kernels and analyticity of non-symmetric jump diffusion semigroups. Preprint. arXiv:1306.5015v2 [math.AP]
- [4] Fedrizzi, E. and Flandoli F.: Pathwise uniqueness and continuous dependence for SDEs with nonregular drift. arXiv:1004.3485v1.

- [5] Flandoli, F., Gubinelli, M. and Priola, E.: Well-posedness of the transport equation by stochastic perturbation. *Inven. Math.* **180** (2010), 1-53.
- [6] Ikeda, N., Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes*, 2nd ed.. North-Holland/Kodansha, Amsterdam/Tokyo, 1989.
- [7] Krylov, N.V.: *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*. Graduate Studies in Mathematics, Vol. 12, AMS, 1996.
- [8] Krylov, N.V. and Röckner, M.: Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Relat. Fields* **131** (2005), 154-196.
- [9] Kunita, H.: Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In *Real and Stochastic Analysis*, 305–373. Birkhäuser, Boston, 2004.
- [10] Priola, E.: Pathwise uniqueness for singular SDEs driven by stable processes. *Osaka Journal of Mathematics*, **49** (2012), 421-447.
- [11] Priola, E.: Stochastic flow for SDEs with jumps and irregular drift term. arXiv:1405.2575v1.
- [12] Protter, P.: *Stochastic Integration and Differential Equations*. 2nd ed., Springer-Verlag, Berlin, 2004.
- [13] Silvestre, L.: On the differentiability of the solution to an equation with drift and fractional diffusion. *Indiana Univ. Math. J.* **61** (2012), 557-584.
- [14] Stein, E.M.: *Singular Integrals and Differentiability Properties of Functions*. Princeton, N.J., Princeton University Press, 1970.
- [15] Tanaka, H., Tsuchiya, M. and Watanabe, S.: Perturbation of drift-type for Lévy processes. *J. Math. Kyoto Univ.* **14** (1974), 73-92.
- [16] Veretennikov, A. Ju.: On the strong solutions of stochastic differential equations. *Theory Probab. Appl.* **24** (1979), 354-366.
- [17] Wang, F., Xu, L. and Zhang, X.: Gradient estimates for SDEs driven by multiplicative Lévy noise. arXiv:1301.4528.
- [18] Zhang, X.: Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.* **16** (2011), 1096-1116.
- [19] Zhang, X.: Stochastic functional differential equations driven by Lévy processes and quasi-linear partial integro-differential equations. *Ann. Appl. Probab.* **22** (2012), 2505-2538.
- [20] Zhang, X.: Stochastic differential equations with Sobolev drifts and driven by  $\alpha$ -stable processes. *Ann. Inst. H. Poincaré Probab. Statist.* **49** (2013), 1057-1079.
- [21] Zhang, X.: Derivative formula and gradient estimate for SDEs driven by  $\alpha$ -stable processes. *Stoch. Proc. Appl.* **123** (2013), 1213-1228.
- [22] Zvonkin, A.K.: A transformation of the phase space of a diffusion process that removes the drift. *Mat. Sbornik*, **93 (135)** (1974), 129-149.

ZHEN-QING CHEN: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195, USA  
 EMAIL: ZQCHEN@UW.EDU

RENMING SONG: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS, URBANA, IL 61801, USA  
 EMAIL: RSONG@MATH.UIUC.EDU

XICHENG ZHANG: SCHOOL OF MATHEMATICS AND STATISTICS, WUHAN UNIVERSITY, WUHAN, HUBEI 430072, P.R.CHINA  
 EMAIL: XICHENGZHANG@GMAIL.COM