The Annals of Probability
2017, Vol. 45, No. 3, 1799–1841
DOI: 10.1214/16-AOP1102

FUNDAMENTAL SOLUTIONS OF NONLOCAL HÖRMANDER'S OPERATORS II[1]

BY XICHENG ZHANG

Dedicated to the memory of Professor Paul Malliavin

Wuhan University

Consider the following nonlocal integro-differential operator: for $\alpha \in (0,2)$:

$$\mathcal{L}^{(\alpha)}_{\sigma,b} f(x) := \text{p.v.} \int_{|z|<\delta} \frac{f(x+\sigma(x)z) - f(x)}{|z|^{d+\alpha}} \, \mathrm{d}z + b(x) \cdot \nabla f(x) + \mathscr{L} f(x),$$

where $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ are smooth functions and have bounded partial derivatives of all orders greater than 1, δ is a small positive number, p.v. stands for the Cauchy principal value and $\mathscr{L}$ is a bounded linear operator in Sobolev spaces. Let $B_1(x) := \sigma(x)$ and $B_{j+1}(x) := b(x) \cdot \nabla B_j(x) - \nabla b(x) \cdot B_j(x)$ for $j \in \mathbb{N}$. Suppose $B_j \in C_b^\infty(\mathbb{R}^d; \mathbb{R}^d \otimes \mathbb{R}^d)$ for each $j \in \mathbb{N}$. Under the following uniform Hörmander's type condition: for some $j_0 \in \mathbb{N}$,

$$\inf_{x \in \mathbb{R}^d} \inf_{|u|=1} \sum_{j=1}^{j_0} |u B_j(x)|^2 > 0,$$

by using Bismut's approach to the Malliavin calculus with jumps, we prove the existence of fundamental solutions to operator $\mathcal{L}^{(\alpha)}_{\sigma,b}$. In particular, we answer a question proposed by Nualart [*Sankhyā A* **73** (2011) 46–49] and Varadhan [*Sankhyā A* **73** (2011) 50–51].

1. Introduction and main results. Consider the following nonlocal (integro-differential) operator: for $\alpha \in (0,2)$:

$$\mathcal{L}^{(\alpha)}_{\sigma,b} f(x) := \text{p.v.} \int_{\mathbb{R}^d_0} \frac{f(x+\sigma(x)z) - f(x)}{|z|^{d+\alpha}} \, \mathrm{d}z + b(x) \cdot \nabla f(x), \tag{1.1}$$

where $\mathbb{R}^d_0 := \mathbb{R}^d - \{0\}$, and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ are two smooth functions and have bounded k-order partial derivatives for all $k \geq 1$. Define a family of matrix-valued functions $B_j(x)$, $j \in \mathbb{N}$ recursively as follows:

$$B_1(x) := \sigma(x) \quad \text{and} \quad B_{j+1}(x) := b(x) \cdot \nabla B_j(x) - \nabla b(x) \cdot B_j(x) \qquad \text{for } j \in \mathbb{N}.$$

Received October 2014; revised January 2016.

[1]Supported by NNSFs of China (Nos. 11271294, 11325105).

MSC2010 subject classifications. 60H07, 60H10, 60H30.

Key words and phrases. Hörmander condition, fundamental solution, Malliavin calculus with jumps, nonlocal kinetic Fokker–Planck operator, Poisson functional.

Recently, in previous works [25] and [26], we have proved that if for each $x \in \mathbb{R}^d$, there is a $n(x) \in \mathbb{N}$ such that

$$\operatorname{Rank}\big[B_1(x), \ldots, B_{n(x)}(x)\big] = d,$$

then the heat kernel of operator $\mathcal{L}^{(\alpha)}_{\sigma,b}$ denoted by $\rho_t(x, y)$ exists, and as a function of y, the mapping

$$(0, \infty) \times \mathbb{R}^d \ni (t, x) \mapsto \rho_t(x, \cdot) \in L^1\big(\mathbb{R}^d\big)$$

is continuous. Moreover, when $\sigma(x) = \sigma$ is *constant* and $B_j \in C^\infty_b$ for each $j \in \mathbb{N}$, under the following uniform Hörmander's type condition: for some $j_0 \in \mathbb{N}$,

$$\inf_{x\in\mathbb{R}^d} \inf_{|u|=1} \sum_{j=1}^{j_0} \big|uB_j(x)\big|^2 > 0,$$

we also show the smoothness of $(t, x, y) \mapsto \rho_t(x, y)$. The proofs in [25] and [26] are based on the Malliavin calculus to the subordinate Brownian motion (cf. [15]). More precisely, let us consider the following stochastic differential equation (abbreviated as SDE):

$$\mathrm{d}X_t(x) = b\big(X_t(x)\big)\,\mathrm{d}t + \sigma\big(X_{t-}(x)\big)\,\mathrm{d}W_{S_t}, \qquad X_0(x) = x, \tag{1.2}$$

where W_t is a d-dimensional Brownian motion and S_t is an independent $\alpha/2$-stable subordinator. It is well known that the generator of Markov process $X_t(x)$ is given by $\mathcal{L}^{(\alpha)}_{\sigma,b}$. Thus, the main purpose is to study the existence and smoothness of the distribution density $\rho_t(x, y)$ of $X_t(x)$, where the key point in [26] is to regard $X_t(x)$ as a functional of Brownian motion due to the independence of W and S. If $\sigma(x)$ depends on x, since the solution $\{X_t(x), x \in \mathbb{R}^d\}$ of SDE (1.2) does not form a stochastic diffeomorphism flow in general (cf. [19]), it seems hard to prove the smoothness of $\rho_t(x, y)$ in the framework of [26]. In this work, we shall study the smoothness of $\rho_t(x, y)$ for *nonconstant* coefficient $\sigma(x)$ in a different framework.

As far as we know, Bismut [4] first used Girsanov's transformation to study the smoothness of distribution densities to SDEs with jumps. Later, in the monograph [3], Bichteler, Gravereaux and Jacod systematically developed the Malliavin calculus with jumps and studied the smooth density for SDEs driven by nondegenerate jump noises. In [18], Picard used difference operator to give another criterion for the smoothness of the distribution density of Poisson functionals, and also applied it to SDEs driven by pure jump Lévy processes. By combining the classical Malliavin calculus and Picard's difference operator argument, Ishikawa and Kunita [10] obtained a new criterion for the smooth density of Wiener–Poisson functionals (see also [13]). On the other hand, Cass [5] established a Hörmander's-type theorem for SDEs with jumps by proving a Norris' type lemma for discontinuous semimartingales, but the Brownian diffusion term cannot disappear. In the pure jump degenerate case, by using a Komatsu–Takeuchi's estimate proven in [11] for discontinuous semimartingales, Takeuchi [22] and Kunita [12, 14] also obtained similar Hörmander's theorems. However, their results do not cover operator (1.1). More discussions about their results can be found in [26].

1.1. *Statement of main results*. For $\delta > 0$, define $\Gamma_0^\delta := \{z \in \mathbb{R}^d : 0 < |z| < \delta\}$. Let us now consider the following nonlocal (integro-differential) operator:

$$\mathcal{L}_0 f(x) := \text{p.v.} \int_{\Gamma_0^\delta} \big(f\big(x + \sigma(x,z)\big) - f(x)\big)\nu(\mathrm{d}z) + b(x) \cdot \nabla f(x), \tag{1.3}$$

where $\sigma(x,z) : \mathbb{R}^d \times \Gamma_0^\delta \to \mathbb{R}^d$ and $b(x) : \mathbb{R}^d \to \mathbb{R}^d$ are Borel measurable functions, and $\nu(\mathrm{d}z)$ is a Lévy measure on Γ_0^δ and satisfies

$$\int_{\varepsilon<|z|<\delta} \sigma(x,z)\nu(\mathrm{d}z) = 0 \qquad \forall \varepsilon \in (0,\delta), x \in \mathbb{R}^d. \tag{1.4}$$

Here, p.v. stands for the Cauchy principal value. Notice that (1.4) is a common assumption in the study of nonlocal operators, which is some symmetric requirement (cf. [7, 8]). For example, if ν is symmetric and $\sigma(x,-z) = -\sigma(x,z)$, then (1.4) holds.

Let $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$. We make the following assumptions:

(H_b^σ) b and σ are smooth and for any $k \in \mathbb{N}$, $m, j \in \mathbb{N}_0$ and some constants $C_k, C_{mj} \geq 1$,

$$\big|\nabla^k b(x)\big| \leq C_k, \qquad \big|\nabla_x^m \nabla_z^j \sigma(x,z)\big| \leq C_{mj}|z|^{(1-j)\vee 0}. \tag{1.5}$$

(H^ν) $\nu(\mathrm{d}z)|_{\Gamma_0^\delta} = \kappa(z)\,\mathrm{d}z|_{\Gamma_0^\delta}$, where $\kappa \in C^\infty(\Gamma_0^\delta; (0,\infty))$ satisfies the following order condition: for some $\alpha \in (0,2)$,

$$\lim_{\varepsilon \downarrow 0} \varepsilon^{\alpha-2} \int_{|z|\leq\varepsilon} |z|^2 \kappa(z)\,\mathrm{d}z =: c_1 > 0, \tag{1.6}$$

and bounded condition: for any $m \in \mathbb{N}$ and some $C_m \geq 1$,

$$\big|\nabla^m \log \kappa(z)\big| \leq C_m |z|^{-m}, \qquad z \in \Gamma_0^{\frac{\delta}{2}}. \tag{1.7}$$

(H^{j_0}) Let $B_1(x) := \nabla_z \sigma(x,0)$ and define $B_{j+1}(x) := b(x) \cdot \nabla B_j(x) - \nabla b(x) \cdot B_j(x)$ for $j \in \mathbb{N}$. Assume that for some $j_0 \in \mathbb{N}$, $m \in \mathbb{N}_0$, $c_0 > 0$ and all $x \in \mathbb{R}^d$,

$$\inf_{|u|=1} \sum_{j=1}^{j_0} \big|u B_j(x)\big|^2 \geq c_0 / \big(1 + |x|^m\big). \tag{1.8}$$

The first aim of this paper is to prove the following Hörmander's type theorem.

THEOREM 1.1. *Under* (H_b^σ), (H^ν) *and* (H^{j_0}), *if* $\delta < \frac{1}{2C_{10}}$, *where* C_{10} *is the same as in* (1.5), *then there exists a nonnegative smooth function* $\rho_t(x,y)$ *on* $(0,\infty) \times \mathbb{R}^d \times \mathbb{R}^d$ *called fundamental solution of operator* $\mathcal{L}_0$ *such that*

$$\partial_t \rho_t(x,y) = \mathcal{L}_0 \rho_t(\cdot,y)(x) \qquad \forall (t,x,y) \in (0,\infty) \times \mathbb{R}^d \times \mathbb{R}^d.$$

Let us briefly introduce the strategy of proving this theorem. Let $N(\mathrm{d}t, \mathrm{d}z)$ be a Poisson random measure with intensity $\mathrm{d}t\nu(\mathrm{d}z)$, and $\tilde{N}(\mathrm{d}t, \mathrm{d}z) := N(\mathrm{d}t, \mathrm{d}z) - \mathrm{d}t\nu(\mathrm{d}z)$ the compensated Poisson random measure. Consider the following SDE:

$$X_t(x) = x + \int_0^t b(X_s(x))\,\mathrm{d}s + \int_0^t \int_{\Gamma_0^\delta} \sigma(X_{s-}(x), z)\tilde{N}(\mathrm{d}s, \mathrm{d}z). \tag{1.9}$$

Under (H_b^σ), it is well known that the above SDE has a unique solution $X_t(x)$, which defines a Markov process with generator $\mathcal{L}_0$. Let $\mathcal{T}_t^0 f(x) := \mathbb{E} f(X_t(x))$. Our aim is to show that under (1.5)–(1.8), $X_t(x)$ admits a smooth density, which, by Sobolev's embedding theorem, will be a consequence of the following gradient-type estimate: for any $m, k \in \mathbb{N}_0$ and $f \in C_b^\infty(\mathbb{R}^d)$,

$$\left|\nabla^m \mathcal{T}_t^0 \nabla^k f(x)\right| \le C(x)(t \wedge 1)^{-\gamma_{mk}} \|f\|_\infty, \tag{1.10}$$

where ∇ stands for the gradient operator and $\gamma_{mk} > 0$. This will be achieved by using Bismut's approach to the Malliavin calculus with jumps, where the core task is to prove the L^p-integrability of the inverse of the reduced Malliavin matrix (see Section 3 and [16]).

However, the above result requires that δ is smaller than $1/(2C_{10})$ so that SDE (1.9) defines a C^∞-stochastic diffeomorphism flow. In other words, the large jump is not allowed in (1.9) or $\mathcal{L}_0$. The following result provides a way to treat large jumps under stronger assumptions.

THEOREM 1.2. *Let $\mathscr{L}$ be a bounded linear operator in Sobolev space $\mathbb{W}^{k,p}(\mathbb{R}^d)$ for any $p > 1$ and $k \in \mathbb{N}_0$. Under (H_b^σ), (H^ν) and (H^{j_0}) with $m = 0$ and $B_j \in C_b^2$ for each $j = 1, \dots, j_0 + 1$, if $\delta < \frac{1}{2C_{10}}$, where C_{10} is the same as in (1.5), then there exists a continuous function $\rho_t(x, y)$ on $(0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d$ called fundamental solution of operator $\mathcal{L}_0 + \mathscr{L}$ with the properties that*:

(i) *For each $t > 0$ and $y \in \mathbb{R}^d$, the mapping $x \mapsto \rho_t(x, y)$ is smooth, and for any $T > 0$, there is a $\gamma = \gamma(\alpha, j_0, d) > 0$ such that for any $p \in (1, \infty)$ and $k \in \mathbb{N}_0$,*

$$\left\|\nabla_x^k \rho_t(x, \cdot)\right\|_p \le C t^{-(k+d)\gamma} \qquad \forall (t, x) \in (0, T] \times \mathbb{R}^d. \tag{1.11}$$

(ii) *For any $p \in (1, \infty)$ and $\varphi \in L^p(\mathbb{R}^d)$, $\mathcal{T}_t\varphi(x) := \int_{\mathbb{R}^d} \varphi(y)\rho_t(x, y)\,\mathrm{d}y \in \bigcap_k \mathbb{W}^{k,p}(\mathbb{R}^d)$ satisfies*

$$\partial_t \mathcal{T}_t\varphi(x) = (\mathcal{L}_0 + \mathscr{L})\mathcal{T}_t\varphi(x) \qquad \forall (t, x) \in (0, \infty) \times \mathbb{R}^d. \tag{1.12}$$

The idea of proving this result is as follows: Let $\mathcal{T}_t$ be the semigroup associated with operator $\mathcal{L}_0 + \mathscr{L}$. By Duhamel's formula, we can formally write

$$\mathcal{T}_t\varphi = \mathcal{T}_t^0\varphi + \int_0^t \mathcal{T}_{t-s}^0 \mathscr{L}\mathcal{T}_s\varphi\,\mathrm{d}s.$$

Using similar short-time estimate as in (1.10) and suitable interpolation techniques, we shall prove some gradient estimates in Soboelv spaces for $\mathcal{T}_t\varphi$, which in turn yields the desired results by Sobolev's embedding theorem as above. In order to obtain the gradient estimate for $\mathcal{T}_t^0$ in L^p-spaces, we need to assume $m=0$ and $B_j \in C_b^2$ for each $j=1,\dots,j_0+1$ in (H^{j_0}).

In this result, the operator $\mathscr{L}$ is designed to treat the large jump part as expected. To illustrate this point, we present an application of Theorem 1.2. Consider operator $\mathcal{L}_{\sigma,b}^{(\alpha)}$ in (1.1) with σ taking the following special form:

$$\sigma(x) = \begin{pmatrix} 0_{d_1\times d_1}, & 0_{d_1\times d_2} \\ 0_{d_2\times d_1}, & \sigma_0(x) \end{pmatrix}, \tag{1.13}$$

where $d_1+d_2=d$ and $\sigma_0(x)$ is an invertible $d_2\times d_2$-matrix-valued function.

COROLLARY 1.3. *Assume that σ_0, b are smooth and have bounded partial derivatives of all orders greater than* 1, *and* (H^{j_0}) *holds with $m=0$ and $B_j \in C_b^2$ for each $j=1,\dots,j_0+1$. If σ_0 is invertible and satisfies*

$$\|\sigma_0^{-1}\|_\infty < \infty, \tag{1.14}$$

then the conclusions in Theorem 1.2 *hold for operator $\mathcal{L}_{\sigma,b}^{(\alpha)}$.*

PROOF. For $\delta \in (0, \frac{1}{2\|\nabla\sigma_0\|_\infty})$, let $\chi_\delta : [0,\infty) \to [0,1]$ be a smooth function with

$$\chi_\delta(x) = 1, \qquad x \in \left[0, \frac{\delta}{2}\right], \qquad \chi_\delta(x) = 0, \qquad x \in [\delta, \infty).$$

We make the following decomposition:

$$\mathcal{L}_{\sigma,b}^{(\alpha)} f(x) = \mathcal{L}_0 f(x) + \mathscr{L} f(x),$$

where

$$\mathcal{L}_0 f(x) := \mathrm{p.v.} \int_{\Gamma_0^\delta} \frac{f(x+\sigma(x)z) - f(x)}{|z|^{d+\alpha}} \chi_\delta(|z|)\,\mathrm{d}z + b(x)\cdot\nabla f(x)$$

and

$$\mathscr{L} f(x) := \int_{\mathbb{R}_0^d} \frac{f(x+\sigma(x)z) - f(x)}{|z|^{d+\alpha}} \big(1-\chi_\delta(|z|)\big)\,\mathrm{d}z. \tag{1.15}$$

CLAIM. *$\mathscr{L}$ is a bounded linear operator in Sobolev spaces $\mathbb{W}^{k,p}(\mathbb{R}^d)$ for each $p>1$ and $k \in \mathbb{N}_0$.*

PROOF. Let $z=(z_1,z_2)$ with $z_1 \in \mathbb{R}^{d_1}$ and $z_2 \in \mathbb{R}^{d_2}$. Define

$$\xi(x,z) := \big(z_1, \sigma_0^{-1}(x) z_2\big) \in \mathbb{R}^d.$$

Notice that by (1.14), there is a positive constant $c_0>0$ such that for all x, z,

$$c_0|z| \le \big|\xi(x,z)\big| \le c_0^{-1}|z|.$$

By the change of variables, we have

$$\begin{aligned}\mathscr{L}f(x) &= \int_{\mathbb{R}^d_0}\big(f\big(x+(0,z_2)\big)-f(x)\big)\frac{1-\chi_\delta(|\xi(x,z)|)}{|\xi(x,z)|^{d+\alpha}}\det\big(\sigma_0^{-1}(x)\big)\,\mathrm{d}z \\ &= \int_{|z|>\frac{\delta}{2c_0}}\big(f\big(x+(0,z_2)\big)-f(x)\big)\frac{1-\chi_\delta(|\xi(x,z)|)}{|\xi(x,z)|^{d+\alpha}}\det\big(\sigma_0^{-1}(x)\big)\,\mathrm{d}z.\end{aligned} \tag{1.16}$$

Thus, by Minkovskii's inequality, we have

$$\|\mathscr{L}f\|_p \le \int_{|z|>\frac{\delta}{2c_0}}\big\|f\big(\cdot+(0,z_2)\big)-f(\cdot)\big\|_p\frac{\|\det(\sigma_0^{-1})\|_\infty}{(c_0|z|)^{d+\alpha}}\,\mathrm{d}z \le C\|f\|_p,$$

which shows the claim for $k=0$. For $k\ge 1$, starting from (1.16) and by the chain rule and cumbersome calculations, one sees that

$$\big\|\nabla^k\mathscr{L}f\big\|_p \le C\sum_{j=0}^k\big\|\nabla^j f\big\|_p.$$

The proof of the claim is thus complete. $\square$

Moreover, if we let $\kappa(z):=\chi_\delta(|z|)|z|^{-d-\alpha}$, then it is easy to check that (1.7) is true. Thus, we can use Theorem 1.2 to deduce the result. $\square$

It is noticed that in Corollary 1.3, σ is required to take a special form (1.13), which plays a crucial role in showing the boundedness of $\mathscr{L}$ defined by (1.15) in $\mathbb{W}^{k,p}$-space. Without assuming (1.13), due to the non-invertibility of $x\mapsto x+\sigma(x)z$, it seems hard to show that the operator $\mathscr{L}$ in (1.15) is bounded in $\mathbb{W}^{k,p}$-space. Consider the following operator:

$$\mathcal{L}^\nu_{\sigma,b}f(x) := \text{p.v.}\int_{\mathbb{R}^d_0}\big(f\big(x+\sigma(x,z)\big)-f(x)\big)\nu(\mathrm{d}z)+b(x)\cdot\nabla f(x),$$

where σ, b and ν are as above. Let $\mathcal{T}_t$ be the corresponding semigroup associated to $\mathcal{L}^\nu_{\sigma,b}$. Instead of working in the Sobolev space, if we consider the usual Hölder space $\mathbb{C}^\beta := \mathbb{H}^{\beta,\infty}$ [see (5.1) below for a definition], then we have the following.

THEOREM 1.4. *Under* (H^σ_b), (H^ν) *and* (H^{j_0}) *with* $m=0$ *and* $B_j\in C^2_b$ *for each* $j=1,\dots,j_0+1$, *if* $\int_{|z|\ge 1}|z|^q\nu(\mathrm{d}z)<\infty$ *for some* $q>0$, *then there exists a probability density function* $\rho_t(x,y)$ *such that for any* $\varphi\in L^\infty(\mathbb{R}^d)$, $\mathcal{T}_t\varphi(x)=\int_{\mathbb{R}^d}\varphi(y)\rho_t(x,y)\,\mathrm{d}y$ *belongs to* $\mathbb{C}^{q+\varepsilon}$, *where* $\varepsilon\in(0,1)$ *only depends on* α, j_0, d *with* α *from* (1.6) *and* j_0 *from* (1.8). *Moreover, if* $\alpha<q+\varepsilon$, *then* $\partial_t\mathcal{T}_t\varphi(x)=\mathcal{L}^\nu_{\sigma,b}\mathcal{T}_t\varphi(x)$ *for all* $(t,x)\in(0,\infty)\times\mathbb{R}^d$.

As a corollary, we have the following.

COROLLARY 1.5. *Consider operator $\mathcal{L}^{(\alpha)}_{\sigma,b}$ in (1.1). Assume that σ and b are smooth and have bounded partial derivatives of all orders greater than* 1, *and* (H^{j_0}) *holds with $m=0$ and $B_j \in C^2_b$ for each $j=1,\dots,j_0+1$. For some $\varepsilon>0$ and any $\varphi \in L^\infty(\mathbb{R}^d)$, $\mathcal{T}_t\varphi(x)=\mathbb{E}\varphi(X_t(x)) \in \mathbb{C}^{\alpha+\varepsilon}$ is a classical solution of equation*

$$\partial_t f(t,x)=\mathcal{L}^{(\alpha)}_{\sigma,b} f(t,x), \qquad (t,x)\in(0,\infty)\times\mathbb{R}^d.$$

Compared with Corollary 1.3, in this corollary, σ is not assume to take form (1.13). The price we have to pay is that the regularity of $\mathcal{T}_t\varphi$ depends on the moment of the Lévy measure $\nu(\mathrm{d}z)=\mathrm{d}z/|z|^{d+\alpha}$.

1.2. *Examples.* In this subsection, we provide several examples to illustrate our main results.

EXAMPLE 1.6. Let $b:\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma(x):\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^d$ be two C^∞_b-functions. Suppose that σ is uniformly nondegenerate. By Corollay 1.3, the law of solutions to SDE (1.2) has a continuous density, which is smooth in the first variable. Even in this case, this result seems to be new as all of the well-known results require that $x\mapsto x+\sigma(x)z$ is invertible (cf. [2, 18]).

EXAMPLE 1.7. Consider the following second-order stochastic differential equation:

$$\mathrm{d}\dot{X}_t = F(X_t,\dot{X}_t)\,\mathrm{d}t+\sigma_0(X_{t-},\dot{X}_{t-})\,\mathrm{d}W_{S_t}, \qquad (X_0,\dot{X}_0)=(x,\mathbf{v}),$$

where $\dot{X}_t$ denotes the first order derivative of X_t with respect to the time variable, $(x,\mathbf{v})$ stands for the position and velocity, $F:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^d$ are two C^∞_b-functions. Notice that if we let $Z_t:=(X_t,\dot{X}_t)$, then Z_t solves the following degenerate SDE:

$$\mathrm{d}Z_t=\big(\dot{X}_t, F(Z_t)\big)\,\mathrm{d}t+\sigma_0(Z_{t-})\,\mathrm{d}W_{S_t},$$

whose generator is given by

$$\begin{aligned}\mathcal{L}^{(\alpha)}_{\mathbf{v}} f(x,\mathbf{v}) := \mathrm{p.v.}\int_{\mathbb{R}^d_0}\frac{f(x,\mathbf{v}+\sigma_0(x,\mathbf{v})\mathbf{v}')-f(x,\mathbf{v})}{|\mathbf{v}'|^{d+\alpha}}\,\mathrm{d}\mathbf{v}'\\ +\mathbf{v}\cdot\nabla_x f(x,\mathbf{v})+F(x,\mathbf{v})\cdot\nabla_{\mathbf{v}} f(x,\mathbf{v}).\end{aligned}$$

Suppose that for any $k,m\in\mathbb{N}_0$,

$$\begin{aligned}\big|\nabla^k_x\nabla^m_{\mathbf{v}} F(x,\mathbf{v})\big| &\le C_1/\big(1+|\mathbf{v}|^2\big),\\ \big|\nabla^{k+1}_x\nabla^m_{\mathbf{v}}\sigma_0(x,\mathbf{v})\big| &\le C_1/\big(1+|\mathbf{v}|^2\big)\end{aligned}\tag{1.17}$$

and

$$c_0|\xi|\le\big|\sigma^*_0(x,\mathbf{v})\xi\big|\le C_0|\xi|, \qquad \xi\in\mathbb{R}^d.\tag{1.18}$$

Let $b(x,\mathbf{v}) := (\mathbf{v}, F(x,\mathbf{v}))$ and σ be defined by (1.13) with $d_1 = d_2 = d$. We now check that (1.8) holds for $m = 0$ and $j_0 = 2$. First of all, by (1.17) and (1.18), it is easy to see $B_1, B_2, B_3 \in C_b^2$. For $u = (u^1, u^2) \in \mathbb{R}^{2d}$, by definition we have

$$|uB_1(x,\mathbf{v})|^2 + |uB_2(x,\mathbf{v})|^2 = |u_2\sigma_0(x,\mathbf{v})|^2 + |u_2 A(x,\mathbf{v}) - u_1\sigma_0(x,\mathbf{v})|^2,$$

where $A(x,\mathbf{v}) = \mathbf{v}\cdot\nabla_x\sigma_0(x,\mathbf{v}) + F(x,\mathbf{v})\cdot\nabla_{\mathbf{v}}\sigma_0(x,\mathbf{v}) - \nabla_{\mathbf{v}}F(x,\mathbf{v})\cdot\sigma_0(x,\mathbf{v})$ is a bounded function by (1.17). Hence, for any $\varepsilon\in(0,1)$, by (1.18) we have

$$\begin{aligned}
&|uB_1(x,\mathbf{v})|^2 + |uB_2(x,\mathbf{v})|^2 \\
&\quad\ge |u_2\sigma_0(x,\mathbf{v})|^2 + \varepsilon|u_1\sigma_0(x,\mathbf{v})|^2 - \frac{\varepsilon}{1-\varepsilon}|u_2A(x,\mathbf{v})|^2 \\
&\quad\ge \left(c_0 - \frac{\varepsilon}{1-\varepsilon}\|A\|_\infty\right)|u_2|^2 + \varepsilon c_0|u_1|^2.
\end{aligned}$$

By taking ε small enough, we get (1.8) with $m = 0$. Thus, by Corollary 1.3, $Z_t(x,\mathbf{v})$ admits a continuous density $\rho_t(x,\mathbf{v};x',\mathbf{v}')$ so that for each $p>1$ and $\varphi\in L^p(\mathbb{R}^{2d})$, the function

$$f_t(x,\mathbf{v}) := \mathbb{E}\varphi(Z_t(x,\mathbf{v})) = \int_{\mathbb{R}^{2d}}\varphi(x',\mathbf{v}')\rho_t(x,\mathbf{v};x',\mathbf{v}')\,\mathrm{d}x'\,\mathrm{d}\mathbf{v}'$$

belongs to $C^\infty((0,\infty)\times\mathbb{R}^{2d})$ and satisfies the following nonlocal kinetic Fokker–Planck equation:

$$\partial_t f_t(x,\mathbf{v}) = \mathcal{L}_{\mathbf{v}}^{(\alpha)} f_t(x,\mathbf{v}), \qquad f_0(x,\mathbf{v}) = \varphi(x,\mathbf{v}). \tag{1.19}$$

It should be emphasized that the initial value φ is not necessary to be smooth, and the following special case has been studied in [1, 6]:

$$\partial_t f_t(x,\mathbf{v}) = a(x,\mathbf{v})\tilde\Delta_{\mathbf{v}}^{\frac{\alpha}{2}} f_t(x,\mathbf{v}) + \mathbf{v}\cdot\nabla_x f_t(x,\mathbf{v}), \qquad f_0(x,\mathbf{v}) = \varphi(x,\mathbf{v}),$$

where $a\in C_b^\infty$ has a positive lower bound, and $\tilde\Delta^{\frac{\alpha}{2}}$ is some cutoff fractional Laplacian.

EXAMPLE 1.8. More generally, consider the following n-order stochastic differential equation:

$$\mathrm{d}X_t^{(n)} = F(X_t^{(0)}, X_t^{(1)},\dots,X_t^{(n)})\,\mathrm{d}t + \sigma_0(X_{t-}^{(0)}, X_{t-}^{(1)},\dots,X_{t-}^{(n)})\,\mathrm{d}W_{S_t}$$

with initial value $(X_0^{(0)},\dots,X_0^{(n)}) = (x^0,x^1,\dots,x^n) =: x\in\mathbb{R}^{(n+1)d}$, where $X_t^{(k)}$ denotes the kth order derivative of X_t with respect to t, $F:\mathbb{R}^{(n+1)d}\to\mathbb{R}^d$ and $\sigma_0:\mathbb{R}^{(n+1)d}\to\mathbb{R}^d\otimes\mathbb{R}^d$ are C_b^∞-functions. Let $\nabla_{(0)} := \nabla_{x^0}$ and $\nabla_{(n)} := (\nabla_{x^1},\dots,\nabla_{x^n})$. Suppose that for any $k,m\in\mathbb{N}_0$,

$$\begin{aligned}
&|\nabla_{(0)}^k\nabla_{(n)}^m F(x)| \le C_1/(1+|(x^1,\dots,x^n)|^{n+1}),\\
&|\nabla_{(0)}^{k+1}\nabla_{(n)}^m \sigma_0(x)| \le C_1/(1+|(x^1,\dots,x^n)|^{n+1})
\end{aligned}\tag{1.20}$$

and

$$c_0|\xi| \le \left|\sigma_0^*(x)\xi\right| \le C_0|\xi|, \qquad \xi \in \mathbb{R}^d. \tag{1.21}$$

Under the above conditions, one can verify as above that (H^{j_0}) with $j_0 = n+1$ and $m = 0$ holds for $b(x) = (x^1, \dots, x^n, F(x))$ and σ in (1.13) with $d_1 = nd$ and $d_2 = d$. Thus, we can apply Corollary 1.3 to this situation.

EXAMPLE 1.9. Consider the following random ODE:

$$\mathrm{d}X_t/\mathrm{d}t = F(X_t) + L_t, \qquad X_0 = x \in \mathbb{R}^d,$$

where $F \in C_b^\infty(\mathbb{R}^d; \mathbb{R}^d)$, and L_t is a Lévy process with Lévy measure ν satisfying (H^ν). Let $Z_t = (X_t, L_t)$. Then Z_t satisfies SDE (1.2) with $b(x, \mathbf{v}) := (\mathbf{v} + F(x), 0)$ and σ as in (1.13) with $\sigma_0 = \mathbb{I}$. In this case, it is easy to see that (H^{j_0}) holds for $j_0 = 2$ and $m = 0$. In particular, Corollary 1.3 is applicable.

1.3. *Layout and notations.* This paper is organized as follows: In Section 2, we first recall Bismut's approach to the Malliavin calculus with jumps, and an inequality for discontinuous semimaringales proven in [26] which is originally due to Komatsu–Takeuchi [11]. Moreover, we also prove an estimate for exponential Poisson random integrals. In Section 3, we prove a quantitive estimate for the Laplace transform of the reduced Malliavin matrix, which is the key step in our proofs and can be read independently. In Section 4, we prove Theorem 1.1. In Section 5, we treat big jump part and prove our main Theorems 1.2 and 1.4 by interpolation and bootstrap arguments. In the Appendix, two technical lemmas are proven.

Before concluding this Introduction, we collect some notation and make some conventions for later use.

- Write $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, $\mathbb{R}_0^d := \mathbb{R}^d - \{0\}$.
- $\nabla := (\partial_1, \dots, \partial_d)$ denotes the gradient operator.
- For a càdlàg function $f : \mathbb{R}_+ \to \mathbb{R}^d$, $\Delta f_s := f_s - f_{s-}$.
- The inner product in Euclidean spaces is denoted by $\langle x, y\rangle$ or $x \cdot y$.
- For $p \in [1, \infty]$, $(L^p(\mathbb{R}^d), \|\cdot\|_p)$ is the L^p-space with respect to the Lebesgue measure.
- $\mathbb{W}^{k,p}$: Sobolev space; $\mathbb{H}^{\beta,p}$: Bessel potential space; $\mathbb{H}^{\beta,\infty} = \mathbb{C}^\beta$: Hölder space.
- For a smooth function $f : \mathbb{R}^d \to \mathbb{R}^d$, $(\nabla f)_{ij} := (\partial_j f^i)$ denotes the Jacobian matrix of f.
- C_0: The space of all continuous functions with values vanishing at infinity.
- C_b^k: The space of all bounded continuous functions with bounded continuous partial derivatives up to k-order. Here, k can be infinity.
- C_{p}^k: The space of all continuous functions with all partial derivatives up to k-order being of polynomial growth. Here k can be infinity.
- The letters c and C with or without indices will denote unimportant constants, whose values may change in different places.

2. Preliminaries.

2.1. *Bismut's approach to the Malliavin calculus with jumps*. In this subsection, we recall some basic facts about Bismut's approach to the Malliavin calculus with jumps (cf. [3, 4] and [20], Section 2). Let $\Gamma \subset \mathbb{R}^d$ be an open set containing the origin. Let us define

$$\Gamma_0 := \Gamma \setminus \{0\}, \qquad \varrho(z) := 1 \vee \mathrm{d}\big(z, \Gamma_0^c\big)^{-1}, \tag{2.1}$$

where $\mathrm{d}(z, \Gamma_0^c)$ is the distance of z to the complement of Γ_0. Notice that $\varrho(z) = \frac{1}{|z|}$ near 0.

Let Ω be the canonical space of all integer-valued measure ω on $[0, 1] \times \Gamma_0$ with $\mu(A) < +\infty$ for any compact set $A \subset [0, 1] \times \Gamma_0$. Define the canonical process on Ω as follows:

$$N(\omega; \mathrm{d}t, \mathrm{d}z) := \omega(\mathrm{d}t, \mathrm{d}z).$$

Let $(\mathscr{F}_t)_{t\in[0,1]}$ be the smallest right-continuous filtration on Ω such that N is optional. In the following, we write $\mathscr{F} := \mathscr{F}_1$, and endow $(\Omega, \mathscr{F})$ with a unique probability measure $\mathbb{P}$ such that N is a Poisson random measure with intensity $\mathrm{d}t\,\nu(\mathrm{d}z)$, where $\nu(\mathrm{d}z) = \kappa(z)\,\mathrm{d}z$ with

$$\begin{gathered} \kappa \in C^1\big(\Gamma_0; (0, \infty)\big), \qquad \int_{\Gamma_0} \big(1 \wedge |z|^2\big)\kappa(z)\,\mathrm{d}z < +\infty, \\ \big|\nabla \log \kappa(z)\big| \le C\varrho(z), \end{gathered} \tag{2.2}$$

where $\varrho(z)$ is defined by (2.1). In the following, we write

$$\tilde{N}(\mathrm{d}t, \mathrm{d}z) := N(\mathrm{d}t, \mathrm{d}z) - \mathrm{d}t\,\nu(\mathrm{d}z).$$

Let $p \ge 1$ and $k \in \mathbb{N}$. We introduce the following spaces for later use:

- $\mathbb{L}^1_p$: The space of all predictable processes: $\xi : \Omega \times [0, 1] \times \Gamma_0 \to \mathbb{R}^k$ with finite norm:

$$\|\xi\|_{\mathbb{L}^1_p} := \left[\mathbb{E}\left(\int_0^1 \int_{\Gamma_0} \big|\xi(s, z)\big|\nu(\mathrm{d}z)\mathrm{d}s\right)^p\right]^{\frac{1}{p}} + \left[\mathbb{E}\int_0^1 \int_{\Gamma_0} \big|\xi(s, z)\big|^p \nu(\mathrm{d}z)\,\mathrm{d}s\right]^{\frac{1}{p}} < \infty.$$

- $\mathbb{L}^2_p$: The space of all predictable processes: $\xi : \Omega \times [0, 1] \times \Gamma_0 \to \mathbb{R}^k$ with finite norm:

$$\|\xi\|_{\mathbb{L}^2_p} := \left[\mathbb{E}\left(\int_0^1 \int_{\Gamma_0} \big|\xi(s, z)\big|^2 \nu(\mathrm{d}z)\,\mathrm{d}s\right)^{\frac{p}{2}}\right]^{\frac{1}{p}} + \left[\mathbb{E}\int_0^1 \int_{\Gamma_0} \big|\xi(s, z)\big|^p \nu(\mathrm{d}z)\,\mathrm{d}s\right]^{\frac{1}{p}} < \infty.$$

- $\mathbb{V}_p$: The space of all predictable processes $\mathbf{v} : \Omega \times [0, 1] \times \Gamma_0 \to \mathbb{R}^d$ with finite norm:

$$\|\mathbf{v}\|_{\mathbb{V}_p} := \|\nabla_z \mathbf{v}\|_{\mathbb{L}^1_p} + \|\mathbf{v}\varrho\|_{\mathbb{L}^1_p} < \infty,$$

where $\varrho(z)$ is defined by (2.1). Below we shall write

$$\mathbb{V}_{\infty-} := \bigcap_{p\geq 1} \mathbb{V}_p.$$

- $\mathbb{V}_0$: The space of all predictable processes $\mathbf{v} : \Omega \times [0, 1] \times \Gamma_0 \to \mathbb{R}^d$ with the following properties: (i) $\mathbf{v}$ and $\nabla_z \mathbf{v}$ are bounded; (ii) there exists a compact subset $U \subset \Gamma_0$ such that

$$\mathbf{v}(t, z) = 0 \qquad \forall z \notin U.$$

Moreover, $\mathbb{V}_0$ is dense in $\mathbb{V}_p$ for all $p \geq 1$ (cf. [20], Lemma 2.1).

Let $C^\infty_{\mathrm{p}}(\mathbb{R}^m)$ be the class of all smooth functions on $\mathbb{R}^m$ which together with all the derivatives are of at most polynomial growth. Let $\mathcal{F}C^\infty_{\mathrm{p}}$ be the class of all Poisson functionals on Ω with the following form:

$$F(\omega) = f\big(\omega(g_1), \dots, \omega(g_m)\big),$$

where $f \in C^\infty_{\mathrm{p}}(\mathbb{R}^m)$ and $g_1, \dots, g_m \in \mathbb{V}_0$ are nonrandom, and

$$\omega(g_j) := \int_0^1 \int_{\Gamma_0} g_j(s, z)\omega(\mathrm{d}s, \mathrm{d}z).$$

Notice that

$$\mathcal{F}C^\infty_{\mathrm{p}} \subset \bigcap_{p\geq 1} L^p(\Omega, \mathscr{F}, \mathbb{P}).$$

For $\mathbf{v} \in \mathbb{V}_{\infty-}$ and $F \in \mathcal{F}C^\infty_{\mathrm{p}}$, let us define

$$D_{\mathbf{v}} F =: \sum_{j=1}^m (\partial_j f)(\cdot) \int_0^1 \int_{\Gamma_0} \mathbf{v}(s, z) \cdot \nabla_z g_j(s, z)\omega(\mathrm{d}s, \mathrm{d}z),$$

where "$(\cdot)$" stands for $(\omega(g_1), \dots, \omega(g_m))$.

We have the following integration by parts formula (cf. [20], Theorem 2.9).

THEOREM 2.1. *Let $\mathbf{v} \in \mathbb{V}_{\infty-}$ and $p > 1$. The linear operator $(D_{\mathbf{v}}, \mathcal{F}C^\infty_{\mathrm{p}})$ is closable in $L^p(\Omega)$. The closure is denoted by $\mathbb{W}^{1,p}_{\mathbf{v}}(\Omega)$, which is a Banach space with respect to the norm*:

$$\|F\|_{\mathbf{v};1,p} := \|F\|_{L^p} + \|D_{\mathbf{v}} F\|_{L^p}.$$

Moreover, for any $F \in \mathbb{W}^{1,p}_{\mathbf{v}}(\Omega)$, we have

$$\mathbb{E}(D_{\mathbf{v}} F) = -\mathbb{E}\big(F \operatorname{div}(\mathbf{v})\big), \tag{2.3}$$

where $\mathrm{div}(\mathbf{v})$ *is defined by*

$$\mathrm{div}(\mathbf{v}) := \int_0^1 \int_{\Gamma_0} \frac{\mathrm{div}(\kappa \mathbf{v})(s,z)}{\kappa(z)} \tilde{N}(\mathrm{d}s, \mathrm{d}z). \tag{2.4}$$

Here, with a little of abuse, the div *in the right integral denotes the usual divergence of a vector field in Euclidean space.*

Below, we shall write

$$\mathbb{W}_{\mathbf{v}}^{1,\infty-}(\Omega) := \bigcap_{p>1} \mathbb{W}_{\mathbf{v}}^{1,p}(\Omega).$$

The following Kusuoka and Stroock's formula is proven in [20], Proposition 2.11.

PROPOSITION 2.2. *Fix* $\mathbf{v} \in \mathbb{V}_{\infty-}$. *Let* $\eta(\omega, s, z) : \Omega \times [0,1] \times \Gamma_0 \to \mathbb{R}$ *be a measurable map and satisfy that for each* $(s,z) \in \Omega \times [0,1] \times \Gamma_0$, $\eta(\cdot, s, z) \in \mathbb{W}_{\mathbf{v}}^{1,\infty-}(\Omega)$, *and for each* $(\omega, s) \in \Omega \times [0,1]$, $\eta(\omega, s, \cdot) \in C^1(\Gamma_0)$, *and for each* $z \in \Gamma_0$,

$$s \mapsto \eta(s,z),\ D_{\mathbf{v}}\eta(s,z),\ \nabla_z \eta(s,z) \textit{ are left-continuous and } \mathscr{F}_s\textit{-adapted}, \tag{2.5}$$

and for any $p \geq 1$,

$$\mathbb{E}\left[\sup_{s\in[0,1]} \sup_{z\in\Gamma_0} \left(\frac{|\eta(s,z)|^p + |D_{\mathbf{v}}\eta(s,z)|^p}{(1 \wedge |z|)^p} + \left|\nabla_z \eta(s,z)\right|^p \right)\right] < +\infty. \tag{2.6}$$

Then $\mathscr{I}(\eta) := \int_0^1 \int_{\Gamma_0} \eta(s,z) \tilde{N}(\mathrm{d}s, \mathrm{d}z) \in \mathbb{W}_{\mathbf{v}}^{1,\infty-}(\Omega)$ *and*

$$\begin{aligned} D_{\mathbf{v}} \mathscr{I}(\eta) = {} & \int_0^1 \int_{\Gamma_0} D_{\mathbf{v}} \eta(s,z) \tilde{N}(\mathrm{d}s, \mathrm{d}z) \\ & + \int_0^1 \int_{\Gamma_0} \langle \nabla_z \eta(s,z), \mathbf{v}(s,z) \rangle N(\mathrm{d}s, \mathrm{d}z). \end{aligned} \tag{2.7}$$

We also need the following Burkholder's type inequalities (cf. [20], Lemma 2.3).

LEMMA 2.3. (i) *For any* $p > 1$, *there is a constant* $C_p > 0$ *such that for any* $\xi \in \mathbb{L}_p^1$,

$$\mathbb{E}\left(\sup_{t\in[0,1]} \left| \int_0^t \int_{\Gamma_0} \xi(s,z) N(\mathrm{d}s, \mathrm{d}z) \right|^p \right) \leq C_p \|\xi\|_{\mathbb{L}_p^1}^p. \tag{2.8}$$

(ii) *For any* $p \geq 2$, *there is a constant* $C_p > 0$ *such that for any* $\xi \in \mathbb{L}_p^2$,

$$\mathbb{E}\left(\sup_{t\in[0,1]} \left| \int_0^t \int_{\Gamma_0} \xi(s,z) \tilde{N}(\mathrm{d}s, \mathrm{d}z) \right|^p \right) \leq C_p \|\xi\|_{\mathbb{L}_p^2}^p. \tag{2.9}$$

2.2. *Two lemmas*. We first recall the following important Komatsu–Takeuchi's type estimate proven in [26], Theorem 4.2, which will be used in Section 3.

LEMMA 2.4. *Let $(f_t)_{t\geq 0}$ and $(f_t^0)_{t\geq 0}$ be two m-dimensional semimartingales given by*

$$f_t = f_0 + \int_0^{t\wedge\tau} (f_s^0 + h_s^\delta)\,\mathrm{d}s + \int_0^t \int_{|z|\leq\delta} g_{s-}(z)\tilde{N}(\mathrm{d}s, \mathrm{d}z),$$

$$f_t^0 = f_0^0 + \int_0^t f_s^{00}\,\mathrm{d}s + \int_0^t \int_{|z|\leq\delta} g_{s-}^0(z)\tilde{N}(\mathrm{d}s, \mathrm{d}z),$$

where $\delta \in (0,1]$, τ is a stopping time and $f_t, f_t^0, h_t^\delta, f_t^{00}$ and $g_t(z), g_t^0(z)$ are càdlàg $\mathscr{F}_t$-adapted processes. Assume that for some $\kappa \geq 1$ and $\beta \geq 0$,

$$|f_t|^2 \vee |f_t^0|^2 \vee |f_t^{00}|^2 \vee (\delta^{-\beta} h_t^\delta) \vee \left(\sup_z \frac{|g_t(z)|^2 \vee |g_t^0(z)|^2}{1 \wedge |z|^2}\right) \leq \kappa, \qquad \text{a.s.}$$

Then for any $\varepsilon, T \in (0,1]$, there exists a positive random variable ζ with $\mathbb{E}\zeta \leq 1$ such that

$$\begin{aligned} c_0 \int_0^{T\wedge\tau} |f_t^0|^2\,\mathrm{d}t \leq{}& (\delta^{-\frac{3}{2}} + \varepsilon^{-\frac{3}{2}}) \int_0^T |f_t|^2\,\mathrm{d}t + \kappa\delta^{\frac{1}{2}}\log\zeta \\ & + \kappa(\varepsilon\delta^{-\frac{1}{2}} + \varepsilon^{\frac{1}{2}} + T\delta^{\frac{1}{2}\wedge\beta}), \end{aligned} \tag{2.10}$$

where $c_0 \in (0,1)$ only depends on $\int_{|z|\leq 1} |z|^2 \nu(\mathrm{d}z)$.

The following result will be used in Section 4.

LEMMA 2.5. *Let $g_s(z), \eta_s$ be two left continuous $\mathscr{F}_s$-adapted processes satisfying*

$$0 \leq g_s(z) \leq \eta_s, \qquad |g_s(z) - g_s(0)| \leq \eta_s|z| \qquad \forall |z| \leq 1, \tag{2.11}$$

and for any $p \geq 2$,

$$\mathbb{E}\Big(\sup_{s\in[0,1]} |\eta_s|^p\Big) < +\infty.$$

If for some $\alpha \in (0,2)$,

$$\lim_{\varepsilon\to 0} \varepsilon^{\alpha-2} \int_{|z|\leq\varepsilon} |z|^2 \nu(\mathrm{d}z) =: c_1 > 0, \tag{2.12}$$

then for any $\delta \in (0,1)$, there exist $c_2, \theta \in (0,1)$, $C_2 \geq 1$ such that for all $\lambda, p \geq 1$ and $t \in (0,1)$,

$$\begin{aligned} &\mathbb{E}\exp\left\{-\lambda \int_0^t \int_{\mathbb{R}_0^d} g_s(z)\zeta(z)N(\mathrm{d}s, \mathrm{d}z)\right\} \\ &\quad \leq C_2\left(\mathbb{E}\exp\left\{-c_2\lambda^\theta \int_0^t g_s(0)\,\mathrm{d}s\right\}\right)^{\frac{1}{2}} + C_p\lambda^{-p}, \end{aligned} \tag{2.13}$$

where $\zeta(z) = \zeta_\delta(z)$ is a nonnegative smooth function with

$$\zeta_\delta(z) = |z|^3, \qquad |z| \le \delta/4, \qquad \zeta_\delta(z) = 0, \qquad |z| > \delta/2.$$

PROOF. For $\lambda \ge 1$ and $\beta > 0$, define a stopping time

$$\tau := \inf\{s > 0 : \eta_s \ge \lambda^\beta\} \wedge 1.$$

Set

$$h_t^\lambda := \int_{\mathbb{R}_0^d} \big(1 - \mathrm{e}^{-\lambda g_t(z)\zeta(z)}\big)\nu(\mathrm{d}z)$$

and

$$M_t^\lambda := -\lambda \int_0^{t\wedge\tau} \int_{\mathbb{R}_0^d} g_s(z)\zeta(z) N(\mathrm{d}s, \mathrm{d}z) + \int_0^{t\wedge\tau} h_s^\lambda \, \mathrm{d}s. \tag{2.14}$$

By Itô's formula, we have

$$\mathrm{e}^{M_t^\lambda} = 1 + \int_0^{t\wedge\tau} \int_{\mathbb{R}_0^d} \mathrm{e}^{M_{s-}^\lambda} \big(\mathrm{e}^{-\lambda g_s(z)\zeta(z)} - 1\big)\tilde{N}(\mathrm{d}s, \mathrm{d}z).$$

Since for any $x \ge 0$,

$$1 - \mathrm{e}^{-x} \le 1 \wedge x,$$

by (2.11) and the definition of τ, we have

$$\begin{aligned} M_t^\lambda &\le \int_0^{t\wedge\tau} h_s^\lambda \, \mathrm{d}s \le \int_0^{t\wedge\tau} \int_{\mathbb{R}_0^d} \big(1 \wedge \big(\lambda g_s(z)\zeta(z)\big)\big)\nu(\mathrm{d}z)\,\mathrm{d}s \\ &\le \int_{\mathbb{R}_0^d} \big(1 \wedge \big(\lambda^{1+\beta}\zeta(z)\big)\big)\nu(\mathrm{d}z) < \infty. \end{aligned}$$

Hence, $\mathbb{E}\mathrm{e}^{M_t^\lambda} = 1$, and by (2.14) and Hölder's inequality,

$$\begin{aligned} &\mathbb{E}\exp\left\{-\frac{\lambda}{2}\int_0^{t\wedge\tau} \int_{\mathbb{R}_0^d} g_s(z)\zeta(z) N(\mathrm{d}s, \mathrm{d}z)\right\} \\ &\quad \le \left(\mathbb{E}\exp\left\{-\int_0^{t\wedge\tau} h_s^\lambda \, \mathrm{d}s\right\}\right)^{\frac{1}{2}}. \end{aligned} \tag{2.15}$$

Since by (2.11) and definition of τ, $1_{s<\tau} g_s(z) \le \lambda^\beta$, and

$$1 - \mathrm{e}^{-x} \ge \frac{x}{\mathrm{e}}, \qquad x \le 1,$$

for any $q \geq 1+\beta$, there exists a $c \in (0, 1)$ small enough such that for all $\lambda \geq 1$ and $s < \tau$,

$$\begin{aligned} h_s^\lambda &\geq \int_{|z|^3 \leq c\lambda^{-q}} (1 - \mathrm{e}^{-\lambda g_s(z)|z|^3}) \nu(\mathrm{d}z) \geq \frac{\lambda}{\mathrm{e}} \int_{|z|^3 \leq c\lambda^{-q}} g_s(z)|z|^3 \nu(\mathrm{d}z) \\ &= \frac{\lambda g_s(0)}{\mathrm{e}} \int_{|z|^3 \leq c\lambda^{-q}} |z|^3 \nu(\mathrm{d}z) + \frac{\lambda}{\mathrm{e}} \int_{|z|^3 \leq c\lambda^{-q}} \big(g_s(z) - g_s(0)\big)|z|^3 \nu(\mathrm{d}z). \end{aligned} \tag{2.16}$$

Notice that by (2.12), for any $p \geq 2$, there exist constants $c_0, C_0 > 0$ such that for all $\varepsilon \in (0, 1)$ (cf. [20], Lemma 5.2),

$$c_0 \varepsilon^{p-\alpha} \leq \int_{|z| \leq \varepsilon} |z|^p \nu(\mathrm{d}z) \leq C_0 \varepsilon^{p-\alpha}. \tag{2.17}$$

If we choose

$$\beta \in \left(0, \frac{\alpha \wedge 1}{3-\alpha}\right), \qquad q = \begin{cases} 1+\beta, & \alpha \in (0, 1], \\ \dfrac{3(1+\beta)}{4-\alpha}, & \alpha \in (1, 2), \end{cases}$$

then by (2.16), (2.11) and (2.17), for all $\lambda \geq 1$ and $s < \tau$, we further have

$$h_s^\lambda \geq c_2 g_s(0) \lambda^{1-\frac{(3-\alpha)q}{3}} - C_1 \lambda^{1+\beta-\frac{(4-\alpha)q}{3}} \geq c_2 g_s(0) \lambda^{1-\frac{(3-\alpha)q}{3}} - C_1. \tag{2.18}$$

On the other hand, by Chebyshev's inequality, for any $p \geq 2$ we have

$$\mathbb{P}(\tau \leq t) = \mathbb{P}\Big(\sup_{s \in [0,t]} \eta_s > \lambda^\beta\Big) \leq \lambda^{-\beta p} \mathbb{E}\Big(\sup_{s \in [0,t]} |\eta_s|^p\Big),$$

which together with (2.15) and (2.18) yields the desired estimate (2.13). $\square$

3. Estimate of Laplace transform of reduced Malliavin matrix. The current section is written to be independent of the settings in Section 2.1 so that it can be used to other framework such as the one used by Picard in [18]. Let L_t be a d-dimensional pure jump Lévy process with Lévy measure ν. We assume that the Lévy measure ν satisfies the following conditions: for some $\alpha \in (0, 2)$,

$$\begin{aligned} &\int_{|z|<\delta} |z|^2 \nu(\mathrm{d}z) \leq C\delta^{2-\alpha} \qquad \forall \delta \in (0, 1), \\ &\int_{|z|\geq 1} |z|^m \nu(\mathrm{d}z) < \infty \qquad \forall m \in \mathbb{N}. \end{aligned} \tag{3.1}$$

Let $N(\mathrm{d}t, \mathrm{d}z)$ be the Poisson random measure associated with L_t, that is,

$$N\big((0, t] \times E\big) = \sum_{s \leq t} 1_E(\Delta L_s), \qquad E \in \mathscr{B}(\mathbb{R}_0^d).$$

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. Consider the following SDE:

$$X_t(x) = x + \int_0^t b\big(X_s(x)\big)\,\mathrm{d}s + \int_0^t \int_{\mathbb{R}_0^d} \sigma\big(X_{s-}(x), z\big)\tilde{N}(\mathrm{d}s, \mathrm{d}z), \tag{3.2}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ are two smooth functions satisfying that for any $k \in \mathbb{N}$, $m \in \mathbb{N}_0$ and $j = 0, 1$,

$$\big|\nabla^k b(x)\big| \le C, \qquad \big|\nabla_x^m \nabla_z^j \sigma(x, z)\big| \le C|z|^{1-j} \tag{3.3}$$

and

$$\int_{r<|z|<R} \sigma(x, z)\nu(\mathrm{d}z) = 0, \qquad 0 < r < R < \infty. \tag{3.4}$$

Under (3.3), it is well known that SDE (3.2) has a unique solution denoted by $X_t := X_t(x)$, which defines a C^∞-stochastic flow (cf. [9] and [19]). Let $J_t := J_t(x) := \nabla X_t(x) = (\partial_j X_t^i(x))_{ij}$ be the Jacobian matrix of $X_t(x)$, which solves the following linear matrix-valued SDE:

$$J_t = \mathbb{I} + \int_0^t \nabla b(X_s) J_s\,\mathrm{d}s + \int_0^t \int_{\mathbb{R}_0^d} \nabla_x \sigma(X_{s-}, z) J_{s-} \tilde{N}(\mathrm{d}s, \mathrm{d}z). \tag{3.5}$$

If we further assume

$$\inf_{x\in\mathbb{R}^d} \inf_{z\in\mathbb{R}^d} \det\big(\mathbb{I} + \nabla_x \sigma(x, z)\big) > 0, \tag{3.6}$$

then the matrix $J_t(x)$ is invertible (cf. [9]). Let $K_t = K_t(x)$ be the inverse matrix of $J_t(x)$. By Itô's formula, it is easy to see that K_t solves the following linear matrix-valued SDE (cf. [26], Lemma 3.2):

$$\begin{aligned} K_t = \mathbb{I} &- \int_0^t K_s \nabla b(X_s)\,\mathrm{d}s + \int_0^t \int_{\mathbb{R}_0^d} K_{s-} Q(X_{s-}, z)\tilde{N}(\mathrm{d}s, \mathrm{d}z) \\ &- \int_0^t \int_{\mathbb{R}_0^d} K_{s-} Q(X_{s-}, z) \nabla_x \sigma(X_{s-}, z)\nu(\mathrm{d}z)\,\mathrm{d}s, \end{aligned} \tag{3.7}$$

where

$$Q(x, z) := \big(\mathbb{I} + \nabla_x \sigma(x, z)\big)^{-1} - \mathbb{I}. \tag{3.8}$$

First of all, we have the following easy estimate. Since the proof is standard by Burkholder's inequality (see Lemma 2.3) and Gronwall's inequality, we omit the details.

LEMMA 3.1. *Under* (3.1), (3.3) *and* (3.6), *for any* $p \ge 1$, *we have*

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}\bigg(\sup_{t\in[0,1]}\bigg(\frac{|X_t(x)|^p}{1+|x|^p} + \big|J_t(x)\big|^p + \big|K_t(x)\big|^p\bigg)\bigg) < +\infty. \tag{3.9}$$

We now prove the following crucial lemma.

LEMMA 3.2. *Let $V:\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^d$ be a $C^\infty_{\rm p}$-function. Under* (3.1), (3.3), (3.4) *and* (3.6), *there exist $\beta_1,\beta_3\in(0,1)$, $\beta_2\geq 1$ only depending on α, and constants $C_1\geq 1$, $c_1\in(0,1)$ such that for all $\delta,t\in(0,1)$, $x\in\mathbb{R}^d$, $|u|=1$ and $p\geq 1$,*

$$\begin{aligned}&\mathbb{P}\left(\begin{aligned}&\int_0^t \big|uK_s(x)[b,V]\big(X_s(x)\big)\big|^2\,\mathrm{d}s\geq t\delta^{\beta_1}\\&\text{and}\int_0^t\big|uK_s(x)V\big(X_s(x)\big)\big|^2\,\mathrm{d}s\leq t\delta^{\beta_2}\end{aligned}\right)\\&\quad\leq C_1\mathrm{e}^{-c_1t\delta^{-\beta_3}}+C_p(x)\delta^p,\end{aligned}\tag{3.10}$$

where $[b,V]:=b\cdot\nabla V-\nabla b\cdot V$, and $C_p(x)$ continuously depends on x. Moreover, if we assume that $V,[b,V],[b,[b,V]]\in C^2_b$, then $C_p(x)$ can be independent of x.

PROOF. We divide the proof into four steps.

(1) Fixing $\delta\in(0,1)$, we decompose the Lévy process L_t as the small and large jump parts, that is, $L_t=L^\delta_t+\hat{L}^\delta_t$, where

$$L^\delta_t:=\int_{|z|\leq\delta} z\tilde{N}\big((0,t],\mathrm{d}z\big),\ \hat{L}^\delta_t:=\int_{|z|>\delta} zN\big((0,t],\mathrm{d}z\big).$$

Clearly,

$$L^\delta_t \text{ and } \hat{L}^\delta_t \text{ are independent.}$$

Let us fix a path $\hbar$ with finitely many jumps on any finite time interval. Let $X^\delta_t(x;\hbar)$ solve the following SDE:

$$\begin{aligned}X^\delta_t(x;\hbar)&=x+\int_0^t b\big(X^\delta_s(x;\hbar)\big)\,\mathrm{d}s+\sum_{s\leq t}\sigma\big(X^\delta_{s-}(x;\hbar),\Delta\hbar_s\big)\\&\quad+\int_0^t\int_{|z|\leq\delta}\sigma\big(X^\delta_{s-}(x;\hbar),z\big)\tilde{N}(\mathrm{d}s,\mathrm{d}z).\end{aligned}\tag{3.11}$$

Let $K^\delta_t(x;\hbar):=[\nabla X^\delta_t(x;\hbar)]^{-1}$. Clearly, by (3.4) we have

$$X_t(x)=X^\delta_t(x;\hbar)|_{\hbar=\hat{L}^\delta},\qquad K_t(x)=K^\delta_t(x;\hbar)|_{\hbar=\hat{L}^\delta}.\tag{3.12}$$

Moreover, $K^\delta_t:=K^\delta_t(x;0)$ solves the following equation:

$$\begin{aligned}K^\delta_t&=\mathbb{I}-\int_0^t K^\delta_s\nabla b\big(X^\delta_s\big)\,\mathrm{d}s+\int_0^t\int_{|z|\leq\delta}K^\delta_{s-}Q\big(X^\delta_{s-},z\big)\tilde{N}(\mathrm{d}s,\mathrm{d}z)\\&\quad-\int_0^t\int_{|z|\leq\delta}K^\delta_{s-}Q\big(X^\delta_{s-},z\big)\nabla_x\sigma\big(X^\delta_{s-},z\big)\nu(\mathrm{d}z)\,\mathrm{d}s.\end{aligned}\tag{3.13}$$

(2) Let $V : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ be a C_{p}^∞-function. Define

$$H_V(x,z) := V\big(x+\sigma(x,z)\big) - V(x) + Q(x,z)V\big(x+\sigma(x,z)\big), \tag{3.14}$$

$$G_V(x,z) := H_V(x,z) + \nabla_x\sigma(x,z)\cdot V(x) - \sigma(x,z)\cdot\nabla V(x) \tag{3.15}$$

and

$$V_0(x) := [b,V](x), \qquad G_V^\delta(x) := \int_{|z|\le\delta} G_V(x,z)\nu(\mathrm{d}z),$$
$$V_1(x) := [b,V_0](x) + \int_{|z|\le\delta} G_{V_0}(x,z)\nu(\mathrm{d}z).$$

Since $V \in C_{\mathrm{p}}^\infty$, by (3.1) and (3.3), there is an $m \in \mathbb{N}_0$ such that for all $|z| \le 1$,

$$\begin{aligned} &\big|H_V(x,z)\big| \le C\big(1+|x|^m\big)|z|, \qquad \big|G_V(x,z)\big| \le C\big(1+|x|^m\big)|z|^2,\\ &\big|G_V^\delta(x)\big|| \le C\big(1+|x|^m\big)\delta^{2-\alpha}, \qquad \big|V_0(x)\big| + \big|V_1(x)\big| \le C\big(1+|x|^m\big). \end{aligned} \tag{3.16}$$

For a row vector $u \in \mathbb{R}^d$, we introduce the processes:

$$f_t := uK_t^\delta V\big(X_t^\delta\big), \qquad f_t^0 := uK_t^\delta V_0\big(X_t^\delta\big), \qquad h_t^\delta := uK_t^\delta G_V^\delta\big(X_t^\delta\big),$$
$$f_t^{00} := uK_t^\delta V_1\big(X_t^\delta\big), \qquad g_t(z) := uK_t^\delta H_V\big(X_t^\delta, z\big), \qquad g_t^0(z) := uK_t^\delta H_{V_0}\big(X_t^\delta, z\big),$$

where

$$X_t^\delta := X_t^\delta(x;0), \qquad K_t^\delta := K_t^\delta(x;0).$$

By equations (3.13) and (3.11) with $\hbar = 0$, using Itô's formula, we have

$$\begin{aligned} f_t &= uV(x) + \int_0^t uK_s^\delta[b,V]\big(X_s^\delta\big)\,\mathrm{d}s + \int_0^t\int_{|z|\le\delta} g_{s-}(z)\tilde N(\mathrm{d}s,\mathrm{d}z)\\ &\quad + \int_0^t\int_{|z|\le\delta} uK_s^\delta G_V\big(X_s^\delta,z\big)\nu(\mathrm{d}z)\,\mathrm{d}s\\ &= uV(x) + \int_0^t \big(f_s^0 + h_s^\delta\big)\,\mathrm{d}s + \int_0^t\int_{|z|\le\delta} g_{s-}(z)\tilde N(\mathrm{d}s,\mathrm{d}z) \end{aligned}$$

and

$$f_t^0 = uV_0(x) + \int_0^t f_s^{00}\,\mathrm{d}s + \int_0^t\int_{|z|\le\delta} g_{s-}^0(z)\tilde N(\mathrm{d}s,\mathrm{d}z).$$

For $\gamma \in (0, 4-2\alpha)$, define a stopping time

$$\tau := \tau_u(x) := \inf\big\{s \ge 0 : \big|uK_s^\delta(x;0)\big|^2 \vee \big|X_s^\delta(x;0)\big|^{2m} > \delta^{-\gamma/2}\big\}.$$

By (3.16) and (3.1), there is a constant $\kappa_0 \ge 1$ (not depending on x) such that for all $t \in [0,\tau)$ and $|z| \le 1$,

$$|f_t|^2, \big|f_t^0\big|^2, \big|f_t^{00}\big|^2 \le \kappa_0\delta^{-\gamma}, \qquad \big|h_t^\delta\big|^2 \le \kappa_0\delta^{4-2\alpha-\gamma},$$
$$\big|g_t(z)\big|^2, \big|g_t^0(z)\big|^2 \le \kappa_0\delta^{-\gamma}|z|^2.$$

If we make the following replacement in Theorem 2.4:

$$f_t, g_t(z), f_t^0, h_t^\delta, g_t^0(z) \Rightarrow f_{t\wedge\tau}, 1_{t<\tau} g_t(z), f_{t\wedge\tau}^0, h_{t\wedge\tau}^\delta, 1_{t<\tau} g_t^0(z),$$

then by (2.10) with $\varepsilon = \delta^5$ and $\kappa = \kappa_0 \delta^{-\gamma}$, we obtain

$$\begin{aligned} c_0 \int_0^{t\wedge\tau} |f_s^0|^2 \,\mathrm{d}s &\le \big(\delta^{-\frac{3}{2}} + \delta^{-\frac{15}{2}}\big) \int_0^t |f_{s\wedge\tau}|^2 \,\mathrm{d}s + \kappa_0 \delta^{\frac{1}{2}-\gamma} \log\zeta \\ &\quad + \kappa_0 \delta^{-\gamma} \big(\delta^{5-\frac{1}{2}} + \delta^{\frac{5}{2}} + t\delta^{\frac{1}{2}\wedge(4-2\alpha-\gamma)}\big) \\ &\le 2\kappa_0 \delta^{-\frac{15}{2}} \int_0^t |f_{s\wedge\tau}|^2 \,\mathrm{d}s + \kappa_0 \delta^{\frac{1}{2}-\gamma} \log\zeta \\ &\quad + 2\kappa_0 \big(\delta^{\frac{5}{2}-\gamma} + t\delta^{\frac{1}{2}\wedge(4-2\alpha-\gamma)-\gamma}\big) \qquad \text{a.s.,} \end{aligned}$$

where $c_0 \in (0,1)$ only depends on $\int_{|z|\le 1} |z|^2 \nu(\mathrm{d}z)$ and $\zeta > 0$ with $\mathbb{E}\zeta \le 1$. From this, dividing both sides by $2\kappa_0 \delta^{\frac{1}{2}-\gamma}$ and taking exponential, then multiplying $1_{\tau\ge t}$ and taking expectations, we derive that for $c_1 := \frac{c_0}{2\kappa_0}$ and $\beta_0 := 0 \wedge (\frac{7}{2} - 2\alpha - \gamma)$,

$$\begin{aligned} &\sup_{u\in\mathbb{R}^d} \mathbb{E}\bigg(\exp\bigg\{c_1 \delta^{\gamma-\frac{1}{2}} \int_0^t \big|uK_s^\delta[b,V]\big(X_s^\delta\big)\big|^2 \,\mathrm{d}s \\ &\qquad - \delta^{\gamma-8} \int_0^t \big|uK_s^\delta V\big(X_s^\delta\big)\big|^2 \,\mathrm{d}s\bigg\} 1_{\tau\ge t}\bigg) \\ &\quad\le \mathbb{E}\big(1_{\tau\ge t}\zeta^{1/2}\big) \exp\big\{\delta^2 + t\delta^{\frac{1}{2}\wedge(4-2\alpha-\gamma)-\frac{1}{2}}\big\} \le \exp\big\{\delta^2 + t\delta^{\beta_0}\big\}. \end{aligned} \tag{3.17}$$

(3) For $t \in (0,1)$ and $u \in \mathbb{R}^d$, define a random set

$$\Omega_t^u(x;\hbar) := \Big\{\omega : \sup_{s\in[0,t]} \big(\big|uK_s^\delta(\omega,x;\hbar)\big|^2 \vee \big|X_s^\delta(\omega,x;\hbar)\big|^{2m}\big) \le \delta^{-\gamma/2}\Big\},$$

and let

$$\begin{aligned} \mathcal{J}_t^u(x;\hbar) := \exp\bigg\{&c_1 \delta^{\gamma-\frac{1}{2}} \int_0^t \big|uK_s^\delta(x;\hbar)[b,V]\big(X_s^\delta(x;\hbar)\big)\big|^2 \,\mathrm{d}s \\ &- \delta^{\gamma-8} \int_0^t \big|uK_s^\delta(x;\hbar) V\big(X_s^\delta(x;\hbar)\big)\big|^2 \,\mathrm{d}s\bigg\} 1_{\Omega_t^u(x;\hbar)}. \end{aligned} \tag{3.18}$$

Since $\Omega_t^u(x;0) \subset \{\tau_u(x) \ge t\}$, by (3.17) we have

$$\sup_{x\in\mathbb{R}^d} \sup_{u\in\mathbb{R}^d} \mathbb{E}\mathcal{J}_t^u(x;0) \le \exp\big\{\delta^2 + t\delta^{\beta_0}\big\}. \tag{3.19}$$

Let $0 = t_0 < t_1 < \cdots < t_n \le t_{n+1} = t$ be the jump times of $\hbar$. If we set

$$\phi_{t_j}(x;\hbar) := X_{t_j-}^\delta(x;\hbar) + \sigma\big(X_{t_j-}^\delta(x;\hbar), \Delta\hbar_{t_j}\big),$$

then for $s \in [0, t_{j+1} - t_j)$,

$$X_{s+t_j}^\delta(x;\hbar) = X_s^\delta\big(\phi_{t_j}(x;\hbar);0\big) \Rightarrow K_{s+t_j}^\delta(x;\hbar) = \big[\nabla\phi_{t_j}(x;\hbar)\big]^{-1} K_s^\delta\big(\phi_{t_j}(x;\hbar);0\big)$$

and

$$\Omega^u_{t_{j+1}}(x;\hbar) = \Omega^u_{t_j}(x;\hbar) \cap \Big\{ \sup_{s\in(0,t_{j+1}-t_j]} \big(|uK^\delta_{s+t_j}(x;\hbar)|^2 \vee |X^\delta_{s+t_j}(x;\hbar)|^{2m}\big) \le \delta^{-\gamma/2}\Big\}.$$

Thus, by the Markovian property, we have for all $u \in \mathbb{R}^d$,

$$\begin{aligned}\mathbb{E}\mathcal{J}^u_{t_{n+1}}(x;\hbar) &= \mathbb{E}\big(\mathcal{J}^u_{t_n}(x;\hbar)\cdot(\mathbb{E}\mathcal{J}^{u'}_{t_{n+1}-t_n}(y;0))|_{u'=u[\nabla\phi_{t_n}(x;\hbar)]^{-1},y=\phi_{t_n}(x;\hbar)}\big)\\ &\overset{(3.19)}{\le} \mathbb{E}\mathcal{J}^u_{t_n}(x;\hbar)\exp\{\delta^2 + (t_{n+1}-t_n)\delta^{\beta_0}\}\\ &\le \cdots\\ &\le \prod_{j=0}^n \exp\{\delta^2 + (t_{j+1}-t_j)\delta^{\beta_0}\}\\ &= \exp\{\delta^2(n+1) + t_{n+1}\delta^{\beta_0}\}.\end{aligned}\tag{3.20}$$

Let N^δ_t be the jump number of $\hat{L}^\delta_\cdot$ before time t, that is,

$$N^\delta_t = \sum_{s\in(0,t]} 1_{|\Delta\hat{L}^\delta_s|>0} = \int_{|z|>\delta} N((0,t],\mathrm{d}z) = \sum_{s\in(0,t]} 1_{|\Delta L_s|>\delta},$$

which is a Poisson process with intensity $\int_{|z|>\delta}\nu(\mathrm{d}z) =: \lambda_\delta$. We estimate λ_δ as follows: letting $m = [\log\delta^{-1}/\log 2]$, by (3.1) we have

$$\begin{aligned}\lambda_\delta &\le \int_{|z|\ge 1}\nu(\mathrm{d}z) + \sum_{k=0}^m \int_{2^k\delta\le|z|\le 2^{k+1}\delta}\nu(\mathrm{d}z)\\ &\le C + \sum_{k=0}^m (2^k\delta)^{-2}\int_{2^k\delta\le|z|\le 2^{k+1}\delta}|z|^2\nu(\mathrm{d}z)\\ &\le C + C\sum_{k=0}^m (2^k\delta)^{-2}(2^{k+1}\delta)^{2-\alpha}\\ &= C + C2^{2-\alpha}\sum_{k=0}^m (2^k\delta)^{-\alpha} \le C\delta^{-\alpha}.\end{aligned}\tag{3.21}$$

Recalling (3.12), (3.18) and the independence of L^δ and $\hat{L}^\delta$, we have for any $x, u \in \mathbb{R}^d$,

$$\mathbb{E}\bigg(\exp\bigg\{c_1\delta^{\gamma-\frac{1}{2}}\int_0^t |uK_s(x)[b,V](X_s(x))|^2\,\mathrm{d}s - \delta^{\gamma-8}\int_0^t |uK_s(x)V(X_s(x))|^2\,\mathrm{d}s\bigg\}1_{\Omega^u_t(x;\hat{L}^\delta)}\bigg)$$

$$
\begin{aligned}
&= \mathbb{E}\big(\mathbb{E}\mathcal{J}_t^u(x;\hbar)|_{\hbar=\hat{L}^\delta}\big) = \sum_{n=0}^\infty \mathbb{E}\big(\big(\mathbb{E}\mathcal{J}_t^u(x;\hbar)\big)_{\hbar=\hat{L}^\delta}; N_t^\delta = n\big) \\
&\overset{(3.20)}{\le} \sum_{n=0}^\infty \exp\{\delta^2(n+1) + t\delta^{\beta_0}\}\mathbb{P}\big(N_t^\delta = n\big) \\
&= \exp\{\delta^2 + t\delta^{\beta_0}\} \sum_{n=0}^\infty \mathrm{e}^{\delta^2 n} \frac{(t\lambda_\delta)^n}{n!} \mathrm{e}^{-t\lambda_\delta} \\
&= \exp\{\delta^2 + t\delta^{\beta_0} + (\mathrm{e}^{\delta^2} - 1)t\lambda_\delta\} \\
&\overset{(3.21)}{\le} \exp\{1 + t\delta^{\beta_0} + C_2 t\delta^{2-\alpha}\} \qquad \forall t, \delta \in (0,1),
\end{aligned}
\tag{3.22}
$$

where in the last step we have used that $\mathrm{e}^s - 1 \le 3s$ for $s \in (0,1)$.

(4) By (3.22) and Chebyshev's inequality, we have for any $\beta \in (0,1)$,

$$
\begin{aligned}
\mathbb{P}\bigg(\bigg\{ & c_1 \delta^{\gamma - \frac{1}{2}} \int_0^t \big|uK_s(x)[b,V]\big(X_s(x)\big)\big|^2 \,\mathrm{d}s \\
& - \delta^{\gamma-8} \int_0^t \big|uK_s(x)V\big(X_s(x)\big)\big|^2 \,\mathrm{d}s \ge t\delta^{-\frac{\beta}{2}} \bigg\} \cap \Omega_t^u\big(x;\hat{L}^\delta\big)\bigg) \\
& \le \exp\{1 + t\delta^{\beta_0} + C_2 t\delta^{2-\alpha} - t\delta^{-\frac{\beta}{2}}\},
\end{aligned}
$$

where $\beta_0 = 0 \wedge (\frac{7}{2} - 2\alpha - \gamma)$, and by (3.9),

$$
\begin{aligned}
\mathbb{P}\big(\big[\Omega_t^u\big(x;\hat{L}^\delta\big)\big]^c\big) &= \mathbb{P}\Big\{ \sup_{s\in[0,t]} \big(|uK_s(x)|^2 \vee |X_s^\delta(x)|^{2m}\big) > \delta^{-\gamma/2} \Big\} \\
&\le \delta^p \mathbb{E}\Big(\sup_{s\in[0,t]} \big(|uK_s(x)|^2 \vee |X_s^\delta(x)|^{2m}\big)^{2p/\gamma}\Big) \\
&\le \big(C_p(x) + C|u|^{4p/\gamma}\big)\delta^p \qquad \forall p \ge 1.
\end{aligned}
$$

In particular, for $\beta \in (0 \vee (4\alpha + 2\gamma - 7), 1)$, there exists a $\delta_0 \in (0,1)$ such that for all $\delta \in (0, \delta_0)$, $t \in (0,1)$, $x \in \mathbb{R}^d$, $|u| = 1$ and $p \ge 1$,

$$
\mathbb{P}\left(\begin{array}{l}
\displaystyle\int_0^t \big|uK_s(x)[b,V]\big(X_s(x)\big)\big|^2 \,\mathrm{d}s \ge \frac{2t\delta^{\frac{1-\beta}{2}-\gamma}}{c_1} \\
\displaystyle\text{and} \int_0^t \big|uK_s(x)V\big(X_s(x)\big)\big|^2 \,\mathrm{d}s \le t\delta^{8-\frac{\beta}{2}-\gamma}
\end{array}\right) \le \exp\{2 - t\delta^{-\frac{\beta}{2}}\} + C_p(x)\delta^p,
$$

which then gives the desired estimate (3.10) by adjusting the constants and rescaling δ.

(5) Finally, if $V, [b,V], [b,[b,V]] \in C_b^2$, then the m in (3.16) can be zero so that the constant $C_p(x)$ does not depend on the starting point x. □

The reduced Malliavin matrix is defined by

$$\hat{\Sigma}_t(x) := \int_0^t K_s(x)\big[(\nabla_z\sigma)(\nabla_z\sigma)^*\big]\big(X_s(x), 0\big)K_s^*(x)\,\mathrm{d}s. \tag{3.23}$$

We are now in a position to prove the following main result of this section.

THEOREM 3.3. *Let $B_1(x) := \nabla_z\sigma(x, 0)$ and define for $j \in \mathbb{N}$,*

$$B_{j+1}(x) := [b, B_j] := b(x)\cdot\nabla B_j(x) - \nabla b(x)\cdot B_j(x).$$

Assume that for some $j_0 \in \mathbb{N}$, $m \in \mathbb{N}_0$ and $c_0 > 0$,

$$\inf_{|u|=1}\sum_{j=1}^{j_0}\big|uB_j(x)\big|^2 \geq c_0/\big(1+|x|^m\big), \qquad x \in \mathbb{R}^d. \tag{3.24}$$

Under (3.1), (3.3), (3.4) and (3.6), there exist $\gamma = \gamma(\alpha, j_0) \in (0,1)$ and constants $C_2 \geq 1$, $c_2 \in (0,1)$ such that for all $t \in (0,1)$, $x \in \mathbb{R}^d$, $\lambda \geq 1$ and $p \geq 1$,

$$\sup_{|u|=1}\mathbb{E}\exp\big\{-\lambda u\hat{\Sigma}_t(x)u^*\big\} \leq C_2\exp\big\{-c_2 t\lambda^\gamma\big\} + C_p(x)(\lambda t)^{-p}, \tag{3.25}$$

where $C_p(x)$ continuously depends on x. Moreover, if $m = 0$ in (3.24) and $B_j \in C_b^2$ for each $j = 1, \dots, j_0+1$, then $C_p(x)$ can be independent of x.

PROOF. Let $\beta_1, \beta_2, \beta_3$ be as in (3.10). Set $a := \frac{\beta_1}{\beta_2} \leq 1$ and define for $j = 1, \dots, j_0$,

$$E_j := \left\{\int_0^t \big|uK_s(x)B_j\big(X_s(x)\big)\big|^2\,\mathrm{d}s \leq t\delta^{a^j\beta_2}\right\}.$$

Since $a^{j+1}\beta_2 = a^j\beta_1$ and $B_{j+1} = [b, B_j]$, by (3.10) with δ replaced by δ^{a^j}, we have for any $p \geq 1$,

$$\mathbb{P}\big(E_j \cap E_{j+1}^c\big) \leq C_1\exp\big\{-c_1 t\delta^{-a^j\beta_3}\big\} + C_p(x)\delta^{a^j p}. \tag{3.26}$$

Noticing that

$$E_1 \subset \left(\bigcap_{j=1}^{j_0} E_j\right) \cup \left(\bigcup_{j=1}^{j_0-1}\big(E_j \cap E_{j+1}^c\big)\right),$$

we have

$$\mathbb{P}(E_1) \leq \mathbb{P}\left(\bigcap_{j=1}^{j_0} E_j\right) + \sum_{j=1}^{j_0-1}\mathbb{P}\big(E_j \cap E_{j+1}^c\big). \tag{3.27}$$

On the other hand, if we define

$$\tau := \inf\big\{t \geq 0 : \big|J_t(x)\big| \wedge \big(1+|X_t(x)|^m\big) \geq \delta^{-a^{j_0}\beta_2/4}\big\},$$

then for any $s \leq \tau$ and $|u| = 1$,

$$|uK_s(x)|^2 \geq |J_s(x)|^{-2} \geq \delta^{a^{j_0}\beta_2/2}.$$

Thus, by (3.24) we have

$$\begin{aligned}\bigcap_{j=1}^{j_0} E_j \cap \{\tau \geq t\} &\subset \left\{\sum_{j=1}^{j_0} \int_0^t |uK_s(x)B_j(X_s(x))|^2 \, \mathrm{d}s \leq t \sum_{j=1}^{j_0} \delta^{a^j\beta_2}; \tau \geq t\right\} \\ &\subset \left\{c_0 \int_0^t |uK_s(x)|^2/(1+|X_s(x)|^m) \, \mathrm{d}s \leq t \sum_{j=1}^{j_0} \delta^{a^{j_0}\beta_2}; \tau \geq t\right\} \qquad (3.28) \\ &\subset \{tc_0\delta^{3a^{j_0}\beta_2/4} \leq tj_0\delta^{a^{j_0}\beta_2}\} = \varnothing,\end{aligned}$$

provided $\delta < \delta_1 = (c_0/j_0)^{4/(a^{j_0}\beta_2)}$. Moreover, by (3.9), we have for any $p \geq 2$,

$$\mathbb{P}(\tau < t) \leq \mathbb{P}\Big(\sup_{s\in[0,t]} (|J_s(x)| \wedge (1+|X_s(x)|^m)) \geq \delta^{-a^{j_0}\beta_2/4}\Big) \leq C_p(x)\delta^{pa^{j_0}\beta_2/4}.$$

Now, combining this with (3.26)–(3.28) and resetting $\varepsilon = \delta^{\beta_1}$ and $\theta = a^{j_0}\beta_3/\beta_1$, we obtain that for all $\varepsilon \in (0,1)$, $t \in (0,1)$, $x \in \mathbb{R}^d$ and $p \geq 1$,

$$\sup_{|u|=1} \mathbb{P}\left\{\int_0^t |uK_s(x)B_1(X_s(x))|^2 \, \mathrm{d}s \leq t\varepsilon\right\} \leq C_2 \exp\{-c_1 t\varepsilon^{-\theta}\} + C_p(x)\varepsilon^p.$$

For $\lambda \geq t$, setting $r := (\lambda/t)^{\frac{-1}{1+\theta}}$ and $\xi := \frac{1}{t}\int_0^t |uK_s(x)B_1(X_s(x))|^2 \, \mathrm{d}s$, we have

$$\begin{aligned}\mathbb{E}\mathrm{e}^{-\lambda\xi} &= \int_0^\infty \lambda \mathrm{e}^{-\lambda\varepsilon}\mathbb{P}(\xi \leq \varepsilon) \, \mathrm{d}\varepsilon \\ &\leq \int_r^\infty \lambda \mathrm{e}^{-\lambda\varepsilon} \, \mathrm{d}\varepsilon + \int_0^r \lambda \mathrm{e}^{-\lambda\varepsilon}(C_2\mathrm{e}^{-c_1 t\varepsilon^{-\theta}} + C_p(x)\varepsilon^p) \, \mathrm{d}\varepsilon \\ &= \mathrm{e}^{-\lambda r} + C_2 \int_0^{\lambda r} \mathrm{e}^{-s-c_1 t\lambda^\theta s^{-\theta}} \, \mathrm{d}s + C_p(x)\lambda^{-p} \int_0^{\lambda r} \mathrm{e}^{-s}s^p \, \mathrm{d}s \\ &\leq \mathrm{e}^{-\lambda r} + C_2\mathrm{e}^{-c_1 tr^{-\theta}} \int_0^{\lambda r} \mathrm{e}^{-s} \, \mathrm{d}s + C_p(x)\lambda^{-p} \\ &\leq \mathrm{e}^{-t(\lambda/t)^{\frac{\theta}{1+\theta}}} + C_2\mathrm{e}^{-c_1 t(\lambda/t)^{\frac{\theta}{1+\theta}}} + C_p(x)\lambda^{-p}.\end{aligned}$$

By replacing λ with λt and recalling definition (3.23), we obtain the desired estimate (3.25).

If $m = 0$ in (3.24) and $B_j \in C_b^2$ for each $j = 1, \ldots, j_0 + 1$, from the above proof and by Lemma 3.2, it is easy to see that $C_p(x)$ is independent of x. $\square$

4. Proof of Theorem 1.1. In the remainder of this paper, we assume (H_b^σ) and (H^ν) and choose δ being small enough so that

$$|\nabla_x \sigma(x,z)| \le \frac{1}{2}, \qquad |z| \le \delta, \tag{4.1}$$

and set

$$\Gamma_0^\delta := \{z \in \mathbb{R}^d : 0 < |z| < \delta\}.$$

Let $X_t(x) = X_t$ solve the following SDE:

$$X_t = x + \int_0^t b(X_s)\,\mathrm{d}s + \int_0^t \int_{\Gamma_0^\delta} \sigma(X_{s-}, z)\tilde{N}(\mathrm{d}s, \mathrm{d}z). \tag{4.2}$$

It is well known that the generator of $X_t(x)$ is given by $\mathcal{L}_0$ as in (1.3).

This section is based on Section 2.1, Lemma 2.5 and Theorem 3.3. We first prove the following Malliavin differentiability of X_t with respect to ω in the sense of Theorem 2.1.

LEMMA 4.1. *Fix* $\mathbf{v} \in \mathbb{V}_{\infty-}$. *For any* $t \in [0,1]$, *we have* $X_t \in \mathbb{W}_{\mathbf{v}}^{1,\infty-}(\Omega)$ *and*

$$\begin{aligned} D_{\mathbf{v}} X_t = {} & \int_0^t \nabla b(X_s) D_{\mathbf{v}} X_s \,\mathrm{d}s + \int_0^t \int_{\Gamma_0^\delta} \nabla_x \sigma(X_{s-}, z) D_{\mathbf{v}} X_{s-} \tilde{N}(\mathrm{d}s, \mathrm{d}z) \\ & + \int_0^t \int_{\Gamma_0^\delta} \langle \nabla_z \sigma(X_{s-}, z), \mathbf{v}(s,z)\rangle N(\mathrm{d}s, \mathrm{d}z). \end{aligned} \tag{4.3}$$

Moreover, for any $p \ge 2$, *we have*

$$\sup_{x \in \mathbb{R}^d} \mathbb{E}\Big(\sup_{t\in[0,1]} |D_{\mathbf{v}} X_t(x)|^p\Big) < \infty. \tag{4.4}$$

PROOF. (1) Consider the following Picard's iteration: $X_t^0 \equiv x$ and for $n \in \mathbb{N}$,

$$X_t^n := x + \int_0^t b(X_s^{n-1})\,\mathrm{d}s + \int_0^t \int_{\Gamma_0^\delta} \sigma_2(X_{s-}^{n-1}, z)\tilde{N}(\mathrm{d}s, \mathrm{d}z).$$

Since b and σ are Lipschitz continuous, it is by now standard to prove that for any $p \ge 2$,

$$\sup_{n\in\mathbb{N}} \mathbb{E}\Big(\sup_{t\in[0,1]} |X_t^n|^p\Big) < \infty \quad \text{and} \quad \lim_{n\to\infty} \mathbb{E}\Big(\sup_{t\in[0,1]} |X_t^n - X_t|^p\Big) = 0. \tag{4.5}$$

(2) Next, we use induction to show that for each $n \in \mathbb{N}$,

$$X_t^n \in \mathbb{W}_{\mathbf{v}}^{1,\infty-}(\Omega) \quad \text{and} \quad \mathbb{E}\Big(\sup_{t\in[0,1]} |D_{\mathbf{v}} X_t^n|^p\Big) < +\infty \qquad \forall p \ge 2. \tag{4.6}$$

First of all, it is clear that (4.6) holds for $n=0$. Suppose now that (4.6) holds for some $n\in\mathbb{N}$. By (4.5) and the induction hypothesis, it is easy to check that the assumptions of Proposition 2.2 are satisfied. Thus, $X_t^{n+1}\in\mathbb{W}_{\mathbf{v}}^{1,\infty-}(\Omega)$ and

$$D_{\mathbf{v}}X_t^{n+1}=\int_0^t\nabla b(X_s^n)D_{\mathbf{v}}X_s^n\,\mathrm{d}s+\int_0^t\int_{\Gamma_0^\delta}\nabla_x\sigma(X_{s-}^n,z)D_{\mathbf{v}}X_{s-}^n\tilde{N}(\mathrm{d}s,\mathrm{d}z)$$
$$+\int_0^t\int_{\Gamma_0^\delta}\langle\nabla_z\sigma(X_{s-}^n,z),\mathbf{v}(s,z)\rangle N(\mathrm{d}s,\mathrm{d}z).$$

By Lemma 2.3, we have for any $p\geq 2$,

$$\begin{aligned}\mathbb{E}\Big(\sup_{s\in[0,t]}|D_{\mathbf{v}}X_s^{n+1}|^p\Big)&\leq C\int_0^t\mathbb{E}|D_{\mathbf{v}}X_s^n|^p\,\mathrm{d}s\\&\quad+C\mathbb{E}\Big(\int_0^t\int_{\Gamma_0^\delta}|\langle\nabla_z\sigma(X_{s-}^n,z),\mathbf{v}(s,z)\rangle|\nu(\mathrm{d}z)\,\mathrm{d}s\Big)^p\\&\quad+C\mathbb{E}\Big(\int_0^t\int_{\Gamma_0^\delta}|\langle\nabla_z\sigma(X_{s-}^n,z),\mathbf{v}(s,z)\rangle|^p\nu(\mathrm{d}z)\,\mathrm{d}s\Big).\end{aligned}$$

Since $\mathbf{v}\in\mathbb{V}_{\infty-}$, by (H$\sigma_b$) we further have

$$\begin{aligned}\mathbb{E}\Big(\sup_{s\in[0,t]}|D_{\mathbf{v}}X_s^{n+1}|^p\Big)&\leq C\int_0^t\mathbb{E}|D_{\mathbf{v}}X_s^n|^p\,\mathrm{d}s+C\\&\leq C\int_0^t\mathbb{E}\Big(\sup_{r\in[0,s]}|D_{\mathbf{v}}X_r^n|^p\Big)\mathrm{d}s+C,\end{aligned}$$

where C is independent of n and the starting point x. Thus, we have proved (4.6) by the induction hypothesis. Moreover, by Gronwall's inequality, we also have

$$\sup_{n\in\mathbb{N}}\mathbb{E}\Big(\sup_{s\in[0,1]}|D_{\mathbf{v}}X_s^n|^p\Big)<+\infty.\tag{4.7}$$

(3) Let Y_t solve the following linear matrix-valued SDE:

$$Y_t=\int_0^t\nabla b(X_s)Y_s\,\mathrm{d}s+\int_0^t\int_{\Gamma_0^\delta}\nabla_x\sigma(X_{s-},z)Y_{s-}\tilde{N}(\mathrm{d}s,\mathrm{d}z)$$
$$+\int_0^t\int_{\Gamma_0^\delta}\langle\nabla_z\sigma(X_{s-},z),\mathbf{v}(s,z)\rangle N(\mathrm{d}s,\mathrm{d}z).$$

By Fatou's lemma and (4.5), (4.7), for any $p\geq 2$, we have

$$\varlimsup_{n\to\infty}\mathbb{E}|D_{\mathbf{v}}X_t^n-Y_t|^p\leq C\int_0^t\varlimsup_{n\to\infty}\mathbb{E}|D_{\mathbf{v}}X_s^{n-1}-Y_s|^p\,\mathrm{d}s,$$

which then gives

$$\varlimsup_{n\to\infty}\mathbb{E}|D_{\mathbf{v}}X_t^n-Y_t|^p=0.$$

Thus, $X_t \in \mathbb{W}^{1,p}_{\mathbf{v}}(\Omega)$ and $D_{\mathbf{v}}X_t = Y_t$. Moreover, the estimate (4.4) follows by (4.5) and (4.7). $\square$

Let $J_t = J_t(x)$ be the Jacobian matrix of $x \mapsto X_t(x)$, and $K_t(x)$ be the inverse of $J_t(x)$. Recalling equations (3.5) and (4.3), by the formula of constant variation, we have for any $\mathbf{v} \in \mathbb{V}_{\infty-}$,

$$D_{\mathbf{v}}X_t = J_t \int_0^t \int_{\Gamma_0^\delta} K_s \nabla_z \sigma(X_{s-}, z)\mathbf{v}(s, z)N(\mathrm{d}s, \mathrm{d}z). \tag{4.8}$$

Here, the integral is the Lebesgue–Stieltjes integral.

Next, we want to choose special direction $\Theta = (\mathbf{v}_1, \dots, \mathbf{v}_d)$ so that the Malliavin matrix $D_\Theta X_t := (D_{\mathbf{v}_1}X_t, \dots, D_{\mathbf{v}_d}X_t)$ is invertible. Let

$$U(x, z) := \big(\mathbb{I} + \nabla_x \sigma(x, z)\big)^{-1} \nabla_z \sigma(x, z), \qquad x \in \mathbb{R}^d, z \in \Gamma_0^\delta,$$

and define

$$\mathbf{v}_j(x; s, z) := \big[K_{s-}(x)U\big(X_{s-}(x), z\big)\big]^*_{\cdot j}\zeta(z),$$

where $\zeta(z) = \zeta_\delta(z)$ is a nonnegative smooth function with

$$\zeta_\delta(z) = |z|^3, \qquad |z| \le \delta/4, \qquad \zeta_\delta(z) = 0, \qquad |z| > \delta/2.$$

The following lemma is easily verified by definitions and (3.9).

LEMMA 4.2. *For any $m \in \mathbb{N}_0$, there is a constant $C > 0$ such that for all $x \in \mathbb{R}^d$ and $z \in \Gamma_0^\delta$,*

$$\big|\nabla_x^m U(x, z)\big|, \qquad \big|\nabla_z^m U(x, z)\big| \le C, \qquad \big|U(x, z) - U(x, 0)\big| \le C|z|. \tag{4.9}$$

Moreover, for each $j = 1, \dots, d$ and $x \in \mathbb{R}^d$, $\mathbf{v}_j(x) \in \mathbb{V}_{\infty-}$.

Write

$$\Theta(s, z) := \Theta(x; s, z) := \big(\mathbf{v}_1(x; s, z), \dots, \mathbf{v}_d(x; s, z)\big)$$

and

$$(D_\Theta X_t)_{ij} := D_{\mathbf{v}_j} X_t^i.$$

Noticing that by equation (3.7),

$$K_s = K_{s-}\big(\mathbb{I} + \nabla_x \sigma(X_{s-}, \Delta L_s)\big)^{-1},$$

by (4.8) we have

$$D_\Theta X_t(x) = J_t(x)\Sigma_t(x), \tag{4.10}$$

where

$$\Sigma_t(x) := \int_0^t \int_{\Gamma_0^\delta} K_{s-}(x)\big(UU^*\big)\big(X_{s-}(x), z\big)K^*_{s-}(x)\zeta(z)N(\mathrm{d}s, \mathrm{d}z). \tag{4.11}$$

LEMMA 4.3. *For any $p \geq 2$ and $m, k \in \mathbb{N}_0$ with $m + k \geq 1$, we have*

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}\Big(\sup_{t\in[0,1]} \big|D_{\mathbf{v}_{j_1}} \cdots D_{\mathbf{v}_{j_m}} \nabla^k X_t(x)\big|^p\Big) < \infty, \tag{4.12}$$

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}\Big(\sup_{t\in[0,1]} \big|D_{\mathbf{v}_{j_1}} \cdots D_{\mathbf{v}_{j_m}} \operatorname{div}\big(\mathbf{v}_i(x)\big)\big|^p\Big) < \infty, \tag{4.13}$$

where $j_1, \dots, j_m$ and i runs in $\{1, 2, \dots, d\}$.

PROOF. For $m + k = 1$, (4.12) has been proven in (3.9) and (4.4). For general k and m, it follows by induction. Let us look at (4.13) with $m = 1$. Notice that by (2.4),

$$\operatorname{div}(\mathbf{v}_i) = \int_0^1 \int_{\Gamma_0^\delta} \big[\langle \nabla \log \kappa(z), \mathbf{v}_i(s, z)\rangle + \operatorname{div}_z(\mathbf{v}_i)(s, z)\big] \tilde{N}(\mathrm{d}s, \mathrm{d}z).$$

By Proposition 2.2, we have

$$\begin{aligned} D_{\mathbf{v}_j} \operatorname{div}(\mathbf{v}_i) = & \int_0^1 \int_{\Gamma_0^\delta} \big[\langle \nabla \log \kappa(z), D_{\mathbf{v}_j} \mathbf{v}_i(s, z)\rangle + D_{\mathbf{v}_j} \operatorname{div}_z(\mathbf{v}_i)(s, z)\big] \tilde{N}(\mathrm{d}s, \mathrm{d}z) \\ & + \int_0^1 \int_{\Gamma_0^\delta} \langle \mathbf{v}_j(s, z), \nabla_z \langle \nabla \log \kappa(z), \mathbf{v}_i(s, z)\rangle \\ & + \nabla_z \operatorname{div}_z(\mathbf{v}_i)(s, z)\rangle N(\mathrm{d}s, \mathrm{d}z). \end{aligned}$$

In view of $\operatorname{supp} \mathbf{v}_i(s, \cdot) \subset \Gamma_0^{\frac{\delta}{2}}$, by Lemma 2.3 and (1.7), (4.9), (4.12), one obtains (4.13) with $m = 1$. For general m, it follows by similar calculations. □

Write

$$\mathcal{T}_t^0 f(x) := \mathbb{E} f\big(X_t(x)\big). \tag{4.14}$$

The following lemma is proven in the Appendix.

LEMMA 4.4. *Under (H_b^σ), there exists a constant $C > 0$ such that for any $f \in L^1(\mathbb{R}^d)$,*

$$\sup_{t\in[0,1]} \int_{\mathbb{R}^d} \big|\mathcal{T}_t^0 f(x)\big| \,\mathrm{d}x \leq C \int_{\mathbb{R}^d} \big|f(x)\big| \,\mathrm{d}x. \tag{4.15}$$

Now we can prove the following main result of this section.

THEOREM 4.5. *Assume that (H_b^σ), (H^ν) and (H^{j_0}) hold. Let δ be as in (4.1). For any $k, m, n \in \mathbb{N}_0$ with $n \leq k + m$, $p \in (1, \infty]$ and $R > 0$, there exist $\gamma_{kmn} \geq 0$*

only depending on k, m, n, α, j_0, d and a constant $C_R \geq 1$ such that for all $f \in \bigcap_k \mathbb{W}^{k,p}(\mathbb{R}^d)$ and $t \in (0,1)$,

$$\|\nabla^k \mathcal{T}_t^0 \nabla^m f\|_{p;R} \leq C_R t^{-\gamma_{kmn}} \|f\|_{n,p}, \tag{4.16}$$

where $\|\cdot\|_{p;R}$ denotes the norm in $L^p(B_R)$, B_R is the ball in $\mathbb{R}^d$ with radius R, γ_{kmn} is increasing with respect to k, m and decreasing in n, and $\gamma_{kmn} = 0$ for $n = k + m$. In particular, $X_t(x)$ admits a smooth density $\rho_t(x, y)$ such that

$$\partial_t \rho_t(x, y) = \mathcal{L}_0 \rho_t(\cdot, y)(x) \qquad \forall (t, x, y) \in (0, 1) \times \mathbb{R}^d \times \mathbb{R}^d.$$

Moreover, if $m = 0$ and $B_j \in C_b^2$ in (H^{j_0}) *for each $j = 1, \ldots, j_0 + 1$, then R can be infinity in* (4.16) *so that the local norm in* (4.16) *becomes a global norm.*

PROOF. Below we only prove (4.16) for $p \in (1, \infty)$. For $p = \infty$, it is similar and simpler. We assume $f \in C_0^\infty(\mathbb{R}^d)$ and divide the proof into four steps.

(1) Let $\Sigma_t(x)$ be defined by (4.11). In view of $U(x, 0) = \nabla_z \sigma(x, 0)$, by (3.9), (4.9), (3.25) and Lemma 2.5, there are constants $C_3 \geq 1$, $c_3, \theta \in (0, 1)$ and $\gamma = \gamma(\alpha, j_0) \in (0, 1)$ such that for all $t \in (0, 1)$, $x \in \mathbb{R}^d$, $\lambda \geq 1$ and $p \geq 1$,

$$\sup_{|u|=1} \mathbb{E} \exp\{-\lambda u \Sigma_t(x) u^*\} \leq C_3 \exp\{-c_3 t \lambda^\gamma\} + C_p(x)(\lambda^\theta t)^{-p}, \tag{4.17}$$

where $\Sigma_t(x)$ is defined by (4.11). As in [26], Lemma 5.3, for any $p \geq 1$ and $x \in \mathbb{R}^d$, there exist constant $C_p(x) \geq 1$ an $\gamma' = \gamma'(\alpha, j_0, d) > 0$ such that for all $t \in (0, 1)$,

$$\mathbb{E}\big((\det \Sigma_t(x))^{-p}\big) \leq C_p(x) t^{-\gamma' p},$$

which in turn gives that for all $p \geq 1$,

$$\|\Sigma_t^{-1}(x)\|_{L^p(\Omega)} \leq C_p(x) t^{-\gamma'}. \tag{4.18}$$

(2) For $t \in (0, 1)$ and $x \in \mathbb{R}^d$, let $\mathscr{C}_t(x)$ be the class of all polynomial functionals of

$$\operatorname{div}\Theta, \Sigma_t^{-1}, K_t, (\nabla^k X_t)_{k=1}^{\ell_1}, \big(D_{\mathbf{v}_{j_1}} \cdots D_{\mathbf{v}_{j_m}}(X_t, \ldots, \nabla^{\ell_2} X_t, K_t, \operatorname{div}\Theta, \Sigma_t)\big)_{m=1}^{\ell_3},$$

where $\ell_1, \ell_2, \ell_3 \in \mathbb{N}$, $j_i \in \{1, \ldots, d\}$, and the starting point x is dropped in the above random variables. By (4.18) and Lemma 4.3, for any $H_t(x) \in \mathscr{C}_t(x)$, there exists a $\gamma(H) \geq 0$ only depending on the degree of Σ_t^{-1} appearing in H and α, j_0, d such that for all $t \in (0, 1)$ and $p \geq 1$,

$$\|H_t(x)\|_{L^p(\Omega)} \leq C_p(x) t^{-\gamma(H)}, \tag{4.19}$$

where $C_p(x)$ continuously depends on x. Notice that if H_t does not contain Σ_t^{-1}, then $\gamma(H) = 0$.

(3) Let $\xi \in \mathscr{C}_t(x)$. Since $D_\Theta X$ is an invertible matrix, by (4.10) and the integration by parts formula (2.3), we have

$$\begin{aligned}\mathbb{E}\big((\nabla f)(X_t)\xi\big) &= \mathbb{E}\big(\nabla f(X_t) D_\Theta X_t \cdot (D_\Theta X_t)^{-1}\xi\big) = \mathbb{E}\big(D_\Theta f(X_t)\Sigma_t^{-1}K_t\xi\big) \\ &= \mathbb{E}\left(-f(X_t)\left(\operatorname{div}\Theta\cdot\Sigma_t^{-1}K_t\xi - \sum_{i=1}^d D_{\mathbf{v}_i}\big[(\Sigma_t^{-1}K_t)_{i\cdot}\xi\big]\right)\right) \\ &= \mathbb{E}\big(f(X_t)\xi'\big),\end{aligned}$$

where $\xi' \in \mathscr{C}_t(x)$. Starting from this formula, by the chain rule and induction, we have for $n \le k+m$,

$$\begin{aligned}\nabla^k\mathbb{E}\big((\nabla^m f)(X_t)\big) &= \sum_{j=0}^k \mathbb{E}\big((\nabla^{m+j} f)(X_t)G_j(\nabla X_t,\dots,\nabla^k X_t)\big) \\ &= \sum_{j=0}^n \mathbb{E}\big((\nabla^j f)(X_t)H_j\big),\end{aligned}$$

where $\{G_j, j=1,\dots,k\}$ are real polynomial functions and $H_j \in \mathscr{C}_t(x)$. Notice that if $n = k+m$, then H_j will not contain Σ_t^{-1}.

(4) Now, for any $p \in (1,\infty)$, by Hölder's inequality, we have

$$\begin{aligned}\|\nabla^k \mathcal{T}_t^0 \nabla^m f\|_{p;R} &\le \sum_{j=0}^n \left(\int_{B_R} \big|\mathbb{E}\big((\nabla^j f)(X_t(x))H_j(x)\big)\big|^p \,\mathrm{d}x\right)^{\frac{1}{p}} \\ &\le \sum_{j=0}^n \left(\int_{B_R} \mathbb{E}\big(|\nabla^j f|^p(X_t(x))\big)\big(\mathbb{E}|H_j(x)|^{\frac{p}{p-1}}\big)^{p-1}\,\mathrm{d}x\right)^{\frac{1}{p}} \\ &\overset{(4.19)}{\le} C_R \sum_{j=0}^n t^{-\gamma(H_j)}\left(\int_{\mathbb{R}^d} \mathbb{E}\big(|\nabla^j f|^p(X_t(x))\big)\,\mathrm{d}x\right)^{\frac{1}{p}} \\ &\overset{(4.15)}{\le} C_R t^{-\max\{\gamma(H_j), j=1,\dots,n\}}\|f\|_{n,p}, \quad t\in(0,1).\end{aligned}$$

(5) Finally, if $m = 0$ and $B_j \in C_b^2$ in (H^{j_0}) for each $j = 1,\dots,j_0+1$, then by Theorem 3.3, all the constants appearing in the above estimates will be independent of the starting point x. Thus, we can take $R = \infty$ in step (4). The proof is complete. $\square$

5. Proofs of Theorems 1.2 and 1.4. We first recall some definitions about the Sobolev and Hölder spaces. For $k \in \mathbb{N}_0$ and $p \in [1,\infty]$, let $\mathbb{W}^{k,p} = \mathbb{W}^{k,p}(\mathbb{R}^d)$ be the usual Sobolev space with the norm:

$$\|\varphi\|_{k,p} := \sum_{j=0}^k \|\nabla^j \varphi\|_p.$$

For $\beta \geq 0$ and $p \in [1, \infty)$, let $\mathbb{H}^{\beta,p} := (I - \Delta)^{-\frac{\beta}{2}}(L^p(\mathbb{R}^d))$ be the usual Bessel potential space. For $p = \infty$, let $\mathbb{H}^{\beta,\infty}$ be the usual Hölder space, that is, if $\beta = k + \theta$ with $\theta \in [0, 1)$, then

$$\|\varphi\|_{\beta,\infty} := \|\varphi\|_{k,\infty} + \big[\nabla^k \varphi\big]_\theta < \infty, \tag{5.1}$$

where $[\nabla^k \varphi]_0 := 0$ by convention and for $\theta \in (0, 1)$,

$$\big[\nabla^k \varphi\big]_\theta := \sup_{x \neq y} \frac{|\nabla^k \varphi(x) - \nabla^k \varphi(y)|}{|x - y|^\theta}.$$

It is well known that for any $k \in \mathbb{N}_0$ and $p \in (1, \infty)$ (cf. [21]),

$$\mathbb{H}^{k,p} = \mathbb{W}^{k,p},$$

and for any $\beta_1, \beta_2 \geq 0$, $p \in (1, \infty)$ and $\theta \in [0, 1]$,

$$\big[\mathbb{H}^{\beta_1,p}, \mathbb{H}^{\beta_2,p}\big]_\theta = \mathbb{H}^{\beta_1 + \theta(\beta_2 - \beta_1),p}, \tag{5.2}$$

and if $\beta_1 + \theta(\beta_2 - \beta_1)$ is not an integer, then

$$\big(\mathbb{H}^{\beta_1,\infty}, \mathbb{H}^{\beta_2,\infty}\big)_{\theta,\infty} = \mathbb{H}^{\beta_1 + \theta(\beta_2 - \beta_1),\infty}, \tag{5.3}$$

where $[\cdot, \cdot]_\theta$ [resp., $(\cdot, \cdot)_{\theta,\infty}$] stands for the complex (resp., real) interpolation space.

We recall the following interpolation theorem (cf. [23], page 59, Theorem (a)).

THEOREM 5.1. *Let $A_i \subset B_i, i = 0, 1$ be Banach spaces. Let $\mathscr{T} : A_i \to B_i, i = 0, 1$ be bounded linear operators. For $\theta \in [0, 1]$, we have*

$$\|\mathscr{T}\|_{A_\theta \to B_\theta} \leq \|\mathscr{T}\|^{1-\theta}_{A_0 \to B_0} \|\mathscr{T}\|^\theta_{A_1 \to B_1},$$

where $A_\theta := [A_0, A_1]_\theta$, $B_\theta := [B_0, B_1]_\theta$, and $\|\mathscr{T}\|_{A_\theta \to B_\theta}$ denotes the operator norm of $\mathscr{T}$ mapping A_θ to B_θ. The same is true for real interpolation spaces.

In what follows, we always assume that (H^σ_b) and (H^ν) hold and (H^{j_0}) holds with $m = 0$ and $B_j \in C^2_b$ for each $j = 1, \dots, j_0 + 1$. Let $\mathcal{T}^0_t$ be the semigroup defined by (4.14), whose infinitesimal generator is given by $\mathcal{L}_0$. We have the following.

LEMMA 5.2. *Let γ_{100} be the same as in Theorem 4.5. For any $p \in (1, \infty)$, $\theta \in [0, 1)$ and $\beta \geq 0$, there exist constants $C_1, C_2 > 0$ such that for all $t \in (0, 1)$,*

$$\big\|\mathcal{T}^0_t \varphi\big\|_{\theta+\beta,p} \leq C_1 t^{-\theta\gamma_{100}} \|\varphi\|_{\beta,p}, \tag{5.4}$$

and if β and $\theta + \beta$ are not integers, then

$$\big\|\mathcal{T}^0_t \varphi\big\|_{\theta+\beta,\infty} \leq C_2 t^{-\theta\gamma_{100}} \|\varphi\|_{\beta,\infty}. \tag{5.5}$$

PROOF. Let $\theta \in [0,1)$ and $\beta \geq 0$. For any $p \in (1,\infty]$, by Theorem 4.5 and the interpolation Theorem 5.1, there exists a constant $C>0$ such that for all $t \in (0,1)$,

$$\|\mathcal{T}_t^0\varphi\|_{\frac{\beta}{1-\theta},p} \leq C\|\varphi\|_{\frac{\beta}{1-\theta},p}$$

and

$$\|\mathcal{T}_t^0\varphi\|_{1,p} \leq Ct^{-\gamma_{100}}\|\varphi\|_p.$$

On the other hand, noticing that by (5.2),

$$[\mathbb{H}^{\frac{\beta}{1-\theta},p}, \mathbb{H}^{1,p}]_\theta = \mathbb{H}^{\beta+\theta,p}, \qquad [\mathbb{H}^{\frac{\beta}{1-\theta},p}, \mathbb{H}^{0,p}]_\theta = \mathbb{H}^{\beta,p},$$

and if β and $\theta+\beta$ are not integers, then by (5.3),

$$(\mathbb{H}^{\frac{\beta}{1-\theta},\infty}, \mathbb{H}^{1,\infty})_{\theta,\infty} = \mathbb{H}^{\beta+\theta,\infty}, \qquad (\mathbb{H}^{\frac{\beta}{1-\theta},\infty}, \mathbb{H}^{0,\infty})_{\theta,\infty} = \mathbb{H}^{\beta,\infty},$$

by the interpolation Theorem 5.1 again, we obtain the desired estimate. $\square$

LEMMA 5.3. *Let γ_{010} be the same as in Theorem 4.5. For any $p \in (1,\infty]$ and $\theta \in (0,1)$, there exists a constant $C>0$ such that for all $\varphi \in L^p(\mathbb{R}^d) \cap C_b^\infty(\mathbb{R}^d)$ and $t \in (0,1)$,*

$$\|\mathcal{T}_t^0\Delta^{\frac{\theta}{2}}\varphi\|_p \leq Ct^{-\theta\gamma_{010}}\|\varphi\|_p. \tag{5.6}$$

PROOF. Notice that

$$\mathcal{T}_t^0\Delta^{\frac{\theta}{2}}\varphi(x) = \mathbb{E}\int_{\mathbb{R}^d} \frac{\varphi(X_t(x)+z)-\varphi(X_t(x))}{|z|^{d+\theta}}\,\mathrm{d}z = I_1(x)+I_2(x), \tag{5.7}$$

where

$$I_1(x) := \mathbb{E}\int_{|z|\leq t^{\gamma_{010}}} \frac{\varphi(X_t(x)+z)-\varphi(X_t(x))}{|z|^{d+\theta}}\,\mathrm{d}z,$$

$$I_2(x) := \mathbb{E}\int_{|z|> t^{\gamma_{010}}} \frac{\varphi(X_t(x)+z)-\varphi(X_t(x))}{|z|^{d+\theta}}\,\mathrm{d}z.$$

For $I_1(x)$, setting $\varphi_{sz}(x) := \varphi(x+sz)$, we have

$$\begin{aligned} I_1(x) &= \mathbb{E}\int_{|z|\leq t^{\gamma_{010}}} \left(\int_0^1 z\cdot\nabla\varphi\big(X_t(x)+sz\big)\,\mathrm{d}s\right)\frac{\mathrm{d}z}{|z|^{d+\theta}} \\ &= \int_{|z|\leq t^{\gamma_{010}}} \left(\int_0^1 z\cdot\mathcal{T}_t^0\nabla\varphi_{sz}(x)\,\mathrm{d}s\right)\frac{\mathrm{d}z}{|z|^{d+\theta}}. \end{aligned}$$

Hence,

$$\begin{aligned} \|I_1\|_p &\leq \int_{|z|\leq t^{\gamma_{010}}} \left(\int_0^1 \|\mathcal{T}_t^0\nabla\varphi_{sz}\|_p\,\mathrm{d}s\right)\frac{\mathrm{d}z}{|z|^{d+\theta-1}} \\ &\overset{(4.16)}{\leq} Ct^{-\gamma_{010}}\|\varphi\|_p \int_{|z|\leq t^{\gamma_{010}}} \frac{\mathrm{d}z}{|z|^{d+\theta-1}} \leq Ct^{-\theta\gamma_{010}}\|\varphi\|_p. \end{aligned} \tag{5.8}$$

For $I_2(x)$, we have

$$
\begin{aligned}
\|I_2\|_p &\leq \int_{|z|>t^{\gamma_{010}}} \frac{\|\mathcal{T}_t^0\varphi_z\|_p + \|\mathcal{T}_t^0\varphi\|_p}{|z|^{d+\theta}}\,\mathrm{d}z \\
&\overset{(4.15)}{\leq} C\|\varphi\|_p \int_{|z|>t^{\gamma_{010}}} \frac{1}{|z|^{d+\theta}}\,\mathrm{d}z \leq Ct^{-\theta\gamma_{010}}\|\varphi\|_p.
\end{aligned} \tag{5.9}
$$

Combining (5.7)–(5.9), we obtain (5.6). □

5.1. *Proof of Theorem* 1.2. Let $\mathscr{L}$ be a bounded linear operator in Sobolev space $\mathbb{W}^{k,p}(\mathbb{R}^d)$ for any $p > 1$ and $k \in \mathbb{N}_0$. Let $\mathcal{T}_t$ be the semigroup in $L^p(\mathbb{R}^d)$ associated with $\mathcal{L}_0 + \mathscr{L}$, that is, for any $\varphi \in L^p(\mathbb{R}^d)$,

$$\partial_t \mathcal{T}_t\varphi = \mathcal{L}_0\mathcal{T}_t\varphi + \mathscr{L}\mathcal{T}_t\varphi.$$

By Duhamel's formula, we have

$$\mathcal{T}_t\varphi = \mathcal{T}_t^0\varphi + \int_0^t \mathcal{T}_{t-s}^0\mathscr{L}\mathcal{T}_s\varphi\,\mathrm{d}s. \tag{5.10}$$

LEMMA 5.4. *Let $\gamma_{100}, \gamma_{010}$ be as in Theorem* 4.5. *Fix $\theta \in (0, \frac{1}{\gamma_{100}} \wedge 1)$ and $\theta' \in [0, \frac{1}{\gamma_{010}} \wedge 1)$. For any $m \in \mathbb{N}$ and $p \in (1, \infty)$, there exists a constant $C > 0$ such that for all $t \in (0, 1)$ and $\varphi \in L^p(\mathbb{R}^d) \cap C_b^\infty(\mathbb{R}^d)$,*

$$\|\mathcal{T}_t\Delta^{\frac{\theta'}{2}}\varphi\|_{m\theta,p} \leq Ct^{-m\theta\gamma_{100}-\theta'\gamma_{010}}\|\varphi\|_p. \tag{5.11}$$

PROOF. First of all, since $\mathscr{L}$ is a bounded linear operator in $\mathbb{W}^{k,p}$, by interpolation Theorem 5.1, we have for all $\beta \geq 0$ and $p \in (1, \infty)$,

$$\|\mathscr{L}\varphi\|_{\beta,p} \leq C\|\varphi\|_{\beta,p}.$$

Let $\theta \in (0, \frac{1}{\gamma_{100}} \wedge 1)$ and $m \in \mathbb{N}$. By (5.10) and Lemma 5.2, we have

$$
\begin{aligned}
\|\mathcal{T}_t\varphi\|_{m\theta,p} &\leq \|\mathcal{T}_t^0\varphi\|_{m\theta,p} + \int_0^t \|\mathcal{T}_{t-s}^0\mathscr{L}\mathcal{T}_s\varphi\|_{m\theta,p}\,\mathrm{d}s \\
&\leq Ct^{-\theta\gamma_{100}}\|\varphi\|_{(m-1)\theta,p} + C\int_0^t \|\mathcal{T}_s\varphi\|_{m\theta,p}\,\mathrm{d}s,
\end{aligned}
$$

which, by Gronwall's inequality, yields that for all $t \in (0, 1)$,

$$\|\mathcal{T}_t\varphi\|_{m\theta,p} \leq Ct^{-\theta\gamma_{100}}\|\varphi\|_{(m-1)\theta,p}.$$

Thus, by the semigroup property of $\mathcal{T}_t$ and iteration, we obtain

$$\|\mathcal{T}_{(m+1)t}\varphi\|_{m\theta,p} \leq Ct^{-\theta\gamma_{100}}\|\mathcal{T}_{mt}\varphi\|_{(m-1)\theta,p} \leq \cdots \leq Ct^{-m\theta\gamma_{100}}\|\mathcal{T}_t\varphi\|_p. \tag{5.12}$$

On the other hand, by (5.10) and Lemma 5.2, we have

$$\begin{aligned}\left\|\mathcal{T}_t\Delta^{\frac{\theta'}{2}}\varphi\right\|_p &\leq \left\|\mathcal{T}^0_t\Delta^{\frac{\theta'}{2}}\varphi\right\|_p+\int_0^t\left\|\mathcal{T}^0_{t-s}\mathscr{L}\mathcal{T}_s\Delta^{\frac{\theta'}{2}}\varphi\right\|_p\mathrm{d}s\\ &\leq Ct^{-\theta'\gamma_{010}}\|\varphi\|_p+C\int_0^t\left\|\mathcal{T}_s\Delta^{\frac{\theta'}{2}}\varphi\right\|_p\mathrm{d}s,\end{aligned}$$

which, by Gronwall's inequality, yields that for all $t\in(0,1)$,

$$\left\|\mathcal{T}_t\Delta^{\frac{\theta'}{2}}\varphi\right\|_p\leq Ct^{-\theta'\gamma_{010}}\|\varphi\|_p. \tag{5.13}$$

Combining (5.12) with (5.13), we obtain the desired estimate. □

Now we can give the following.

PROOF OF THEOREM 1.2. By a standard time-shifting argument, we can assume $t\in(0,1)$. Let $\mathscr{L}$ be a bounded linear operator in Sobolev space $\mathbb{W}^{k,p}(\mathbb{R}^d)$ for any $p>1$ and $k\in\mathbb{N}_0$. Let $\mathcal{T}_t$ be the semigroup in $L^p(\mathbb{R}^d)$ associated with $\mathcal{L}_0+\mathscr{L}$. For any $p\in(1,\infty)$ and $\varphi\in L^p(\mathbb{R}^d)$, by Lemma 5.4 and Sobolev's embedding theorem, we have $\mathcal{T}_t\varphi\in C^\infty_b(\mathbb{R}^d)$ and for any $k\in\mathbb{N}_0$, $t\in(0,1)$ and $\theta'\in[0,\frac{1}{\gamma_{010}}\wedge 1)$,

$$\left\|\mathcal{T}_t\Delta^{\frac{\theta'}{2}}\varphi\right\|_{k,\infty}\leq C\left\|\mathcal{T}_t\Delta^{\frac{\theta'}{2}}\varphi\right\|_{k+d,p}\leq Ct^{-(k+d)\gamma_{100}-\theta'\gamma_{0101}}\|\varphi\|_p. \tag{5.14}$$

In particular, there is a function $\rho_t(x,\cdot)\in L^{\frac{p}{p-1}}(\mathbb{R}^d)$ such that for any $\varphi\in L^p(\mathbb{R}^d)$,

$$\mathcal{T}_t\varphi(x)=\int_{\mathbb{R}^d}\varphi(y)\rho_t(x,y)\,\mathrm{d}y.$$

By (5.14), for each $k\in\mathbb{N}_0$, we have

$$\begin{aligned}&\operatorname*{ess.\,sup}_{x\in\mathbb{R}^d}\left\|\nabla^k_x\Delta_y^{\frac{\theta'}{2}}\rho_t(x,\cdot)\right\|_{\frac{p}{p-1}}\\ &\quad=\operatorname*{ess.\,sup}_{x\in\mathbb{R}^d}\sup_{\varphi\in C^\infty_0(\mathbb{R}^d),\|\varphi\|_p\leq1}\left|\int\varphi(y)\nabla^k_x\Delta_y^{\frac{\theta'}{2}}\rho_t(x,y)\,\mathrm{d}y\right|\\ &\quad=\operatorname*{ess.\,sup}_{x\in\mathbb{R}^d}\sup_{\varphi\in C^\infty_0(\mathbb{R}^d),\|\varphi\|_p\leq1}\left|\int\Delta_y^{\frac{\theta'}{2}}\varphi(y)\nabla^k_x\rho_t(x,y)\,\mathrm{d}y\right|\\ &\quad=\operatorname*{ess.\,sup}_{x\in\mathbb{R}^d}\sup_{\varphi\in C^\infty_0(\mathbb{R}^d),\|\varphi\|_p\leq1}\left|\nabla^k_x\mathcal{T}_t\Delta^{\frac{\theta'}{2}}\varphi(x)\right|\\ &\quad=\sup_{\varphi\in C^\infty_0(\mathbb{R}^d),\|\varphi\|_p\leq1}\left\|\nabla^k_x\mathcal{T}_t\Delta^{\frac{\theta'}{2}}\varphi\right\|_\infty\leq Ct^{-(k+d)\gamma_{100}-\theta'\gamma_{0101}},\end{aligned}$$

where ∇^k_x stands for the distributional derivative. Therefore, by Fubini's theorem, we have for any $R > 0$ and $p > 1$,

$$\int_{\mathbb{R}^d}\int_{B_R} \left|\nabla^k_x \Delta_y^{\frac{\theta'}{2}} \rho_t(x, y)\right|^{\frac{p}{p-1}} \mathrm{d}x\, \mathrm{d}y < \infty,$$

which, by Sobolev's embedding theorem again, produces that $(x, y) \mapsto \rho_t(x, y)$ is continuous and for each $y \in \mathbb{R}^d$,

$$x \mapsto \rho_t(x, y) \text{ is smooth.}$$

As for equation (1.12), it follows by (5.10). □

5.2. *Proof of Theorem* 1.4. Let δ be as in (4.1). We decompose the operator $\mathcal{L}^\nu_{\sigma,b}$ as

$$\mathcal{L}^\nu_{\sigma,b} f(x) = \mathcal{L}_0 f(x) + \mathcal{L}_1 f(x),$$

where

$$\mathcal{L}_0 f(x) := \text{p.v.} \int_{|z|<\delta} f\big((x + \sigma(x, z)\big) - f(x)\big)\nu(\mathrm{d}z) + b(x) \cdot \nabla f(x)$$

and

$$\mathcal{L}_1 f(x) := \int_{|z|\geq\delta} f\big((x + \sigma(x, z)\big) - f(x)\big)\nu(\mathrm{d}z).$$

LEMMA 5.5. *If* $\int_{|z|\geq 1} |z|^q \nu(\mathrm{d}z) < \infty$ *for some* $q > 0$, *then for any* $\beta \in [0, q]$, *there exists a constant* $C > 0$ *such that for all* $f \in \mathbb{H}^{\beta,\infty}$,

$$\|\mathcal{L}_1 f\|_{\beta,\infty} \leq C\|f\|_{\beta,\infty}. \tag{5.15}$$

PROOF. First of all, (5.15) is clearly true for $\beta = 0$. By interpolation Theorem 5.1, it suffices to prove (5.15) for $\beta \in [0, q] \cap \mathbb{N}$ and $\beta = q$. Setting $\phi_z(x) := x + \sigma(x, z)$, by (1.5), we have

$$\left\|\nabla^m_x \phi_z\right\|_\infty \leq C\big(1 + |z|\big) \qquad \forall m \in \mathbb{N}. \tag{5.16}$$

If $q \in (0, 1)$, then

$$\begin{aligned}
[f \circ \phi_z]_q &= \sup_{x\neq y} \frac{|f \circ \phi_z(x) - f \circ \phi_z(y)|}{|x - y|^q} \\
&\leq [f]_q \sup_{x\neq y} \frac{|\phi_z(x) - \phi_z(y)|^q}{|x - y|^q} \\
&\leq [f]_q \|\nabla\phi_z\|^q_\infty \overset{(5.16)}{\leq} C[f]_q\big(1 + |z|^q\big).
\end{aligned}$$

Hence,

$$[\mathcal{L}_1 f]_q \le C[f]_q \int_{|z|\ge 1} (1+|z|^q)\nu(\mathrm{d}z).$$

For $q=1$, it is easy to see that (5.15) is true by the chain rule. Now assume $q\in(1,2)$. By the chain rule, we have

$$\begin{aligned}
\big[\nabla(f\circ\phi_z)\big]_{q-1} &= \sup_{x\neq y}\frac{|(\nabla f)\circ\phi_z(x)\cdot\nabla\phi_z(x)-(\nabla f)\circ\phi_z(y)\cdot\nabla\phi_z(y)|}{|x-y|^{q-1}}\\
&\le \sup_{x\neq y}\frac{|(\nabla f)\circ\phi_z(x)-(\nabla f)\circ\phi_z(y)|\cdot\|\nabla\phi_z\|_\infty}{|x-y|^{q-1}}\\
&\quad+\sup_{x\neq y}\frac{\|\nabla f\|_\infty|\nabla\phi_z(x)-\nabla\phi_z(y)|}{|x-y|^{q-1}}\\
&\overset{(5.16)}{\le} C\|\nabla f\|_{q-1}\big(1+|z|\big)\sup_{x\neq y}\frac{|\phi_z(x)-\phi_z(y)|^{q-1}}{|x-y|^{q-1}}\\
&\quad+C\|\nabla f\|_\infty\big(1+|z|\big)^{2-q}\sup_{x\neq y}\frac{|\nabla\phi_z(x)-\nabla\phi_z(y)|^{q-1}}{|x-y|^{q-1}}\\
&\overset{(5.16)}{\le} C[\nabla f]_{q-1}\big(1+|z|\big)^q+C\|\nabla f\|_\infty\big(1+|z|\big).
\end{aligned}$$

Thus,

$$[\nabla\mathcal{L}_1 f]_{q-1}\le C\|\nabla f\|_{q-1,\infty}\int_{|z|\ge 1}(1+|z|^q)\nu(\mathrm{d}z).$$

For $q\ge 2$, it follows by similar calculations. □

Let $\mathcal{T}_t$ be the semigroup associated with $\mathcal{L}^\nu_{\sigma,b}$. For any $\varphi\in C^\infty_b(\mathbb{R}^d)$, as above by Duhamel's formula, we have

$$\mathcal{T}_t\varphi(x)=\mathcal{T}^0_t\varphi(x)+\int_0^t\mathcal{T}^0_{t-s}\mathcal{L}_1\mathcal{T}_s\varphi(x)\,\mathrm{d}s. \tag{5.17}$$

LEMMA 5.6. *Let γ_{100} be as in Theorem 4.5. If $\int_{|z|\ge 1}|z|^q\nu(\mathrm{d}z)<\infty$ for some $q>0$, then for any $\beta\in(0,\frac{1}{\gamma_{100}}\wedge 1)$, there exists a constant $C>0$ such that for all $t\in(0,1)$ and $\varphi\in L^\infty(\mathbb{R}^d)$,*

$$\|\mathcal{T}_t\varphi\|_{q+\beta,\infty}\le Ct^{-(q+\beta)\gamma_{100}}\|\varphi\|_\infty.$$

PROOF. Without loss of generality, we assume that $q+\beta$ is not an integer. Fix an irrational number $q_0\in(0,q]$ and choose $m\in\mathbb{N}$ being large so that

$$\theta:=\frac{q_0}{m}<\frac{1}{\gamma_{100}}\wedge 1.$$

By (5.17), (5.15) and Lemma 5.2, we have

$$\|\mathcal{T}_t\varphi\|_{m\theta,\infty} \le \|\mathcal{T}_t^0\varphi\|_{m\theta,\infty} + \int_0^t \|\mathcal{T}_{t-s}^0\mathcal{L}_1\mathcal{T}_s\varphi\|_{m\theta,\infty}\,\mathrm{d}s$$
$$\le Ct^{-\theta\gamma_{100}}\|\varphi\|_{(m-1)\theta,\infty} + C\int_0^t \|\mathcal{T}_s\varphi\|_{m\theta,\infty}\,\mathrm{d}s,$$

which, by Gronwall's inequality, yields that for all $t \in (0, 1)$,

$$\|\mathcal{T}_t\varphi\|_{m\theta,\infty} \le Ct^{-\theta\gamma_{100}}\|\varphi\|_{(m-1)\theta,\infty}.$$

Since $j\theta$ is not an integer for any $j \in \mathbb{N}$, by iteration we obtain

$$\|\mathcal{T}_t\varphi\|_{q_0,\infty} = \|\mathcal{T}_t\varphi\|_{m\theta,\infty} \le Ct^{-m\theta\gamma_{100}}\|\varphi\|_\infty = Ct^{-q_0\gamma_{100}}\|\varphi\|_\infty.$$

Next, we choose $\theta_0 \in (0, \frac{1}{\gamma_{100}} \wedge 1)$ and an irrational number $q_0 \le q$ so that $q_0+\theta_0 = q + \beta$. As above, we have

$$\|\mathcal{T}_t\varphi\|_{q_0+\theta_0,\infty} \le Ct^{-\theta_0\gamma_{100}}\|\varphi\|_{q_0,\infty} + C\int_0^t (t-s)^{-\theta_0\gamma_{100}}\|\mathcal{L}_1\mathcal{T}_s\varphi\|_{q_0,\infty}\,\mathrm{d}s$$
$$\le Ct^{-\theta_0\gamma_{100}}\|\varphi\|_{q_0,\infty} + C\|\varphi\|_{q_0,\infty}\int_0^t (t-s)^{-\theta_0\gamma_{100}}\,\mathrm{d}s$$
$$\le C\big(t^{-\theta_0\gamma_{100}} + t^{1-\theta_0\gamma_{100}}\big)\|\varphi\|_{q_0,\infty}.$$

Thus,

$$\|\mathcal{T}_{2t}\varphi\|_{q_0+\theta_0,\infty} \le Ct^{-\theta_0\gamma_{100}}\|\mathcal{T}_t\varphi\|_{q_0,\infty} \le Ct^{-(q_0+\theta_0)\gamma_{100}}\|\varphi\|_\infty.$$

The proof is complete. □

LEMMA 5.7. *Let γ_{010} be as in Theorem 4.5. For any $\theta \in (0, \frac{1}{\gamma_{010}} \wedge 1)$, there exists a constant $C > 0$ such that for all $t \in (0, 1)$ and $\varphi \in C_b^\infty(\mathbb{R}^d)$,*

$$\|\mathcal{T}_t\Delta^{\frac{\theta}{2}}\varphi\|_\infty \le Ct^{-\theta\gamma_{010}}\|\varphi\|_\infty.$$

PROOF. By (5.17) and (5.6), we have

$$\|\mathcal{T}_t\Delta^{\frac{\theta}{2}}\varphi\|_\infty \le \|\mathcal{T}_t^0\Delta^{\frac{\theta}{2}}\varphi\|_\infty + \int_0^t \|\mathcal{T}_{t-s}^0\mathcal{L}_1\mathcal{T}_s\Delta^{\frac{\theta}{2}}\varphi\|_\infty\,\mathrm{d}s$$
$$\le Ct^{-\theta\gamma_{010}}\|\varphi\|_\infty + C\int_0^t \|\mathcal{T}_s\Delta^{\frac{\theta}{2}}\varphi\|_\infty\,\mathrm{d}s,$$

which in turn gives the desired estimate by Gronwall's inequality. □

PROOF OF THEOREM 1.4. Let $X_t(x)$ solve SDE (3.2). By Lemmas 5.7 and A.4 in the Appendix, there exists a function $\rho_t(x, y) \in (L^1 \cap L^p)(\mathbb{R}^d)$ for some $p > 1$ such that for all $\varphi \in C_0(\mathbb{R}^d)$,

$$\mathcal{T}_t\varphi(x) = \mathbb{E}\varphi\big(X_t(x)\big) = \int_{\mathbb{R}^d} \varphi(y)\rho_t(x, y)\,\mathrm{d}y.$$

By a further approximation, the above equality also holds for any $\varphi \in L^\infty(\mathbb{R}^d)$. The $q+\varepsilon$-order Hölder continuity of $x \mapsto \mathcal{T}_t\varphi(x)$ follows by Lemma 5.6. $\square$

APPENDIX

A.1. Proof of Lemma 4.4. Let δ be as in (4.1). For $0<\varepsilon<\delta$, let $X_t^\varepsilon(x) = X_t^\varepsilon$ solve the following SDE:

$$X_t^\varepsilon = x + \int_0^t b(X_s^\varepsilon)\,\mathrm{d}s + \int_0^t \int_{\Gamma_\varepsilon^\delta} \sigma(X_{s-}^\varepsilon, z)\tilde{N}(\mathrm{d}s, \mathrm{d}z), \tag{A.1}$$

where $\Gamma_\varepsilon^\delta := \{z \in \mathbb{R}^d : \varepsilon \le |z| < \delta\}$. We first prove the following limit theorem.

LEMMA A.1. *Under* (H_b^σ), *there exist a subsequence* $\varepsilon_k \to 0$ *and a null set* Ω_0 *such that for all* $\omega \notin \Omega_0$,

$$\lim_{k\to\infty} \sup_{|x|\le R} \sup_{t\in[0,1]} \big|X_t^{\varepsilon_k}(x,\omega) - X_t(x,\omega)\big| = 0 \qquad \forall R \in \mathbb{N}.$$

PROOF. Set $Z_t^\varepsilon := X_t^\varepsilon - X_t$. By Burkholder's inequality (2.9) and (H_b^σ), we have for any $p \ge 2$,

$$\begin{aligned}
\mathbb{E}\Big(\sup_{s\in[0,t]} |Z_s^\varepsilon|^p\Big) \le{}& C\mathbb{E}\Big(\int_0^t \big|b(X_s^\varepsilon) - b(X_s)\big|\,\mathrm{d}s\Big)^p \\
&+ C\mathbb{E}\Big(\sup_{t'\in[0,t]} \Big|\int_0^{t'} \int_{\Gamma_0^\varepsilon} \sigma(X_{s-}, z)\tilde{N}(\mathrm{d}s, \mathrm{d}z)\Big|^p\Big) \\
&+ C\mathbb{E}\Big(\sup_{t'\in[0,t]} \Big|\int_0^{t'} \int_{\Gamma_\varepsilon^\delta} \big(\sigma(X_{s-}^\varepsilon, z) - \sigma(X_{s-}, z)\big)\tilde{N}(\mathrm{d}s, \mathrm{d}z)\Big|^p\Big) \\
\le{}& C\int_0^t \mathbb{E}|Z_s^\varepsilon|^p\,\mathrm{d}s + C\int_{\Gamma_0^\varepsilon} |z|^p \nu(\mathrm{d}z) + C\Big(\int_{\Gamma_0^\varepsilon} |z|^2 \nu(\mathrm{d}z)\Big)^{\frac{p}{2}},
\end{aligned}$$

where C is independent of $\varepsilon, t \in (0,1)$ and $x \in \mathbb{R}^d$. By Gronwall's inequality, we obtain

$$\lim_{\varepsilon\to 0} \sup_{x\in\mathbb{R}^d} \mathbb{E}\Big(\sup_{t\in[0,1]} \big|X_t^\varepsilon(x) - X_t(x)\big|^p\Big) = 0.$$

Similarly, we can prove that for any $p \ge 2$,

$$\lim_{\varepsilon\to 0} \sup_{x\in\mathbb{R}^d} \mathbb{E}\Big(\sup_{t\in[0,1]} \big|\nabla X_t^\varepsilon(x) - \nabla X_t(x)\big|^p\Big) = 0.$$

Thus, for any $R > 0$, by Sobolev's embedding theorem, we have

$$\lim_{\varepsilon\to 0}\mathbb{E}\Big(\sup_{|x|<R}\sup_{t\in[0,1]}\big|X_t^\varepsilon(x)-X_t(x)\big|^p\Big)$$
$$\le C\lim_{\varepsilon\to 0}\mathbb{E}\Big(\sup_{t\in[0,1]}\big\|X_t^\varepsilon(\cdot)-X_t(\cdot)\big\|^p_{\mathbb{W}^{1,p}(B_R)}\Big)=0,$$

where $p > d$ and $\mathbb{W}^{1,p}(B_R)$ is the first order Sobolev space over $B_R := \{x \in \mathbb{R}^d : |x| < R\}$. The desired limit follows by a suitable choice of subsequence ε_k. □

Define $\phi(x,z) := x + \sigma(x,z)$. By (4.1), the mapping $x \mapsto \phi(x,z)$ is invertible for each $|z| \le \delta$. Let $\phi^{-1}(x,z)$ be the inverse of $x \mapsto \phi(x,z)$. Write

$$\hat\sigma(x,z) := \sigma\big(\phi^{-1}(x,z),z\big),\qquad |z|\le\delta \tag{A.2}$$

and

$$\hat b(x) := b(x) + \int_{\Gamma_0^\delta}\big[\sigma\big(\phi^{-1}(x,z),z\big)-\sigma(x,z)\big]\nu(\mathrm{d}z), \tag{A.3}$$

$$\hat b_\varepsilon(x) := b(x) + \int_{\Gamma_\varepsilon^\delta}\big[\sigma\big(\phi^{-1}(x,z),z\big)-\sigma(x,z)\big]\nu(\mathrm{d}z). \tag{A.4}$$

By the chain rule, the following lemma is easy.

LEMMA A.2. *Under* (H_b^σ), *there exists a constant* $C > 0$ *such that for all* $x \in \mathbb{R}^d$ *and* $|z| \le \delta$,

$$\big|\hat\sigma(x,z)\big|\le C|z|,\qquad \big|\nabla_x\hat\sigma(x,z)\big|\le C|z|.$$

Moreover, $\hat b, \hat b_\varepsilon \in C_b^1(\mathbb{R}^d)$ *and for some* $C > 0$,

$$\|\hat b_\varepsilon-\hat b\|_\infty+\|\nabla\hat b_\varepsilon-\nabla\hat b\|_\infty\le C\int_{\Gamma_0^\varepsilon}|z|^2\nu(\mathrm{d}z).$$

Fix $T \in [0,1]$. For $t \in [0,T]$, define

$$\hat L_t^T := L_{T-}-L_{T-t+}\qquad\text{with } L_{T-t+} := \lim_{s\downarrow t}L_{T-s}.$$

In particular, $(\hat L_t^T)_{t\in[0,T]}$ is still a Lévy process with the same Lévy measure ν and

$$\Delta\hat L_t^T=\Delta L_{T-t}. \tag{A.5}$$

Let $\hat N^T(\mathrm{d}s,\mathrm{d}z)$ be the Poisson random measure associated with $\hat L_t^T$, that is,

$$\hat N^T\big((0,t]\times E\big) := \sum_{0<s\le t}1_E\big(\Delta\hat L_s^T\big),\qquad E\in\mathscr{B}(\mathbb{R}^d),$$

and $\tilde{\hat N}^T(\mathrm{d}s,\mathrm{d}z) := \hat N^T(\mathrm{d}s,\mathrm{d}z) - \mathrm{d}s\,\nu(\mathrm{d}z)$ the compensated Poisson random measure. We have

LEMMA A.3. *Let $\hat{X}_t^T(x) = \hat{X}_t^T$ solve the following SDE*:

$$\hat{X}_t^T = x - \int_0^t \hat{b}(\hat{X}_s^T)\,\mathrm{d}s - \int_0^t \int_{\Gamma_0^\delta} \hat{\sigma}(\hat{X}_{s-}^T, z)\tilde{\hat{N}}^T(\mathrm{d}s, \mathrm{d}z), \tag{A.6}$$

where $\hat{\sigma}$ and $\hat{b}$ are defined by (A.2) *and* (A.3), *respectively. Then*

$$\hat{X}_T^T(x) = X_T^{-1}(x) \qquad \forall x \in \mathbb{R}^d, \ a.s. \tag{A.7}$$

PROOF. For $\varepsilon \in (0, \delta)$, let $X_t^\varepsilon(x) = X_t^\varepsilon$ solve the following random ODE:

$$\begin{aligned} X_t^\varepsilon &= x + \int_0^t b(X_s^\varepsilon)\,\mathrm{d}s + \int_0^t \int_{\Gamma_\varepsilon^\delta} \sigma(X_{s-}^\varepsilon, z)\tilde{N}(\mathrm{d}s, \mathrm{d}z) \\ &= x + \int_0^t \tilde{b}_\varepsilon(X_s^\varepsilon)\,\mathrm{d}s + \sum_{0<s\le t} \sigma(X_{s-}^\varepsilon, \Delta L_s) 1_{\Gamma_\varepsilon^\delta}(\Delta L_s), \end{aligned}$$

where

$$\tilde{b}_\varepsilon(x) = b(x) - \int_{\Gamma_\varepsilon^\delta} \sigma(x, z)\nu(\mathrm{d}z).$$

By the change of variables, we have

$$\begin{aligned} X_{T-t}^\varepsilon &= X_T^\varepsilon - \int_{T-t}^T \tilde{b}_\varepsilon(X_s^\varepsilon)\,\mathrm{d}s - \sum_{T-t<s\le T} \sigma(X_{s-}^\varepsilon, \Delta L_s) 1_{\Gamma_\varepsilon^\delta}(\Delta L_s) \\ &= X_T^\varepsilon - \int_0^t \tilde{b}_\varepsilon(X_{T-s}^\varepsilon)\,\mathrm{d}s - \sum_{0\le s<t} \sigma(X_{(T-s)-}^\varepsilon, \Delta L_{T-s}) 1_{\Gamma_\varepsilon^\delta}(\Delta L_{T-s}). \end{aligned}$$

Noticing that if $\Delta L_t \in \Gamma_\varepsilon^\delta$, then

$$X_t^\varepsilon - X_{t-}^\varepsilon = \sigma(X_{t-}^\varepsilon, \Delta L_t) \Rightarrow X_{t-} = \phi^{-1}(X_t^\varepsilon, \Delta L_t),$$

and since $\Delta L_T = 0$ almost surely, we further have

$$\begin{aligned} X_{T-t}^\varepsilon &= X_T^\varepsilon - \int_0^t \tilde{b}_\varepsilon(X_{T-s}^\varepsilon)\,\mathrm{d}s - \sum_{0<s<t} \hat{\sigma}(X_{T-s}^\varepsilon, \Delta L_{T-s}) 1_{\Gamma_\varepsilon^\delta}(\Delta L_{T-s}) \\ &\overset{(A.5)}{=} X_T^\varepsilon - \int_0^t \tilde{b}_\varepsilon(X_{T-s}^\varepsilon)\,\mathrm{d}s - \sum_{0<s<t} \hat{\sigma}(X_{T-s}^\varepsilon, \Delta \hat{L}_s^T) 1_{\Gamma_\varepsilon^\delta}(\Delta \hat{L}_s^T), \end{aligned}$$

where $\hat{\sigma}(x, z)$ is defined by (A.2). In particular,

$$\begin{aligned} X_{T-t+}^\varepsilon &= X_T^\varepsilon - \int_0^t \tilde{b}_\varepsilon(X_{T-s}^\varepsilon)\,\mathrm{d}s - \int_0^t \int_{\Gamma_\varepsilon^\delta} \hat{\sigma}(X_{T-s}^\varepsilon, z)\hat{N}(\mathrm{d}s, \mathrm{d}z) \\ &= X_T^\varepsilon - \int_0^t \hat{b}_\varepsilon(X_{T-s}^\varepsilon)\,\mathrm{d}s - \int_0^t \int_{\Gamma_\varepsilon^\delta} \hat{\sigma}(X_{T-s}^\varepsilon, z)\tilde{\hat{N}}(\mathrm{d}s, \mathrm{d}z), \end{aligned}$$

where $\hat{b}_\varepsilon(x)$ is defined by (A.4). On the other hand, let $\hat{X}_t^{T,\varepsilon}(x) = \hat{X}_t^{T,\varepsilon}$ solve the following SDE:

$$\hat{X}_t^{T,\varepsilon} = x - \int_0^t \hat{b}_\varepsilon(\hat{X}_s^{T,\varepsilon})\,\mathrm{d}s - \int_0^t \int_{\Gamma_\varepsilon^\delta} \hat{\sigma}(\hat{X}_{s-}^{T,\varepsilon}, z)\tilde{\hat{N}}^T(\mathrm{d}s, \mathrm{d}z).$$

By the uniqueness of solutions to random ODEs, we have

$$X_{T-t+}^\varepsilon(x) = \hat{X}_t^{T,\varepsilon}(X_T^\varepsilon(x)) \qquad \forall x \in \mathbb{R}^d, \text{ a.s.}$$

In particular,

$$x = \hat{X}_T^{T,\varepsilon}(X_T^\varepsilon(x)) \qquad \forall x \in \mathbb{R}^d, \text{ a.s.} \tag{A.8}$$

By Lemmas A.2, A.1 and taking limits for (A.8), we obtain

$$x = \hat{X}_T^T(X_T(x)) \qquad \forall x \in \mathbb{R}^d, \text{ a.s.}$$

The proof is complete. □

Now we can give

PROOF OF LEMMA 4.4. By equation (A.6) and a standard calculation, we have for any $p \geq 2$,

$$\sup_{T\in[0,1]} \sup_{x\in\mathbb{R}^d} \mathbb{E}|\nabla \hat{X}_T^T(x)|^p < \infty,$$

which, together with (A.7), implies that

$$\sup_{T\in[0,1]} \sup_{x\in\mathbb{R}^d} \mathbb{E}(\det(\nabla X_T^{-1}(x))) < \infty.$$

The desired estimate (4.15) then follows by the change of variables and the above estimate. □

A.2. A criterion for the existence of density.

LEMMA A.4. *Let $\mathcal{T}$ be a bounded linear operator in $C_b(\mathbb{R}^d)$. Assume that for some $\theta \in (0, 1)$ and any $\varphi \in C_b^\infty(\mathbb{R}^d)$,*

$$\|\mathcal{T}\Delta^{\frac{\theta}{2}}\varphi\|_\infty \leq C_\theta \|\varphi\|_\infty. \tag{A.9}$$

Then there exists a measurable function $\rho(x, y)$ with $\rho(x, \cdot) \in (L^1 \cap L^p)(\mathbb{R}^d)$ for some $p > 1$ and such that for any $\varphi \in C_0(\mathbb{R}^d)$,

$$\mathcal{T}\varphi(x) = \int_{\mathbb{R}^d} \varphi(y)\rho(x, y)\,\mathrm{d}y. \tag{A.10}$$

PROOF. By Riesz's representation theorem, there exists a family of finite signed measures $\mu_x(\mathrm{d}y)$ such that $x \mapsto \mu_x(\mathrm{d}y)$ is weakly continuous and for any $\varphi \in C_0(\mathbb{R}^d)$,

$$\mathcal{T}\varphi(x) = \int_{\mathbb{R}^d} \varphi(y)\mu_x(\mathrm{d}y). \tag{A.11}$$

Let ϱ be a nonnegative symmetric smooth function with compact support and $\int_{\mathbb{R}^d} \varrho(y)\,\mathrm{d}y = 1$. Let $\varrho_\varepsilon(y) := \varepsilon^{-d}\varrho(\varepsilon^{-1}y)$ be a family of mollifies. For $R > 0$, let $\chi_R : \mathbb{R}^d \to [0, 1]$ be a smooth cutoff function with

$$\chi_R(x) = 1, \qquad |x| \le R, \qquad \chi_R(x) = 0, \qquad |x| \ge 2R.$$

For $\varphi \in L^\infty(\mathbb{R}^d)$, set

$$\varphi_\delta(x) := \varphi * \varrho_\delta(x), \qquad \varphi^R_{\delta,\varepsilon}(x) := (\varphi_\delta \chi_R) * \varrho_\varepsilon(x)$$

and

$$\mu^\varepsilon_x(z) := \int_{\mathbb{R}^d} \varrho_\varepsilon(y - z)\mu_x(\mathrm{d}y).$$

It is easy to see that $\mu^\varepsilon_x \in \bigcap_k \mathbb{W}^{k,1}(\mathbb{R}^d)$ and $\Delta^{\frac{\theta}{2}}\varphi^R_{\delta,\varepsilon} \in C_0(\mathbb{R}^d)$. Thus, by (A.11) we have

$$\begin{aligned}
\mathcal{T}\Delta^{\frac{\theta}{2}}\varphi^R_{\delta,\varepsilon}(x) &= \int_{\mathbb{R}^d} \big(\Delta^{\frac{\theta}{2}}(\varphi_\delta \chi_R)\big) * \varrho_\varepsilon(y)\mu_x(\mathrm{d}y) \\
&= \int_{\mathbb{R}^d} \Delta^{\frac{\theta}{2}}(\varphi_\delta \chi_R)(z)\mu^\varepsilon_x(z)\,\mathrm{d}z = \int_{\mathbb{R}^d} \varphi_\delta \chi_R(z)\Delta^{\frac{\theta}{2}}\mu^\varepsilon_x(z)\,\mathrm{d}z,
\end{aligned}$$

which yields by (A.9) that

$$\left|\int_{\mathbb{R}^d} \varphi_\delta \chi_R(z)\Delta^{\frac{\theta}{2}}\mu^\varepsilon_x(z)\,\mathrm{d}z\right| \le C_\theta \|\varphi^R_{\delta,\varepsilon}\| \le C_\theta \|\varphi\|_\infty.$$

Letting $R \to \infty$ and $\delta \to 0$, by the dominated convergence theorem, we obtain that for all $\varphi \in L^\infty(\mathbb{R}^d)$,

$$\left|\int_{\mathbb{R}^d} \varphi(z)\Delta^{\frac{\theta}{2}}\mu^\varepsilon_x(z)\,\mathrm{d}z\right| \le C_\theta \|\varphi\|_\infty,$$

which gives

$$\sup_{x\in\mathbb{R}^d} \sup_{\varepsilon\in(0,1)} \|\Delta^{\frac{\theta}{2}}\mu^\varepsilon_x\|_1 \le C_\theta.$$

Moreover, we also have

$$\sup_{x\in\mathbb{R}^d} \sup_{\varepsilon\in(0,1)} \|\mu^\varepsilon_x\|_1 \le \sup_{\|\varphi\|_\infty \le 1} \|\mathcal{T}\varphi\|_\infty.$$

By Sobolev's embedding theorem, there is a $p > 1$ such that

$$\sup_{x\in\mathbb{R}^d} \sup_{\varepsilon\in(0,1)} \|\mu^\varepsilon_x\|_p < \infty.$$

Since $L^p(\mathbb{R}^d)$ is weakly compact, for each fixed $x \in \mathbb{R}^d$, there is a subsequence $\varepsilon_k \to 0$ and $\rho(x, \cdot) \in (L^1 \cap L^p)(\mathbb{R}^d)$ such that for any $\varphi \in C_0(\mathbb{R}^d) \subset L^{\frac{p}{p-1}}(\mathbb{R}^d)$,

$$\mathcal{T}\varphi_{\varepsilon_k}(x) = \int_{\mathbb{R}^d} \mu_x^{\varepsilon_k}(z)\varphi(z)\,\mathrm{d}z \overset{k\to\infty}{\to} \int_{\mathbb{R}^d} \rho(x,z)\varphi(z)\,\mathrm{d}z. \tag{A.12}$$

On the other hand, for any $\varphi \in C_0(\mathbb{R}^d)$,

$$\|\mathcal{T}\varphi_\varepsilon - \mathcal{T}\varphi\|_\infty \le C\|\varphi_\varepsilon - \varphi\|_\infty \overset{\varepsilon\to 0}{\to} 0,$$

which together with (A.12) yields (A.10). □

Acknowledgements. The author is grateful to Linlin Wang and Longjie Xie for their useful conversations.

REFERENCES

[1] ALEXANDRE, R. (2012). Fractional order kinetic equations and hypoellipticity. *Anal. Appl.* (*Singap.*) **10** 237–247. MR2948893

[2] BALLY, V. and CLÉMENT, E. (2011). Integration by parts formula and applications to equations with jumps. *Probab. Theory Related Fields* **151** 613–657. MR2851695

[3] BICHTELER, K., GRAVEREAUX, J.-B. and JACOD, J. (1987). *Malliavin Calculus for Processes with Jumps*. *Stochastics Monographs* **2**. Gordon and Breach Science Publishers, New York. MR1008471

[4] BISMUT, J.-M. (1983). Calcul des variations stochastique et processus de sauts. *Z. Wahrsch. Verw. Gebiete* **63** 147–235. MR0701527

[5] CASS, T. (2009). Smooth densities for solutions to stochastic differential equations with jumps. *Stochastic Process. Appl.* **119** 1416–1435. MR2513114

[6] CHEN, H., LI, W.-X. and XU, C.-J. (2011). Gevrey hypoellipticity for a class of kinetic equations. *Comm. Partial Differential Equations* **36** 693–728. MR2763329

[7] CHEN, Z.-Q. (2016). Heat kernels and analyticity of non-symmetric jump diffusion semigroups. *Probab. Theory Related Fields* **165** 267–312.

[8] CHEN, Z.-Q. (2009). Symmetric jump processes and their heat kernel estimates. *Sci. China Ser. A* **52** 1423–1445. MR2520585

[9] FUJIWARA, T. and KUNITA, H. (1985). Stochastic differential equations of jump type and Lévy processes in diffeomorphisms group. *J. Math. Kyoto Univ.* **25** 71–106. MR0777247

[10] ISHIKAWA, Y. and KUNITA, H. (2006). Malliavin calculus on the Wiener–Poisson space and its application to canonical SDE with jumps. *Stochastic Process. Appl.* **116** 1743–1769. MR2307057

[11] KOMATSU, T. and TAKEUCHI, A. (2001). On the smoothness of PDF of solutions to SDE of jump type. *Int. J. Differ. Equ. Appl.* **2** 141–197. MR1930241

[12] KUNITA, H. (2009). Smooth density of canonical stochastic differential equation with jumps. *Astérisque* **327** 69–91. MR2642353

[13] KUNITA, H. (2011). Analysis of nondegenerate Wiener–Poisson functionals and its applications to Itô's SDE with jumps. *Sankhyā A* **73** 1–45. MR2887083

[14] KUNITA, H. (2013). Nondegenerate SDE's with jumps and their hypoelliptic properties. *J. Math. Soc. Japan* **65** 993–1035. MR3084987

[15] KUSUOKA, S. (2010). Malliavin calculus for stochastic differential equations driven by subordinated Brownian motions. *Kyoto J. Math.* **50** 491–520. MR2723861

[16] NUALART, D. (2006). *The Malliavin Calculus and Related Topics*, 2nd ed. Springer, Berlin. MR2200233

[17] NUALART, D. (2011). Discussion of Hiroshi Kunita's article: Analysis of nondegenerate Wiener–Poisson functionals and its applications to Itô's SDE with jumps [MR2887083]. *Sankhyā A* **73** 46–49. MR2887084

[18] PICARD, J. (1996). On the existence of smooth densities for jump processes. *Probab. Theory Related Fields* **105** 481–511. MR1402654

[19] PROTTER, P. E. (2004). *Stochastic Integration and Differential Equations*, 2nd ed. *Applications of Mathematics* (*New York*) **21**. Springer, Berlin. MR2020294

[20] SONG, Y. and ZHANG, X. (2015). Regularity of density for SDEs driven by degenerate Lévy noises. *Electron*. *J*. *Probab*. **20** 1–27. MR3325091

[21] STEIN, E. M. (1970). *Singular Integrals and Differentiability Properties of Functions*. *Princeton Mathematical Series* **30**. Princeton Univ. Press, Princeton, NJ. MR0290095

[22] TAKEUCHI, A. (2002). The Malliavin calculus for SDE with jumps and the partially hypoelliptic problem. *Osaka J*. *Math*. **39** 523–559. MR1932281

[23] TRIEBEL, H. (1978). *Interpolation Theory*, *Function Spaces*, *Differential Operators*. *North-Holland Mathematical Library* **18**. North-Holland, Amsterdam. MR0503903

[24] VARADHAN, S. R. S. (2011). Discussion of Hiroshi Kunita's article: Analysis of nondegenerate Wiener–Poisson functionals and its applications to Itô's SDE with jumps [MR2887083]. *Sankhyā A* **73** 50–51. MR2887085

[25] ZHANG, X. (2014). Densities for SDEs driven by degenerate α-stable processes. *Ann*. *Probab*. **42** 1885–1910. MR3262494

[26] ZHANG, X. (2014). Fundamental solutions of nonlocal Hörmander's operators. Available at arXiv:1306.5016.

SCHOOL OF MATHEMATICS AND STATISTICS
WUHAN UNIVERSITY
WUHAN, HUBEI 430072
P.R. CHINA
E-MAIL: XichengZhang@gmail.com