HÖRMANDER'S HYPOELLIPTIC THEOREM FOR NONLOCAL OPERATORS

ZIMO HAO, XUHUI PENG* AND XICHENG ZHANG

Abstract. In this paper we show the Hörmander hypoelliptic theorem for nonlocal operators by a purely probabilistic method: the Malliavin calculus. Roughly speaking, under general Hörmander's Lie bracket conditions, we show the regularization effect of discontinuous Lévy noises for possibly degenerate stochastic differential equations with jumps. To treat the large jumps, we use the perturbation argument together with interpolation techniques and some short time asymptotic estimates of the semigroup. As an application, we show the existence of fundamental solutions for operator $\partial_t - \mathscr{K}$, where $\mathscr{K}$ is the nonlocal kinetic operator:

$$\mathscr{K} f(x, \mathrm{v}) = \mathrm{p.v}\int_{\mathbb{R}^d}(f(x, \mathrm{v}+w) - f(x, \mathrm{v}))\frac{\kappa(x, \mathrm{v}, w)}{|w|^{d+\alpha}}\mathrm{d}w + \mathrm{v}\cdot\nabla_x f(x, \mathrm{v}) + b(x, \mathrm{v})\cdot\nabla_{\mathrm{v}} f(x, \mathrm{v}).$$

Here $\kappa_0^{-1} \leqslant \kappa(x, \mathrm{v}, w) \leqslant \kappa_0$ belongs to $C_b^\infty(\mathbb{R}^{3d})$ and is symmetric in w, p.v. stands for the Cauchy principal value, and $b \in C_b^\infty(\mathbb{R}^{2d}; \mathbb{R}^d)$.

Keywords: Hörmander's conditions, Malliavin calculus, Hypoellipticity, Nonlocal operators

AMS 2010 Mathematics Subject Classification: 60H07, 60H10, 60H30.

1. Introduction

1.1 Introduction. Let $\mathscr{A}$ be a differential operator in $\mathbb{R}^d$ with smooth coefficients. Hypoellipticity in the theory of PDEs means that for any distribution u and open subset $U \subset \mathbb{R}^d$,

$$\mathscr{A} u|_U \in C^\infty(U) \Rightarrow u \in C^\infty(U).$$

Let $A_0, A_1, \cdots A_d$ be $d+1$-differential operators of first order (or vector fields) with smooth coefficients and c a smooth function. The classical Hörmander's hypoelliptic theorem tells us that if

$$\mathscr{V} := \mathrm{Lie}(A_1, \cdots, A_d, [A_0, A_1], \cdots, [A_0, A_d]) = \mathbb{R}^d,$$

where $[A_0, A_k] := A_0A_k - A_kA_0$ is the Lie bracket, and $\mathscr{V}$ stands for the Lie algebra generated by vector fields $A_k, [A_0, A_k], k = 1, \cdots, d$, then $\mathscr{A} := \sum_{k=1}^d A_k^2 + A_0 + c - \frac{\partial}{\partial t}$ is hypoelliptic in $\mathbb{R}^{d+1}$ (cf. [8], [24], [7] and [13], etc.).

Consider the following Itô's type SDE

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma_k(X_t)\mathrm{d}W_t^k,\ X_0 = x \in \mathbb{R}^d, \tag{1.1}$$

where W is a d-dimensional standard Brownian motion and $b, \sigma_k : \mathbb{R}^d \to \mathbb{R}^d, k = 1, \cdots, d$ are C_b^∞-functions. Here and below we use Einstein's summation convention: If an index appears

*Corresponding Author.

X. Peng is partially supported by NSFC (No.11501195) and a Scientific Research Fund of Hunan Provincial Education Department (No. 17C0953). X.Zhang is partially supported by NNSFC grant of China (No. 11731009) and the DFG through the CRC 1283 "Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications".

twice in a product, then it will be summed automatically. We now define $d+1$-vector fields by

$$A_0 := \left(b^i - \tfrac{1}{2}\sigma^j_k\partial_j\sigma^i_k\right)\partial_i,\ \ A_k := \sigma^i_k\partial_i,\ k=1,\cdots,d, \tag{1.2}$$

where $\partial_i := \partial_{x_i} = \frac{\partial}{\partial x_i}$. Let $\mu_t(x,\mathrm{d}y)$ be the distributional density of the unique solution $X_t(x)$ of SDE (1.1). By Itô's formula, one sees that in the distributional sense,

$$\partial_t\mu_t(x,\cdot) = \left(\tfrac{1}{2}A_k^2 + A_0\right)^*\mu_t(x,\cdot),$$

where the asterisk stands for the adjoint operator. Notice that

$$\left(\tfrac{1}{2}A_k^2 + A_0\right)^* = \tfrac{1}{2}A_k^2 + \widehat{A}_0 + c,$$

where $\widehat{A_0} := (\sigma^i_k\partial_j\sigma^j_k)\partial_i - A_0$ and $c := \partial_i\partial_j(\sigma^i_k\sigma^j_k)/2 - \mathrm{div}b$. By Hörmander's hypoelliptic theorem, if

$$\mathscr{V} = \mathrm{Lie}(A_1,\cdots,A_d,[\widehat{A_0},A_1],\cdots,[\widehat{A_0},A_d]) = \mathbb{R}^d,$$

then $\mu_t(x,\cdot)$ admits a smooth density (see [24]). In [14], Malliavin provides a purely probabilistic proof for the above result by infinitely dimensional stochastic calculus of variations invented by him, which is now called the Malliavin calculus (see [18]). Since then, the Malliavin calculus has been developed very well, and emerged in many fields such as financial, control, filtering, and so on. Notice that in [7], Hairer presents a short and self-contained proof for Hörmander's theorem based on Malliavin's idea.

In this paper we are concerned with the following SDE with jumps:

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma_k(X_t)\mathrm{d}W^k_t + \int_{\mathbb{R}^d_0} g(X_{t-},z)\widetilde{N}(\mathrm{d}t,\mathrm{d}z),\ X_0 = x\in\mathbb{R}^d, \tag{1.3}$$

where $\mathbb{R}^d_0 := \mathbb{R}^d\setminus\{0\}$, and $N(\mathrm{d}t,\mathrm{d}z)$ is a Poisson random measure with intensity $\mathrm{d}t\nu(\mathrm{d}z)$,

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

and ν is a symmetric Lévy measure over $\mathbb{R}^d_0$, and $b,\sigma_k:\mathbb{R}^d\to\mathbb{R}^d, k=1,\cdots,d$ and $g:\mathbb{R}^d\times\mathbb{R}^d_0\to\mathbb{R}^d$ are smooth Lipschitz functions. It is well known that SDE (1.3) admits a unique strong solution $X_t(x)$ for each initial value $x\in\mathbb{R}^d$ (for example, see [20]). Suppose $g(x,-z) = -g(x,z)$. By Itô's formula, one sees that the generator of SDE (1.3) is given by

$$\mathscr{A}\varphi(x) := \frac{1}{2}A_k^2\varphi(x) + A_0\varphi(x) + \mathrm{p.v.}\int_{\mathbb{R}^d_0}\Big(\varphi(x+g(x,z)) - \varphi(x)\Big)\nu(\mathrm{d}z), \tag{1.4}$$

where p.v. stands for the Cauchy principal value, and A_0, A_k are defined by (1.2). More precisely, for $\varphi\in C^\infty_b(\mathbb{R}^d)$, if we define

$$\mathcal{T}_t\varphi(x) := \mathbb{E}\varphi(X_t(x)), \tag{1.5}$$

then

$$\partial_t\mathcal{T}_t\varphi = \mathscr{A}\mathcal{T}_t\varphi = \mathcal{T}_t\mathscr{A}\varphi. \tag{1.6}$$

The aim of this work is to show that under full Hörmander's conditions, the solution $X_t(x)$ of SDE (1.3) admits a smooth density.

The smoothness of the distribution density of the solutions to SDEs with jumps has been studied extensively since Malliavin's initiated work. We mention some of them. In [3], Bismut put forward a simple argument: Girsanov's transformation to study the smoothness of the distribution densities to SDEs with jumps. In [2], Bichteler, Gravereaux and Jacod give a systematic introducion for the Malliavin calculus with jumps. In [19], Picard used the difference operators to present another criterion for the smoothness of the distribution densities of Poisson functionals, see [9] for recent development for Wiener-Poisson functionals. Under partial Hörmander's conditions, there are also several works to study the smoothnees of degenerate SDEs with jumps. In [4], Cass established a Hörmander's type theorem for SDEs with jumps by proving a Norris' type lemma for discontinuous semimartingales. However, the Brownian noise can not disappear. In the pure jump degenerate case, Komatsu and Takeuchi [11, Theorem 3] introduced a quite useful estimate for discontinuous semimartingales, and then proved a Hörmander's type theorem for SDEs with jumps. Some subsequent results based on Komatsu-Takeuchi's estimate are referred to [21, 12]. Unfortunately, there is a gap in the proof of [11, Theorem 3]. We fill it up in [26] in a slightly different form (see Lemma 3.1 below). Basing on this new form of Komatsu-Takeuchi's type estimate, we prove a Hörmander's type theorem for pure jump SDEs with *nonzero drifts* in [27]. Other works about the regularization of jump noises can be found in [1, 17, 10] and references therein.

In Malliavin's probabilistic proof of Hörmander's hypoelliptic theorem, one of the key steps is to show the L^p-integrability of the inverse of the Malliavin covariance matrix. In the nondegenerate full noise case, it is relatively easy to obtain. However, under Hörmander's Lie bracket conditions, it is a quite challenge problem. In particular, Norris [16] provides an important estimate for general continuous semimartingales to treat this (see [18]). Now it is usually called Norris' lemma (see [7, Lemma 4.11] for an elegant proof), which can be considered as a quantitative version of Doob-Meyer's decomposition theorem. For general discontinuous semimartingales, Komatsu-Takeuchi's estimate should be regarded as a substitution of Norris' lemma. We shall use it to prove a full Hörmander's theorem for SDEs with jumps, see Theorem 1.1 below.

One of the motivations of studying nonlocal Hörmander's hypoelliptic theorem comes from the study of spatial inhomeogenous Boltzmann's equations. It is well known that the linearized spatial inhomeogenous Boltzmann's equation can be written as the following form that involves non-local operator of fractional Laplacian type (cf. [23] and [5]):

$$\partial_t f + \mathrm{v} \cdot \nabla_x f = \mathrm{p.v.} \int_{\mathbb{R}^d} (f(\cdot + w) - f(\cdot)) \frac{K_g(\cdot, w)}{|w|^{\alpha+d}} \mathrm{d}w + f H_g, \tag{1.7}$$

where f and g are functions of x, v and w, and

$$K_g(\mathrm{v}, w) := 2 \int_{\{h \cdot w = 0\}} g(\mathrm{v} - h)|h - w|^{\gamma+1+\alpha} \mathrm{d}h,$$

and

$$H_g(\mathrm{v}) := 2 \int_{\mathbb{R}^d} \int_{\{h \cdot w = 0\}} (g(\mathrm{v} - h) - g(\mathrm{v} - h + w)) \frac{|h - w|^{\gamma+1+\alpha}}{|w|^{\alpha+d}} \mathrm{d}h \mathrm{d}w.$$

Here $\gamma + \alpha \in (-1, 1)$. Note that K_g is a symmetric kernel in w, i.e., $K_g(\cdot, w) = K_g(\cdot, -w)$, and $f H_g$ is a zero order term in f. We shall see in Section 7 that the principal part of (1.7) can be written as the form of (1.4).

1.2 Main results. To make our statement of main results as simple and apparent as possible, throughout this paper we assume that for some $\alpha \in (0, 2)$,

$$\nu(\mathrm{d}z) = \mathrm{d}z/|z|^{d+\alpha}.$$

We also introduce the following assumptions about b, σ_k and g: for some $\ell \in \mathbb{N} \cup \{\infty\}$,

($\mathbf{H}_\ell$) For any $i \in \mathbb{N}$ and $j = 0, \cdots, \ell$, there are $C_i, C_{ij} \geqslant 1$ such that for all $x \in \mathbb{R}^d$ and $|z| < 1$,

$$|\nabla^i b(x)| + |\nabla^i \sigma_k(x)| \leqslant C_i, \quad |\nabla^i_x \nabla^j_z g(x, z)| \leqslant C_{ij} |z|^{1-j}.$$

Moreover, we require $g(x, -z) = -g(x, z)$ and for some $\beta \in (0, 1]$,

$$|\nabla_z g(x, z) - \nabla_z g(x, 0)| \leqslant C_x |z|^\beta, \quad |z| \leqslant 1,$$

where $C_x > 0$ continuously depends on $x \in \mathbb{R}^d$.

($\mathbf{H}^{\mathrm{o}}_g$) It holds that $\inf_{x,z\in\mathbb{R}^d} \det(\mathbb{I} + \nabla_x g(x, z)) > 0$ and $\mathrm{supp}\{g(x, \cdot)\} \subset B_1$.

Remark 1.1. *It should be kept in mind that $g(x, z) = \widetilde{\sigma}(x)z$ with $\widetilde{\sigma} : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ satisfying $\|\nabla^i \widetilde{\sigma}\|_\infty \leqslant C_i$ for $i \in \mathbb{N}$, fullfills the assumptions about g in* **($\mathbf{H}_\ell$)**. *Moreover, in order to make* **($\mathbf{H}^{\mathrm{o}}_g$)** *hold, one needs to assume $g(x, z) = \widetilde{\sigma}(x)z \cdot 1_{|z|\leqslant\delta}$ with δ being small enough so that SDE* (1.3) *defines a stochastic diffeomorphism flows in $\mathbb{R}^d$.*

Let A_0, A_k be as in (1.2) and $\widetilde{A}_k(x) := \partial_{z_k} g^i(x, 0)\partial_i$. Define

$$\mathscr{V}_0 := \{A_k, \widetilde{A}_k, k = 1, \cdots, d\},$$

and for $j = 1, 2, \cdots$,

$$\mathscr{V}_j := \Big\{[A_k, V], [\widetilde{A}_k, V], [A_0, V] : V \in \mathscr{V}_{j-1}, k = 1, \cdots, d\Big\}. \tag{1.8}$$

The following strong Hörmander's condition is imposed:

($\mathbf{H}^{\mathrm{str}}_{\mathbf{or}}$) For some $j_0 \in \mathbb{N}_0 := \{0\} \cup \mathbb{N}$, $\mathrm{span}\{\cup_{j=0}^{j_0} \mathscr{V}_j\} = \mathbb{R}^d$ at each point $x \in \mathbb{R}^d$.

We aim to prove the following result.

Theorem 1.1. *Under* **($\mathbf{H}_\infty$)**+**($\mathbf{H}^{\mathrm{o}}_g$)**+**($\mathbf{H}^{\mathrm{str}}_{\mathbf{or}}$)**, *there is a nonnegative smooth function $\rho_t(x, y)$ on $(0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d$ so that*

$$\mathbb{P} \circ X_t^{-1}(x)(\mathrm{d}y) = \rho_t(x, y)\mathrm{d}y,$$

where $X_t(x)$ is the solution of SDE (1.3) *with starting point $X_0(x) = x$, and*

$$\partial_t \rho_t(x, y) = \mathscr{A}\rho_t(\cdot, y)(x) = \mathscr{A}^*\rho_t(x, \cdot)(y), \quad \lim_{t\downarrow 0} \rho_t(x, y) = \delta_x(\mathrm{d}y), \tag{1.9}$$

where $\mathscr{A}^*$ is the adjoint operator of $\mathscr{A}$ (see (1.4)*), and δ_x is the Dirac measure concentrated at x.*

To treat the large jumps, we make the following stronger assumptions:

($\mathbf{H}'_\ell$) In addition to **($\mathbf{H}_\ell$)**, we assume that $\cup_{j=0}^\infty \mathscr{V}_j \subset C^\infty_b(\mathbb{R}^d)$ and

$$|\nabla^j_z g(x, z)| \leqslant C_j |z|^{1-j}, \ |z| < 1, \ j \in \mathbb{N}_0.$$

($\mathbf{H}^{\mathrm{uni}}_{\mathbf{or}}$) The following uniform Hörmander's condition holds: for some $j_0 \in \mathbb{N}_0$ and $c_0 > 0$,

$$\inf_{x\in\mathbb{R}^d} \inf_{|u|=1} \sum_{j=0}^{j_0} \sum_{V\in\mathscr{V}_j} |uV(x)|^2 \geqslant c_0. \tag{1.10}$$

Remark 1.2. *In* $(\mathbf{H}'_\ell)$, *the drift* b *may be linear growth, but* σ *and* g *are bounded in* x. *If* b *is also bounded, then for each* $V \in \cup_{j=0}^\infty \mathscr{V}_j$, *it automatically holds that* $V \in C_b^\infty(\mathbb{R}^d)$.

By a perturbation argument, we can prove the following result. Since its proof is completely the same as in [26, Theorem 1.2], we omit the details.

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* $(\mathbf{H}'_2)$+$(\mathbf{H}^{\rm o}_g)$+$(\mathbf{H}^{\rm uni}_{\mathbf{or}})$, *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* $\mathscr{A}+\mathscr{L}$ *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 there is a* $\gamma = \gamma(\alpha, j_0, d) > 0$ *such that for any* $p \in (1,\infty)$, $T > 0$ *and* $k \in \mathbb{N}_0$,

$$\|\nabla^k_x \rho_t(x,\cdot)\|_p \leqslant C t^{-(k+d)\gamma}, \quad \forall (t,x) \in (0,T]\times\mathbb{R}^d. \tag{1.11}$$

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

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

The above result provides a way of treating the large jumps. In applications, we usually take $\mathscr{L}$ as the large jump operator, for example,

$$\mathscr{L}\varphi(x) := \int_{|z|\geqslant\delta}\Big(\varphi(x+z)-\varphi(x)\Big)\kappa(x,z)\nu(\mathrm{d}z), \ \delta > 0.$$

In fact, we shall apply Theorem 1.2 to the nonlocal kinetic operators in Section 7. However, sometimes it is not easy to verify the boundedness of the large jump operator in $\mathbb{W}^{k,p}$. The following theorem provides part results for general SDE (1.3) without assuming $(\mathbf{H}^{\rm o}_g)$, which is still based on the perturbation argument and suitable interpolation techniques as in [26].

Theorem 1.3. *Under* $(\mathbf{H}'_2)$+$(\mathbf{H}^{\rm uni}_{\mathbf{or}})$ *and* $g \in C_b^\infty(\mathbb{R}^d\times B_1^c)$, *where* B_1 *is the unit ball, there is a nonnegative measurable function* $\rho_t(x,y)$ *on* $(0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d$ *so that*

(i) For each $t > 0$ *and* $x \in \mathbb{R}^d$, $\mathbb{P}\circ X_t^{-1}(x)(\mathrm{d}y) = \rho_t(x,y)\mathrm{d}y$, *where* $X_t(x)$ *is the solution of SDE* (1.3) *with starting point* $X_0(x) = x$.

(ii) There are $\varepsilon_0, \vartheta_0, q_0$ *and* γ *such that for all* $\varepsilon \in [0,\varepsilon_0)$, $\vartheta \in [0,\vartheta_0)$ *and* $q \in [1, q_0)$,

$$\sup_{x\in\mathbb{R}^d}\|(\mathbb{I}-\Delta)_x^{\frac{\alpha+\varepsilon}{2}}\Delta_y^{\frac{\vartheta}{2}}\rho_t(x,\cdot)\|_q \leqslant C t^{-\gamma}, \ t\in(0,1).$$

(iii) If the support of $g(x,\cdot)$ *is contained in a ball* B_R *for all* $x \in \mathbb{R}^d$, *where* $R \geqslant 1$, *then for any* $k \in \mathbb{N}_0$, *there are* ϑ_0, q_0 *and* γ_k *such that for all* $\vartheta \in [0,\vartheta_0)$ *and* $q \in [1,q_0)$,

$$\sup_{x\in\mathbb{R}^d}\|\nabla_x^k\Delta_y^{\frac{\vartheta}{2}}\rho_t(x,\cdot)\|_q \leqslant C t^{-\gamma_k}, \ t\in(0,1).$$

Remark 1.3. *The above (ii) implies that for any* $\varphi \in L^\infty(\mathbb{R}^d)$,

$$\mathcal{T}_t\varphi(x) := \int_{\mathbb{R}^d}\varphi(y)\rho_t(x,y)\mathrm{d}y \in \mathbb{C}^{\alpha+\varepsilon},$$

where $\mathbb{C}^{\alpha+\varepsilon}$ *is the usual Hölder space. In particular, the strong Feller property holds for* $\mathcal{T}_t$. *Moreover, if* $\sigma_k \equiv 0$ *and* $\alpha \in [1,2)$, *then* $\mathcal{T}_t\varphi$ *satisfies the following nonlocal equation in the classical sense*

$$\partial_t\mathcal{T}_t\varphi = \mathscr{A}\mathcal{T}_t\varphi, \ t > 0.$$

1.3 Examples. Below we provide several simple examples to illustrate our results.

Example 1. (A standard nonlinear example) Let L_t be an one dimensional Lévy process with Lévy measure $\nu(\mathrm{d}z) = \mathrm{d}z/|z|^{1+\alpha}$, where $\alpha \in (0, 2)$. Consider the following SDE:

$$\mathrm{d}X_t = -\sin(X_t)\mathrm{d}t + \cos(X_t)\mathrm{d}L_t, \quad X_0 = x.$$

In this case, $A_0 = -\sin(x)\partial_x$ and $\widetilde{A}_1 = \cos(x)\partial_x$. The generator of X_t is given by

$$\mathscr{A}\varphi(x) := -\sin(x)\varphi'(x) + \mathrm{p.v.}\int_{\mathbb{R}}\Big(\varphi(x + \cos(x)z) - \varphi(x)\Big)\nu(\mathrm{d}z).$$

Clearly, $[A_0, \widetilde{A}_1] = \partial_x$ and $(\mathbf{H}^{\mathrm{uni}}_{\mathbf{or}})$ holds with $c_0 = 1$ in (1.10).

Example 2. (Nonlocal Grushin's type operator) Let $L_t = (L^1_t, L^2_t)$ be a two-dimensional Lévy process with Lévy measure $\nu(\mathrm{d}z) = 1_{|z|\leqslant 1}|z|^{-2-\alpha}\mathrm{d}z$, where $\alpha \in (0, 2)$. Let $X_t = (X^1_t, X^2_t)$ solve the following SDE:

$$\begin{cases} \mathrm{d}X^1_t = \mathrm{d}L^1_t, & X^1_0 = x_1, \\ \mathrm{d}X^2_t = X^1_t\mathrm{d}L^2_t, & X^2_0 = x_2. \end{cases}$$

In this case, $A_k = 0$ for $k = 0, 1, 2$, $\widetilde{A}_1 = \partial_{x_1}$, $\widetilde{A}_2 = x_1\partial_{x_2}$, and the generator of X_t is given by

$$\mathscr{A}\varphi(x) := \mathrm{p.v.}\int_{\mathbb{R}^2}\Big(\varphi(x_1 + z_1, x_2 + x_1 z_2) - \varphi(x)\Big)\nu(\mathrm{d}z).$$

Clearly, $[\widetilde{A}_1, \widetilde{A}_2] = \partial_{x_2}$ and $(\mathbf{H}^{\mathrm{uni}}_{\mathbf{or}})$ holds with $c_0 = 2$ in (1.10).

Example 3. (Local and nonlocal Grushin's type operator) Let L_t be an one-dimensional Lévy process with Lévy measure $\nu(\mathrm{d}z) = |z|^{-1-\alpha}\mathrm{d}z$, where $\alpha \in (0, 2)$ and W_t is an one-dimensional Brownian motion independent of the process L_t. Let $X_t = (X^1_t, X^2_t)$ solve the following SDE:

$$\begin{cases} \mathrm{d}X^1_t = \mathrm{d}L_t, & X^1_0 = x_1, \\ \mathrm{d}X^2_t = X^1_t\mathrm{d}W_t, & X^2_0 = x_2. \end{cases}$$

In this case, $A_0 = A_1 = 0, A_2 = x_1\partial_{x_2}, \widetilde{A}_1 = \partial_{x_1}, \widetilde{A}_2 = 0$, and the generator of X_t is given by

$$\mathscr{A}\varphi(x) := \frac{1}{2}x_1^2\partial^2_{x_2}\varphi(x) + \mathrm{p.v.}\int_{\mathbb{R}}\Big(\varphi(x_1 + z_1, x_2) - \varphi(x)\Big)\nu(\mathrm{d}z_1).$$

Clearly, $[\widetilde{A}_1, A_2] = \partial_{x_2}$ and $(\mathbf{H}^{\mathrm{uni}}_{\mathbf{or}})$ holds with $c_0 = 2$ in (1.10).

Example 4. (Nonlocal Kolmogorov's type operator) Let L_t be an one-dimensional Lévy process with Lévy measure $\nu(\mathrm{d}z) = |z|^{-1-\alpha}\mathrm{d}z$, where $\alpha \in (0, 2)$. Let $X_t = (X^1_t, X^2_t)$ solve the following SDE:

$$\begin{cases} \mathrm{d}X^1_t = X^2_t\mathrm{d}t, & X^1_0 = x_1, \\ \mathrm{d}X^2_t = \mathrm{d}L_t - X^1_t\mathrm{d}t, & X^2_0 = x_2. \end{cases}$$

In this case, $A_1 = A_2 = 0, A_0 = x_2\partial_{x_1} - x_1\partial_{x_2}, \widetilde{A}_1 = 0, \widetilde{A}_2 = \partial_{x_2}$, and the generator of X_t is given by

$$\mathscr{A}\varphi(x) := x_2\partial_{x_1}\varphi(x) - x_1\partial_{x_2}\varphi(x) + \mathrm{p.v.}\int_{\mathbb{R}}\Big(\varphi(x_1, x_2 + z_2) - \varphi(x)\Big)\nu(\mathrm{d}z_2).$$

Clearly, $[A_0, \widetilde{A}_2] = \partial_{x_1}$ and $(\mathbf{H}^{\mathrm{uni}}_{\mathbf{or}})$ holds with $c_0 = 2$ in (1.10).

Example 5. (Nonlocal relativistic operator) Let L_t be an one-dimensional Lévy process with Lévy measure $\nu(\mathrm{d}z) = |z|^{-1-\alpha}\mathrm{d}z$, where $\alpha \in (0,2)$. Let $Z_t = (X_t, V_t)$ solve the following SDE:

$$\begin{cases} \mathrm{d}X_t = V_t/\sqrt{1+|V_t|^2}\mathrm{d}t, & X_0 = x, \\ \mathrm{d}V_t = \mathrm{d}L_t - X_t\mathrm{d}t, & V_0 = \mathrm{v}. \end{cases}$$

In this case, $A_1 = A_2 = 0, A_0 = \mathrm{v}/\sqrt{1+|\mathrm{v}|^2}\partial_x - x\partial_{\mathrm{v}}, \widetilde{A}_1 = 0, \widetilde{A}_2 = \partial_{\mathrm{v}}$, and the generator of Z_t is given by

$$\mathscr{A}\varphi(x,\mathrm{v}) := \frac{\mathrm{v}}{\sqrt{1+|\mathrm{v}|^2}}\partial_x\varphi(x,\mathrm{v}) - x\partial_{\mathrm{v}}\varphi(x,\mathrm{v}) + \mathrm{p.v.}\int_{\mathbb{R}}\Big(\varphi(x,\mathrm{v}+\mathrm{v}') - \varphi(x,\mathrm{v})\Big)\nu(\mathrm{d}w).$$

Clearly, $[A_0, \widetilde{A}_2] = (1+|\mathrm{v}|^2)^{-3/2}\partial_x$ and $(\mathbf{H}^{\mathrm{str}}_{\mathbf{or}})$ holds.

1.4 Structure. This paper is organized as follows: In Section 2 we recall Bismut's approach to the Malliavin calculus of Wiener-Poisson functionals. In Section 3, we recall and prove an improved Komatsu-Takeuchi's type estimate. In Section 4, we show the key estimate of the Laplace transform of the reduced Malliavin matrix. In Section 5, we prove Theorem 1.1. In Section 6 we prove Theorem 1.3. Finally, in Section 7, we apply our main result to the nonlocal kinetic operators and show the existence of smooth fundamental solutions, where the key point is to write the nonlocal operator as the generator of an SDE. For this aim, we need to solve a relaxed Jacobi equation.

2. Preliminaries

In this subsection, we recall some basic facts about Bismut's approach to the Malliavin calculus with jumps (see [22, Section 2]). Let $\Gamma \subset \mathbb{R}^d$ be an open set containing the origin. We define

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

where $\mathbf{d}(z, \Gamma_0^c)$ 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 points $\omega = (w, \mu)$, where

- $w : [0,1] \to \mathbb{R}^d$ is a continuous function with $w(0) = 0$;
- μ is an integer-valued measure 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: for $\omega = (w, \mu)$,

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

Let $(\mathscr{F}_t)_{t\in[0,1]}$ be the smallest right-continuous filtration on Ω such that W and N are optional. In the following, we write $\mathscr{F} := \mathscr{F}_1$, and endow $(\Omega, \mathscr{F})$ with the unique probability measure $\mathbb{P}$ such that

- W is a standard d-dimensional Brownian motion;
- 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

$$\kappa \in C^1(\Gamma_0; (0,\infty)), \quad \int_{\Gamma_0}(1 \wedge |z|^2)\kappa(z)\mathrm{d}z < +\infty, \quad |\nabla \log \kappa(z)| \leqslant C\varrho(z), \tag{2.2}$$

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

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

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

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

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

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

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

- $\mathbb{H}_p$: The space of all measurable adapted processes $h : \Omega \times [0,1] \to \mathbb{R}^d$ with finite norm:

$$\|h\|_{\mathbb{H}_p} := \left[\mathbb{E}\left(\int_0^1 |h(s)|^2\mathrm{d}s\right)^{\frac{p}{2}}\right]^{\frac{1}{p}} < +\infty.$$

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

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

  where $\varrho(z)$ is defined by (3.7). Below we shall write

$$\mathbb{H}_{\infty-} := \cap_{p\geqslant1}\mathbb{H}_p, \quad \mathbb{V}_{\infty-} := \cap_{p\geqslant1}\mathbb{V}_p.$$

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

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

- For any $p \geqslant 1$, $\mathbb{V}_0$ (resp. $\mathbb{H}_0$) is dense in $\mathbb{V}_p$ (resp. $\mathbb{H}_p$).

Let $C^\infty_p(\mathbb{R}^m)$ be the class of all smooth functions on $\mathbb{R}^m$ whose derivatives of all orders have at most polynomial growth. Let $\mathcal{F}C^\infty_p$ be the class of all Wiener-Poisson functionals on Ω with the following form:

$$F(\omega) = f(w(h_1),\cdots,w(h_{m_1}),\mu(g_1),\cdots,\mu(g_{m_2})), \quad \omega = (w,\mu) \in \Omega,$$

where $f \in C^\infty_p(\mathbb{R}^{m_1+m_2})$, $h_1,\cdots,h_{m_1} \in \mathbb{H}_0$ and $g_1,\cdots,g_{m_2} \in \mathbb{V}_0$ are non-random, and

$$w(h_i) := \int_0^1 \langle h_i(s),\mathrm{d}w(s)\rangle_{\mathbb{R}^d}, \quad \mu(g_j) := \int_0^1\int_{\Gamma_0} g_j(s,z)\mu(\mathrm{d}s,\mathrm{d}z).$$

Notice that

$$\mathcal{F}C^\infty_p \text{ is dense in } \cap_{p\geqslant1} L^p(\Omega,\mathscr{F},\mathbb{P}).$$

For $F \in \mathcal{F}C^\infty_p$ and $\Theta = (h,\mathrm{v}) \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$, define

$$D_\Theta F := \sum_{i=1}^{m_1}\partial_i f\int_0^1 \langle h(s),h_i(s)\rangle_{\mathbb{R}^d}\mathrm{d}s + \sum_{j=1}^{m_2}\partial_{j+m_1} f\int_0^1\int_{\Gamma_0} \nabla_{\mathrm{v}} g_j(s,z)\mu(\mathrm{d}s,\mathrm{d}z), \tag{2.3}$$

where $\nabla_{\mathrm{v}} g_j(s,z) := \mathrm{v}_i(s,z)\partial_{z_i} g_j(s,z)$.

We have the following integration by parts formula (cf. [22, Theorem 2.9]).

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

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

Moreover, we have the following consequences:

(i) For any $F \in \mathbb{W}_\Theta^{1,p}(\Omega)$, the following integration by parts formula holds:

$$\mathbb{E}(D_\Theta F) = \mathbb{E}(F\mathrm{div}(\Theta)), \tag{2.4}$$

where $\mathrm{div}(\Theta)$ is defined by

$$\mathrm{div}\Theta := \int_0^1 \langle h(s), \mathrm{d}W_s\rangle_{\mathbb{R}^d} - \int_0^1\int_{\Gamma_0} \frac{\mathrm{div}(\kappa\mathrm{v})(s,z)}{\kappa(z)}\widetilde{N}(\mathrm{d}s, \mathrm{d}z). \tag{2.5}$$

(ii) For $m, k \in \mathbb{N}$ and $F = (F_1, \cdots, F_m) \in (\mathbb{W}_\Theta^{1,\infty-})^m$, $\varphi \in C_p^\infty(\mathbb{R}^m; \mathbb{R}^k)$, we have

$$\varphi(F) \in (\mathbb{W}_\Theta^{1,\infty-})^k \ \text{ and } \ D_\Theta\varphi(F) = D_\Theta F^i \partial_i \varphi(F). \tag{2.6}$$

The following Kusuoka and Stroock's formula is proven in [22, Proposition 2.11].

Proposition 2.1. *Fix $\Theta = (h, \mathrm{v}) \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$. Let $\eta(\omega, s, z) : \Omega \times [0,1] \times \Gamma_0 \to \mathbb{R}$ and $f(\omega, s) : \Omega \times [0,1] \to \mathbb{R}^d$ be measurable maps and satisfy that for each $(s,z) \in [0,1] \times \Gamma_0$,*

$$\eta(s,z), f(s) \in \mathbb{W}_\Theta^{1,\infty-}, \quad \eta(s,\cdot) \in C^1(\Gamma_0),$$

and $s \mapsto f(s), D_\Theta f(s)$ are $\mathscr{F}_s$-adapted,

$$s \mapsto \eta(s,z), D_\Theta\eta(s,z), \nabla_z\eta(s,z) \text{ are left-continuous and } \mathscr{F}_s\text{-adapted}. \tag{2.7}$$

Assume that for any $p > 1$, $\int_0^1 \|f(s)\|_{\Theta;1,p}^p \mathrm{d}s < \infty$ and

$$\mathbb{E}\left[\sup_{s\in[0,1]}\sup_{z\in\Gamma_0}\left(\frac{|\eta(s,z)|^p + |D_\Theta\eta(s,z)|^p}{(1\wedge|z|)^p} + |\nabla_z\eta(s,z)|^p\right)\right] < +\infty. \tag{2.8}$$

Then $\mathscr{I}_1(f) := \int_0^1 f(s)\mathrm{d}W_s$, $\mathscr{I}_2(\eta) := \int_0^1\int_{\Gamma_0} \eta(s,z)\widetilde{N}(\mathrm{d}s,\mathrm{d}z) \in \mathbb{W}_\Theta^{1,\infty-}$ and

$$\begin{aligned} D_\Theta\mathscr{I}_1(f) &= \int_0^1 D_\Theta f(s)\mathrm{d}W_s + \int_0^1 f(s)\dot{h}(s)\mathrm{d}s, \\ D_\Theta\mathscr{I}_2(\eta) &= \int_0^1\int_{\Gamma_0} D_\Theta\eta(s,z)\widetilde{N}(\mathrm{d}s,\mathrm{d}z) + \int_0^1\int_{\Gamma_0} \nabla_{\mathrm{v}}\eta(s,z)N(\mathrm{d}s,\mathrm{d}z), \end{aligned} \tag{2.9}$$

where $\nabla_{\mathrm{v}}\eta(s,z) := \mathrm{v}_i(s,z)\partial_{z_i}\eta(s,z)$.

We also need the following Burkholder's type inequalities (cf. [22, Lemma 2.3]).

Lemma 2.1. *(i) For any $p > 1$, there is a constant $C_p > 0$ such that for any $\xi \in \mathbb{L}_p^1$,*

$$\mathbb{E}\left(\sup_{t\in[0,1]}\left|\int_0^t\int_{\Gamma_0} \xi(s,z)N(\mathrm{d}s,\mathrm{d}z)\right|^p\right) \leqslant C_p\|\xi\|_{\mathbb{L}_p^1}^p. \tag{2.10}$$

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

$$\mathbb{E}\left(\sup_{t\in[0,1]}\left|\int_0^t\int_{\Gamma_0} \xi(s,z)\widetilde{N}(\mathrm{d}s,\mathrm{d}z)\right|^p\right) \leqslant C_p\|\xi\|_{\mathbb{L}_p^2}^p. \tag{2.11}$$

The following result is taken from [27, Lemma 2.5], which is stated in a slightly different form.

Lemma 2.2. *Let $g_s(z), \eta_s$ be two left continuous $\mathscr{F}_s$-adapted processes satisfying that for some $\beta \in (0,1]$,*

$$0 \leqslant g_s(z) \leqslant \eta_s,\ |g_s(z) - g_s(0)| \leqslant \eta_s |z|^\beta,\ \forall |z| \leqslant 1, \tag{2.12}$$

and for any $p \geqslant 2$,

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

Let f_s be a nonnegative measurable adapted process and $\nu(\mathrm{d}z) = \mathrm{d}z/|z|^{d+\alpha}$. For any $\delta \in (0,1)$ and $m \geqslant 2$, there exist $c_2, \theta \in (0,1), C_2 \geqslant 1$ such that for all $\lambda, p \geqslant 1$ and $t \in (0,1)$,

$$\begin{aligned}
&\mathbb{E}\exp\left\{-\lambda\int_0^t\!\!\int_{\mathbb{R}^d_0} g_s(z)\zeta_{m,\delta}(z)N(\mathrm{d}s,\mathrm{d}z) - \lambda\int_0^t f_s\mathrm{d}s\right\}\\
&\quad\leqslant C_2\left(\mathbb{E}\exp\left\{-c_2\lambda^\theta\int_0^t (f_s + g_s(0))\mathrm{d}s\right\}\right)^{\frac{1}{2}} + C_p\lambda^{-p},
\end{aligned}\tag{2.13}$$

where $\zeta_{m,\delta}(z)$ is a nonnegative smooth function with

$$\zeta_{m,\delta}(z) = |z|^m,\ \ |z| \leqslant \delta/4,\ \ \zeta_{m,\delta}(z) = 0,\ \ |z| > \delta/2.$$

3. Improved Komatsu-Takeuchi's type estimate

Let $\mathscr{S}_m$ be the class of all $m-$dimensional semi-martingales with the following form

$$X_t = X_0 + \int_0^t f^0_s\mathrm{d}s + \int_0^t f^k_s\mathrm{d}W^k_s + \int_0^t\int_{\mathbb{R}^d_0} g_s(z)\widetilde{N}(\mathrm{d}s,\mathrm{d}z),$$

where $f^k_s, k = 0,\cdots,d$ and $g_s(z)$ are m-dimensional predictable processes with

$$\|X_\cdot(\omega)\|_{\mathscr{S}_m} := \sup_{s\in[0,1]}\left(|X_s(\omega)|^2 \vee |f^0_s(\omega)|^2 \vee |f^k_s(\omega)|^2 \vee \sup_{z\in\mathbb{R}^d}\frac{|g_s(z,\omega)|^2}{1\wedge|z|^2}\right) < \infty \text{ for } a.a. - \omega.$$

Here and below we use the following convention: If an index appears twice in a product, then it will be summed automatically. For instances,

$$\int_0^t f^k_s\mathrm{d}W^k_s := \sum_{k=1}^d\int_0^t f^k_s\mathrm{d}W^k_s,\ \ |f^k_s(\omega)|^2 := \sum_{k=1}^d |f^k_s(\omega)|^2.$$

For $\kappa > 0$, let $\mathscr{S}^\kappa_m$ be the subclass of $\mathscr{S}_m$ with

$$\|X_\cdot(\omega)\|_{\mathscr{S}_m} \leqslant \kappa \text{ for } a.a. - \omega.$$

We first recall the following estimates from [26, Theorem 4.2].

Lemma 3.1. *For $\kappa > 0$, let $(X_t)_{t\geqslant 0}$ and $(f^0_t)_{t\geqslant 0}$ be two semimartingales in $\mathscr{S}^\kappa_m$ with the form*

$$X_t = X_0 + \int_0^{t\wedge\tau}(f^0_s + h^\delta_s)\mathrm{d}s + \int_0^t f^k_s\mathrm{d}W^k_s + \int_0^t\!\!\int_{|z|\leqslant\delta} g_{s-}(z)\widetilde{N}(\mathrm{d}s,\mathrm{d}z),$$

$$f^0_t = f^0_0 + \int_0^t f^{00}_s\mathrm{d}s + \int_0^t f^{0k}_s\mathrm{d}W^k_s + \int_0^t\!\!\int_{|z|\leqslant\delta} g^0_{s-}(z)\widetilde{N}(\mathrm{d}s,\mathrm{d}z),$$

where $\delta \in (0,1]$ and τ is a stopping time. Assume that for some $\beta \in [0,2]$,

$$|h_t^\delta|^2 \leqslant \kappa\delta^\beta, \quad a.s.$$

For any $\varepsilon, \delta, t \in (0,1]$, there are positive random variables ζ_1, ζ_2 with $\mathbb{E}\zeta_i \leqslant 1, i = 1,2$ such that almost surely

$$c_0 \int_0^t \left(|f_s^k|^2 + \int_{|z|\leqslant\delta} |g_s(z)|^2 \nu(\mathrm{d}z)\right)\mathrm{d}s \leqslant (\delta^{-1} + \varepsilon^{-1}) \int_0^t |X_s|^2 \mathrm{d}s + \kappa\delta \log\zeta_1 + \kappa(\varepsilon + t\delta), \quad (3.1)$$

and

$$c_1 \int_0^{t\wedge\tau} |f_s^0|^2 \mathrm{d}s \leqslant (\delta^{-\frac{3}{2}} + \varepsilon^{-\frac{3}{2}}) \int_0^t |X_s|^2 \mathrm{d}s + \kappa\delta^{\frac{1}{2}} \log\zeta_2 + \kappa(\varepsilon\delta^{-\frac{1}{2}} + \varepsilon^{\frac{1}{2}} + t\delta^{\frac{1}{2}\wedge\beta}), \quad (3.2)$$

where $c_0, c_1 \in (0,1)$ only depend on $\int_{|z|\leqslant 1} |z|^2 \nu(\mathrm{d}z)$.

Below we show a refinement of (3.1) and [11, Lemma 5.1] for our aim.

Lemma 3.2. *For $\kappa > 0$ and $\delta \in (0,1)$, let $(X_t)_{t\geqslant 0}$ be a semi-martingale in $\mathscr{S}_m^\kappa$ with the form*

$$X_t = X_0 + \int_0^t f_s^0 \mathrm{d}s + \int_0^t f_s^k \mathrm{d}W_s^k + \int_0^t \int_{|z|\leqslant\delta} g_s(z)\widetilde{N}(\mathrm{d}s, \mathrm{d}z).$$

For any $\varepsilon, \delta, t \in (0,1)$, there is a positive random variable ζ with $\mathbb{E}\zeta \leqslant 1$ such that for all $\beta \in [0,2]$, it holds almost surely

$$c_0 \int_0^t \left(|f_s^k|^2 + \int_{|z|\leqslant\delta} |g_s(z)|^2 \nu(\mathrm{d}z)\right)\mathrm{d}s \leqslant (\delta^{-\beta} + \varepsilon^{-1}) \int_0^t |X_s|^2 \mathrm{d}s + \kappa\delta^\beta \log\zeta + \kappa\varepsilon, \quad (3.3)$$

where $c_0 \in (0,1)$ only depends on $\int_{|z|\leqslant 1} |z|^2 \nu(\mathrm{d}z)$. In particular, if $\nu(\mathrm{d}z) = \mathrm{d}z/|z|^{d+\alpha}$ for some $\alpha \in (0,2)$ and

$$g_s(z) = \Gamma_s \cdot z + \widetilde{g}_s(z) \text{ with } \sup_{s\in[0,1]} \left(\|\Gamma_s\|_{HS}^2 \vee \sup_{|z|\leqslant 1} \frac{|\widetilde{g}_s(z)|^2}{|z|^4}\right) \leqslant \kappa, \quad a.e. - \omega, \quad (3.4)$$

where $\Gamma_s : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^m \otimes \mathbb{R}^d$ is a matrix valued predictable process, and $\|\cdot\|_{HS}$ denotes the Hilbert-Schdmit norm, then for some $c_0 = c_0(d,\alpha) \in (0,1)$,

$$c_0 \int_0^t \left(|f_s^k|^2 + \|\Gamma_s\|_{HS}^2\right)\mathrm{d}s \leqslant \delta^{2\alpha-6} \int_0^t |X_s|^2 \mathrm{d}s + \kappa\delta^\alpha \log\zeta + \kappa\delta^2. \quad (3.5)$$

Proof. By replacing $(X_t, f_t^k, g_t(z))$ with $(X_t, f_t^k, g_t(z))/\sqrt{\kappa}$, we may assume that

$$|X_t|^2 \vee |f_t^0|^2 \vee |f_t^k|^2 \vee \frac{|g_t(z)|^2}{1 \wedge |z|^2} \leqslant 1. \quad (3.6)$$

Notice that for $t \leqslant 2\varepsilon$,

$$\int_0^t |f_s^k|^2 \mathrm{d}s + \int_0^t \int_{|z|\leqslant\delta} |g_s(z)|^2 \nu(\mathrm{d}z)\mathrm{d}s \leqslant 2\varepsilon\left(1 + \int_{|z|\leqslant 1} |z|^2 \nu(\mathrm{d}z)\right).$$

This implies (3.3) with $\zeta \equiv 1$. Below, without loss of generality, we assume $t > 2\varepsilon$ and $m = 1$. Following the proof in [11], let $\rho(u) : \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing smooth function with

$$\rho(u) = u \text{ for } 0 \leqslant u < 4 \text{ and } \rho(u) = \tfrac{9}{2} \text{ for } u > 6, \rho'(u) \leqslant 1. \quad (3.7)$$

Since $\rho(u) \leqslant u$ for $u \geqslant 0$, letting $\varepsilon_s := (t-\varepsilon) \wedge s - 0 \vee (s-\varepsilon)$, we have

$$
\begin{aligned}
\delta^{-\beta} \int_0^t |X_s|^2 \mathrm{d}s \geqslant \int_0^t \rho(\delta^{-\beta}|X_s|^2)\mathrm{d}s &\geqslant \int_\varepsilon^t \rho(\delta^{-\beta}|X_s|^2)\mathrm{d}s - \int_0^{t-\varepsilon} \rho(\delta^{-\beta}|X_s|^2)\mathrm{d}s \\
&= \int_0^{t-\varepsilon} \mathrm{d}s \int_s^{s+\varepsilon} \mathrm{d}\rho(\delta^{-\beta}|X_r|^2) = \int_0^t \varepsilon_s \mathrm{d}\rho(\delta^{-\beta}|X_s|^2),
\end{aligned}
$$

where the second equality is due to Fubini's theorem. By Itô's formula, we obtain

$$
\begin{aligned}
\delta^{-\beta} \int_0^t |X_s|^2 \mathrm{d}s \geqslant\ & 2\delta^{-\beta} \int_0^t \varepsilon_s \rho_s' X_s f_s^0 \mathrm{d}s + \delta^{-\beta} \int_0^t \varepsilon_s \{\rho_s' + 2\delta^{-\beta}|X_s|^2 \rho_s''\} |f_s^k|^2 \mathrm{d}s \\
&+ \left\{ 2\delta^{-\beta} \int_0^t \varepsilon_s \rho_s' X_s f_s^k \mathrm{d}W_s^k + \int_0^t \int_{|z|\leqslant\delta} \varepsilon_s (\tilde{\rho}_s(z) - \rho_s) \widetilde{N}(\mathrm{d}s, \mathrm{d}z) \right\} \\
&+ \int_0^t \int_{|z|\leqslant\delta} \varepsilon_s \Big((\tilde{\rho}_s(z) - \rho_s) - 2\delta^{-\beta} \rho_s' X_s g_s(z) \Big) \nu(\mathrm{d}z)\mathrm{d}s =: \sum_{i=1}^4 I_i(t),
\end{aligned}
$$

where $\tilde{\rho}_s(z) := \rho(\delta^{-\beta}|X_s + g_s(z)|^2)$ and

$$
\rho_s := \rho(\delta^{-\beta}|X_s|^2), \quad \rho_s' := \rho'(\delta^{-\beta}|X_s|^2), \quad \rho_s'' := \rho''(\delta^{-\beta}|X_s|^2).
$$

For $I_1(t)$, thanks to $|\varepsilon_s| \leqslant \varepsilon$, by (3.6) and (3.7) we have

$$
|I_1(t)| \leqslant \delta^{-\beta} \int_0^t (|X_s|^2 + |\varepsilon_s f_s^0|^2)\mathrm{d}s \leqslant \delta^{-\beta} \int_0^t |X_s|^2 \mathrm{d}s + \delta^{-\beta}\varepsilon^2 t.
$$

For $I_2(t)$, noticing that

$$
|\varepsilon_s - \varepsilon| \leqslant \varepsilon\{1_{(0,\varepsilon)}(s) + 1_{(t-\varepsilon,t)}(s)\} \tag{3.8}
$$

and

$$
\varepsilon_s \rho_s' = \varepsilon + \varepsilon(\rho_s' - 1) + (\varepsilon_s - \varepsilon)\rho_s' \geqslant \varepsilon - \varepsilon\delta^{-\beta}|X_s|^2 - |\varepsilon_s - \varepsilon|, \tag{3.9}
$$

where we have used $|\rho'(u) - 1| \leqslant u$ and $|\rho'(u)| \leqslant 1$, we have

$$
\begin{aligned}
I_2(t) &\geqslant \delta^{-\beta} \int_0^t \varepsilon_s \rho_s' |f_s^k|^2 \mathrm{d}s - 2\|\rho''\|_\infty \delta^{-2\beta} \int_0^t \varepsilon_s |X_s|^2 |f_s^k|^2 \mathrm{d}s \\
&\geqslant \delta^{-\beta} \int_0^t \Big\{ \varepsilon - \varepsilon\delta^{-\beta}|X_s|^2 - |\varepsilon_s - \varepsilon| \Big\} |f_s^k|^2 \mathrm{d}s - 2\|\rho''\|_\infty \delta^{-2\beta}\varepsilon \int_0^t |X_s|^2 \mathrm{d}s \\
&\geqslant \varepsilon\delta^{-\beta} \int_0^t |f_s^k|^2 \mathrm{d}s - (1 + 2\|\rho''\|_\infty)\varepsilon\delta^{-2\beta} \int_0^t |X_s|^2 \mathrm{d}s - \delta^{-\beta} \int_0^t |\varepsilon_s - \varepsilon| \mathrm{d}s \\
&\geqslant \varepsilon\delta^{-\beta} \int_0^t |f_s^k|^2 \mathrm{d}s - (1 + 2\|\rho''\|_\infty)\varepsilon\delta^{-2\beta} \int_0^t |X_s|^2 \mathrm{d}s - 2\delta^{-\beta}\varepsilon^2.
\end{aligned}
$$

For $I_3(t)$, noticing that

$$
|\varepsilon_s(\tilde{\rho}_s(z) - \rho_s)| \leqslant \varepsilon \Big\{ 9 \wedge (\delta^{-\beta} |\, |X_s + g_s(z)|^2 - |X_s|^2 |) \Big\},
$$

by [26, Lemma 4.1] with $R = \frac{1}{9\varepsilon}$, there is a positive random variable ζ_1 with $\mathbb{E}\zeta_1 \leqslant 1$ so that

$$
-I_3(t) - 9\varepsilon \log \zeta_1 \leqslant \frac{2\delta^{-2\beta}}{9\varepsilon} \int_0^t |\varepsilon_s \rho_s' X_s|^2 |f_s^k|^2 \mathrm{d}s + \frac{2}{9\varepsilon} \int_0^t \int_{|z|\leqslant\delta} \varepsilon_s^2 (\tilde{\rho}_s(z) - \rho_s)^2 \nu(\mathrm{d}z)\mathrm{d}s
$$

$$\overset{(3.6)}{\leqslant} \frac{2\varepsilon\delta^{-2\beta}}{9}\left[\int_0^t |X_s|^2 \mathrm{d}s + \int_0^t \int_{|z|\leqslant\delta} (|X_s + g_s(z)|^2 - X_s^2)^2 \nu(\mathrm{d}z)\mathrm{d}s\right]$$
$$\leqslant \frac{2\varepsilon\delta^{-2\beta}}{9}\left[\left(1 + 4\int_{|z|\leqslant\delta} |z|^2 \nu(\mathrm{d}z)\right)\int_0^t |X_s|^2 \mathrm{d}s + \int_0^t\int_{|z|\leqslant\delta} |g_s(z)|^4 \nu(\mathrm{d}z)\mathrm{d}s\right]$$
$$\overset{(3.6)}{\leqslant} C\varepsilon\delta^{-2\beta}\int_0^t |X_s|^2 \mathrm{d}s + \frac{2\varepsilon\delta^{2-2\beta}}{9}\int_0^t\int_{|z|\leqslant\delta} |g_s(z)|^2 \nu(\mathrm{d}z)\mathrm{d}s.$$

For $I_4(t)$, noticing that by Taylor's expansion for $x \mapsto \rho(\delta^{-\beta}x)$,

$$\rho(\delta^{-\beta}|X+g|^2) - \rho(\delta^{-\beta}|X|^2) - 2\delta^{-\beta}Xg\rho'(\delta^{-\beta}|X|^2)$$
$$= \delta^{-\beta}\rho'(\delta^{-\beta}|X|^2)g^2 + \tfrac{1}{2}\delta^{-2\beta}\rho''(\vartheta)(2Xg + |g|^2)^2,$$

where $\vartheta := \delta^{-\beta}(r|X|^2 + (1-r)|X+g|^2)$ for some $r \in [0,1]$, we can write

$$I_4(t) = \frac{\delta^{-2\beta}}{2}\int_0^t\int_{|z|\leqslant\delta} \varepsilon_s\rho''(\vartheta_s(z))(2X_s g_s(z) + |g_s(z)|^2)^2 \nu(\mathrm{d}z)\mathrm{d}s$$
$$+ \delta^{-\beta}\int_0^t\int_{|z|\leqslant\delta} \varepsilon_s\rho_s'|g_s(z)|^2 \nu(\mathrm{d}z)\mathrm{d}s =: I_{41}(t) + I_{42}(t),$$

where $\vartheta_s(z) := \delta^{-\beta}(r|X_s|^2 + (1-r)|X_s + g_s(z)|^2)$. For $I_{41}(t)$, noticing that

$$|X_s| \vee |g_s(z)| \leqslant \delta \Rightarrow \vartheta_s(z) \leqslant \delta^{-\beta}(r\delta^2 + 4(1-r)\delta^2) \leqslant 4\delta^{2-\beta} \leqslant 4,$$

in virtue of $\rho''(u) = 0$ for $u \leqslant 4$, by (3.6) we have

$$|I_{41}(t)| = \frac{\delta^{-2\beta}}{2}\left|\int_0^t\int_{|z|\leqslant\delta<|X_s|} \varepsilon_s\rho''(\vartheta_s(z))(2X_s g_s(z) + |g_s(z)|^2)^2 \nu(\mathrm{d}z)\mathrm{d}s\right|$$
$$\leqslant \frac{\delta^{-2\beta}}{2}\varepsilon\|\rho''\|_\infty\left(\int_{|z|\leqslant\delta} |z|^2\nu(\mathrm{d}z)\right)\int_0^t (3|X_s|)^2 \mathrm{d}s.$$

For $I_{42}(t)$, as in the treatment of $I_2(t)$, by (3.9) we have

$$I_{42}(t) \geqslant \delta^{-\beta}\int_0^t\int_{|z|\leqslant\delta} \left\{\varepsilon - \varepsilon\delta^{-\beta}|X_s|^2 - |\varepsilon_s - \varepsilon|\right\}|g_s(z)|^2 \nu(\mathrm{d}z)\mathrm{d}s$$
$$\geqslant \varepsilon\delta^{-\beta}\int_0^t\int_{|z|\leqslant\delta} |g_s(z)|^2 \nu(\mathrm{d}z)\mathrm{d}s - \delta^{-\beta}\int_0^t (\varepsilon\delta^{-\beta}|X_s|^2 + |\varepsilon_s - \varepsilon|)\mathrm{d}s\int_{|z|\leqslant\delta} |z|^2\nu(\mathrm{d}z)$$
$$\geqslant \varepsilon\delta^{-\beta}\int_0^t\int_{|z|\leqslant\delta} |g_s(z)|^2 \nu(\mathrm{d}z)\mathrm{d}s - \left(\varepsilon\delta^{-2\beta}\int_0^t |X_s|^2\mathrm{d}s + 2\varepsilon^2\delta^{-\beta}\right)\int_{|z|\leqslant 1} |z|^2\nu(\mathrm{d}z).$$

Combining the above estimates, we get for some $C \geqslant 9$ only depending on $\int_{|z|\leqslant 1} |z|^2\nu(\mathrm{d}z)$,

$$\delta^{-\beta}\left\{2 + C\varepsilon\delta^{-\beta}\right\}\int_0^t |X_s|^2\mathrm{d}s + 9\varepsilon\log\zeta_1 + C\varepsilon^2\delta^{-\beta}$$
$$\geqslant \varepsilon\delta^{-\beta}\int_0^t |f_s^k|^2\mathrm{d}s + \left(\varepsilon\delta^{-\beta} - \tfrac{2}{9}\varepsilon\delta^{2-2\beta}\right)\int_0^t\int_{|z|\leqslant\delta} |g_s(z)|^2\nu(\mathrm{d}z)\mathrm{d}s,$$

which gives (3.3) with $\zeta = \zeta_1^{9/C}$ by dividing both sides by $C\varepsilon\delta^{-\beta}$.

If $\nu(\mathrm{d}z)=\mathrm{d}z/|z|^{d+\alpha}$, then by $g_s(z)=\Gamma_s z+\widetilde{g}_s(z)$ and (3.4), we have

$$\begin{aligned}\int_0^t\int_{|z|\leqslant\delta}|g_s(z)|^2\nu(\mathrm{d}z)&\geqslant\frac{1}{2}\int_0^t\int_{|z|\leqslant\delta}|\Gamma_s z|^2\nu(\mathrm{d}z)-\int_0^t\int_{|z|\leqslant\delta}|\widetilde{g}_s(z)|^2\nu(\mathrm{d}z)\\&\geqslant c\delta^{2-\alpha}\int_0^t\|\Gamma_s\|_{HS}^2\mathrm{d}s-\kappa\int_0^t\int_{|z|\leqslant\delta}|z|^4\nu(\mathrm{d}z)\\&\geqslant c\delta^{2-\alpha}\int_0^t\|\Gamma_s\|_{HS}^2\mathrm{d}s-C\kappa\delta^{4-\alpha},\end{aligned}\tag{3.10}$$

where in the second inequality we have used that

$$\int_{|z|\leqslant\delta}z_iz_j\mathrm{d}z/|z|^{d+\alpha}=c\mathbf{1}_{i=j}\delta^{2-\alpha}.$$

Substituting this into (3.3) and taking $\beta=2$ and $\varepsilon=\delta^{4-\alpha}$, we get

$$c_0\int_0^t|f_s^k|^2\mathrm{d}s+c\delta^{2-\alpha}\int_0^t\|\Gamma_s\|_{HS}^2\mathrm{d}s\leqslant(\delta^{\alpha-4}+\delta^{-2})\int_0^t|X_s|^2\mathrm{d}s+\kappa\delta^2\log\zeta+\kappa\delta^{4-\alpha},$$

which gives (3.5) by dividing both sides by $\delta^{2-\alpha}$. □

Remark 3.1. *If we take $\beta=1$ and $\varepsilon=\delta$ in* (3.3)*, then* (3.3) *reduces to* (3.1) *with $\varepsilon=\delta$. Here the key point for us is the estimate* (3.5)*, which can not be derived from* (3.1)*.*

4. Estimates of Laplace transform of reduced Malliavin matrix

The result of this section is independent of the framework in Section 2. Consider the following SDE with jumps:

$$X_t=x+\int_0^t b(X_s)\mathrm{d}s+\int_0^t\sigma_k(X_s)\mathrm{d}W_s^k+\int_0^t\int_{|z|<1}g(X_{s-},z)\widetilde{N}(\mathrm{d}s,\mathrm{d}z).\tag{4.1}$$

It is well known that under $(\mathbf{H}_1)$, SDE (4.1) admits a unique solution denoted by $X_t=X_t(x)$, and $x\mapsto X_t(x)$ are smooth. Let $J_t:=J_t(x):=\nabla X_t(x)$ be the Jacobian matrix of $X_t(x)$. It is also well known that J_t solves the following linear matrix-valued SDE:

$$J_t=\mathbb{I}+\int_0^t\nabla b(X_s)J_s\mathrm{d}s+\int_0^t\nabla\sigma_k(X_s)J_s\mathrm{d}W_s^k+\int_0^t\int_{|z|<1}\nabla_x g(X_{s-},z)J_{s-}\widetilde{N}(\mathrm{d}s,\mathrm{d}z).\tag{4.2}$$

Moreover, under $(\mathbf{H}_g^{\circ})$, the matrix $J_t(x)$ is invertible. Let $K_t:=K_t(x)$ be the inverse matrix of $J_t(x)$. By Itô's formula, K_t solves the following linear matrix-valued SDE:

$$\begin{aligned}K_t=\mathbb{I}&-\int_0^t K_s[\nabla b(X_s)-(\nabla\sigma_k)^2(X_s)]\mathrm{d}s-\int_0^t K_s\nabla\sigma_k(X_s)\mathrm{d}W_s^k\\&+\int_0^t\int_{|z|<1}K_{s-}Q(X_{s-},z)\widetilde{N}(\mathrm{d}s,\mathrm{d}z)-\int_0^t\int_{|z|<1}K_{s-}(Q\cdot\nabla_x g)(X_{s-},z)\nu(\mathrm{d}z)\mathrm{d}s,\end{aligned}\tag{4.3}$$

where

$$Q(x,z):=(\mathbb{I}+\nabla_x g(x,z))^{-1}-\mathbb{I}=-(\mathbb{I}+\nabla_x g(x,z))^{-1}\cdot\nabla_x g(x,z).$$

Below we introduce some notations for later use.

- For $k=1,\cdots,d$, let $A_0,A_k,\widetilde{A}_k$ be defined as in the introduction. For simplicity, we write

$$A:=(A_1,\cdots,A_d),\quad \widetilde{A}:=(\widetilde{A}_1,\cdots,\widetilde{A}_d).\tag{4.4}$$

- For a vector field $V(x) := \mathrm{v}_i(x)\partial_i$, we also identify
$$V := (\mathrm{v}_1, \cdots, \mathrm{v}_d).$$
- For a smooth vector field $V : \mathbb{R}^d \to \mathbb{R}^d$, define
$$\overline{V} := [A_0, V] + \tfrac{1}{2}[A_k, [A_k, V]],$$
and for $\delta \in (0,1)$,
$$\begin{aligned} G_V(x,z) &:= V(x+g(x,z)) - V(x) + Q(x,z)V(x+g(x,z)), \\ H^\delta_V(x) &:= \int_{|z|\leqslant\delta} [G_V(x,z) + \nabla_x g(x,z)\cdot V(x) - g(x,z)\cdot\nabla V(x)]\nu(\mathrm{d}z). \end{aligned} \tag{4.5}$$
- For a row vector $u \in \mathbb{R}^d$ and a smooth vector field $V : \mathbb{R}^d \to \mathbb{R}^d$, define
$$\begin{aligned} \mathscr{H}_t(u,x) &:= \int_0^t |uK_s(x)V(X_s(x))|^2 \mathrm{d}s, \\ \mathscr{W}_t(u,x) &:= \int_0^t |uK_s(x)([A,V],[\widetilde{A},V],\overline{V})(X_s(x))|^2 \mathrm{d}s, \end{aligned} \tag{4.6}$$

where $([A,V],[\widetilde{A},V],\overline{V})$ is a matrix given by
$$([A_1,V],\cdots,[A_d,V],[\widetilde{A}_1,V],\cdots,[\widetilde{A}_d,V],\overline{V}).$$

We need the following easy lemma.

Lemma 4.1. *Let $V \in C^\infty_p(\mathbb{R}^d;\mathbb{R}^d)$. Suppose that* $(\mathbf{H}_1)$ *and* $(\mathbf{H}^\circ_g)$ *hold.*

(i) There are $m \in \mathbb{N}_0$ and $C > 0$ such that for all $\delta \in (0,1)$, $x \in \mathbb{R}^d$ and $|z| \leqslant 1$,
$$|Q(x,z)| \leqslant C|z|, \quad |G_V(x,z)| \leqslant C(1+|x|^m)|z|, \quad |H^\delta_V(x)| \leqslant C(1+|x|^m)\delta^{2-\alpha}.$$
(ii) $[A_k,V],[\widetilde{A}_k,V],\overline{V} \in C^\infty_p(\mathbb{R}^d;\mathbb{R}^d)$.

(iii) $\nabla_z G_V(x,0) = [\widetilde{A},V](x)$.

Proof. (i) The first two inequalities are obvious by definitions. For the third one, by Taylor's formula, we have
$$|G_V(x,z) + \nabla_x g(x,z)\cdot V(x) - g(x,z)\cdot\nabla V(x)| \leqslant C(1+|x|^m)|z|^2.$$
Combining this inequality with $\int_{|z|\leqslant\delta} |z|^2\mathrm{d}z/|z|^{d+\alpha} \leqslant C\delta^{2-\alpha}$, we obtain the third inequality.

(ii) It follows by definition.

(iii) Let $R(x,z) := (\mathbb{I} + \nabla_x g(x,z))^{-1}$. Note that
$$G_V(x,z) = R(x,z)V(x+g(x,z)) - V(x)$$
and
$$\nabla_z G_V(x,0) = \nabla_z R(x,0)V(x+g(x,0)) + R(x,0)\nabla V(x+g(x,0))\nabla_z g(x,0).$$
Since $g(x,0) = 0$, $R(x,0) = \mathbb{I}$ and $\nabla_z R(x,0) = -\nabla_x\nabla_z g(x,0)$, we get
$$\nabla_z G_V(x,0) = \nabla V(x)\nabla_z g(x,0) - \nabla_x\nabla_z g(x,0)V(x).$$
The proof is complete. □

Remark 4.1. *Under* $(\mathbf{H}'_1)$ *and $V \in C^\infty_b(\mathbb{R}^d;\mathbb{R}^d)$, the m in the above (i) can be zero.*

The following lemma is standard by BDG's inequality (see Lemma 2.1). We omit the details.

Lemma 4.2. *Under* $(\mathbf{H}_1)$ *and* $(\mathbf{H}^{\rm o}_g)$*, for any* $p \geqslant 1$*, we have*

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

Now we can show the following crucial lemma.

Lemma 4.3. *Let* $\beta := \frac{\alpha\wedge(2-\alpha)}{8}$ *and* $d_0 \in \mathbb{N}$*. Under* $(\mathbf{H}_1)$ *and* $(\mathbf{H}^{\rm o}_g)$*, for any* C^∞_p*-function* $V:\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^{d_0}$*, there exist constants* $c\in(0,1), C\geqslant 1$ *such that for all* $\delta, t\in(0,1), x\in\mathbb{R}^d$ *and* $p\geqslant 1$*,*

$$\sup_{|u|=1}\mathbb{P}\Big\{\mathscr{H}_t(u,x)\leqslant\delta^7 t, \mathscr{W}_t(u,x)\geqslant\delta^\beta t\Big\}\leqslant C_p(x)\delta^p+C\mathrm{e}^{-c\delta^{-\beta}t}, \tag{4.7}$$

where $\mathscr{H}_t(u,x)$ *and* $\mathscr{W}_t(u,x)$ *are defined by* (4.6)*. Moreover, under* $(\mathbf{H}'_1)$ *and* $V\in C^\infty_b$*, the constant* $C_p(x)$ *can be independent of* x*.*

Proof. We divide the proof into four steps.

(1) Fixing $\delta\in(0,1)$, we make the following decomposition:

$$L_t := \int_0^t\int_{|z|\leqslant\delta} z\widetilde{N}(\mathrm{d}s,\mathrm{d}z)+\int_0^t\int_{\delta<|z|<1} zN(\mathrm{d}s,\mathrm{d}z) =: L^\delta_t+\hat{L}^\delta_t,$$

where L^δ and $\hat{L}^\delta$ are the small jump part and large jump part of L, respectively. Clearly,

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

Let us fix a cádlág path $\hbar:\mathbb{R}_+\to\mathbb{R}^d$ with finitely many jumps on the finite time interval. Let $X^\delta_t(x;\hbar)$ solve the following SDE:

$$\begin{aligned} X^\delta_t(x;\hbar) &= x+\int_0^t b(X^\delta_s(x;\hbar))\mathrm{d}s+\int_0^t A_k(X^\delta_s(x;\hbar))\mathrm{d}W^k_s\\ &\quad+\int_0^t\int_{|z|\leqslant\delta} g(X^\delta_{s-}(x;\hbar),z)\widetilde{N}(\mathrm{d}s,\mathrm{d}z)+\sum_{0<s\leqslant t} g(X^\delta_{s-}(x;\hbar),\Delta\hbar_s). \end{aligned}$$

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

$$X_t(x)=X^\delta_t(x;\hbar)|_{\hbar=\hat{L}^\delta},\quad K_t(x)=K^\delta_t(x;\hbar)|_{\hbar=\hat{L}^\delta}, \tag{4.8}$$

which implies that

$$\mathscr{H}_t(u,x)=\mathscr{H}^\delta_t(u,x;\hbar)|_{\hbar=\hat{L}^\delta},\ \mathscr{W}_t(u,x)=\mathscr{W}^\delta_t(u,x;\hbar)|_{\hbar=\hat{L}^\delta}, \tag{4.9}$$

where for a row vector $u\in\mathbb{R}^d$,

$$\begin{aligned} \mathscr{H}^\delta_t(u,x;\hbar) &:= \int_0^t |uK^\delta_s(x;\hbar)V(X^\delta_s(x;\hbar))|^2\mathrm{d}s,\\ \mathscr{W}^\delta_t(u,x;\hbar) &:= \int_0^t |uK^\delta_s(x;\hbar)([A,V],[\widetilde{A},V],\overline{V})(X^\delta_s(x;\hbar))|^2\mathrm{d}s. \end{aligned} \tag{4.10}$$

(2) Below, we first consider the case $\hbar=0$. For simplicity, we drop the superscript δ and write

$$K_t = K_t(x;0),\quad X_t = X_t(x;0).$$

Moreover, for fixed $u \in \mathbb{R}^d$, we introduce the following processes: for $k = 1, \cdots, d$,

$$
\begin{aligned}
&f_t := uK_tV(X_t), \quad f_t^0 := uK_t\overline{V}(X_t), \quad h_t := uK_tH_V^\delta(X_t), \\
&f_t^k := uK_t[A_k, V](X_t), \quad g_t(z) := uK_{t-}G_V(X_{t-}, z), \\
&f_t^{0k} := uK_t[A_k, \overline{V}](X_t), \quad g_t^0(z) := uK_{t-}G_{\overline{V}}(X_{t-}, z), \\
&f_t^{00} := uK_t\Big([A_0, \overline{V}] + \tfrac{1}{2}[A_j, [A_j, \overline{V}]] + H_{\overline{V}}^\delta\Big)(X_t),
\end{aligned}
$$

where H_V^δ and G_V are defined by (4.5). Note that K_t solves the following equation (see (4.3)):

$$
\begin{aligned}
K_t = \mathbb{I} &- \int_0^t K_s[\nabla b(X_s) - (\nabla A_k)^2(X_s)]\mathrm{d}s - \int_0^t K_s\nabla A_k(X_s)\mathrm{d}W_s^k \\
&+ \int_0^t\int_{|z|\leqslant\delta} K_{s-}Q(X_{s-}, z)\widetilde{N}(\mathrm{d}s, \mathrm{d}z) - \int_0^t\int_{|z|\leqslant\delta} K_{s-}(Q \cdot \nabla_x g)(X_{s-}, z)\nu(\mathrm{d}z)\mathrm{d}s.
\end{aligned}
$$

Using Itô's formula, one finds that

$$
\begin{aligned}
f_t &= uV(x) + \int_0^t (f_s^0 + h_s)\mathrm{d}s + \int_0^t f_s^k\mathrm{d}W_s^k + \int_0^t\int_{|z|\leqslant\delta} g_s(z)\widetilde{N}(\mathrm{d}s, \mathrm{d}z), \\
f_t^0 &= u\overline{V}(x) + \int_0^t f_s^{00}\mathrm{d}s + \int_0^t f_s^{0k}\mathrm{d}W_s^k + \int_0^t\int_{|z|\leqslant\delta} g_s^0(z)\widetilde{N}(\mathrm{d}s\mathrm{d}z).
\end{aligned}
\tag{4.11}
$$

Let $\Gamma_t := uK_{t-}(\nabla_z G_V)(X_{t-}, 0)$. By (iii) of Lemma 4.1, we have

$$
\Gamma_t = uK_{t-}[\widetilde{A}, V](X_{t-}).
$$

Since $V \in C_p^\infty$, by Lemma 4.1, there exist an $m \in \mathbb{N}_0$ and $C \geqslant 1$ independent of $u, x \in \mathbb{R}^d$ such that for all $t \in [0, 1]$, $|z| \leqslant 1$ and $k = 0, \cdots, d$,

$$
\begin{aligned}
&|f_t| + |f_t^k| + |f_t^{0k}| \leqslant C|uK_t|(1 + |X_t|^m), \\
&\|\Gamma_t\|_{HS} \leqslant C|uK_{t-}|(1 + |X_{t-}|^m) \\
&|g_t(z)| + |g_t^0(z)| \leqslant C|uK_{t-}|(1 + |X_{t-}|^m)|z|, \\
&|\widetilde{g}_t(z)| := |g_t(z) - \Gamma_t z| \leqslant C|uK_{t-}|(1 + |X_{t-}|^m)|z|^2,
\end{aligned}
\tag{4.12}
$$

and

$$
|h_t| \leqslant |uK_t| \cdot |H_V^\delta(X_t)| \leqslant C|uK_t|(1 + |X_t|^m)\delta^{2-\alpha}. \tag{4.13}
$$

For $\gamma := \frac{\alpha\wedge(2-\alpha)}{4}$ and m being as above, define a stopping time

$$
\tau := \inf\Big\{s > 0 : |uK_s|^2 \vee |X_s|^{2m} > \delta^{-\gamma/2}\Big\}.
$$

By (4.12) and (4.13), there exists $\kappa_0 > 0$ such that for all $t \in [0, \tau)$, $|z| \leqslant 1$ and $k = 0, \cdots, d$

$$
\begin{aligned}
&|f_t|^2 + |f_t^k|^2 + |f_t^{0k}|^2 + |\Gamma_t|^2 \leqslant \kappa_0\delta^{-\gamma}, \qquad |g_t(z)|^2 + |g_t^0(t)|^2 \leqslant \kappa_0\delta^{-\gamma}|z|^2, \\
&|\widetilde{g}_t(z)|^2 \leqslant \kappa_0\delta^{-\gamma}|z|^4, \qquad |h_t|^2 \leqslant \kappa_0\delta^{4-2\alpha-\gamma}.
\end{aligned}
\tag{4.14}
$$

Moreover, by (4.11) we also have

$$
\begin{aligned}
f_{t\wedge\tau} &= uV(x) + \int_0^{t\wedge\tau} (f_s^0 + h_s)\mathrm{d}s + \int_0^t 1_{s<\tau}f_s^k\mathrm{d}W_s^k + \int_0^t\int_{|z|\leqslant\delta} 1_{s<\tau}g_s(z)\widetilde{N}(\mathrm{d}s, \mathrm{d}z), \\
f_{t\wedge\tau}^0 &= u\overline{V}(x) + \int_0^t 1_{s<\tau}f_s^{00}\mathrm{d}s + \int_0^t 1_{s<\tau}f_s^{0k}\mathrm{d}W_s^k + \int_0^t\int_{|z|\leqslant\delta} 1_{s<\tau}g_s^0(z)\widetilde{N}(\mathrm{d}s\mathrm{d}z).
\end{aligned}
$$

Fix $t \in (0,1)$. By (3.5) with $\kappa = \kappa_0\delta^{-\gamma}$, there are constant $c_0 = c_0(d,\alpha) \in (0,1)$ and random variable $\zeta_1 > 0$ with $\mathbb{E}\zeta_1 \leqslant 1$ such that

$$c_0 \int_0^{t\wedge\tau} (|f_s^k|^2 + \|\Gamma\|_{HS}^2)\mathrm{d}s \leqslant \delta^{2\alpha-6} \int_0^t |f_{s\wedge\tau}|^2 \mathrm{d}s + \kappa_0\delta^{\alpha-\gamma}\log\zeta_1 + \kappa_0\delta^{2-\gamma}. \tag{4.15}$$

By (3.2) with $\varepsilon = \delta^4$, $\kappa = \kappa_0\delta^{-\gamma}$ and $\beta = 4-2\alpha$, there are constant $c_1 = c_1(d,\alpha) > 0$ and random variable $\zeta_2 > 0$ with $\mathbb{E}\zeta_2 \leqslant 1$ such that

$$c_1 \int_0^{t\wedge\tau} |f_s^0|^2 \mathrm{d}s \leqslant 2\delta^{-6} \int_0^t |f_{s\wedge\tau}|^2 \mathrm{d}s + \kappa_0\delta^{\frac12-\gamma}\log\zeta_2 + \kappa_0\delta^{-\gamma}(2\delta^2 + t\delta^{\frac12\wedge(4-2\alpha)}). \tag{4.16}$$

Combining (4.15), (4.16) with definition (4.10), one finds that for some $c_2 = c_2(\kappa_0,d,\alpha) > 0$,

$$c_2\mathscr{W}_t^\delta(u,x;0) \leqslant \delta^{-6}\mathscr{H}_t^\delta(u,x;0) + \tfrac14\log(\zeta_1^{\delta^{\alpha-\gamma}}\zeta_2^{\delta^{\frac12-\gamma}}) + \delta^{2-\gamma} + t\delta^{\frac12\wedge(4-2\alpha)-\gamma},\ t<\tau.$$

Multiplying both sides by $\delta^{-\gamma}$ and taking exponential, we obtain

$$\mathbf{1}_{\{t<\tau\}}\exp\left\{c_2\delta^{-\gamma}\mathscr{W}_t^\delta(u,x;0) - \delta^{-6-\gamma}\mathscr{H}_t^\delta(u,x;0)\right\} \leqslant (\zeta_1^{\delta^{\alpha-2\gamma}}\zeta_2^{\delta^{\frac12-2\gamma}})^{\frac14}\mathrm{e}^{\delta^{2-2\gamma}+t\delta^{\frac12\wedge(4-2\alpha)-2\gamma}}.$$

Recalling $\gamma = \frac{\alpha\wedge(2-\alpha)}{4}$, taking expectations and by $\mathbb{E}\zeta_1 \leqslant 1$ and $\mathbb{E}\zeta_2 \leqslant 1$, we derive

$$\sup_{u,x\in\mathbb{R}^d} \mathbb{E}\Big(\mathbf{1}_{\{t<\tau\}}\exp\{c_2\delta^{-\gamma}\mathscr{W}_t^\delta(u,x;0) - \delta^{-7}\mathscr{H}_t^\delta(u,x;0)\}\Big) \leqslant \mathrm{e}^{\delta^\alpha+t},\ \forall \delta,t\in(0,1). \tag{4.17}$$

(3) Let m be as in (4.12). We introduce the following random set for later use:

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

We use (4.17) to show that there is a $C \geqslant 1$ such that for all $t,\delta\in(0,1)$ and $u,x\in\mathbb{R}^d$,

$$\mathbf{II} := \mathbb{E}\Big(\mathbf{1}_{\Omega_t^\delta(u,x;\hat{L}^\delta)}\cdot\exp\{c_2\delta^{-\gamma}\mathscr{W}_t(u,x) - \delta^{-7}\mathscr{H}_t(u,x)\}\Big) \leqslant C. \tag{4.18}$$

Let

$$\mathcal{J}_t^u(x;\hbar) := \mathbf{1}_{\Omega_t^\delta(u,x;\hbar)}\cdot\exp\Big\{c_2\delta^{-\gamma}\mathscr{W}_t^\delta(u,x;\hbar) - \delta^{-7}\mathscr{H}_t^\delta(u,x;\hbar)\Big\}.$$

Since $\Omega_t^\delta(u,x;0) \subset \{t<\tau\}$, by (4.17) we have

$$\sup_{u,x\in\mathbb{R}^d}\mathbb{E}\mathcal{J}_t^u(x;0) \leqslant \mathrm{e}^{\delta^\alpha+t}. \tag{4.19}$$

Let $0 = t_0 < t_1 < \cdots < t_n \leqslant t_{n+1} = t$ be the jump times of $\hbar$ before time t. For $j = 0,1,\cdots,n$ and $s\in[0,t_{j+1}-t_j)$, noting that

$$X_{s+t_j}^\delta(x,\omega;\hbar) = X_s^\delta(X_{t_j}^\delta(x,\omega;\hbar),\theta_{t_j}\omega;0)$$
$$\Rightarrow K_{s+t_j}^\delta(x,\omega;\hbar) = K_{t_j}^\delta(x,\omega;\hbar)K_s^\delta(X_{t_j}^\delta(x,\omega;\hbar),\theta_{t_j}\omega;0),$$

where θ_{t_j} is the usual shift operator in Ω, by definition we have

$$\mathbf{1}_{\Omega_{t_{j+1}}^\delta(u,x;\hbar)}(\omega) = \mathbf{1}_{\Omega_{t_j}^\delta(u,x;\hbar)}(\omega)\cdot\mathbf{1}_{\Omega_{t_{j+1}-t_j}^\delta(u',x';0)}(\theta_{t_j}\omega)\Big|_{u'=uK_{t_j}^\delta(x,\omega;\hbar),x'=X_{t_j}^\delta(x,\omega;\hbar)}.$$

Thus, by the Markov property and (4.19), we have for all $u,x\in\mathbb{R}^d$,

$$\mathbb{E}\mathcal{J}_{t_{n+1}}^u(x;\hbar) = \mathbb{E}\Big(\mathcal{J}_{t_n}^u(x;\hbar)\cdot\mathbb{E}\mathcal{J}_{t_{n+1}-t_n}^{u'}(x';0)|_{u'=uK_{t_n}^\delta(x;\hbar),x'=X_{t_n}^\delta(x;\hbar)}\Big)$$
$$\leqslant \mathrm{e}^{\delta^\alpha+t_{n+1}-t_n}\mathbb{E}\mathcal{J}_{t_n}^u(x;\hbar) \leqslant \cdots \leqslant \mathrm{e}^{n\delta^\alpha+t}.$$

Now we can give a proof of (4.18). Let

$$N_t^\delta := \sum_{s\in[0,t]} \mathbf{1}_{|\Delta \hat{L}_s|>\delta}, \quad \lambda_\delta := \int_{\delta<|z|<1} \mathrm{d}z/|z|^{d+\alpha}.$$

Then N^δ is a Poisson process with intensity λ_δ. By the independence of L^δ and $\hat{L}^\delta$, we have

$$\begin{aligned}\mathbf{II} &\overset{(4.9)}{=} \mathbb{E}\mathcal{J}_t^u(x;\hat{L}^\delta) = \sum_{n=0}^\infty \mathbb{E}(\mathbb{E}\mathcal{J}_t^u(x;\hbar)|_{\hbar=\hat{L}^\delta}; N_t^\delta = n)\\ &\leqslant \sum_{n=0}^\infty \mathrm{e}^{n\delta^\alpha + t}\mathbb{P}(N_t^\delta = n) = \mathrm{e}^t \sum_{n=0}^\infty \mathrm{e}^{n\delta^\alpha}\frac{(t\lambda_\delta)^n}{n!}\mathrm{e}^{-t\lambda_\delta}\\ &= \mathrm{e}^t \exp\{(\mathrm{e}^{\delta^\alpha}-1)\lambda_\delta t\} \leqslant \mathrm{e}^t \exp\{3\delta^\alpha\lambda_\delta\} \leqslant C,\end{aligned}$$

where we have used that $\mathrm{e}^s - 1 \leqslant 3s$ for $s\in(0,1)$ and $\lambda_\delta \leqslant c\delta^{-\alpha}$.

(4) For any $p>1$, by Chebyshev's inequality we have

$$\begin{aligned}\mathbb{P}((\Omega_t^\delta(u,x;\hat{L}^\delta))^c) &= \mathbb{P}\left(\sup_{s\in[0,t]}\left(|uK_s(x)|^2 \vee |X_s(x)|^{2m}\right) > \delta^{-\gamma/2}\right)\\ &\leqslant \delta^p \mathbb{E}\left[\sup_{s\in[0,t]}\left(|uK_s(x)|^2 \vee |X_s(x)|^{2m}\right)^{2p/\gamma}\right]\\ &\leqslant (C_p(x) + C|u|^{\frac{4p}{\gamma}})\delta^p.\end{aligned} \tag{4.20}$$

Therefore, for all $t\in(0,1)$, $x\in\mathbb{R}^d$, $|u|=1$ and $p\geqslant 1$, by (4.8), (4.20) and (4.18), we have

$$\begin{aligned}&\mathbb{P}\left\{\mathscr{H}_t(u,x) \leqslant \delta^7 t, \mathscr{W}_t(u,x) \geqslant \delta^{\gamma/2} t\right\} \leqslant \mathbb{P}((\Omega_t^\delta(u,x;\hat{L}^\delta))^c)\\ &+\mathbb{P}\left\{c_2\delta^{-\gamma}\mathscr{W}_t(u,x) - \delta^{-7}\mathscr{H}_t(u,x) \geqslant c_2\delta^{-\gamma/2}t - t; \Omega_t^\delta(u,x;\hat{L}^\delta)\right\}\\ &\leqslant (C_p(x) + C|u|^{\frac{4p}{\gamma}})\delta^p + C\mathrm{e}^{t-c_2\delta^{-\gamma/2}t}.\end{aligned}$$

We complete the proof of (4.7) by setting $\beta = \frac{\gamma}{2} = \frac{\alpha\wedge(2-\alpha)}{8}$.

(5) Under $(\mathbf{H}_1')$ and $V\in C_b^\infty$, the m in Lemma 4.1 and (4.12) can be zero so that $C_p(x)$ can be independent of x. $\square$

By a standard chain argument, we have the following lemma, which is the same in spirit as [12, Lemma 3.1].

Lemma 4.4. *Under* $(\mathbf{H}_1)$*, for any* $n\in\mathbb{N}$*, there is a constant* $C_n\geqslant 1$ *such that for all* $R\geqslant 1$*,*

$$\sup_{|x|<R}\mathbb{P}\left(\sup_{s\in[0,\varepsilon]}|X_s(x)| \geqslant C_n R\right) \leqslant C_n\varepsilon^n, \quad \forall \varepsilon\in(0,1). \tag{4.21}$$

Proof. Let $\mathbb{D}$ be the space of all cádlág functions from $[0,1]$ to $\mathbb{R}^d$. Let $\mathbf{P}_x = \mathbb{P}\circ X_\cdot^{-1}(x)$ be the law of $X_\cdot(x)$ in $\mathbb{D}$. With a little of confusion, let X be the coordinate process over $\mathbb{D}$ so that $(X,\mathbf{P}_x)_{x\in\mathbb{R}^d}$ forms a family of strong Markov processes. Let $\tau_0 = 0$ and $R\geqslant 1$. For $j\in\mathbb{N}$, define

$$\tau_j := \inf\left\{t > \tau_{j-1} : |X_t - X_{\tau_{j-1}}| > R\right\}, \; L_t := \int_0^t\int_{|z|<1} z\widetilde{N}(\mathrm{d}s,\mathrm{d}z).$$

Clearly, we have

$$\sup_{t\in[\tau_{j-1},\tau_j)}|X_t-X_{\tau_{j-1}}|\leqslant R,\quad X_{\tau_j}=X_{\tau_j-}+g(X_{\tau_j-},\Delta L_{\tau_j}).$$

Since $|g(x,z)|\leqslant C(1+|x|)$ for all $x\in\mathbb{R}^d$ and $|z|<1$, we have

$$|X_{\tau_j}|\leqslant|X_{\tau_j-}|+C(1+|X_{\tau_j-}|)\leqslant C_1R+C_2|X_{\tau_{j-1}}|.$$

By induction method, one sees that for each $j=1,2,\cdots,$

$$|X_{\tau_j}|\leqslant\frac{(C_2^j-1)C_1}{C_2-1}R+C_2^j|X_0|\leqslant C_nR,\ |X_0|\leqslant R,$$

and therefore,

$$\sup_{s\in[0,\tau_n]}|X_s|\leqslant C_nR,\ |X_0|\leqslant R,\tag{4.22}$$

which implies that

$$\left\{\sup_{s\in[0,\varepsilon]}|X_s|>C_nR\right\}\subset\Big\{\tau_n<\varepsilon\Big\},\ \ |X_0|\leqslant R.\tag{4.23}$$

Noting that $\tau_n=\tau_{n-1}+\tau_1\circ\theta_{\tau_{n-1}}$, where $\theta_\cdot$ is the usual shift operator, by the strong Markov property, we have

$$\begin{aligned}\mathbf{P}_x\Big\{\tau_n<\varepsilon\Big\}&=\mathbf{P}_x\Big\{\tau_n<\varepsilon;\tau_{n-1}<\varepsilon\Big\}\leqslant\mathbf{P}_x\Big\{\tau_1\circ\theta_{\tau_{n-1}}<\varepsilon;\tau_{n-1}<\varepsilon\Big\}\\&=\mathbf{E}_x\left(\mathbf{P}_{X_{\tau_{n-1}}}\Big(\tau_1<\varepsilon\Big);\tau_{n-1}<\varepsilon\right).\end{aligned}\tag{4.24}$$

On the other hand, by BDG's inequality and the linear growth of b,σ and g, we have

$$\begin{aligned}\mathbf{E}_y\left(\sup_{t\in[0,\varepsilon]}|X_t-X_0|^2\right)&\lesssim\mathbf{E}_y\left(\int_0^\varepsilon|b(X_s)|\mathrm{d}s\right)^2+\mathbf{E}_y\left(\int_0^\varepsilon|\sigma(X_s)|^2\mathrm{d}s\right)\\&\quad+\mathbf{E}_y\left(\int_0^\varepsilon\int_{|z|<1}|g(X_s,z)|^2\nu(\mathrm{d}z)\mathrm{d}s\right)\\&\lesssim\varepsilon\sup_{s\in[0,\varepsilon]}\mathbf{E}_y(1+|X_s|^2)\lesssim\varepsilon(1+|y|^2).\end{aligned}$$

Hence,

$$\mathbf{P}_y\Big(\tau_1<\varepsilon\Big)\leqslant\mathbf{P}_y\left(\sup_{t\in[0,\varepsilon]}|X_t-X_0|>R\right)\leqslant R^{-2}\mathbf{E}_y\left(\sup_{t\in[0,\varepsilon]}|X_t-X_0|^2\right)\leqslant CR^{-2}(1+|y|^2)\varepsilon.$$

Thus, by (4.22) and (4.24), we get for $|x|<R$,

$$\mathbf{P}_x\Big\{\tau_n<\varepsilon\Big\}\leqslant CR^{-2}(1+C_n^2R^2)\varepsilon\mathbf{P}_x\Big\{\tau_{n-1}<\varepsilon\Big\}\leqslant\cdots\leqslant C_n\varepsilon^n.$$

The proof is complete by (4.23). □

Remark 4.2. *Note that* (4.21) *can not be obtained by simple applications of Chebyshev's inequality. Intuitively, consider the Poisson process N_t with intensity λ. For $n\in\mathbb{N}$, we clearly have*

$$\mathbb{P}(N_t\geqslant n)=\mathrm{e}^{-\lambda t}\sum_{k=n}^\infty\frac{(\lambda t)^k}{k!}=(\lambda t)^n\mathrm{e}^{-\lambda t}\sum_{k=0}^\infty\frac{(\lambda t)^k}{(k+n)!}\leqslant(\lambda t)^n.$$

If we let τ_n be the n-th jump time of N, then $\{N_t\geqslant n\}=\{\tau_n\leqslant t\}$.

We define the reduced Malliavin matrix by

$$\hat{\Sigma}_t(x) := \int_0^t K_s(x)(AA^* + \widetilde{A}\widetilde{A}^*)(X_s(x))K_s^*(x)\mathrm{d}s,$$

where A and $\widetilde{A}$ are defined in (4.4). Following the proof of [27, Theorem 3.3], we have

Theorem 4.1. *Under* $(\mathbf{H}_1)$, $(\mathbf{H}_g^{\mathrm{o}})$ *and* $(\mathbf{H}_{\mathbf{or}}^{\mathrm{str}})$, *there exist* $\gamma = \gamma(\alpha, j_0) \in (0,1)$ *and constants* $C_2 \geqslant 1, c_2 \in (0,1)$ *such that for all* $t \in (0,1)$ *and* $R, \lambda, p \geqslant 1$,

$$\sup_{|x|<R}\sup_{|u|=1}\mathbb{E}\Big[\exp\{-\lambda u\hat{\Sigma}_t(x)u^*\}\Big] \leqslant C_2\exp\{-c_2 t\lambda^\gamma\} + C_{R,p}(\lambda t)^{-p},$$

where $C_{R,p} > 0$ *continuously depends on* R, p.

Proof. Let A_0, A_k be as in (1.2) and $\widetilde{A}_k := \partial_{z_k} g^i(\cdot, 0)\partial_i$. Define

$$\mathscr{U}_0 := \{A_k, \widetilde{A}_k, k = 1, \cdots, d\},$$

and for $j = 1, 2, \cdots$,

$$\mathscr{U}_j := \Big\{[A_k, V], [\widetilde{A}_k, V], [A_0, V] + \tfrac{1}{2}[A_k, [A_k, V]] : V \in \mathscr{U}_{j-1}, k = 1, \cdots, d\Big\}.$$

Recall the definition of $\mathscr{V}_j$ in (1.8). It is easy to see that

$$\mathrm{span}\{\cup_{j=0}^{j_0}\mathscr{V}_j\} \subset \mathrm{span}\{\cup_{j=0}^{j_0+1}\mathscr{U}_j\}. \tag{4.25}$$

Let $\gamma := \alpha \wedge (2-\alpha)/56$ and $x, u \in \mathbb{R}^d$ with $|u| = 1$. For $j = 0, 1, \cdots, j_0 + 1$, define

$$E_j^x := \left\{\sum_{V \in \mathscr{U}_j}\int_0^t |uK_s(x)V(X_s(x))|^2\mathrm{d}s \leqslant t\varepsilon^{7\gamma^j}\right\},$$

Notice that

$$E_0^x \subset \Big(\cap_{j=0}^{j_0+1}E_j^x\Big) \cup \Big(\cup_{j=0}^{j_0-1}(E_j^x \setminus E_{j+1}^x)\Big)$$

and

$$E_{j+1}^x = \left\{\sum_{V \in \mathscr{U}_j}\int_0^t |uK_s(x)([A, V], [\widetilde{A}, V], \overline{V})(X_s(x))|^2\mathrm{d}s \leqslant t\varepsilon^{7\gamma^{j+1}}\right\}.$$

By (4.7) with $V \in \mathscr{U}_j$ and $\delta = \varepsilon^{\gamma^j}$, we have

$$\mathbb{P}(E_j^x \setminus E_{j+1}^x) \leqslant C_p(x)\varepsilon^p + C\exp\{-ct\varepsilon^{-8\gamma^{j+1}}\}.$$

To estimate $\mathbb{P}(\cap_{j=0}^{j_0+1}E_j^x)$, note that

$$\begin{aligned}\cap_{j=0}^{j_0+1}E_j^x &\subset \left\{\sum_{j=0}^{j_0+1}\sum_{V \in \mathscr{U}_j}\int_0^t |uK_s(x)V(X_s(x))|^2\mathrm{d}s \leqslant t\sum_{j=0}^{j_0+1}\varepsilon^{7\gamma^j}\right\}\\ &\subset \left\{\sum_{j=0}^{j_0+1}\sum_{V \in \mathscr{U}_j}\int_0^t |uK_s(x)V(X_s(x))|^2\mathrm{d}s \leqslant tj_0\varepsilon^{7\gamma^{j_0+1}}\right\}.\end{aligned} \tag{4.26}$$

By $(\mathbf{H}_{\mathbf{or}}^{\mathrm{str}})$ and (4.25), for each $x \in \mathbb{R}^d$, we have

$$\inf_{|u|=1}\sum_{j=0}^{j_0+1}\sum_{V \in \mathscr{U}_j}|uV(x)|^2 > 0,$$

which implies that for any $\kappa \geqslant 1$, there is a $c_0 > 0$ such that

$$\inf_{|x|\leqslant\kappa}\inf_{|u|=1}\sum_{j=0}^{j_0+1}\sum_{V\in\mathscr{U}_j}|uV(x)|^2 \geqslant c_0. \tag{4.27}$$

Let C_n be as in Lemma 4.4. Define stopping times

$$\tau_1^x := \inf\{t>0 : |X_t(x)| \geqslant C_nR\},\ \tau_2^x := \inf\{t>0 : \|J_t(x)\|_{HS} \geqslant \varepsilon^{-7\gamma^{j_0+1}/8}\}.$$

Noticing that for $s < \tau_2^x$,

$$\varepsilon^{7\gamma^{j_0+1}/8} \leqslant \|J_s(x)\|_{HS}^{-1} \leqslant |uK_s(x)|,$$

by (4.27) with $\kappa = C_nR$, we have on $\{\tau_1^x > t\varepsilon^{7\gamma^{j_0+1}/4}\}\cap\{\tau_2^x > t\}$,

$$\sum_{j=0}^{j_0+1}\sum_{V\in\mathscr{U}_j}\int_0^t |uK_s(x)V(X_s(x))|^2\mathrm{d}s \geqslant c_0\int_0^{t\varepsilon^{7\gamma^{j_0+1}/4}}|uK_s(x)|^2\mathrm{d}s \geqslant c_0t\varepsilon^{7\gamma^{j_0+1}/2}.$$

Thus, by (4.26), we have for $\varepsilon \leqslant \varepsilon_0$ small enough,

$$\left(\cap_{j=0}^{j_0+1} E_j^x\right)\cap\{\tau_1^x > t\varepsilon^{7\gamma^{j_0+1}/4}\}\cap\{\tau_2^x > t\} = \emptyset.$$

On the other hand, by Lemma 4.4, we have

$$\sup_{|x|<R}\mathbb{P}\left(\tau_1^x \leqslant t\varepsilon^{7\gamma^{j_0+1}/4}\right) \leqslant C_n(t\varepsilon^{7\gamma^{j_0+1}/4})^n,$$

and by Lemma 4.2 and Chebyshev's inequality, for any $p \geqslant 1$,

$$\sup_{x\in\mathbb{R}^d}\mathbb{P}\left(\tau_2^x \leqslant t\right) \leqslant C_p(\varepsilon^{7\gamma^{j_0+1}/8})^p.$$

Combining the above calculations, we obtain

$$\sup_{|x|<R}\mathbb{P}(E_0^x) \leqslant C_{R,p}\varepsilon^p + C\exp\{-ct\varepsilon^{-8\gamma^{j_0+2}}\} + C_n(t\varepsilon^{7\gamma^{j_0+1}/4})^n. \tag{4.28}$$

Noting that

$$u\hat{\Sigma}_t(x)u^* = \sum_{V\in\mathscr{U}_0}\int_0^t |uK_s(x)V(X_s(x))|^2\mathrm{d}s,$$

by (4.28) with $n \geqslant 4p/(7\gamma^{j_0+1})$, we obtain

$$\sup_{|x|<R}\sup_{|u|=1}\mathbb{P}\Big(u\hat{\Sigma}_t(x)u^* \leqslant t\varepsilon^{7\gamma^j}\Big) \leqslant C_{R,p}\varepsilon^p + C\exp\{-ct\varepsilon^{-8\gamma^{j_0+1}}\}.$$

The desired estimate of the Laplace transform of $u\hat{\Sigma}_t(x)u^*$ follows from this estimate (see [27]). □

Remark 4.3. *Under* $(\mathbf{H}_1')$, $(\mathbf{H}_g^{\mathrm{o}})$ *and* $(\mathbf{H}_{\mathbf{or}}^{\mathrm{uni}})$, *from the above proofs, one sees that* $C_{R,p}$ *can be independent of* R.

5. Proof of Theorem 1.1

For $p, R \in [1, \infty]$ and $\varphi \in C_c^\infty(\mathbb{R}^d)$, define

$$\|\varphi\|_{p;R} := \left(\int_{B_R} |\varphi(x)|^p \mathrm{d}x\right)^{1/p}, \ \mathcal{T}_t^0\varphi(x) := \mathbb{E}\varphi(X_t(x)).$$

We first prepare the following lemma for later use.

Lemma 5.1. *Under* $(\mathbf{H}_1)$ *and* $(\mathbf{H}_g^o)$*, for any* $k, m \in \mathbb{N}_0$ *and* $R, p \geqslant 1$*, there exists a constant* $C_R = C(R, p, k)$ *such that for all* $t \in (0, 1)$ *and* $\varphi \in C_c^\infty(\mathbb{R}^d)$*,*

$$\|\nabla^k \mathcal{T}_t^0 \nabla^m \varphi\|_{p;R} \leqslant C_R \sum_{j=0}^k \|\nabla^{m+j}\varphi\|_p. \tag{5.1}$$

Under $(\mathbf{H}_1')$ *and* $(\mathbf{H}_g^o)$*, the above* R *can be* ∞ *so that the global estimate holds.*

Proof. (i) We first show (5.1) for $k = m = 0$. By the change of variables, we have

$$\|\mathcal{T}_t^0\varphi\|_{p;R}^p = \mathbb{E}\int_{B_R} |\varphi(X_t(x))|^p \mathrm{d}x = \int_{\mathbb{R}^d} |\varphi(y)|^p \, \mathbb{E}\Big(1_{B_R}(X_t^{-1}(y)) \det(\nabla X_t^{-1}(y))\Big)\mathrm{d}y.$$

Noticing that

$$\nabla X_t^{-1}(y) = (\nabla X_t)^{-1} \circ X_t^{-1}(y) = K_t \circ X_t^{-1}(y),$$

we have

$$\|\mathcal{T}_t^0\varphi\|_{p;R} \leqslant \mathbb{E}\left(\sup_{|x|<R} \det(K_t(x))\right)\|\varphi\|_p.$$

On the other hand, from (4.1) and (4.3), it is by now standard to show that for any $p \geqslant 2$,

$$\mathbb{E}|K_t(x) - K_t(y)|^p \leqslant C|x-y|^p, \quad x, y \in \mathbb{R}^d,$$

which implies that

$$\mathbb{E}|\det K_t(x) - \det K_t(y)|^p \leqslant C|x-y|^p, \quad x, y \in \mathbb{R}^d.$$

Hence, by Kolmogorov's theorem,

$$\mathbb{E}\left(\sup_{|x|<R} \det K_t(x)\right) \leqslant C_R.$$

Thus (5.1) holds for $k = m = 0$.

(ii) Next for $k \in \mathbb{N}$ and $m \in \mathbb{N}_0$, by the chain rule, we have

$$\nabla^k \mathcal{T}_t^0 \nabla^m \varphi(x) = \nabla^k \mathbb{E}\Big((\nabla^m\varphi)(X_t(x))\Big) = \sum_{j=0}^k \mathbb{E}\Big((\nabla^{m+j}\varphi)(X_t^x)G_j\Big), \tag{5.2}$$

where $\{G_j, j = 0, \cdots, k\}$ are real polynomial functions of $\nabla X_t(x), \cdots, \nabla^k X_t(x)$. Hence, by Hölder's inequality,

$$\|\nabla^k \mathcal{T}_t^0 \nabla^m \varphi\|_{p;R}^p \leqslant \sum_{j=0}^k \left(\int_{B_R} \mathbb{E}|(\nabla^{m+j}\varphi)(X_t(x))|^p \mathrm{d}x\right) \sup_{x\in B_R} \left(\mathbb{E}|G_j|^{p/(p-1)}\right)^{p-1}. \tag{5.3}$$

By (4.2), it is now standard to show that for any $q \geqslant 1$ and $\ell \in \mathbb{N}$,

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}|\nabla^\ell X_t^x|^q < \infty. \tag{5.4}$$

Estimate (5.1) follows by (5.3), (5.4) and (i).

(iii) Under $(\mathbf{H}_1')$ and $(\mathbf{H}^{\rm o}_g)$, by [27, Lemma 4.4], we have $\|\mathcal{T}^0_t\varphi\|_p \leqslant C\|\varphi\|_p$, which together with (5.3) and (5.4) implies (5.1) with $R=\infty$. □

Now we use the Malliavin calculus introduced in Section 2 to show the following main result (see also [27]), which will automatically produce the conclusions in Theorem 1.1 by Sobolev's embedding theorem.

Theorem 5.1. *Fix $\ell \geqslant 2$. Under $(\mathbf{H}_\ell)$, $(\mathbf{H}^{\rm o}_g)$ and $(\mathbf{H}^{\rm str}_{\mathbf{or}})$, for any $k,m\in\mathbb{N}_0$ with $k+m=\ell-1$, there exists a $\gamma_{km}>0$ such that for all $R\geqslant 1$, $t\in(0,1)$, $p\in(1,\infty]$ and $\varphi\in C^\infty_c(\mathbb{R}^d)$,*

$$\|\nabla^k\mathcal{T}^0_t\nabla^m\varphi\|_{p;R} \leqslant C_{R,p}t^{-\gamma_{km}}\|\varphi\|_p. \tag{5.5}$$

Moreover, under $(\mathbf{H}_\ell')$, $(\mathbf{H}^{\rm o}_g)$ and $(\mathbf{H}^{\rm uni}_{\mathbf{or}})$, one can take $R=\infty$ in (5.5).

First of all, by the chain rule and Proposition 2.1, we have the following Malliavin differentiability of X_t in the sense of Theorem 2.1. Since the proof is completely the same as in [27], we omit the details.

Lemma 5.2. *Let $\Theta=(h,{\rm v})\in\mathbb{H}_{\infty-}\times\mathbb{V}_{\infty-}$. Under $(\mathbf{H}_1)$, for any $t\in[0,1]$, $X_t\in\mathbb{W}^{1,\infty-}_\Theta(\Omega)$ and*

$$\begin{aligned} D_\Theta X_t = &\int_0^t \nabla b(X_s)D_\Theta X_s{\rm d}s + \int_0^t \nabla\sigma_k(X_s)D_\Theta X_s{\rm d}W^k_s + \int_0^t \sigma_k(X_s)\dot h^k_s{\rm d}s\\ &+\int_0^t\!\!\int_{|z|<1}\nabla_x g(X_{s-},z)D_\Theta X_{s-}\widetilde N({\rm d}s,{\rm d}z) + \int_0^t\!\!\int_{|z|<1}\nabla_{\rm v} g(X_{s-},z)N({\rm d}s,{\rm d}z), \end{aligned}\tag{5.6}$$

where $\nabla_{\rm v}g(x,z):=\partial_{z_i}g(x,z){\rm v}_i(s,z)$. Moreover, for any $R,p\geqslant 2$, we have

$$\sup_{|x|<R}\mathbb{E}\left(\sup_{t\in[0,1]}|D_\Theta X_t(x)|^p\right)<\infty. \tag{5.7}$$

Under $(\mathbf{H}_1')$, the above estimate holds for $R=\infty$.

To use the integration by parts formula in Section 2, we need to introduce suitable Malliavin matrix. Let $J_t=J_t(x)=\nabla X_t(x)$ be the Jacobian matrix of $x\mapsto X_t(x)$, and $K_t(x)$ be the inverse of $J_t(x)$. Recalling (4.2) and (5.6), by the formula of constant variation, we have

$$D_\Theta X_t = J_t\int_0^t K_s\sigma_k(X_s)\dot h^k_s{\rm d}s + J_t\int_0^t\!\!\int_{|z|<1}K_s\nabla_{\rm v}g(X_{s-},z)N({\rm d}s,{\rm d}z). \tag{5.8}$$

Here the integral is the Lebesgue-Stieltjes integral.

Next we want to choose special directions $\Theta_j\in\mathbb{H}_{\infty-}\times\mathbb{V}_{\infty-}$, $j=1,\cdots,d$ so that the Malliavin matrix $\mathcal{M}^{ij}_t(x):=(D_{\Theta_j}X^i_t(x))_{i,j=1,\cdots,d}$ is invertible. Let

$$H(x;t):=\int_0^t\sigma^*(X_s(x))K^*_s(x){\rm d}s,$$

where the asterisk stands for the transpose of a matrix, and

$$U(x,z):=(\mathbb{I}+\nabla_x g(x,z))^{-1}\nabla_z g(x,z),\quad x\in\mathbb{R}^d,\quad |z|<1.$$

The following lemma is a direction consequence of the above definition and $(\mathbf{H}_\ell)$.

Lemma 5.3. *Under* $(\mathbf{H}_\ell)$*, for any* $k \in \mathbb{N}_0$ *and* $m = 0, \cdots, \ell - 1$*, there exists* $C > 0$ *such that for all* $x \in \mathbb{R}^d$ *and* $0 < |z| < 1$*,*

$$|\nabla_x^k \nabla_z^m U(x, z)| \leqslant C(1 + |x|)|z|^{-m}, \quad |U(x, z) - U(x, 0)| \leqslant C(1 + |x|)|z|^\beta, \tag{5.9}$$

where β *is the same as in* $(\mathbf{H}_\ell)$*.*

Below we fix $\ell \geqslant 2$ and assume $(\mathbf{H}_\ell)$+$(\mathbf{H}_g^{\text{o}})$. For $j = 1, \cdots, d$, define

$$h_j(x; t) := H(x; t)_{\cdot j}, \quad \mathrm{v}_j(x; s, z) := [K_{s-}(x)U(X_{s-}(x), z)]^*_{\cdot j}\zeta_{\ell,\delta}(z),$$

where $\zeta_{\ell,\delta}(z)$ is a nonnegative smooth function with

$$\zeta_{\ell,\delta}(z) = |z|^{1+\ell}, \quad |z| \leqslant \delta/4, \quad \zeta_{\ell,\delta}(z) = 0, \quad |z| > \delta/2.$$

Let

$$\Theta_j(x) := (h_j(x), \mathrm{v}_j(x)).$$

Noticing that by equation (4.3),

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

by (5.8) we have

$$\mathcal{M}_t^{ij}(x) := D_{\Theta_j} X_t^i(x) = (J_t(x)\Sigma_t(x))_{ij}, \tag{5.10}$$

where $\Sigma_t(x) = \Sigma_t^{(1)}(x) + \Sigma_t^{(2)}(x)$, and

$$\begin{aligned} \Sigma_t^{(1)}(x) &:= \int_0^t K_s(x)(\sigma\sigma^*)(X_s(x))K_s^*(x)\mathrm{d}s, \\ \Sigma_t^{(2)}(x) &:= \int_0^t\int_{|z|<1} K_{s-}(x)(UU^*)(X_{s-}(x), z)K_{s-}^*(x)\zeta_{\ell,\delta}(z)N(\mathrm{d}s, \mathrm{d}z). \end{aligned} \tag{5.11}$$

By Lemma 5.3 and cumbersome calculations (see [27]), we have

Lemma 5.4. *(i) For each* $j = 1, \cdots, d$ *and* $x \in \mathbb{R}^d$*,* $\Theta_j(x) \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$*.*
(ii) For any $R, p \geqslant 1$*,* $k \in \mathbb{N}_0$ *and* $m = 0, \cdots, \ell - 1$*, we have*

$$\sup_{|x|<R} \mathbb{E}\left(\sup_{t\in[0,1]} \left(|D_\Theta^m \nabla^k X_t(x)|^p + |D_\Theta^m \mathcal{M}_t(x)|^p\right)\right) < \infty, \tag{5.12}$$

$$\sup_{|x|<R} \mathbb{E}\left(\sup_{t\in[0,1]} |D_\Theta^m \mathrm{div}(\Theta_i(x))|^p\right) < \infty, \ i = 1, \cdots, d, \tag{5.13}$$

where $D_\Theta := (D_{\Theta_1}, \cdots, D_{\Theta_d})$*.*
(iii) Under $(\mathbf{H}'_\ell)$+$(\mathbf{H}_g^{\text{o}})$*, the* R *in* (5.12)-(5.13) *can be infinity.*

Now we can give

Proof of Theorem 5.1. We divide the proof into three steps. Below we fix $\ell \geqslant 2$ and $k, m \in \mathbb{N}_0$ so that $k + m = \ell - 1$.

(1) Let $\Sigma_t(x) = \Sigma_t^{(1)}(x) + \Sigma_t^{(2)}(x)$ be defined by (5.11). In view of $U(x, 0) = \nabla_z g(x, 0)$, by (5.9), Lemma 2.2 and Theorem 4.1, there are constants $C_3 \geqslant 1$, $c_3, \theta \in (0, 1)$ and $\gamma = \gamma(\alpha, j_0) \in (0, 1)$ such that for all $t \in (0, 1)$ and $R, \lambda, p \geqslant 1$,

$$\sup_{|x|<R} \sup_{|u|=1} \mathbb{E} \exp\{-\lambda u \Sigma_t(x) u^*\} \leqslant C_3 \exp\{-c_3 t \lambda^\gamma\} + C_{R,p}(\lambda^\theta t)^{-p}. \tag{5.14}$$

As in [26, Lemma 5.3], for any $R, p \geqslant 1$, there exist constants $C_{R,p} \geqslant 1$ and $\gamma' = \gamma'(\alpha, j_0, d) > 0$ such that for all $t \in (0,1)$,

$$\sup_{|x|<R} \mathbb{E}\Big((\det \Sigma_t(x))^{-p}\Big) \leqslant C_{R,p} t^{-\gamma' p}. \tag{5.15}$$

Since $\mathcal{M}_t^{-1}(x) = \Sigma_t^{-1}(x) K_t(x)$, by Lemma 4.2, (5.12) and (5.15), we obtain that for all $p \geqslant 1$,

$$\sup_{|x|<R} \|\mathcal{M}_t^{-1}(x)\|_{L^p(\Omega)} \leqslant C_{R,p} t^{-\gamma'}, \ t \in (0,1). \tag{5.16}$$

Under $(\mathbf{H}'_\ell)$+$(\mathbf{H}^{\mathrm{o}}_g)$+$(\mathbf{H}^{\mathrm{uni}}_{\mathbf{or}})$, by Remark 4.3 and (iii) of Lemma 5.4, the above R can be infinity.

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

$$(D^m_\Theta \mathrm{div}\Theta)_{m=0}^{\ell-2}, \mathcal{M}_t^{-1}, (D^m_\Theta \nabla^k X_t)_{k\in\mathbb{N}, m=0,\cdots,\ell}, (D^m_\Theta \mathcal{M}_t)_{m=1}^{\ell-1},$$

where the starting point x is dropped in the above random variables. By (5.16) and Lemma 5.4, for any $H_t(x) \in \mathscr{C}_t(x)$, there exists a $\gamma(H) > 0$ only depending on the degree of $\mathcal{M}_t^{-1}$ appearing in H and α, j_0, d such that for all $t \in (0,1)$ and $p \geqslant 1$,

$$\sup_{|x|<R} \|H_t(x)\|_{L^p(\Omega)} \leqslant C_{R,p} t^{-\gamma(H)}. \tag{5.17}$$

Under $(\mathbf{H}'_\ell)$+$(\mathbf{H}^{\mathrm{o}}_g)$+$(\mathbf{H}^{\mathrm{uni}}_{\mathbf{or}})$, by (5.16) and (iii) of Lemma 5.4, the above R can be infinity.

(3) Since the Malliavin matrix $\mathcal{M}_t = D_\Theta X_t$ is invertible, by the chain rule (2.6), we have

$$D_\Theta\Big(\varphi(X_t)\Big) \cdot \mathcal{M}_t^{-1} = (\nabla\varphi)(X_t).$$

For any $Z \in \mathscr{C}_t(x)$, by (5.10) and the integration by parts formula (2.4), we have

$$\mathbb{E}\Big((\nabla\varphi)(X_t) Z\Big) = \mathbb{E}\Big(D_\Theta(\varphi(X_t)) \cdot \mathcal{M}_t^{-1} Z\Big) = \mathbb{E}\Big(\varphi(X_t) Z'\Big),$$

where

$$Z' := \mathrm{div}\Theta \cdot \mathcal{M}_t^{-1} Z - D_\Theta(\mathcal{M}_t^{-1} Z) \in \mathscr{C}_t(x).$$

Starting from this formula, by (5.2) and induction, there exists $H \in \mathscr{C}_t(x)$ such that

$$\nabla^k \mathbb{E}\Big((\nabla^m\varphi)(X_t)\Big) = \mathbb{E}\Big(\varphi(X_t) H\Big).$$

Therefore, for any $p \in (1,\infty)$, by (5.17), (5.1) and Hölder's inequality, we have

$$\begin{aligned}
\|\nabla^k \mathcal{T}_t^0 \nabla^m \varphi\|_{p;R} &\leqslant \left(\int_{B_R} \Big|\mathbb{E}\Big(\varphi(X_t(x)) H(x)\Big)\Big|^p \mathrm{d}x\right)^{\frac{1}{p}} \\
&\leqslant \left(\int_{B_R} \mathbb{E}\Big(|\varphi|^p(X_t(x))\Big)\Big(\mathbb{E}|H(x)|^{\frac{p}{p-1}}\Big)^{p-1} \mathrm{d}x\right)^{\frac{1}{p}} \\
&\leqslant C_{R,p} t^{-\gamma(H)} \|\varphi\|_p, \quad t \in (0,1).
\end{aligned}$$

Under $(\mathbf{H}'_\ell)$+$(\mathbf{H}^{\mathrm{o}}_g)$+$(\mathbf{H}^{\mathrm{uni}}_{\mathbf{or}})$, by (5.17) and Lemma 5.1, the above R can be infinity. □

6. Proof of Theorem 1.3

Throughout this section we assume $(\mathbf{H}_1')$, $(\mathbf{H}^{\rm uni}_{\mathbf{or}})$ and $g \in C_b^\infty(\mathbb{R}^d \times B_1^c)$. Let $\chi : [0,\infty) \to [0,1]$ be a smooth function with $\chi(r) = 1$ for $r < 1$ and $\chi(r) = 0$ for $r > 2$. For $\delta > 0$, define

$$\chi_\delta(r) := \chi(r/\delta), \quad g_\delta(x,z) := g(x,z)\chi_\delta(z).$$

Choose δ be small enough so that g_δ satisfies $(\mathbf{H}^{\rm o}_g)$. Thus we can write

$$\mathscr{A}\varphi = \mathcal{L}_0\varphi + \mathscr{L}\varphi,$$

where

$$\mathcal{L}_0\varphi(x) := \frac{1}{2}A_k^2\varphi(x) + A_0\varphi(x) + \text{p.v.}\int_{\mathbb{R}^d_0}\Big(\varphi(x+g_\delta(x,z)) - \varphi(x)\Big)\frac{\mathrm{d}z}{|z|^{d+\alpha}},$$

and

$$\mathscr{L}\varphi(x) := \int_{\mathbb{R}^d_0}\Big(\varphi(x+g(x,z)) - \varphi(x+g_\delta(x,z))\Big)\frac{\mathrm{d}z}{|z|^{d+\alpha}}.$$

Let $(\mathcal{T}_t)_{t\geqslant 0}$ (resp. $(\mathcal{T}^0_t)_{t\geqslant 0}$) be the semigroup associated with $\mathscr{A}$ (resp. $\mathcal{L}_0$). Then we have

$$\partial_t\mathcal{T}_t\varphi = \mathscr{A}\mathcal{T}_t\varphi = \mathcal{L}_0\mathcal{T}_t\varphi + \mathscr{L}\mathcal{T}_t\varphi. \tag{6.1}$$

By Duhamel's formula, we have

$$\mathcal{T}_t\varphi = \mathcal{T}^0_t\varphi + \int_0^t \mathcal{T}^0_{t-s}\mathscr{L}\mathcal{T}_s\varphi\mathrm{d}s. \tag{6.2}$$

Notice that under $(\mathbf{H}_1')$ and $(\mathbf{H}^{\rm uni}_{\mathbf{or}})$, (5.1) and (5.5) hold for $R = \infty$.

For $\beta \geqslant 0$ 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. It is well known that for any $k \in \mathbb{N}$ and $p \in (1,\infty)$ (cf. [25]), an equivalent norm in $\mathbb{H}^{k,p}$ is given by

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

For a function $g : \mathbb{R}^d \to \mathbb{R}^d$, we introduce

$$T_g\varphi(x) := \varphi(x+g(x)) - \varphi(x).$$

Lemma 6.1. *Let $m \in \mathbb{N}_0$. Assume $g \in C_b^{m+1}$. For any $\theta \in (0,1)$ and $p > d/\theta$, there is a constant $C = C(p,\theta,m) > 0$ such that*

$$\|T_g\varphi\|_{m,p} \leqslant C\|\varphi\|_{m+\theta,p}\sum_{|\alpha|\leqslant m}\Big(\Pi_{j=1}^m\Big(1+\|\nabla^j g\|_\infty^{\alpha_j}\Big)\Big)\|g\|_\infty^\theta,$$

where $\alpha = (\alpha_1,\cdots,\alpha_m)$ and $|\alpha| = \alpha_1+\cdots+\alpha_m$. In particular, we have

$$\|T_g\varphi\|_{m+\beta,p} \leqslant C\|\varphi\|_{m+\beta+\theta,p}P(\|\nabla g\|_\infty,\cdots,\|\nabla^m g\|_\infty)\Big(1+\|\nabla^{m+1}g\|_\infty^{m+1}\Big)^\beta\|g\|_\infty^\theta,$$

where P is a polynomial function of its arguments.

Proof. Let $\theta \in (0,1)$. First of all, for $m = 0$, we have

$$|T_g\varphi(x)| \leqslant \sup_{y\neq 0}\frac{|\varphi(x+y)-\varphi(x)|}{|y|^\theta}|g(x)|^\theta.$$

Recalling that for $p > d/\theta$ (see [15, Lemma 5]),

$$\left\| \sup_{y \neq 0} \frac{|\varphi(\cdot + y) - \varphi(\cdot)|}{|y|^\theta} \right\|_p \leqslant C\|\varphi\|_{\theta,p},$$

we have

$$\|T_g\varphi\|_p \leqslant C\|\varphi\|_{\theta,p}\|g\|_\infty^\theta.$$

For $m \in \mathbb{N}$, by the chain rule and induction, there is a constant $C > 0$ such that for all $x \in \mathbb{R}^d$,

$$|\nabla^m T_g\varphi(x)| \leqslant C \sum_{|\alpha| \leqslant m} \left(\Pi_{j=1}^n \|\mathbb{I} + \nabla^j g\|_\infty^{\alpha_j}\right) \sum_{j=1}^m |\nabla^j \varphi|(x + g(x)).$$

As above one sees that for any $p > d/\theta$,

$$\|\nabla^m T_g\varphi\|_p \leqslant C \sum_{|\alpha| \leqslant m} \left(\Pi_{j=1}^n (1 + \|\nabla^j g\|_\infty^{\alpha_j})\right) \sum_{j=1}^m \|\nabla^j \varphi\|_{\theta,p}\|g\|_\infty^\theta.$$

The first estimate follows. As for the second estimate, it follows by interpolation. □

Corollary 6.1. *Let $\alpha \in (0, 2)$ and $\beta \in (0, \alpha)$. For $\theta \in (0, 1)$ with $\beta + \theta \in (0, \alpha)$ and $p > d/\theta$, there is a constant $C > 0$ such that for all $\varphi \in \mathbb{H}^{\beta+\theta,p}$,*

$$\|\mathscr{L}\varphi\|_{\beta,p} \leqslant C\|\varphi\|_{\beta+\theta,p}. \tag{6.3}$$

Moreover, if the support of $g(x, \cdot)$ is contained in a ball B_R for all $x \in \mathbb{R}^d$, then the above estimate holds for all $\beta \geqslant 0$.

Proof. Notice that for any $j = 0, 1, \cdots$,

$$\|\nabla_x^j g(\cdot, z)\|_\infty \leqslant C|z|, \quad z \in \mathbb{R}^d,$$

and

$$\mathscr{L}\varphi(x) = \int_{|z|>\delta} \Big(\varphi(x + g(x, z)) - \varphi(x + g_\delta(x, z))\Big)\frac{\mathrm{d}z}{|z|^{d+\alpha}}.$$

By Lemma 6.1 we have

$$\begin{aligned} \|\mathscr{L}\varphi\|_{\beta,p} &\leqslant \int_{|z|>\delta} \Big(\|T_{g(\cdot,z)}\varphi\|_{\beta,p} + \|T_{g_\delta(\cdot,z)}\varphi\|_{\beta,p}\Big)\frac{\mathrm{d}z}{|z|^{d+\alpha}} \\ &\leqslant C\|\varphi\|_{\beta+\theta,p} \int_{|z|>\delta} \frac{|z|^{\beta+\theta}}{|z|^{d+\alpha}}\mathrm{d}z \leqslant C\|\varphi\|_{\beta+\theta,p}, \end{aligned}$$

where the last step is due to $\beta + \theta < \alpha$. If the support of $g(x, \cdot)$ is contained in a ball B_R for all $x \in \mathbb{R}^d$, then

$$\mathscr{L}\varphi(x) = \int_{\delta<|z|\leqslant R} \Big(\varphi(x + g(x, z)) - \varphi(x + g_\delta(x, z))\Big)\frac{\mathrm{d}z}{|z|^{d+\alpha}}.$$

As above, (6.3) is direct by Lemma 6.1. □

The following results are proven in [27, Lemmas 5.2 and 5.3].

Lemma 6.2. *Let γ_{10} and γ_{01} be the same as in Theorem 5.1. Under $(\mathbf{H}_1')$ and $(\mathbf{H}^{\rm uni}_{\mathbf{or}})$, for any $p \in (1,\infty)$, $\theta, \vartheta \in [0,1)$ and $\beta \geqslant 0$, there exists a constant $C > 0$ such that for all $\varphi \in C_0^\infty(\mathbb{R}^d)$ and $t \in (0,1)$,*

$$\|\mathcal{T}^0_t \varphi\|_{\theta+\beta,p} \leqslant C t^{-\theta\gamma_{10}} \|\varphi\|_{\beta,p}, \tag{6.4}$$

$$\|\mathcal{T}^0_t \Delta^{\frac{\vartheta}{2}} \varphi\|_p \leqslant C t^{-\vartheta\gamma_{01}} \|\varphi\|_p. \tag{6.5}$$

Now we can show the following key estimate.

Lemma 6.3. *Let γ_{10}, γ_{01} be as in Theorem 5.1. Fix $\vartheta \in [0, \frac{1}{\gamma_{01}} \wedge 1)$. Under $(\mathbf{H}_1')$, $(\mathbf{H}^{\rm uni}_{\mathbf{or}})$ and $g \in C_b^\infty(\mathbb{R}^d \times B_1^c)$, there exist $\beta > \alpha$, $p_0 \geqslant 1$ and constant $C > 0$ such that for all $t \in (0,1)$, $p > p_0$ and $\varphi \in C_0^\infty(\mathbb{R}^d)$,*

$$\|\mathcal{T}_t \Delta^{\frac{\vartheta}{2}} \varphi\|_{\beta,p} \leqslant C t^{-\beta\gamma_{10} - \vartheta\gamma_{01}} \|\varphi\|_p. \tag{6.6}$$

Moreover, if the support of $g(x,\cdot)$ is contained in a ball B_R for all $x \in \mathbb{R}^d$, then the β in (6.6) can be any positive number.

Proof. Choose $\theta \in (0, \frac{1}{2\gamma_{10}} \wedge 1 \wedge \alpha)$ and $M \in \mathbb{N}$ such that

$$M\theta < \alpha, \quad \beta := (M+1)\theta > \alpha.$$

Let $p > d/\theta$. For $m = 1, \cdots, M$, we have

$$\begin{aligned}
\|\mathcal{T}_t \varphi\|_{(m+1)\theta,p} &\leqslant \|\mathcal{T}^0_t \varphi\|_{(m+1)\theta,p} + \int_0^t \|\mathcal{T}^0_{t-s} \mathscr{L} \mathcal{T}_s \varphi\|_{(m+1)\theta,p} \mathrm{d}s \\
&\overset{(6.4)}{\leqslant} C t^{-\theta\gamma_{10}} \|\varphi\|_{m\theta,p} + C \int_0^t (t-s)^{-2\theta\gamma_{10}} \|\mathscr{L} \mathcal{T}_s \varphi\|_{(m-1)\theta,p} \mathrm{d}s \\
&\overset{(6.3)}{\leqslant} C t^{-\theta\gamma_{10}} \|\varphi\|_{m\theta,p} + C \int_0^t (t-s)^{-2\theta\gamma_{10}} \|\mathcal{T}_s \varphi\|_{m\theta,p} \mathrm{d}s \\
&\leqslant C t^{-\theta\gamma_{10}} \|\varphi\|_{m\theta,p} + C \int_0^t (t-s)^{-2\theta\gamma_{10}} \|\mathcal{T}_s \varphi\|_{(m+1)\theta,p} \mathrm{d}s,
\end{aligned}$$

which, by Gronwall's inequality of Volterra type (cf. [28, Lemma 2.2]), yields that for all $t \in (0,1)$,

$$\|\mathcal{T}_t \varphi\|_{(m+1)\theta,p} \leqslant C t^{-\theta\gamma_{10}} \|\varphi\|_{m\theta,p}. \tag{6.7}$$

Thus, by the semigroup property of $\mathcal{T}_t$ and iteration, we obtain that for $m = 1, \cdots, M$,

$$\|\mathcal{T}_{(m+1)t} \varphi\|_{(m+1)\theta,p} \leqslant C t^{-\theta\gamma_{10}} \|\mathcal{T}_{mt} \varphi\|_{m\theta,p} \leqslant \cdots \leqslant C t^{-(m+1)\theta\gamma_{10}} \|\mathcal{T}_t \varphi\|_p. \tag{6.8}$$

On the other hand, by (6.2), (6.3) and (6.5), we have

$$\begin{aligned}
\|\mathcal{T}_t \Delta^{\frac{\vartheta}{2}} \varphi\|_p &\leqslant \|\mathcal{T}^0_t \Delta^{\frac{\vartheta}{2}} \varphi\|_p + \int_0^t \|\mathcal{T}^0_{t-s} \mathscr{L} \mathcal{T}_s \Delta^{\frac{\vartheta}{2}} \varphi\|_p \mathrm{d}s \\
&\leqslant C t^{-\vartheta\gamma_{01}} \|\varphi\|_p + C \int_0^t \|\mathcal{T}_s \Delta^{\frac{\vartheta}{2}} \varphi\|_{\theta,p} \mathrm{d}s \\
&\leqslant C t^{-\vartheta\gamma_{01}} \|\varphi\|_p + C \int_0^{2t} \|\mathcal{T}_s \Delta^{\frac{\vartheta}{2}} \varphi\|_{\theta,p} \mathrm{d}s \\
&= C t^{-\vartheta\gamma_{01}} \|\varphi\|_p + 2C \int_0^t \|\mathcal{T}_{2s} \Delta^{\frac{\vartheta}{2}} \varphi\|_{\theta,p} \mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&= Ct^{-\vartheta\gamma_{01}}\|\varphi\|_p + 2C\int_0^t \|\mathcal{T}_s\mathcal{T}_s\Delta^{\frac{\vartheta}{2}}\varphi\|_{\theta,p}\mathrm{d}s\\
&\overset{(6.7)}{\leqslant} Ct^{-\vartheta\gamma_{01}}\|\varphi\|_p + C\int_0^t s^{-\theta\gamma_{10}}\|\mathcal{T}_s\Delta^{\frac{\vartheta}{2}}\varphi\|_p\mathrm{d}s,
\end{aligned}
$$

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

$$
\|\mathcal{T}_t\Delta^{\frac{\vartheta}{2}}\varphi\|_p \leqslant Ct^{-\vartheta\gamma_{01}}\|\varphi\|_p. \tag{6.9}
$$

Combining (6.8) with (6.9), we obtain (6.6).

Finally, if the support of $g(x,\cdot)$ is contained in a ball B_R for all $x \in \mathbb{R}^d$, then the m in (6.8) can be any natural number. □

Now we can give

Proof of Theorem 1.3. Without loss of generality, we assume $t \in (0,1)$. Let $\varepsilon \in (0, \beta-\alpha)$ and $p > d/\varepsilon \vee p_0$, where p_0 appears in Lemma 6.3. For any $\varphi \in L^p(\mathbb{R}^d)$, by (6.6) and Sobolev's embedding theorem, we have $(\mathbb{I}-\Delta)^{\frac{\alpha+\varepsilon}{2}}\mathcal{T}_t\varphi \in C_b(\mathbb{R}^d)$ and for any $t \in (0,1)$ and $\vartheta \in [0, \frac{1}{\gamma_{01}} \wedge 1)$,

$$
\|(\mathbb{I}-\Delta)^{\frac{\alpha+\varepsilon}{2}}\mathcal{T}_t\Delta^{\frac{\vartheta}{2}}\varphi\|_\infty \leqslant C\|\mathcal{T}_t\Delta^{\frac{\vartheta}{2}}\varphi\|_{\beta,p} \leqslant Ct^{-\beta\gamma_{10}-\vartheta\gamma_{01}}\|\varphi\|_p. \tag{6.10}
$$

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

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

Moreover, we also have

$$
\begin{aligned}
\sup_{x\in\mathbb{R}^d}\|(\mathbb{I}-\Delta)_x^{\frac{\alpha+\varepsilon}{2}}\Delta_y^{\frac{\vartheta}{2}}\rho_t(x,\cdot)\|_{\frac{p}{p-1}} &= \sup_{x\in\mathbb{R}^d}\sup_{\varphi\in C_0^\infty(\mathbb{R}^d),\|\varphi\|_p\leqslant 1}\left|\int \varphi(y)(\mathbb{I}-\Delta)_x^{\frac{\alpha+\varepsilon}{2}}\Delta_y^{\frac{\vartheta}{2}}\rho_t(x,y)\mathrm{d}y\right|\\
&= \sup_{x\in\mathbb{R}^d}\sup_{\varphi\in C_0^\infty(\mathbb{R}^d),\|\varphi\|_p\leqslant 1}\left|\int \Delta_y^{\frac{\vartheta}{2}}\varphi(y)(\mathbb{I}-\Delta)_x^{\frac{\alpha+\varepsilon}{2}}\rho_t(x,y)\mathrm{d}y\right|\\
&= \sup_{x\in\mathbb{R}^d}\sup_{\varphi\in C_0^\infty(\mathbb{R}^d),\|\varphi\|_p\leqslant 1}|(\mathbb{I}-\Delta)_x^{\frac{\alpha+\varepsilon}{2}}\mathcal{T}_t\Delta^{\frac{\vartheta}{2}}\varphi(x)|\\
&= \sup_{\varphi\in C_0^\infty(\mathbb{R}^d),\|\varphi\|_p\leqslant 1}\|(\mathbb{I}-\Delta)_x^{\frac{\alpha+\varepsilon}{2}}\mathcal{T}_t\Delta^{\frac{\vartheta}{2}}\varphi\|_\infty \leqslant Ct^{-\beta\gamma_{10}-\vartheta\gamma_{01}}.
\end{aligned}
$$

Thus we obtain (ii). As for (iii), it follows by (6.6) for all $\beta \geqslant 0$. □

7. Application to nonlocal kinetic operators

Consider the following nonlocal kinetic operator:

$$
\mathscr{K}u(x,\mathrm{v}) = \mathscr{L}u(x,\mathrm{v}) + \mathrm{v}\cdot\nabla_x u(x,\mathrm{v}) + b(x,\mathrm{v})\cdot\nabla_{\mathrm{v}}u(x,\mathrm{v}),
$$

where

$$
\mathscr{L}u(x,\mathrm{v}) := \mathrm{p.v}\int_{\mathbb{R}^d}(u(x,\mathrm{v}+w) - u(x,\mathrm{v}))\frac{\kappa(x,\mathrm{v},w)}{|w|^{d+\alpha}}\mathrm{d}w.
$$

Here κ and b satisfy the following assumptions:

(A) $b \in C_b^\infty(\mathbb{R}^{2d})$, and $\kappa \in C_b^\infty(\mathbb{R}^{3d})$ satisfies that for some $\kappa_0 > 1$,

$$\kappa_0^{-1} \leqslant \kappa(x, \mathrm{v}, w) \leqslant \kappa_0, \quad \kappa(x, \mathrm{v}, -w) = \kappa(x, \mathrm{v}, w).$$

Moreover, for any $i, j \in \mathbb{N}_0$, there is a constant $C_{ij} > 0$ such that

$$|\nabla_x^{i+1} \nabla_{\mathrm{v}}^j \kappa(x, \mathrm{v}, w)| \leqslant C_{ij}/(1 + |\mathrm{v}|^2).$$

The aim of this section is to use Theorem 1.2 to show the following result.

Theorem 7.1. *Under* **(A)**, *there is a nonnegative continuous function $\rho_t(x, \mathrm{v}, y, w)$ on $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$ such that for each t, y, w, the map $(x, \mathrm{v}) \mapsto \rho_t(x, \mathrm{v}, y, w)$ belongs to $C_b^\infty(\mathbb{R}^{2d})$, and for each $t > 0$ and $y, w \in \mathbb{R}^d$,*

$$\partial_t \rho_t(x, \mathrm{v}, y, w) = \mathscr{K} \rho_t(\cdot, \cdot, y, w)(x, \mathrm{v}).$$

The key point for us is the following transform lemma.

Lemma 7.1. *Given $R \in (0, \infty)$ and $\kappa_0 \in [1, \infty)$, let $\kappa(z) : B_R \to [\kappa_0^{-1}, \kappa_0]$ be a measurable function, where $B_R := \{x \in \mathbb{R}^d : |x| < R\}$. For any $\alpha \in (0, 2)$, there is a homeomorphism $\Phi : B_R \to B_R$ such that for any nonnegative measurable function f,*

$$\int_{B_R} f \circ \Phi(z) \frac{\mathrm{d}z}{|z|^{d+\alpha}} = \int_{B_R} f(z) \frac{\kappa(z)}{|z|^{d+\alpha}} \mathrm{d}z. \tag{7.1}$$

Moreover, we have the following properties about Φ:

(i) $\Phi(0) = 0$ *and if* $\kappa(-z) = \kappa(z)$, *then* $\Phi(-z) = -\Phi(z)$.

(ii) If κ is continuous at 0, *then Φ is differentiable at point zero and*

$$\nabla \Phi(0) = \kappa(0)^{1/\alpha} \mathbb{I}.$$

(iii) If $\kappa \in C^1(B_R)$, then for some $C = C(\kappa_0, \|\nabla \kappa\|_\infty, \alpha, R) > 0$ and any $z \in B_R$,

$$|\nabla \Phi(z) - \nabla \Phi(0)| \leqslant \begin{cases} C|z|^\alpha, & \alpha \in (0, 1), \\ C|z| \log^+ |z|, & \alpha = 1, \\ C|z|, & \alpha \in (1, 2), \end{cases}$$

where $\log^+ |z| = \max\{-\log |z|, 1\}$.

(iv) If $\kappa \in C^j(B_R)$ for some $j \in \mathbb{N}$, then for some $C_j = C_j(R)$ and any $z \in B_R$,

$$|\nabla^j \Phi(z)| \leqslant C_j |z|^{1-j}.$$

Proof. Using the spherical coordinate transform, (7.1) is equivalent to

$$\int_0^R \int_{\mathbb{S}^{d-1}} f \circ \Phi(t\omega) \mathrm{d}\omega \frac{\mathrm{d}t}{|t|^{1+\alpha}} = \int_0^R \int_{\mathbb{S}^{d-1}} f(t\omega) \kappa(t\omega) \mathrm{d}\omega \frac{\mathrm{d}t}{|t|^{1+\alpha}}, \tag{7.2}$$

where $\mathbb{S}^{d-1} := \{\omega : |\omega| = 1\}$ is the unit sphere in $\mathbb{R}^d$. Given $\omega \in \mathbb{S}^{d-1}$, let $\phi(\cdot, \omega)$ and $\psi(\cdot, \omega)$ be defined by the following identity respectively:

$$\int_{\phi(r,\omega)}^R \frac{\mathrm{d}t}{t^{1+\alpha}} = \int_r^R \frac{\kappa(t\omega) \mathrm{d}t}{t^{1+\alpha}}, \quad \int_r^R \frac{\mathrm{d}t}{t^{1+\alpha}} = \int_{\psi(r,\omega)}^R \frac{\kappa(t\omega) \mathrm{d}t}{t^{1+\alpha}}. \tag{7.3}$$

Since $\kappa_0^{-1} \leqslant \kappa(z) \leqslant \kappa_0$, it is easy to see that $\phi(\cdot, \omega), \psi(\cdot, \omega) : [0, R] \to [0, R]$ are strictly increasing continuous functions and have the following properties:

$$\phi(0, \omega) = \psi(0, \omega) = 0, \ \phi(R, \omega) = \psi(R, \omega) = R$$

and

$$\phi(\psi(r,\omega),\omega) = \psi(\phi(r,\omega),\omega) = r. \tag{7.4}$$

Moreover,

$$\kappa_0^{-1/\alpha} r \leqslant \phi(r,\omega) \leqslant \kappa_0^{1/\alpha} r. \tag{7.5}$$

Indeed, by (7.3) and $\kappa_0^{-1} \leqslant \kappa(t\omega) \leqslant \kappa_0$ with $\kappa_0 \geqslant 1$, we have

$$\phi(r,\omega)^{-\alpha} \leqslant R^{-\alpha} + \kappa_0(r^{-\alpha} - R^{-\alpha}) \leqslant \kappa_0 r^{-\alpha} \Rightarrow \phi(r,\omega) \geqslant \kappa_0^{-1/\alpha} r$$

and

$$\phi(r,\omega)^{-\alpha} \geqslant R^{-\alpha} + \kappa_0^{-1}(r^{-\alpha} - R^{-\alpha}) \geqslant \kappa_0^{-1} r^{-\alpha} \Rightarrow \phi(r,\omega) \leqslant \kappa_0^{1/\alpha} r.$$

In particular, by a standard monotone class argument, it holds that for all nonnegative measurable function $g : [0,R) \to [0,\infty)$,

$$\int_0^R g(\psi(t,\omega)) \frac{\mathrm{d}t}{t^{1+\alpha}} = \int_0^R g(t) \frac{\kappa(t\omega)}{t^{1+\alpha}} \mathrm{d}t. \tag{7.6}$$

Now let us define

$$a(z) := \psi(|z|, z/|z|)/|z|, \quad \Phi(z) := a(z) \cdot z.$$

By (7.6), one sees that (7.2) holds for the above Φ, and by (7.4) and (7.5),

$$\kappa_0^{-1/\alpha} \leqslant a(z) \leqslant \kappa_0^{1/\alpha}, \quad \forall z \in B_R. \tag{7.7}$$

(i) From the above construction of Φ, it is easy to see that $\Phi(0) = 0$ and

$$\kappa(-z) = \kappa(z) \Rightarrow \Phi(-z) = -\Phi(z).$$

(ii) We assume $\kappa(z)$ is continuous at 0. We have the following claim:

$$a(0) := \lim_{z \to 0} a(z) = \kappa(0)^{1/\alpha}, \ \nabla\Phi(0) = a(0)\mathbb{I}. \tag{7.8}$$

For $i = 1, \cdots, d$, let $\mathrm{e}_i = (0, \cdots, 1, \cdots, 0)$. Noticing that $\Phi(0) = 0$, we have

$$\partial_i \Phi^j(0) = \lim_{\varepsilon \to 0} \Phi^j(\varepsilon \mathrm{e}_i)/\varepsilon = 1_{i=j} \lim_{\varepsilon \to 0} a(\varepsilon \mathrm{e}_i).$$

To prove (7.8), it suffices to show that

$$0 = \lim_{z \to 0} |a(z)^\alpha - \kappa(0)| = \lim_{r \to 0} \sup_{|\omega|=1} |(\psi(r,\omega)/r)^\alpha - \kappa(0)|. \tag{7.9}$$

From (7.3), one sees that

$$\phi(r,\omega) = \left[\alpha \int_r^R \frac{\kappa(t\omega)}{t^{1+\alpha}} \mathrm{d}t + R^{-\alpha}\right]^{-1/\alpha}, \tag{7.10}$$

which implies that

$$(r/\phi(r,\omega))^\alpha = \alpha r^\alpha \int_r^R \frac{\kappa(t\omega)}{t^{1+\alpha}} \mathrm{d}t + r^\alpha R^{-\alpha}.$$

Hence, by (7.4),

$$\begin{aligned}
\lim_{r \to 0} \sup_{|\omega|=1} |\,(\psi(r,\omega)/r)^\alpha - \kappa(0)| &= \lim_{r \to 0} \sup_{|\omega|=1} |\,(r/\phi(r,\omega))^\alpha - \kappa(0)| \\
&= \lim_{r \to 0} \sup_{|\omega|=1} \left(\alpha r^\alpha \int_r^R \frac{|\kappa(t\omega) - \kappa(0)|}{t^{1+\alpha}} \mathrm{d}t\right)
\end{aligned}$$

$$= \lim_{r\to 0}\sup_{|\omega|=1}\int_{(r/R)^\alpha}^1 |\kappa(r\omega/s^{1/\alpha}) - \kappa(0)|\mathrm{d}s = 0,$$

where the last step is due to the dominated convergence theorem. Thus we get (7.9).

(iii) By definition (7.3) and the change of variable $s = t/|z|$, we get

$$\int_{a(z)|z|}^R \frac{\kappa(tz/|z|)\mathrm{d}t}{t^{1+\alpha}} = \int_{|z|}^R \frac{\mathrm{d}t}{t^{1+\alpha}} \Rightarrow \int_{a(z)}^{R/|z|} \frac{\kappa(sz)\mathrm{d}s}{s^{1+\alpha}} = \frac{1}{\alpha}\left(1 - \frac{|z|^\alpha}{R^\alpha}\right).$$

Assume $\kappa \in C^1(B_R)$. Taking the gradient for both sides with respect to z, we obtain

$$\nabla a(z) = \frac{a(z)^{1+\alpha}}{\kappa(a(z)z)}\left(\int_{a(z)}^{R/|z|} \frac{\nabla\kappa(sz)}{s^\alpha}\mathrm{d}s + \frac{|z|^{\alpha-2}z}{R^\alpha}(1 - \kappa(Rz/|z|))\right).$$

Since $\kappa \in [\kappa_0^{-1}, \kappa_0]$, by (7.7) and elementary calculations, we have

$$|\nabla a(z)| \leqslant \begin{cases} \kappa_0^{2+1/\alpha}\left(\|\nabla\kappa\|_\infty R^{1-\alpha}/(1-\alpha) + (1+\kappa_0)R^{-\alpha}\right)|z|^{\alpha-1}, & \alpha \in (0,1), \\ \kappa_0^3\left(\|\nabla\kappa\|_\infty \log(R\kappa_0/|z|) + (1+\kappa_0)R^{-1}\right), & \alpha = 1, \\ \kappa_0^3\|\nabla\kappa\|_\infty/(\alpha-1) + \kappa_0^{2+1/\alpha}(1+\kappa_0)R^{-1}, & \alpha \in (1,2). \end{cases} \tag{7.11}$$

Finally, noticing that

$$\partial_i\Phi^j(z) = z_j\partial_i a(z) + 1_{i=j}a(z), \quad \partial_i\Phi^j(0) = \kappa(0)^{1/\alpha}1_{i=j} = 1_{i=j}a(0),$$

we have

$$|\partial_i\Phi^j(z) - \partial_i\Phi^j(0)| \leqslant |z_j|\cdot|\partial_i a(z)| + 1_{i=j}|a(z) - a(0)|,$$

which together with (7.11) gives the desired estimate.

(iv) Let Φ^{-1} be the inverse of Φ. We have

$$(\nabla\Phi)\circ\Phi^{-1}\cdot\nabla\Phi^{-1} = \mathbb{I} \Rightarrow \nabla\Phi = (\nabla\Phi^{-1})^{-1}\circ\Phi.$$

Therefore,

$$\nabla\Phi^{-1}\cdot(\nabla^2\Phi)\circ\Phi^{-1}\cdot\nabla\Phi^{-1} + (\nabla\Phi)\circ\Phi^{-1}\cdot\nabla^2\Phi^{-1} = 0,$$

and

$$|\nabla^2\Phi| \leqslant |\nabla\Phi|^3\cdot|\nabla^2\Phi^{-1}\circ\Phi|.$$

By induction method, there is a $C = C(j,d) > 0$ such that

$$|\nabla^j\Phi| \leqslant C\sum_{\gamma\in\mathscr{A}}|\nabla\Phi|^{1+\sum_{i=2}^j i\gamma_i}\prod_{i=2}^j\left|(\nabla^i\Phi^{-1})\circ\Phi\right|^{\gamma_i}, \tag{7.12}$$

where

$$\mathscr{A} := \left\{\gamma = (\gamma_2,\cdots,\gamma_j) : \sum_{i=2}^j (i-1)\gamma_i = j-1\right\}.$$

By (ii), one sees that

$$\|\nabla\Phi\|_\infty \leqslant C_R. \tag{7.13}$$

On the other hand, by definition, it is easy to see that

$$\Phi^{-1}(z) = \phi(|z|, z/|z|)z/|z| =: h(z)g(z),$$

where $g(z) := z/|z|$ and by (7.10),

$$h(z) := \phi(|z|, z/|z|) = \left[\alpha \int_{|z|}^{R} \frac{\kappa(tz/|z|)}{t^{1+\alpha}} \mathrm{d}t + R^{-\alpha}\right]^{-1/\alpha}.$$

By elementary calculations, one finds that

$$|\nabla^j h(z)| \leqslant C|z|^{1-j}, \quad |\nabla^j g(z)| \leqslant C|z|^{-j},$$

and so,

$$|\nabla^j \Phi^{-1}(z)| \leqslant C|z|^{1-j}. \tag{7.14}$$

Substituting (7.14) and (7.13) into (7.12), and by (7.7) and $\Phi(z) = a(z) \cdot z$, we obtain (iii). □

Corollary 7.2. *Given $R \in (0, \infty)$, $\kappa_0 \in [1, \infty)$ and $d_0 \in \mathbb{N}$, let $\kappa(x, z) : \mathbb{R}^{d_0} \times B_R \to [\kappa_0^{-1}, \kappa_0]$ be a measurable function. For any $\alpha \in (0, 2)$, there is a map $\Phi(x, z) : \mathbb{R}^{d_0} \times B_R \to B_R$ such that for any nonnegative measurable function f,*

$$\int_{B_R} f \circ \Phi(x, z) \frac{\mathrm{d}z}{|z|^{d+\alpha}} = \int_{B_R} f(z) \frac{\kappa(x, z)}{|z|^{d+\alpha}} \mathrm{d}z.$$

Moreover, Φ enjoys the following properties:

(i) $\Phi(x, 0) = 0$ and if $\kappa(x, -z) = \kappa(x, z)$, then $\Phi(x, -z) = -\Phi(x, z)$.

(ii) For $x \in \mathbb{R}^{d_0}$, if $\kappa(x, \cdot)$ is continuous at 0, then $\Phi(x, \cdot)$ is differentiable at point zero and

$$\nabla_z \Phi(x, 0) = \kappa(x, 0)\mathbb{I}.$$

(iii) If $\kappa(x, \cdot) \in C^1(B_R)$ and $\|\nabla_z \kappa\|_\infty < \infty$, then there are $\beta \in (0, 1)$ and $C > 0$ such that for all $x \in \mathbb{R}^{d_0}$ and $z \in B_R$,

$$|\nabla_z \Phi(x, z) - \nabla_z \Phi(x, 0)| \leqslant C|z|^\beta.$$

(iv) If $\kappa \in C_b^\infty(\mathbb{R}^{d_0} \times B_R)$, then for all $i, j \in \mathbb{N}_0$, there is a $C_{ij} > 0$ such that for all $x \in \mathbb{R}^{d_0}$ and $z \in B_R$,

$$|\nabla_x^i \nabla_z^j \Phi(x, z)| \leqslant C_{ij}|z|^{1-j},$$

where C_{ij} is a polynomial of $\|\nabla_x^m \nabla_z^n \kappa\|_\infty$, $m = 1, \cdots, i, n = 0, \cdots, j$.

Proof. (i), (ii) and (iii) follow by (i), (ii) and (iii) of Lemma 7.1. As for (iv), it follows by similar calculations as in the proof of (iv) of Lemma 7.1. □

Now we can give the proof of Theorem 7.1.

Proof of Theorem 7.1. Fix $\delta \in (0, 1)$ being small. Define

$$\mathscr{L}_0 u(x, \mathrm{v}) := \mathrm{p.v} \int_{|w|<\delta} (u(x, \mathrm{v} + w) - u(x, \mathrm{v})) \frac{\kappa(x, \mathrm{v}, w)}{|w|^{d+\alpha}} \mathrm{d}w,$$

and

$$\mathscr{K}_0 u(x, \mathrm{v}) := \mathscr{L}_0 u(x, \mathrm{v}) + \mathrm{v} \cdot \nabla_x u(x, \mathrm{v}) + b(x, \mathrm{v}) \cdot \nabla_\mathrm{v} u(x, \mathrm{v}).$$

Then we can write

$$\mathscr{K} u(x, \mathrm{v}) = \mathscr{K}_0 u(x, \mathrm{v}) + \mathscr{L}_1 u(x, \mathrm{v}),$$

where

$$\mathscr{L}_1 u(x, \mathrm{v}) := \int_{|w|\geqslant\delta} (u(x, \mathrm{v} + w) - u(x, \mathrm{v})) \frac{\kappa(x, \mathrm{v}, w)}{|w|^{d+\alpha}} \mathrm{d}w.$$

For each m and $p \geqslant 1$, by the chain rule, it is easy to see that

$$\|\mathscr{L}_1 u\|_{m,p} \leqslant C\|u\|_{m,p}.$$

On the other hand, by Corollary 7.2, there exists a function $g(x, \mathrm{v}, \cdot) : B_\delta \to B_\delta$ so that

$$\mathscr{L}_0 u(x, \mathrm{v}) = \mathrm{p.v}\int_{|w|<\delta} (u(x, \mathrm{v} + g(x, \mathrm{v}, w)) - u(x, \mathrm{v}))\frac{\mathrm{d}w}{|w|^{d+\alpha}},$$

and operator $\mathscr{K}_0$ satisfies $(\mathbf{H}_2')$+$(\mathbf{H}_g^{\mathrm{o}})$+$(\mathbf{H}_{\mathbf{or}}^{\mathrm{uni}})$. The desired result follows by Theorem 1.2. □

References

[1] Bally V. and Clément E.: Integration by parts formula and applications to equations with jumps. *Prob. Theory and Rela. Fields*, 151, no.3-4, 613-657(2011).

[2] Bichteler K., Gravereaux J.B. and Jacod J.: *Malliavin calculus for processes with jumps*. Gordan and Breach Science Publishers, 1987.

[3] Bismut J.M.: Calcul des variations stochastiques et processus de sauts. *Z. Wahrsch. Verw. Gebiete*, **63**, 147-235(1983).

[4] Cass T.: Smooth densities for stochastic differential equations with jumps. *Stoch. Proc. Appl.*, 119, no.5, 1416-1435(2009).

[5] Chen Z.Q. and Zhang X.: L^p-maximal hypoelliptic regularity of nonlocal kinetic Fokker-Planck operators. *J. Math. Pures Appl.* 116 (2018) 52-87.

[6] Fujiwara T. and Kunita H.: Stochastic differential equations of jump type and Lévy processes in diffeomorphism groups. *J. Math. Kryoto Univ.* 25, 71-106(1985).

[7] Harier M.: On Malliavin's proof of Hörmander's theorem. *Bull. Sci. Math.*, 135 (2011), pp. 650-666.

[8] Hörmander L.: Hypoelliptic second order differential equations. *Acta Math.* 119, 147-171(1967).

[9] Ishikawa Y. and Kunita H.: Malliavin calculus on the Wiener-Poisson space and its application to canonical SDE with jumps. *Stoch. Proc. Appl.*, 116, 1743-1769(2006).

[10] Ishikawa Y., Kunita H. and Tsuchiya M.: Smooth density and its short time estimate for jump process determined by SDE. *Stoch. Proc. Appl.*, 128, 3181-3219(2018).

[11] Komatsu T. and Takeuchi, A. (2001). On the smoothness of pdf of solutions to sde of jump type. *International Journal of Differential Equations and Applications*, 2(2), 141-197.

[12] Kunita H.: Nondegenerate sde's with jumps and their hypoelliptic properties. *J. Math. Soc. Japan*, 65(3), 993-1035(2013).

[13] Krylov N.: Hörmander's theoreom for parabolic equations with measurable in the time variable. *SIAM J. Math. Anal*, Vol. 46, No. 1, pp. 854-870(2014).

[14] Malliavin P.: Stochastic calculus of variations and hypoelliptic operators. In: Proc Inter. Symp. on Stoch. Diff. Equations, Kyoto, 195-263 (1976).

[15] Mikulevicius R. and Pragarauskas H.: On the Cauchy problem for certain integro- differential operators in Sobolev and Hölder spaces. *Lithuanian Math. Journal*, Vol.32, No.2, 1992.

[16] Norris J.: Simplified Malliavin calculus. In: Seminaire de Probabilités XX, *Lecture Notes in Math.* 1204, Springer, Berlin, 101-130(1986).

[17] Nourdin I., Simon T.: On the absolute continuity of Lévy processes with drift. *Ann. Probab.* 34(3), 1035-1051 (2006).

[18] Nualart D.: *The Malliavin calculus and related topics*. Springer-Verlag, New York, 2006.

[19] Picard J.: On the existence of smooth densities for jump processes. *Prob. Theory Rela. Fields*, Vol. 105, 481-511(1996).

[20] Protter P. E.: *Stochastic integration and differential equations*. Second Edition, Springer-Verlag, Berlin, (2004).

[21] Takeuchi A.: The Malliavin calculus for SDE with jumps and the partially hypoelliptic problem. *Osaka J. Math.* 39, 523-559(2002).

[22] Song Y. and Zhang X.: Regularity of density for sdes driven by degenerate lévy noises. *Electronic Journal of Probability*, 20 (2015), no. 21, 1-27.

[23] Villani C.: *A review of mathematical topics in collisional kinetic theory*. Handbook of Fluid Mechanics. Ed. S. Friedlander, D.Serre, 2002.
[24] Williams D.: To begin at the beginning. In Stochastic Integrals. *Lecture Notes in Math.* 851 (1981), 1-55.
[25] Stein E.M.: *Singular integrals and differentiability properties of functions.* Princeton University Press, 1970.
[26] Zhang X.: Fundamental solutions of nonlocal hörmander's operators. *Communications in Mathematics and Statistics*, 4(3), 1-44(2016).
[27] Zhang X.: Fundamental solutions of nonlocal hörmander's operators II. *Annals of Probability* , 45, 1799-1841(2017).
[28] Zhang X.: Stochastic Volterra equations in Banach spaces and stochastic partial differential equation. *J. Funct. Anal.*, 258, 1361-1425 (2010).

Zimo Hao: School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, P.R.China, Email: zimohao@whu.edu.cn

Xuhui Peng: School of Mathematics and Statistics, Hunan Normal University, Changsha, Hunan, P.R.China, Email: xhpeng@hunnu.edu.cn

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