NONLOCAL HÖRMANDER’S HYPOELLIPTICITY THEOREM

XICHENG ZHANG

Dedicated to the memory of Professor Paul Malliavin

Abstract. Consider the following nonlocal integro-differential operator: for \( \alpha \in (0, 2) \),
\[
L^{(\alpha)} f(x) := \text{p.v.} \int_{\mathbb{R}^d \setminus \{0\}} \frac{f(x + \sigma(x)z) - f(x)}{|z|^{d+\alpha}} \, dz + b(x) \cdot \nabla f(x),
\]
where \( \sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d \) and \( b : \mathbb{R}^d \to \mathbb{R}^d \) are two \( C^\infty \)-functions, and p.v. stands for the Cauchy principal value.

Let \( B_1(x) := \sigma(x) \) and \( B_{j+1}(x) := b(x) \cdot \nabla B_j(x) - \nabla b(x) \cdot B_j(x) \) for \( j \in \mathbb{N} \). Under the following uniform Hörmander’s type condition: for any \( x \in \mathbb{R}^d \) and some \( n = n(x) \in \mathbb{N} \),
\[
\text{Rank}[B_1(x), B_2(x), \cdots, B_n(x)] = d,
\]
by using the Malliavin calculus, we prove the existence of the heat kernel \( \rho_t(x, y) \) to the operator \( L^{(\alpha)}_0 \) as well as the continuity of \( x \mapsto \rho_t(x, \cdot) \) in \( L^1(\mathbb{R}^d) \) for each \( t > 0 \). Moreover, when \( \sigma(x) = \sigma \) is constant, under the following uniform Hörmander’s type condition: for some \( j_0 \in \mathbb{N} \),
\[
\inf_{x \in \mathbb{R}^d} \inf_{|\beta|=1} \sum_{j=1}^{j_0} |u B_j(x)|^2 > 0,
\]
we also prove the smoothness of \( (t, x, y) \mapsto \rho_t(x, y) \) with \( \rho_t(\cdot, \cdot) \in C^\infty_b(\mathbb{R}^d \times \mathbb{R}^d) \) for each \( t > 0 \).

1. Introduction

1.1. Background. Let \( A_0, A_1, \cdots, A_d \) be \( d + 1 \)-smooth vector fields on \( \mathbb{R}^d \). Consider the following second order differential operator
\[
L^{(2)} := \sum_{k=1}^{d} A_k^2 + A_0.
\] (1.1)
Let \( \mathcal{A}_0 := \{A_1, \cdots, A_d\} \) and for \( j \in \mathbb{N} \),
\[
\mathcal{A}_j := \{[A, A_k] : A \in \mathcal{A}_{j-1}, k = 0, \cdots, d\},
\] (1.2)
where \( [A, A_k] := AA_k - A_A \) is the usual Lie bracket. The celebrated Hörmander’s hypoellipticity theorem asserts that if
\[
\bigcup_{j=0}^{\infty} \mathcal{A}_j \text{ spans } \mathbb{R}^d \text{ at each point } x,
\] (1.3)
then \( \partial_t - L^{(2)} \) has a smooth fundamental solution. In 1967, Hörmander [14] first proved this result by using the analytic argument. Later, in 1976, Malliavin [25] provided a probabilistic proof. Malliavin’s idea was further developed by Bismut [6], Kusuoka-Stroock [22, 23], Watanabe [36], Norris [27] etc. Notice that Hörmander’s condition (1.3) is a local condition in the sense that if \( \bigcup_{j=0}^{\infty} \mathcal{A}_j \) spans \( \mathbb{R}^d \) at point \( x \), then \( \bigcup_{j=0}^{\infty} \mathcal{A}_j \) also spans \( \mathbb{R}^d \) at a neighbourhood of \( x \).

In contrary to the local operator (1.1), we consider the following nonlocal operator
\[
L^{(\alpha)}_{\sigma, b} f(x) := \text{p.v.} \int_{\mathbb{R}^d \setminus \{0\}} \frac{f(x + \sigma(x)z) - f(x)}{|z|^{d+\alpha}} \, dz + b(x) \cdot \nabla f(x), \quad \alpha \in (0, 2),
\] (1.4)
where $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ are two smooth functions, and p.v. stands for the Cauchy principal value. Notice that if $\sigma(x) = \mathbb{I}$ is the identity matrix and $b = 0$, then

$$L_{d,\alpha}^{(a)} f = c_{d,\alpha} \Delta_{d,\alpha} f,$$

where $\Delta_{d,\alpha}$ is the usual fractional Laplacian and $c_{d,\alpha}$ is a constant. Recently, there are increasing interests of studying such type of integro-differential operators. For example, in [26], Morimoto and Xu first studied the hypoelliptic effects to the following fractional kinetic Fokker-Planck equation

$$\partial_t f(t, x, v) + v \cdot \nabla_x f(t, x, v) = a(x, v) \Delta_{d,\alpha} f(t, x, v) + g,$$  \hspace{1cm} (1.5)

where $a$ is a bounded smooth and strictly positive function. See also [2, 11, 1] etc. for some related works about this equation. The physical motivation of studying equation (1.5) is refereed to [38]. Another motivation is linked with a linear model of spatially inhomogeneous Boltzmann equation without an angular cutoff (see [26, 2]). Notice that the operator in equation (1.5) is degenerate, and can be written as (1.4) with nonlinear $a$ and linear $b$ by the change of variables.

On the other hand, let us consider the following stochastic nonlinear oscillators studied in [13, 32, 9] etc.:

$$\begin{cases}
    \text{d}z_i(t) = u_i(t)\text{d}t, & i = 1, \cdots, d, \\
    \text{d}u_i(t) = -\partial_{z_i} H(z(t), u(t))\text{d}t, & i = 2, \cdots, d - 1, \\
    \text{d}u_i(t) = -[\partial_{z_i} H(z(t), u(t)) + \gamma_i u_i(t)]\text{d}t + \sqrt{T_i}\text{d}L_i, & i = 1, d,
\end{cases}$$  \hspace{1cm} (1.6)

where $L_i^1$ and $L_i^d$ are two independent of $\alpha$-stable processes, $d \geq 3$, $\gamma_1, \gamma_d \in \mathbb{R}$, $T_1, T_d > 0$, and

$$H(z, u) := \sum_{i=1}^{d} \left( \frac{1}{2} |u_i|^2 + V(z_i) \right) + \sum_{i=1}^{d-1} U(z_{i+1} - z_i).$$

The typical examples of potential functions $V$ and $U$ are

$$V(z) = \frac{|z|^2}{2}, \quad U(z) = \frac{|z|^2}{2} + \frac{|z|^4}{4},$$

The Hamiltonian $H$ describes a chain of particles with nearest-neighbor interaction. We remark that the generator of stochastic equation (1.6) also takes the form (1.4) with constant $a$ and nonlinear $b$, which is a typical Hörmander’s type operator if we assume $U'' > 0$ (cf. [12]).

In the present work, we are interested in the following problem: Under what degenerate conditions on $\sigma$ and $b$ appearing in (1.4), we can prove a similar theorem for integro-differential operator $L_{\sigma, b}^{(a)}$ as in Hörmander’s theorem. The purpose of this paper is to give partial results under Hörmander’s type conditions.

1.2. Main results. Let $B_1(x) := \sigma(x)$ and for $j \in \mathbb{N}$, define

$$B_{j+1}(x) := b(x) \cdot \nabla B_j(x) - B_j(x) \cdot \nabla b(x).$$  \hspace{1cm} (1.7)

Our first result is the following existence of the heat kernel to $L_{\sigma, b}^{(a)}$ and the $L^1$-continuity of the heat kernel with respect to the first variable.

**Theorem 1.1.** Let $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ be $C^\infty$-functions with bounded partial derivatives of first order. Assume that for each $x \in \mathbb{R}^d$ and some $n = n(x) \in \mathbb{N}$,

$$\text{Rank}[B_1(x), B_2(x), \cdots, B_n(x)] = d,$$  \hspace{1cm} (1.8)

then there exists a function $\rho_\tau(x, y)$ such that for any $f \in C^2_b(\mathbb{R}^d)$ and $t > 0, x \in \mathbb{R}^d$,

$$\partial_t \int_{\mathbb{R}^d} f(y)\rho_\tau(x, y)\text{d}y = \int_{\mathbb{R}^d} L_{\sigma, b}^{(a)} f(y)\rho_\tau(x, y)\text{d}y.$$  \hspace{1cm} (1.9)
Moreover, \((t, x) \mapsto \rho_t(x, \cdot)\) is continuous in \(L^1(\mathbb{R}^d)\) on \((0, \infty) \times \mathbb{R}^d\). In particular, the semigroup defined by \(P_t f(x) := \int_{\mathbb{R}^d} f(y) \rho_t(x, y) dy\) has the strong Feller property.

**Remark 1.2.** If \(\sigma(x) = \sigma\) is constant, in \([40]\) and \([12]\), we have already proved this theorem. Moreover, if we define vector fields

\[
A_0(x) := \sum_{i=1}^{d} b_i(x) \partial_i, \quad A_k(x) := \sum_{i=1}^{d} \sigma_{ki}(x) \partial_i,
\]

and set \(\mathcal{A}_0 := \{A_1, \cdots, A_d\}\) and

\[
\mathcal{A}_j := \left\{ [A, A_0] : A \in \mathcal{A}_{j-1} \right\}, \quad j \in \mathbb{N},
\]

then (1.8) is equivalent that

\[
\bigcup_{j=0}^{\infty} \mathcal{A}_j \text{ spans } \mathbb{R}^d \text{ at each point } x \in \mathbb{R}^d.
\]

Compared with (1.3), in the definition of \(\mathcal{A}_j\), all the Lie bracket is taken only with the drift vector field \(A_0\).

About the smoothness of \(\rho_t(x, y)\), we have the following partial result.

**Theorem 1.3.** Assume that \(\sigma(x) = \sigma\) is constant, \(b\) has bounded partial derivatives of all orders and for some \(j_0 \in \mathbb{N}\),

\[
\inf_{x \in \mathbb{R}^d} \inf_{|\alpha|=1} \sum_{j=1}^{j_0} |\alpha B_j(x)|^2 =: \kappa_1 > 0.
\]

(1.10)

Then there exists a probability density function \(\rho_t(x, y) \in C^\infty(\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d)\) with

\[
\rho_t(\cdot, \cdot) \in C^\infty_b(\mathbb{R}^d \times \mathbb{R}^d) \quad \forall t > 0,
\]

(1.11)

and such that

\[
\partial_t \rho_t(x, y) = L^{(s)}_{\sigma, b} \rho_t(\cdot, y)(x) \quad \forall t > 0,
\]

(1.12)

with \(\lim_{t \downarrow 0} \rho(t, x, y) = \delta(x)\) in the distributional sense.

**Remark 1.4.** In the linear case of \(b(x) = Bx\), Priola and Zabczyk \([30]\) first established this result under (1.10), and Kulik \([18]\) also showed that (1.10) is necessary for the existence of smooth fundamental solutions (see also \([34]\) for the related studies). In the nonlinear case, when \(j_0 = 2\), this result was proven in \([40]\). Here an open question is the smoothness of \(\rho_t(x, y)\) for the general non-constant \(\sigma(x)\). We will study this problem in a subsequent work \([42]\).

1.3. Arguments and Related results. Our proofs of Theorems 1.1 and 1.3 are based on the Malliavin calculus. Let \((L_t)_{t \geq 0}\) be a rotationally invariant \(\alpha\)-stable process. Consider the following stochastic differential equation (abbreviated as SDE) driven by \(L_t\):

\[
dX_t(x) = b(X_t(x)) dt + \sigma(X_t(x)) dL_t, \quad X_0(x) = x.
\]

It is well known that the generator of this SDE is given by \(L^{(s)}_{\sigma, b}\). Our task is thus to show that \(X_t(x)\) admits a \(L^1\)-continuity density \(\rho_t(x, y)\) under (1.8), and which is smooth under (1.10) when \(\sigma(x) = \sigma\) is constant. Let \((W_t)_{t \geq 0}\) be a \(d\)-dimensional standard Brownian motion and \((S_t)_{t \geq 0}\) an independent \(\alpha/2\)-stable subordinator. Then \((W_{S_t})_{t \geq 0}\) (also called subordinated Brownian motion) is a rotationally invariant \(\alpha\)-stable process, and we shall assume

\[
L_t = W_{S_t}.
\]

As in \([21, 39, 40]\), by taking regular conditional probability with respect to \(S_t\), the solution \(X_t(x)\) can be regarded as a Brownian functional so that we can use the classical Malliavin calculus to study the density of \(X_t(x)\).
Let us now recall some well known results about the smooth density for SDEs with jumps. In [7], Bismut first introduced the analogue of the Malliavin calculus for SDEs with jumps by the perturbation of jump sizes and using Girsanov’s transformation. In [5], Bichteler, Gravereaux and Jacod systematically studied the smooth density for SDEs driven by nondegenerate jump noises. In [29], Picard used the difference operator to give another criterion for the smoothness of the distribution density of Poisson functionals, and also applied it to SDEs driven by pure jump Lévy processes. By combining the classical Malliavin calculus and Picard’s difference operator argument, Ishikawa and Kunita in [15] obtained a new criterion for the smooth density of Wiener-Poisson functionals. In [21], Kusuoka developed the Malliavin calculus for subordinated Brownian functionals as mentioned above. The advantage of this method lies in the fact that it is not necessary to develop a new Malliavincalculus for jump processes (see [37, 39, 40, 41] for more applications). It is noticed that in all these results, the jump noises of SDEs are assumed to be nondegenerate. On the other hand, Cass in [10] established a Hörmander’s type theorem for SDEs with jumps by proving a Norris’ type lemma for discontinuous semimartingales, but the Brownian diffusion term can not disappear. In the pure jump degenerate case, by using a Komatsu-Takeuchi’s estimate proven in [17] for discontinuous semimartingales, Takeuchi [35] and Kunita [19, 20] also obtained similar Hörmander’s theorems. Unfortunately, the estimate in [17] seems incomplete, and which will be discussed below.

As usual, in the proof of Theorem 1.1, we only need to check the invertibility of the Malliavin covariance matrix $\Sigma_t(x)$ defined by (2.9) below. To obtain the $L^1$-continuity of the density, we shall use a criterion proved by Bogachev [8, Corollary 9.6.12] about the convergence of a family of Wiener functionals with respect to the total variation distance (see also [24, 4, 12]). While in the proof of Theorem 1.3, the main difficulty is to prove the $L^p$-integrability of the inverse of $\Sigma_t(x)$, where the key point is to prove a Norris’ type lemma for discontinuous semimartingales (cf. [10]). In the continuous diffusion case, Komatsu-Takeuchi [16] used a new estimate for continuous semimartingales to give a simplified probabilistic proof for Hörmander’s theorem. Later, in [17], they also extended it to the case of discontinuous semimartingales so that they can prove a Hörmander’s type theorem for the SDEs driven by general jump processes. However, it seems that there is a gap in their proofs. Here, we shall provide a slightly different and simplified proof for their estimate. In particular, the parameters appearing in the estimate are explicitly computed, which is crucial for our proof. As in [10], the difference with [17] is that the jump size of the discontinuous part has to be small.

1.4. Layout and Notations. This paper is organised as follows: In Section 2, we give some preliminaries about the Malliavin calculus and some necessary lemmas for later use. In particular, the Malliavin covariance matrix is calculated. In Section 3, we prove Theorem 1.1. In Section 4, we prove a Komatsu-Takeuchi’s estimate for discontinuous semimartingales by using different calculations. In Section 5, Theorem 1.3 will be proven. Before concluding this introduction, we collect some notations and make some conventions for later use.

- Write $\mathbb{R}_+ = (0, \infty)$ and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, $\mathbb{R}^d_0 := \mathbb{R}^d - \{0\}$.
- For a càdlàg function $f : \mathbb{R}_+ \to \mathbb{R}^d$, $\Delta f_s := f_s - f_{s-}$.
- The inner product in Euclidean space is denoted by $\langle x, y \rangle$ or $x \cdot y$.
- For $p > 0$, $\Gamma(p) := \int_0^\infty s^{p-1} e^{-s} ds$ denotes the usual Gamma function.
- $S(\mathbb{R}^d)$: The Schwardz space of rapidly decreasing smooth functions.
- $C^\infty_b(\mathbb{R}^d)$: The space of all smooth bounded functions with bounded partial derivatives of all orders.
- $\nabla = (\partial_1, \cdots, \partial_d)$ denotes the gradient operator, and $D$ the Malliavin derivative operator.
• For a smooth function $f : \mathbb{R}^d \to \mathbb{R}^d$, $(\nabla f)_{ij} := (\partial_j f^i)$ denotes the Jacobian matrix of $f$.
• The asterisk * denotes the transpose of a matrix or a column vector, or the dual operator.
• The capital letter $C$ (resp. the small letter $c$) with or without index will denote an unimportant constant with values in $[1, \infty)$ (resp. $(0, 1)$).

2. Preliminaries

We first introduce the canonical space of a subordinated Brownian motion $W_S$. Let $(\mathcal{F}, \mathbb{H}, \mu_\mathcal{F})$ be the classical Wiener space, i.e., $\mathcal{F}$ is the space of all continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^d$ with vanishing values at starting point 0, $\mathbb{H} \subset \mathcal{F}$ is the Cameron-Martin space consisting of all absolutely continuous functions with square integrable derivatives, $\mu_\mathcal{F}$ is the Wiener measure so that the coordinate process

$$W_t(w) := w_t$$

is a $d$-dimensional standard Brownian motion. Let $\mathcal{S}$ be the space of all càdlàg increasing functions from $\mathbb{R}_+$ to $\mathbb{R}_+$ with $\ell_0 = 0$. Suppose that $\mathcal{S}$ is endowed with the Skorohod metric and a probability measure $\mu_\mathcal{S}$ so that the coordinate process

$$S_t(\ell) := \ell_t$$

is a pure jump subordinator with Lévy measure $\nu_\mathcal{S}$ satisfying

$$\int_0^\infty (1 + u)\nu_\mathcal{S}(du) < +\infty.$$ 

Consider now the following product probability space

$$(\Omega, \mathcal{F}, \mathbb{P}) := \left(\mathcal{F} \times \mathcal{S}, \mathcal{B}(\mathcal{F}) \times \mathcal{B}(\mathcal{S}), \mu_\mathcal{F} \times \mu_\mathcal{S}\right),$$

and define for $\omega = (w, \ell) \in \mathcal{F} \times \mathcal{S}$,

$$L_\omega(\omega) := w_{\ell_t}.$$ 

Then $(L_\omega)_{t \geq 0}$ is a $d$-dimensional pure jump Lévy process with Lévy measure $\nu_L$ given by

$$\nu_L(E) = \int_0^\infty \int_E (2\pi u)^{-\frac{d}{2}} e^{-\frac{|u|^2}{2u}} d\nu_\mathcal{S}(du), \quad E \in \mathcal{B}(\mathbb{R}^d). \quad (2.1)$$

In particular, if $\nu_\mathcal{S}(du) = u^{-1-\frac{d}{2}} du$, then

$$\nu_L(dz) = 2^{-\frac{d}{2}+1}(2\pi)^{-\frac{d}{2}} \Gamma(\frac{d+1}{2}) |z|^{-d-\alpha} dz.$$ 

2.1. Conditional Malliavin calculus on $\Omega$. Let us now recall some basic notions and facts about the Malliavin calculus (cf. [28]). Let $\mathcal{U}$ be a real separable Hilbert space. Let $\mathcal{C}(\mathcal{F}; \mathcal{U})$ be the class of all $\mathcal{U}$-valued smooth cylindrical functionals on $\mathcal{F}$ with the form:

$$F = \sum_{i=1}^m f_i(W(h_1), \ldots, W(h_n))u_i,$$

where $f_i \in C^\infty_b(\mathbb{R}^n)$, $u_i \in \mathcal{U}$, $h_1, \ldots, h_n \in \mathbb{H}$, and

$$W(h) := \sum_{k=1}^d \int_0^\infty h^k_s dW^k_s.$$ 

The Malliavin derivative of $F$ is defined by

$$DF := \sum_{i=1}^m \sum_{j=1}^n (\partial_j f_i)(W(h_1), \ldots, W(h_n))u_i \otimes h_j \in \mathcal{U} \otimes \mathbb{H}.$$ 

By an iteration argument, for any $k \in \mathbb{N}$, the higher order Malliavin derivative $D^k F$ of $F$ can be defined as a random variable in $\mathcal{U} \otimes \mathbb{H}^\otimes k$. It is well known that the operator $(D^k, \mathcal{C}(\mathcal{F}; \mathcal{U}))$ is
closable from $L^p(\mathcal{W}; \mathbb{U})$ to $L^p(\mathcal{W}; \mathbb{U} \otimes \mathbb{H}^{\otimes k})$ for each $p \geq 1$ (cf. [28, p.26, Proposition 1.2.1]). For every $p > 1$ and $k \in \mathbb{N}$, we introduce a norm on $C(\mathcal{W}; \mathbb{U})$ by

$$
\|F\|_{k,p} := \left( \mathbb{E}^{\mathcal{W}} |F|^p + \sum_{i=1}^k \mathbb{E}^{\mathcal{W}} \left( \|D^i F\|_{U^{\otimes k}\mathbb{U}}^p \right) \right)^{\frac{1}{p}}.
$$

The Wiener-Sobolev space $\mathbb{D}^{k,p}(\mathcal{W}; \mathbb{U})$ is defined as the closure of $C(\mathcal{W}; \mathbb{U})$ with respect to the above norm. The dual operator $D^*$ of $D$ (also called divergence operator) is defined by relation

$$
\mathbb{E}^{\mathcal{W}} (DF, U)_{\mathbb{H}} = \mathbb{E}^{\mathcal{W}} (FD^* U), \quad U \in \mathbb{D}^{1,2}(\mathcal{W}; \mathbb{H}).
$$

The following Meyer’s inequality holds (cf. [28, p.75, Proposition 1.5.4]). For any $p > 1$ and $U \in \mathbb{D}^{1,p}(\mathcal{W}; \mathbb{H})$, we have

$$
\|D^* U\|_p \leq C_p \|U\|_{1,p}.
$$

(2.2)

Now, let $F(w, \ell)$ be a $\mathbb{U}$-valued functional on $\Omega = \mathcal{W} \times \mathbb{S}$. Let us denote

$$
F^\ell(w) := F(w, \ell),
$$

and define

$$
\mathbb{D}^{k,p}(\mathbb{U}) := \mathbb{D}^{k,p}(\Omega; \mathbb{U}) := \left\{ F \in L^p(\Omega; \mathbb{U}) : \text{for } \mu_{\mathbb{S}}\text{-a.a. } \ell \in \mathbb{S}, F^\ell \in \mathbb{D}^{k,p}(\mathcal{W}; \mathbb{U}), \right\}
$$

with $\|F\|_{k,p} := \int_{\Omega} |D^k F^\ell(w)|_{U^{\otimes k}\mathbb{U}}^p \mu(dw) < \infty$.

and

$$
(D^k F)(w, \ell) := D^k F^\ell(w), \quad F \in \mathbb{D}^{k,p}(\mathbb{U}).
$$

(2.3)

Below we simply write

$$
\mathbb{D}^{\infty}(\mathbb{U}) := \mathbb{D}^{\infty}(\mathbb{U}) := \cap_{k \in \mathbb{N}, p \geq 1} \mathbb{D}^{k,p}(\mathbb{U})
$$

and

$$
\mathbb{D}^{k,p} := \mathbb{D}^{k,p}(\mathbb{R}^d), \quad \mathbb{D}^{\infty} := \mathbb{D}^{\infty}(\mathbb{R}^d).
$$

Let $F = (F_1, \cdots, F_d)$ be a random vector on $\Omega$ with $F^\ell \in \mathbb{D}^{1,2}(\mathcal{W}; \mathbb{R}^d)$ for each $\ell \in \mathbb{S}$. The Malliavin covariance matrix of $F$ is defined by

$$
(S_F)_{ij}(w, \ell) := \langle D^i F^\ell(w), D^j F^\ell(w) \rangle_{\mathbb{H}}.
$$

(2.4)

The following criterion about the $L^1$-continuity of the densities of a random field can be found in [8, Corollary 9.6.12] (see also [24, 4, 12]).

**Theorem 2.1.** Let $\Lambda$ be a metric space and $(X_\lambda)_{\lambda \in \Lambda}$ a family of $\mathbb{R}^d$-valued functionals over $\Omega$. Suppose that for some $p \geq d$,

(H1) For each $\lambda \in \Lambda$ and $\mu_\mathbb{S}$-almost all $\ell \in \mathbb{S}$, $X^\ell_\lambda \in \mathbb{D}^{1,p}(\mathcal{W}; \mathbb{R}^d)$, and if $\lambda_n \to \lambda$, then

$$
\lim_{n \to \infty} \|X^\ell_{\lambda_n} - X^\ell_\lambda\|_{1,p} = 0.
$$

(H2) For each $\lambda \in \Lambda$, the Malliavin covariance matrix $\Sigma_{X^\lambda}$ of $X^\lambda$ is invertible almost surely. Then the law of $X^\lambda$ in $\mathbb{R}^d$ admits a density $\rho_\lambda(x)$ so that $\Lambda \mapsto \rho_\lambda(\cdot)$ is continuous in $L^1(\mathbb{R}^d)$.

**Proof.** First of all, by (H2) and Fubini’s theorem, for $\mu_\mathbb{S}$-almost all $\ell \in \mathbb{S}$, $\Sigma_{X^\lambda}(w)$ is invertible for $\mu_{\mathcal{W}}$-almost all $w$. Thus, by (H1) and Bouleau-Hirsch’s criterion [28, p.92, Theorem 2.1.1], the law of $X^\lambda$ is absolutely continuous with respect to the Lebesgue measure, and so does $X^\lambda$. 


Here and below, the Malliavin covariance matrix of $X_n$ motion $\sigma$ is defined by $(\Sigma n(x))_{ij} := \mathbb{E}[\sigma_i(X_n(x))\sigma_j(X_n(x))]$, where $\sigma$ is a $C^\infty$-function with bounded partial derivatives of first order. It is well known that $\{X_n(x), t \geq 0, x \in \mathbb{R}^d\}$ forms a $C^1$-stochastic flow (cf. [31]). For $0 < r < t$, let $J_{r,t}(x)$ solve the following linear matrix-valued SDE:

$$J_{r,t}(x) = \mathbb{1} + \int_r^t \nabla b(X_s(x))J_{s,t}(x)ds + \sum_{k=1}^d \int_r^t \nabla \sigma_k(X_{t-s}(x))J_{s,t}(x)ds dW^k_S,$$  

where $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ are $C^\infty$-functions with bounded partial derivatives of first order. Let $\lambda_n \to \lambda$. By (H1) and Bogachev’s criterion [8, Corollary 9.6.12] (see also [24, Corollary 2.2]), we obtain that for $\mu_\Sigma$-almost all $\ell \in \mathbb{S}$,

$$\lim_{n \to \infty} \sup_{\|f\|_{\infty} \leq 1} |\mathbb{E}^\mu(f(X_{t+n}^\ell) - f(X_t^\ell))| = 0,$$

which, by the dominated convergence theorem, implies that

$$\lim_{n \to \infty} \sup_{\|f\|_{\infty} \leq 1} |\mathbb{E}(f(X_{t+n}) - f(X_t))| \leq \int \lim_{n \to \infty} \sup_{\|f\|_{\infty} \leq 1} |\mathbb{E}^\mu(f(X_{t+n}^\ell) - f(X_t^\ell))| \mu_\Sigma(d\ell) = 0. \quad (2.5)$$

The desired continuity follows by noticing

$$\int_{\mathbb{R}^d} |\rho_{n}(x) - \rho_{\ell}(x)|dx = \sup_{\|f\|_{\infty} \leq 1} |\mathbb{E}(f(X_{t+n}) - f(X_t))|$$

and (2.20).

The following theorem about the criterion that a random vector admits a smooth density in the Malliavin calculus can be found in [28, p.100-103].

**Theorem 2.2.** Assume that $F = (F^1, \cdots, F^d) \in \mathbb{D}^\infty$ is a smooth functional on $\Omega$ and satisfies that for all $p \geq 2$,

$$\mathbb{E}[(\det \Sigma_F)^{-p}] < \infty.$$  

Let $G \in \mathbb{D}^\infty$ and $\varphi \in C^\infty_b(\mathbb{R}^d)$. Then for any multi-index $\alpha = (\alpha_1, \cdots, \alpha_m) \in \{1, 2, \cdots, d\}^m$,

$$\mathbb{E}[\partial_\alpha \varphi(F)G] = \mathbb{E}[\varphi(F)H_\alpha(F, G)],$$

where $\partial_\alpha = \partial_{\alpha_1} \cdots \partial_{\alpha_m}$, and $H_\alpha(F, G)$ are recursively defined by

$$H_{(i)}(F, G) := \sum_j D^i(G(\Sigma_F^{-1})_{ij}DF_j), \quad (2.6)$$

$$H_\alpha(F, G) := H_{(\alpha_1)}(F, H_{(\alpha_2, \cdots, \alpha_m)}(F, G)).$$

As a consequence, for any $p > 1$, there exist $p_1, p_2, p_3 > 1$ and $n_1, n_2 \in \mathbb{N}$ such that

$$\|H_{\alpha}(F, G)\|_{p} \leq C(\|\det \Sigma_F\|^{-p_1}_{p_1} \|DF\|^{n_2}_{\infty, p_2} \|G\|_{n_2, p_3}). \quad (2.7)$$

In particular, the law of $F$ possesses an infinitely differentiable density $\rho \in \mathcal{S}(\mathbb{R}^d)$.

### 2.2. Malliavin Covariance Matrix

Consider the following SDE driven by subordinated Brownian motion $W_S$:

$$dX_t(x) = b(X_t(x))dt + \sigma(X_{t-s}(x))dW_S, \quad X_0(x) = x,$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ are $C^\infty$-functions with bounded partial derivatives of first order. Let $(\Sigma_n(x))_{ij}$ be the Malliavin covariance matrix of $X_n(x)$ defined by (2.4). We have the following explicit expression about the Malliavin covariance matrix of $X_n(x)$.

**Lemma 2.3.** Let $(\Sigma_n(x))_{ij}$ be the Malliavin covariance matrix of $X_n(x)$ defined by (2.4). We have

$$\Sigma_n(x) = \int_0^\infty J_{s,t}(x)\sigma(X_{t-s}(x))\sigma^*(X_{t-s}(x))J_{s,t}^*(x)ds, \quad (2.9)$$

where $J_{r,t}(x)$ solves the following linear matrix-valued SDE:

$$J_{r,t}(x) = \mathbb{1} + \int_r^t \nabla b(X_s(x))J_{s,t}(x)ds + \sum_{k=1}^d \int_r^t \nabla \sigma_k(X_{t-s}(x))J_{s,t}(x)ds dW^k_s.$$  

Here and below, $\int_r^t \sigma$ always means $\int_{(r,t)} \sigma$. We have the following explicit expression about the Malliavin covariance matrix of $X_n(x)$.
Proof. Fix a path \( \ell \in \mathcal{S} \) and \( x \in \mathbb{R}^d \). Let \( X^\ell_s \) solve the following SDE:
\[
dX^\ell_s = b(X^\ell_s)dt + \sigma(X^\ell_s)dW_t, \quad X^\ell_0 = x, \tag{2.10}
\]
and for \( 0 \leq r < t \), let \( J^\ell_{r,t} \) solve the following linear matrix-valued SDE:
\[
J^\ell_{r,t} = \mathbb{I} + \int_r^t \nabla b(X^\ell_s)J^\ell_{r,s} ds + \sum_{k=1}^d \int_r^t \nabla \sigma_k(X^\ell_{s-})J^\ell_{r,s-} \, dW^k_t. \tag{2.11}
\]
By definitions (2.3) and (2.4), it suffices to prove that
\[
\Sigma^\ell_t = \int_0^t J^\ell_{s,t} \sigma(X^\ell_{s-}) \sigma^*(X^\ell_{s-})(J^\ell_{s,t})^* \, ds, \tag{2.12}
\]
where \( (\Sigma^\ell_{t})_{ij} := \langle D(X^\ell_t)^i, D(X^\ell_t)^j \rangle_{\mathbb{H}} \). For any \( h \in \mathbb{H} \), it is by now standard to prove that
\[
D_hX^\ell_t = \int_0^t \nabla b(X^\ell_s)D_hX^\ell_s ds + \sum_{k=1}^d \int_0^t \nabla \sigma_k(X^\ell_{s-})D_hX^\ell_{s-} \, dW^k_s + \int_0^t \sigma(X^\ell_{s-})h \, \, dW_s. \tag{2.13}
\]
By the formula of constant variations or Fubini’s theorem, we have
\[
D_hX^\ell_t = \int_0^t J^\ell_{s,t} \sigma(X^\ell_{s-}) \, h \, \, ds.
\]
Let \( \{h^n, n \in \mathbb{N} \} \) be an orthonormal basis of \( \mathbb{H} \). We have
\[
\Sigma^\ell_t = \sum_n D_{h^n}X^\ell_t \otimes D_{h^n}X^\ell_t = \sum_n \int_0^t J^\ell_{s,t} \sigma(X^\ell_{s-}) \, dW^n_s \otimes \int_0^t J^\ell_{s,t} \sigma(X^\ell_{s-}) \, dW^n_s,
\]
where for a vector \( x \in \mathbb{R}^d \), \( x \otimes x := (x_i x_j)_{ij} \). From this, we then obtain (2.12) by noticing that (cf. [43, Lemma 4.4])
\[
\sum_n \left( \int_0^t \langle f, dW^n_s \rangle \right) \left( \int_0^t \langle g, dW^n_s \rangle \right) = \int_0^t \langle f, g \rangle \, ds.
\]
The proof is complete. \( \square \)

2.3. **Three useful lemmas.** In the remainder of this paper, we make the following assumption on \( \nu_S \): For some \( \alpha \in (0, 2) \),
\[
0 < \lim_{\varepsilon \downarrow 0} \varepsilon^{\frac{\alpha}{2}-1} \int_0^\infty \nu_S(du) \leq \lim_{\varepsilon \downarrow 0} \varepsilon^{\frac{\alpha}{2}-1} \int_0^\infty \nu_S(du) < \infty, \tag{AS}^\alpha
\]
We use the following filtration:
\[
\mathcal{F}_t := \sigma(W_S, S_s : 0 \leq s \leq t).
\]
Clearly, for \( t > s \), \( W_S - W_{S_s} \) and \( S_t - S_s \) are independent of \( \mathcal{F}_s \).

We need the following lemmas.

**Lemma 2.4.** Let \( f_t: \mathbb{R}_+ \to \mathbb{R}_+ \) be a bounded continuous \( \mathcal{F}_t \)-adapted process. Under (AS)\(^\alpha\), there exist constants \( c_0 = c_0(\nu_S, \alpha, \|f\|_\infty) \in (0, 1) \) and \( \lambda_0 = \lambda_0(\nu_S, \alpha, \|f\|_\infty) > 0 \) such that for any \( \lambda \geq \lambda_0 \),
\[
\mathbb{E} \exp \left( -\lambda \int_0^t f_s dS_s \right) \leq \left( \mathbb{E} \exp \left( -c_0 \lambda^{\frac{\alpha}{2}} \int_0^t f_s ds \right) \right)^{\lambda^2}. \]
Proof. For $\lambda > 0$, set
\[
g_s^1 := \int_0^\infty (1 - e^{-\lambda f_s}) \nu_S(du)
\]
and
\[
M^1_t := -\lambda \int_0^t f_s dS_s + \int_0^t g_s^1 ds.
\]
Let $N_S(dr, du)$ be the Poisson random measure associated with $S$, i.e.,
\[
N_S((0, t] \times U) := \sum_{s \leq t} 1_{U}(S_s - S_{s-}), \quad U \in \mathcal{B}(\mathbb{R}_+).
\]
(2.14)
Let $\tilde{N}_S(dr, du)$ be the compensated Poisson random measure of $N_S(dr, du)$, i.e.,
\[
\tilde{N}_S(dr, du) = N_S(dr, du) - d\nu_S(du).
\]
By Lévy-Itô’s decomposition (cf. [33]), we can write
\[
S_t = t \int_0^1 uv_S(du) + \int_0^t \int_0^1 u\tilde{N}_S(ds, du) + \int_0^\infty \int_1^\infty uN_S(ds, du),
\]
so that
\[
\int_0^t f_s dS_s = \int_0^t f_s \left( \int_0^1 u\nu_S(du) \right) ds + \int_0^t \int_0^1 f_s u\tilde{N}_S(ds, du) + \int_0^\infty \int_1^\infty f_s uN_S(ds, du).
\]
By Itô’s formula (cf. [3]), we have
\[
e^{M^1_t} = 1 + \int_0^t e^{M^1_r \cdot [e^{-\lambda f_s} - 1]} \tilde{N}_S(ds, du).
\]
(2.16)
Since for any $x \geq 0$,
\[
1 - e^{-x} \leq 1 \land x,
\]
we have
\[
M^1_t \leq \int_0^\infty g_s^1 ds \leq t \int_0^\infty (1 \land (\lambda \|f\|_\infty |u|)) \nu_S(du).
\]
Hence, by (2.16) we have
\[
\mathbb{E}e^{M^1_t} = 1,
\]
and by Hörder’s inequality,
\[
\mathbb{E} \exp \left\{ -\frac{\lambda}{2} \int_0^t f_s dS_s \right\} = \mathbb{E} \left( e^{M^1_t/2} \exp \left\{ -\frac{1}{2} \int_0^t g_s^1 ds \right\} \right) \leq \left( \mathbb{E} \exp \left\{ -\int_0^t g_s^1 ds \right\} \right)^{1/2}.
\]
(2.17)
On the other hand, since for any $\kappa \in (0, 1)$ and $x \leq -\log k$,
\[
1 - e^{-x} \geq \kappa x,
\]
letting $\kappa = \frac{1}{e}$, we have
\[
g_s^1 \geq \int_0^{\lambda |f|_{\text{ess}}} (1 - e^{-\lambda f_s}) \nu_S(du) \geq \frac{1}{e} \int_0^{\lambda |f|_{\text{ess}}} (\lambda f_s u) \nu_S(du) = \frac{\lambda f_s}{e} \int_0^{\lambda |f|_{\text{ess}}} u \nu_S(du).
\]
Substituting this into (2.17) and using (AS$^\alpha$), we obtain the desired estimate. \qed

Lemma 2.5. Under (AS$^\alpha$), there exist constants $C_0, C_1 \geq 1$ such that for all $\delta \in (0, 1)$,
\[
\int_{|z| > \delta} \nu_L(dz) \leq C_0 \delta^{-\alpha}, \quad \int_{|z| \leq \delta} |z|^2 \nu_L(dz) \leq C_1 \delta^{2 - \alpha}.
\]
(2.18)
Proof. First of all, by \((\text{AS}^\alpha)\), there exists a constant \(C \geq 1\) such that for all \(\varepsilon \in (0, 1),\)

\[
\int_0^\varepsilon u v_S(du) \leq C \varepsilon^{1-\frac{\alpha}{2}}. \tag{2.19}
\]

For any \(\varepsilon \in (0, 1),\) letting \(m := \lceil \log_2(1/\varepsilon) \rceil + 1,\) we have

\[
v_S([\varepsilon, 1]) \leq \sum_{k=0}^m \int_{2^k \varepsilon}^{2^{k+1} \varepsilon} v_S(du) \leq \sum_{k=0}^m (2^k \varepsilon)^{-\alpha} \int_{2^k \varepsilon}^{2^{k+1} \varepsilon} u v_S(du)
\]

\[
\leq C \sum_{k=0}^m (2^k \varepsilon)^{-\alpha} (2^{k+1} \varepsilon)^{1-\frac{\alpha}{2}} = C 2^{1-\frac{\alpha}{2}} \sum_{k=0}^m (2^k \varepsilon)^{1-\frac{\alpha}{2}},
\]

which implies that for some \(C \geq 1,\)

\[
v_S([\varepsilon, \infty)) \leq C \varepsilon^{\frac{\alpha}{2}}. \tag{2.20}
\]

By \((2.1)\) and the change of variables, we have

\[
\int_{|\cdot|>\delta} v_L(\,d\zeta) = \int_0^\infty \left( \int_{|\cdot|>\delta} (2\pi u)^{-\frac{d}{2}} e^{-\frac{1}{2} \|\zeta\|^2} \,d\zeta \right) v_S(du)
\]

\[
\leq v_S((\delta^2, \infty)) + \pi^{-\frac{d}{2}} \int_0^\delta \left( \int_{|z|<\delta} \frac{2|\zeta|^2}{\delta^2} e^{-\|\zeta\|^2} \,d\zeta \right) v_S(du)
\]

\[
\leq C \delta^{-\alpha} + \pi^{-\frac{d}{2}} \int_0^\delta u \left( \int_{|\cdot|<\delta} \frac{2|\zeta|^2}{\delta^2} e^{-\|\zeta\|^2} \,d\zeta \right) v_S(du)
\]

\[
\leq C \delta^{-\alpha} + 2\pi^{-\frac{d}{2}} \delta^{-2} \int_0^\delta u v_S(du) \left( \int_{\mathbb{R}^d} |x|^2 e^{-\|x\|^2} \,dx \right) \leq \frac{C_0 \delta^{-\alpha}}{2.19} \tag{2.20}
\]

and

\[
\int_{|\cdot|<\delta} |\zeta|^2 v_L(\,d\zeta) = \int_0^\infty \left( \int_{|\cdot|<\delta} (2\pi u)^{-\frac{d}{2}} |\zeta|^2 e^{-\frac{1}{2} \|\zeta\|^2} \,d\zeta \right) v_S(du)
\]

\[
= 2\pi^{-\frac{d}{2}} \int_0^\infty u \left( \int_{|\cdot|<\delta} |x|^2 e^{-\|x\|^2} \,dx \right) v_S(du)
\]

\[
\leq C_d \left( \int_0^\delta u v_S(du) + \delta^2 \int_0^\infty v_S(du) \right) \leq C_1 \delta^{2-\alpha}. \tag{2.19, 2.20}
\]

The proof is complete. \(\square\)

Lemma 2.6. Under \((\text{AS}^\alpha)\), we have

\[
\mathbb{P}(\omega : \{s : \Delta S_s(\omega) > 0\} \text{ is dense in } (0, \infty)) = 1.
\]

Proof. Define a stopping time \(\tau := \inf\{t > 0 : S_t > 0\} = \inf\{t > 0 : \Delta S_t > 0\}.\) As in the proof of \([40\text{, Lemma 2.1}]\), it suffices to prove that

\[
\mathbb{P}(\tau = 0) = 1.
\]

Let \(N_S(ds, du)\) be defined by \((2.14).\) For any \(\varepsilon \in (0, 1),\) we have

\[
\varepsilon \geq \mathbb{E}(\Delta S_\tau, 1_{\Delta S_\tau \leq \varepsilon}) = \mathbb{E}\left( \sum_{0 < t < \tau} \Delta S_t 1_{\Delta S_t \leq \varepsilon} \right) = \mathbb{E}\left( \int_0^\tau \int_0^\varepsilon u N_S(ds, du) \right)
\]

\[
= \mathbb{E}\left( \int_0^\varepsilon \int_0^\tau u v_S(dz)ds \right) = \left( \int_0^\tau u v_S(du) \right) \mathbb{E}\tau,
\]
which, together with (\(\text{AS}^p\)) and letting \(\varepsilon \to 0\), implies that
\[
\mathbb{E}\tau = 0 \Rightarrow \mathbb{P}(\tau = 0) = 1.
\]
The proof is complete. \(\square\)

3. Proof of Theorem 1.1

In order to prove Theorem 1.1, it suffices to check the conditions of Theorem 2.1 and (1.9) follows by Itô’s formula. First of all, we have

**Lemma 3.1.** Assume that \(b, \sigma \in C^\infty\) has bounded partial derivatives of first order. For each \(\ell \in \mathbb{S}, \ p \geq 2\) and all \((t, x) \in (0, \infty) \times \mathbb{R}^d\), we have
\[
X_t^\ell(x) \in \mathbb{D}^{1,p}(\mathbb{W}; \mathbb{R}^d).
\]
Moreover, for fixed \(t > 0\) and \(x \in \mathbb{R}^d\), there exists a \(\mu_\varepsilon\)-null set \(\mathcal{S}_0\) such that for all \(\ell \notin \mathcal{S}_0,
\[
\lim_{(s, y) \to (t, x)} \|X_s^\ell(y) - X_t^\ell(x)\|_{1,p} = 0.
\]
In particular, (H1) of Theorem 2.1 holds.

**Proof.** First of all, by equations (2.12) and (2.13), since \(b, \sigma \in C_b^1\), it is easy to prove that for any \(p \geq 2\) and \(T > 0\), \(x \in \mathbb{R}^d\),
\[
\mathbb{E}\left( \sup_{t \in [0, T]} (|X_t^\ell(x)|^p + \|DX_t^\ell(x)\|_{\mathbb{H}}^p) \right) \leq C(T, p, \ell)(1 + |x|^p),
\]
which gives (3.1). Let us prove (3.2). For fixed \(x, y \in \mathbb{R}^d\), set
\[
Z_t := X_t^\ell(x) - X_t^\ell(y).
\]
By equation (2.13), we have
\[
\langle DZ_t, h \rangle_\mathbb{H} = x - y + \int_0^t \nabla b(X_s(x))(DZ_s, h)_\mathbb{H} ds + \sum_k \int_0^t \nabla \sigma \cdot (X_{s-k}(x))(DZ_{s-k}, h)_\mathbb{H} dW_t^k + \int_0^t |\nabla b(X_s(x)) - \nabla b(X_s(y))|D_t h \|X_s(y)\| ds + \int_0^t |\sigma(X_{s-k}(x)) - \sigma(X_{s-k}(y))|d\ell_t + \sum_k \int_0^t |\nabla \sigma \cdot (X_{s-k}(x)) - \nabla \sigma \cdot (X_{s-k}(y))|D_t h \|X_{s-k}(y)\| dW_t^k.
\]
Let \(\{h_n, n \in \mathbb{N}\}\) be an orthonormal basis of \(\mathbb{H}\). Since
\[
\|DZ_t\|_\mathbb{H}^2 = \sum_n \langle (DZ_t, h_n) \rangle_{\mathbb{H}}^2,
\]
by Burkholder’s inequality (cf. [21] or [39, Lemma 2.3]), we have
\[
\mathbb{E}\|DZ_t\|_\mathbb{H}^p \leq C|x - y|^p + C t^{p-1} \int_0^t \mathbb{E}\|DZ_s\|_{\mathbb{H}}^p ds + C \mathbb{E}\left( \int_0^t \|DZ_{s-k}\|_{\mathbb{H}}^2 d\ell_s \right)^{p/2} + C t^{p-1} \int_0^t \mathbb{E}[|\nabla b(X_s(x)) - \nabla b(X_s(y))|^p] ds + C \mathbb{E}\left( \int_0^t \|Z_{s-k}\|^2 d\ell_s \right)^{p/2} + C \mathbb{E}\left( \int_0^t |\nabla \sigma(X_{s-k}(x)) - \nabla \sigma(X_{s-k}(y))|^2 \|DZ_{s-k}(y)\|_{\mathbb{H}}^2 d\ell_s \right)^{p/2} \leq C|x - y|^p + C t^{p-1} \int_0^t \mathbb{E}\|DZ_s\|_{\mathbb{H}}^p ds + C t^{p/2-1} \int_0^t \mathbb{E}\|DZ_{s-k}\|_{\mathbb{H}}^p d\ell_s + I_t(x, y),
\]
11
where $C$ only depends on $p$, and $I_1(x, y)$ denotes the remaining terms. By Gronwall’s inequality, we derive that

$$ \mathbb{E}[|DZ_t|^p] \leq C_1^p(1 + I_1(x, y)). $$

On the other hand, we similarly have

$$ \mathbb{E}[X^f_t(x) - X^f_t(y)]^p \leq C_1^p|x - y|^p $$

and

$$ \mathbb{E}[|X^f_t(x) - X^f_t(y)|^p] \leq C_1^p(1 + |x|^p). $$

Combining the above estimates and noticing that for fixed $t$ and $\mu_\mathbb{E}$-almost all $\ell$, $s \to \ell_s$ is continuous at $t$, we thus obtain the desired continuity \((3.2)\). \( \square \)

The remaining task is to verify (H2) of Theorem 2.1 under (1.8), i.e., to prove the invertibility of the Malliavin covariance matrix $\Sigma(x)$ given in (2.9).

Let $N(dt, dz)$ be the Poisson random measure associate to $L_t = W_{\Sigma_1}$, i.e.,

$$ N((0, t] \times U) = \sum_{s \leq t} 1_U(\Delta L_s), \quad U \in \mathcal{B}(\mathbb{R}^d). $$

The compensated Poisson random measure of $N(dt, dz)$ is defined by

$$ \tilde{N}(dt, dz) := N(dt, dz) - \nu_{L_t}(dz). $$

Choose $\delta > 0$ being small and fixed so that

$$ 2\delta\|
abla \sigma\|_{\infty} \leq 1. \quad (3.3) $$

Define

$$ L^\delta_t := \int_0^t \int_{|z| \leq \delta} z \tilde{N}(ds, dz), \quad \hat{L}^\delta_t := \int_0^t \int_{|z| > \delta} z \tilde{N}(ds, dz). \quad (3.4) $$

Then $L^\delta_t$ and $\hat{L}^\delta_t$ are two independent Lévy processes with Lévy measures $\sum_{|z| \leq \delta} \nu_L(dz)$ and $\sum_{|z| > \delta} \nu_L(dz)$ respectively, and by Lévy-Itô’s decomposition, we have

$$ L_t = L^\delta_t + \hat{L}^\delta_t. \quad (3.5) $$

Fix $T > 0$ and a càdlàg function $\bar{h} : (0, \infty) \to \mathbb{R}^d$ with

$$ \Delta \bar{h}_T = 0, \quad |\Delta \bar{h}_t| = 0 \text{ or } > \delta, \quad \forall s \geq 0. \quad (3.6) $$

Notice that for almost all $\omega$, the path $t \mapsto L^\delta_t(\omega)$ has this property. Let $X^\delta_t$ solve the following SDE:

$$ X^\delta_t = x + \int_0^t b(X^\delta_s)ds + \int_0^t \int_{|z| \leq \delta} \sigma(X^\delta_s)z\tilde{N}(ds, dz) + \int_0^t \sigma(X^\delta_s)dh_s, $$

and $J^{\delta}_{r,t}$ solve the following linear matrix-valued SDE:

$$ J^{\delta}_{r,t} = \tilde{I} + \int_r^t \nabla b(X^\delta_s)J^{\delta}_{r,s}ds + \int_r^t \int_{|z| \leq \delta} \nabla \sigma(X^\delta_s)zJ^{\delta}_{r,s-}\tilde{N}(ds, dz) + \sum_{k=1}^d \int_r^t \nabla \sigma_{k}(X^\delta_s)J^{\delta}_{r,s-}d\bar{h}^k_s. $$

By \((3.5)\), it is easy to see that

$$ X_t = X^\delta_t \big|_{\bar{h} = \hat{L}^\delta}, \quad J_{r,t} = J^{\delta}_{r,t} \big|_{\bar{h} = \hat{L}^\delta}. \quad (3.7) $$

Define

$$ \lambda := \sup\{t < T : |\Delta \bar{h}_t| > \delta\}. \quad (3.8) $$

By \((3.6)\), we have for all $t \in [\lambda, T]$,

$$ J^{\delta}_{r,t} = \tilde{I} + \int_t^\lambda \nabla b(X^\delta_s)J^{\delta}_{r,s}ds + \int_t^\lambda \int_{|z| \leq \delta} \nabla \sigma(X^\delta_s)zJ^{\delta}_{r,s-}\tilde{N}(ds, dz). $$
Lemma 3.2. For each $t \in [\lambda, T]$, the matrix $J^h_{A, t}$ is invertible and $K^h_{A, t} := (J^h_{A, t})^{-1}$ solves

$$K^h_{A, t} = I - \int_A K_{A, s} \nabla b(X^s_A)ds + \int_A \int_{|z| \leq \delta} K_{A, s} Q(X^s_A, z)\tilde{N}(ds, dz)$$

Moreover, for any $p \geq 2$, we have

$$\sup_{t \in [\lambda, T]} \mathbb{E}|K^h_{A, t}|^p < \infty.$$  \hspace{1cm} (3.10)

Proof. For simplicity of notations, we drop the superscript “$h$”. By Itô’s formula, we have

$$K_{A, t} J_{A, t} = I + \int_A K_{A, s} dJ_{A, s} + \int_A dK_{A, s} J_{A, s} + \sum_{A \leq s \leq t} \Delta K_{A, s} \Delta J_{A, s}$$

$$= I + \int_A K_{A, s} \nabla b(X^s_A) J_{A, s} ds + \int_A \int_{|z| \leq \delta} K_{A, s} \nabla \sigma(X^s_A) z J_{A, s} \tilde{N}(ds, dz)$$

$$- \int_A K_{A, s} \nabla b(X^s_A) J_{A, s} ds + \int_A \int_{|z| \leq \delta} K_{A, s} Q(X^s_A, z) J_{A, s} \tilde{N}(ds, dz)$$

$$- \int_A \int_{|z| \leq \delta} K_{A, s} Q(X^s_A, z) \nabla \sigma(X^s_A) z J_{A, s} \tilde{N}(ds, dz)$$

$$+ \int_A \int_{|z| \leq \delta} K_{A, s} Q(X^s_A, z) \nabla \sigma(X^s_A) z J_{A, s} N(ds, dz).$$

By the definition of $Q(x, z)$, one sees that

$$Q(x, z) + \nabla \sigma(x) z = - Q(x, z) \nabla \sigma(x) z.$$  

Hence,

$$K_{A, t} J_{A, t} = I.$$

Estimate (3.10) is standard by (3.9) and Gronwall’s inequality. The proof is complete. \hfill \Box

Let $V : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ be a matrix-valued smooth function. Let us first introduce two functions:

$$H_V(x, z) := V(x + \sigma(x) z) - V(x) + Q(x, z) V(x + \sigma(x) z),$$ \hspace{1cm} (3.11)

$$G_V(x, z) := H_V(x, z) + \nabla \sigma(x) z \cdot V(x) - \sigma(x) z \cdot \nabla V(x).$$ \hspace{1cm} (3.12)

Define

$$M^h_V(t) := \int_A \int_{|z| \leq \delta} K^h_{A, s} H_V(X^s_A, z) \tilde{N}(ds, dz),$$ \hspace{1cm} (3.13)

$$R^h_V(t) := \int_A \int_{|z| \leq \delta} K^h_{A, s} G_V(X^s_A, z) \nu_L(dz) ds.$$ \hspace{1cm} (3.14)
Lemma 3.3. Let \([b, V] := b \cdot \nabla V - V \cdot \nabla b\). We have \(\mathbb{P}(\Omega_1^V) = 1\), where
\[
\Omega_1^V := \left\{ \omega : K^h_{t,s} V(X^h_s) = V(X^h_t) + \int_t^s K^h_{t,r} [b, V](X^h_r)\, dr + M^h_r(t), \forall t \in [\lambda, T] \right\}.
\]

Proof. For simplicity of notations, we drop the superscript “\(h\)”. By Itô’s formula, we have
\[
V(X_t) = V(X_0) + \int_0^t (b \cdot \nabla) V(X_s)\, ds + \int_0^t \int_{|z| \leq \delta} [V(X_{s^-} + \sigma(X_{s^-})z) - V(X_{s^-})] \tilde{N}(ds, dz)
\]
\[+ \int_0^t \int_{|z| \leq \delta} [V(X_{s^-} + \sigma(X_{s^-})z) - V(X_{s^-}) - \sigma(X_{s^-})z \cdot \nabla V(X_{s^-})] \nu_L(dz)\, ds.
\]
By Itô’s product formula and Lemma 3.2, we have
\[
K_{t,s} V(X_s) = V(X_t) + \int_t^s dK_{t,r} V(X_r) + \int_t^s K_{t,r} dV(X_r) + \sum_{t < r \leq s} \Delta K_{t,r} \Delta V(X_r)
\]
\[= V(X_t) + \int_t^s K_{t,r} [b, V](X_r)\, dr + \int_t^s K_{t,r} Q(X_{t,r}, z)V(X_{s^-})\tilde{N}(ds, dz)
\]
\[- \int_0^t \int_{|z| \leq \delta} K_{t,r} Q(X_{t,r}, z)\nabla \sigma(X_{t,r})z V(X_{s^-}) \nu_L(dz)\, dr
\]
\[+ \int_0^t \int_{|z| \leq \delta} K_{t,r} [V(X_{s^-} + \sigma(X_{s^-})z) - V(X_{s^-})] \tilde{N}(ds, dz)
\]
\[+ \int_0^t \int_{|z| \leq \delta} K_{t,r} [V(X_{s^-} + \sigma(X_{s^-})z) - V(X_{s^-}) - \sigma(X_{s^-})z \cdot \nabla V(X_{s^-})] \nu_L(dz)\, ds
\]
\[+ \int_0^t \int_{|z| \leq \delta} K_{t,r} Q(X_{s^-}, z)[V(X_{s^-} + \sigma(X_{s^-})z) - V(X_{s^-})] \tilde{N}(ds, dz).
\]
Noticing that
\[
V(x + \sigma(x)z) - V(x) - \sigma(x)z \cdot \nabla V(x) + Q(x, z)[V(x + \sigma(x)z) - V(x) - \nabla \sigma(x)z V(x)]
\]
\[= V(x + \sigma(x)z) - V(x) + Q(x, z)V(x + \sigma(x)z) + \nabla \sigma(x)z \cdot V(x) - \sigma(x)z \cdot \nabla V(x)
\]
\[= H_V(x, z) + \nabla \sigma(x)z \cdot V(x) - \sigma(x)z \cdot \nabla V(x) = G_V(x, z),
\]
where \(H_V\) and \(G_V\) are defined by (3.11) and (3.12), we obtain
\[
K_{t,s} V(X_s) = V(X_t) + \int_t^s K_{t,r} [b, V](X_r)\, dr + \int_t^s \int_{|z| \leq \delta} K_{t,r} H_V(X_{s^-}, z)\tilde{N}(ds, dz)
\]
\[+ \int_t^s \int_{|z| \leq \delta} K_{t,r} G_V(X_{s^-}, z) \nu_L(dz)\, ds,
\]
which gives the desired result. \(\square\)

We have

Lemma 3.4. There exists a sequence of numbers \(\varepsilon_m \in (0, \delta) \to 0\) such that \(\mathbb{P}(\Omega_2^V) = 1\), where
\[
\Omega_2^V := \left\{ \omega : M^h_r(t) = \lim_{m \to \infty} \int_t^s \int_{|z| \leq \delta} K^h_{t,r} H_V(X^h_{s^-}, z)\tilde{N}(ds, dz) \text{ uniformly in } t \in [\lambda, T] \right\}.
\]

Proof. Since \(b\) and \(\sigma\) are linear growth, by using a standard stopping time technique, we may assume that \(\sigma\), \(V\) and \(\nabla V\) are bounded so that
\[
|H_V(x, z)| \leq ||\nabla V||_{\infty}||\sigma||_{\infty}|z| + C_d||\nabla \sigma||_{\infty}||V||_{\infty}|z|.
\]
By Doob’s maximal inequality, we have
\[
\mathbb{E}\left( \sup_{t \in [\lambda, T]} \left| M^h_V(t) - \int_{\lambda}^{T} \int_{|z| < \delta} K^h_{\lambda,s} H_V(X^h_{s-}, z) \tilde{N}(ds, dz) \right|^2 \right)
\]
\[
= \mathbb{E}\left( \sup_{t \in [\lambda, T]} \left| \int_{\lambda}^{T} \int_{|z| < \delta} K^h_{\lambda,s} H_V(X^h_{s-}, z) \tilde{N}(ds, dz) \right|^2 \right)
\]
\[
\leq 4 \mathbb{E} \left( \int_{\lambda}^{T} \int_{|z| < \delta} K^h_{\lambda,s} H_V(X^h_{s-}, z) \tilde{N}(ds, dz) \right)^2
\]
\[
\leq 4 \mathbb{E} \left( \int_{\lambda}^{T} \int_{|z| < \delta} |K^h_{\lambda,s} H_V(X^h_{s-}, z)|^2 v_L(dz) ds \right),
\]
which tends to zero as \( \varepsilon \downarrow 0 \) by (3.10) and (3.16). We complete the proof by choosing a suitable sequence \( \varepsilon_n \downarrow 0 \).

By [31] p.68, Theorem 23, we also have

**Lemma 3.5.** For \( n \in \mathbb{N} \), let \( t_k := (kT)/n \lor \lambda \). There exists a subsequence \( n_m \to \infty \) such that \( \mathbb{P}(\Omega^V_n) = 1 \), where

\[
\Omega^V_n := \left\{ \omega : \lim_{n_m \to \infty} \sum_{k=0}^{n_m-1} (M^h_V(t_{k+1} \lor t) - M^h_V(t_k \lor t))_{i,j}(M^h_V(t_{k+1} \lor t) - M^h_V(t_k \lor t))'_{j,i} \right\}
\]
\[
= \sum_{k,k'} \int_{\lambda}^{T} \int_{|z| < \delta} (K^h_{\lambda,s})_{ik} (H_V(X^h_{s-}, z))_{kj} (K^h_{\lambda,s})_{ij'} (H_V(X^h_{s-}, z))_{i'j'} \tilde{N}(ds, dz)
\]

uniformly in \( t \in [\lambda, T] \), \( \forall i, j, i', j' = 1, \ldots, d \).

Now we are in a position to prove the following main result of this section.

**Lemma 3.6.** Under (1.8), for each \( T > 0 \) and \( x \in \mathbb{R}^d \), the Malliavin covariance matrix \( \Sigma_T(x) \) defined by (2.9) is invertible almost surely.

**Proof.** Below we drop the variable “\( \chi \)”, and divide the proof into four steps.

(1) We first prove that
\[
t \mapsto J_{t,T} \text{ has a càdlàg modification on } [0, T].
\]

Let \( \tau_0 := 0 \) and define a sequence of stopping times by
\[
\tau_n := \{ t > \tau_{n-1} : |\Delta L_t| > \delta \}, \quad n \in \mathbb{N}.
\]

Recalling equation (2.8) and by (3.5), and (3.4), we have
\[
J_{t,T} = \mathbb{I} + \int_{r}^{T} \nabla b(X_s) J_{r,s} ds \lor \int_{r}^{T} \int_{|z| < \delta} \nabla \sigma(X_s) z J_{r,s} \tilde{N}(ds, dz)
\]
\[
+ \int_{r}^{T} \int_{|z| > \delta} \nabla \sigma(X_s) z J_{r,s} N(ds, dz).
\]

Thus, by (3.3) and as in Lemma 3.2, we have
\[
J_{\tau_n,t} \text{ is invertible for any } t \in [\tau_n, \tau_{n+1})
\]

15
and the inverse $K_{t_n,t} := J_{t_n,t}^{-1}$ solves

$$K_{t_n,t} = \mathbb{I} - \int_{t_n}^t K_{t_n,s} \nabla b(X_s) ds + \int_{t_n}^t \int_{\lvert z \rvert \leq \delta} K_{t_n,s} Q(X_s, z) \tilde{\mathcal{N}}(ds, dz)$$

$$- \int_{t_n}^t \int_{\lvert z \rvert \leq \delta} K_{t_n,s} Q(X_s, z) \nabla \sigma(X_s) \xi \nu_L(dz) ds.$$

In particular,

$$t \mapsto K_{t_n,t} \text{ is càdlàg on } [\tau_n, \tau_{n+1}).$$

(3.19)

Since for each $t \in [\tau_n, \tau_{n+1})$, by the uniqueness of solutions to equation (3.18),

$$J_{t_n,T} = J_{t,T} J_{t_n,t} \ a.s.,$$

we have

$$J_{t,T} = J_{t_n,T} J_{t_n,t}^{-1} = J_{t_n,T} K_{t_n,t} \ a.s.,$$

which together with (3.19) gives (3.17).

(2) In order to prove the invertibility of $\Sigma_T$, by (2.9) we only need to prove that

$$\mathbb{P} \left( \omega : u \Sigma_T u^* = \int_0^T |u_{J_{t,T}} \sigma(X_{t-})|^2 dS_t = 0, \ \exists u \in \mathbb{S}_{d-1} \right) = 0.$$

Notice that

$$\int_0^T |u_{J_{t,T}} \sigma(X_{t-})|^2 dS_t = \sum_{0 < t \leq T} |u_{J_{t,T}} \sigma(X_{t-})|^2 \Delta S_t.$$

If we set

$$\mathcal{I} := \{ t : \Delta S_t > 0 \},$$

then it suffices to prove that

$$\mathbb{P} \left( \omega : u_{J_{t,T}} \sigma(X_{t-}) = 0, \ \forall t \in \mathcal{I} \cap [0, T], \ \exists u \in \mathbb{S}_{d-1} \right) = 0,$$

which, by Lemma 2.6, (3.17) and the right continuity of $t \mapsto X_t$, is equivalent to

$$\mathbb{P} \left( \omega : u_{J_{t,T}} \sigma(X_t) = 0, \ \forall t \in [0, T], \ \exists u \in \mathbb{S}_{d-1} \right) = 0.$$

Furthermore, by (3.7) and taking regular conditional probability with respect to the large jump $\tilde{L}^\delta$, it is enough to prove that for each càdlàg $\tilde{h}$ satisfying (3.6),

$$\mathbb{P} \left( \omega : u_{J_{t,T}^h} \sigma(X_t^h) = 0, \ \forall t \in [0, T], \ \exists u \in \mathbb{S}_{d-1} \right) = 0.$$

Let $\lambda$ be defined by (3.8). By Lemma 3.2 we have for all $t \in [\lambda, T]$,

$$J^h_{\lambda,T} = J^h_{t,T} J^h_{\lambda,t} \Rightarrow J^h_{t,T} = J^h_{\lambda,T} K^h_{\lambda,t}.$$

Thus, it reduces to prove

$$\mathbb{P} \left( \omega : u_{K^h_{\lambda,t}} \sigma(X_t^h) = 0, \ \forall t \in [\lambda, T], \ \exists u \in \mathbb{S}_{d-1} \right) = 0. \quad (3.20)$$

(3) Let $B_n$ be defined by (1.7), and let $\Omega_1^{B_n}, \Omega_2^{B_n}$ and $\Omega_3^{B_n}$ be defined as in Lemmas 3.3, 3.4 and 3.5 respectively. Set

$$\tilde{\Omega} := \cap_{n \in \mathbb{N}} (\Omega_1^{B_n} \cap \Omega_2^{B_n} \cap \Omega_3^{B_n}).$$

We want to prove that for each $\omega \in \tilde{\Omega}$,

$$u_{K^h_{\lambda,t}(\omega)} \sigma(X_t^h(\omega)) = 0, \ \forall t \in [\lambda, T] \Rightarrow u = 0. \quad (3.21)$$

If this is proven, then by Lemmas 3.3, 3.4 and 3.5, we immediately obtain (3.20).
Below, for simplicity of notations, we drop “ω”. Now suppose that
\[ uK^h_{\lambda,t} \sigma(X^h_\lambda) = uK^h_{\lambda,t} B(X^h_\lambda) = 0, \quad \forall t \in [\lambda, T). \]
We shall use the induction to prove that for any \( n \in \mathbb{N}, \)
\[ uK^h_{\lambda,t} B_n(X^h_\lambda) = 0, \quad \forall t \in [\lambda, T). \tag{3.22} \]
If this is proven, then by letting \( t \downarrow \lambda, \) we obtain
\[ uB_n(X^h_\lambda) = 0, \quad \forall n \in \mathbb{N}, \]
which implies \( u = 0 \) by (1.8), and so (3.21).

(4) Suppose now that (3.22) holds for some \( n \in \mathbb{N}. \) In view of \( \omega \in \Omega^B_1, \) by Lemma 3.3 and recalling \( B_{n+1} = [b, B_n], \) we have
\[ 0 = uK^h_{\lambda,t} B_n(X^h_\lambda) = uB_n(X^h_\lambda) + \int_\lambda^t uK^h_{\lambda,s} B_{n+1}(X^h_s) \, ds + uM_{B_n}^h(t) + uR_{B_n}^h(t), \quad \forall t \in [\lambda, T). \]
From this, we have
\[ uB_n(X^h_\lambda) = 0 \]
and
\[ \int_\lambda^t uK^h_{\lambda,s} B_{n+1}(X^h_s) \, ds + uM_{B_n}^h(t) + uR_{B_n}^h(t) = 0, \quad \forall t \in [\lambda, T). \tag{3.23} \]
Since \( t \mapsto \int_\lambda^t uK^h_{\lambda,s} B_{n+1}(X^h_s) \, ds + uR_{B_n}^h(t) \) is absolutely continuous with respect to the Lebesgue measure, in view of \( \omega \in \Omega^B_3, \) by Lemma 3.5 and (3.23), we have
\[ \int_\lambda^t \int_{|z| \leq \delta} |uK^h_{\lambda,s} - H_{B_n}(X^h_{\lambda,s}, z)|^2 \tilde{N}(ds, dz) = 0, \quad \forall t \in [\lambda, T), \]
which implies that for any \( \varepsilon \in (0, \delta), \)
\[ \int_\lambda^t \int_{|z| \leq \delta} uK^h_{\lambda,s} - H_{B_n}(X^h_{\lambda,s}, z) \tilde{N}(ds, dz) = 0, \quad \forall t \in [\lambda, T). \tag{3.24} \]
On the other hand, since \( \nu_L \) is symmetric, by (3.14) and (3.12), we have for any \( t \in [\lambda, T), \)
\[ uR_{B_n}^h(t) = \lim_{\varepsilon \downarrow 0} \int_\lambda^t \int_{|z| \leq \delta} uK^h_{\lambda,s} - H_{B_n}(X^h_{\lambda,s}, z) \nu_L(dz)ds, \]
and in view of \( \omega \in \Omega^B_3, \) by Lemma 3.4
\[ uM_{B_n}^h(t) = \lim_{m \to \infty} \int_\lambda^t \int_{|z| \leq \delta} uK^h_{\lambda,s} - H_{B_n}(X^h_{\lambda,s}, z) \tilde{N}(ds, dz). \]
Thus, we obtain that for any \( t \in [\lambda, T), \)
\[ uM_{B_n}^h(t) + uR_{B_n}^h(t) = \lim_{m \to \infty} \int_\lambda^t \int_{|z| \leq \delta} uK^h_{\lambda,s} - H_{B_n}(X^h_{\lambda,s}, z) \tilde{N}(ds, dz) \tag{3.24} = 0, \]
which together with (3.23) yields
\[ \int_\lambda^t uK^h_{\lambda,s} B_{n+1}(X^h_s) \, ds = 0, \quad \forall t \in [\lambda, T). \]
By the right continuities of \( s \mapsto K^h_{\lambda,s}, X^h_s, \) we obtain (3.22). The proof is complete. \( \square \)
Moreover, for any $R > 0$, let $W_t$ be a $d$-dimensional Brownian motion and $N(ds, dz)$ a Poisson random measure over $\mathbb{R}_0^d$ with intensity measure $ds\nu(dz)$, where $\nu$ is an infinite measure on $\mathbb{R}_0^d$ with
\[
\int_{\mathbb{R}_0^d} (1 \wedge |z|^2) \nu(dz) < +\infty.
\]
Let $\tilde{N}(ds, dz) := N(ds, dz) - ds\nu(dz)$ be the compensated Poisson random measure.

Let $\mathcal{S}_m$ be the class of $m$-dimensional semimartingales with the following form:
\[
X_t = \xi_0 + \int_0^t \xi_s^0 ds + \int_0^t \xi_s^k dW_s^k + \int_0^t \int_{\mathbb{R}_0^d} \eta_s(z) \tilde{N}(ds, dz),
\]
where $\xi^k_s, k = 0, 1, \ldots, d$ and $\eta_s(z)$ are $m$-dimensional predictable processes with
\[
||X_\infty||_{\mathcal{S}_m} := \sup_{s \in [0, 1]} \left( ||X_s(\omega)||^2 \right) < \infty \text{ a.e. - } \omega.
\]
Here and below, we use the following convention: If an index appears twice in a product, then it will be summed automatically. For example,
\[
\int_0^t \xi_s^k dW_s^k := \sum_{k=1}^d \int_0^t \xi_s^k dW_s^k,
\]
For $\kappa > 1$, let $\mathcal{S}_m^\kappa$ be the subclass of $\mathcal{S}_m$ with
\[
||X_\infty||_{\mathcal{S}_m} \leq \kappa \text{ a.s. - } \omega.
\]

We first prepare the following easy result about the exponential supermartingales.

**Lemma 4.1.** Let $(X_t)_{t \geq 0}$ be a one-dimensional semimartingale in $\mathcal{S}_1$ taking form (4.1) with $\xi_0 = \xi_s^0 = 0$. If we let
\[
\mathcal{E}_t(X) := \exp \left\{ \int_0^t \xi_s^k dW_s^k + \int_0^t \int_{\mathbb{R}_0^d} \eta_s(z) \tilde{N}(ds, dz)
\right.
\]
\[
- \frac{1}{2} \int_0^t |\xi_s|^2 ds - \int_0^t \int_{\mathbb{R}_0^d} \left( (e^{\eta_s(z)} - 1 - \eta_s(z)) \nu(dz) ds \right),
\]
then for any $t \geq 0$,
\[
\mathbb{E}\mathcal{E}_t(X) \leq 1.
\]
Moreover, for any $R > 0$, we have on $|R\eta_s(z)| < 1$,
\[
\left| \int_0^t \xi_s^k dW_s^k + \int_0^t \int_{\mathbb{R}_0^d} \eta_s(z) \tilde{N}(ds, dz) - \frac{1}{R} \log \mathcal{E}_t(RX) \right|
\]
\[
\leq \frac{R}{2} \int_0^t |\xi_s|^2 ds + 2R \int_0^t \int_{\mathbb{R}_0^d} |\eta_s(z)|^2 \nu(dz) ds.
\]

**Proof.** By Itô’s formula, we have
\[
\mathcal{E}_t(X) = 1 + \int_0^t \mathcal{E}_s(X) \xi_s^k dW_s^k + \int_0^t \int_{\mathbb{R}_0^d} \mathcal{E}_{s-}(X)(e^{\eta_s(z)} - 1) \tilde{N}(ds, dz).
\]
Hence,
\[
t \mapsto \mathcal{E}_t(X) \text{ is a positive local martingale (and also a supermartingale),}
\]
which then implies (4.2). As for (4.3), it follows by
\[ |e^x - x - 1| \leq 2x^2, \ \forall |x| \leq 1, \]
and a direct calculation.

We are now in a position to prove the following Komatsu-Takeuchi’s estimate.

**Theorem 4.2.** For \( \kappa \geq 1 \), let \((f_t)_{t \geq 0}\) and \((f_t^0)_{t \geq 0}\) be two \( m \)-dimensional semimartingales in \( \mathcal{S}_m^\kappa \) given by
\[
f_t = f_0 + \int_0^{t \wedge T} f^k_s ds + \int_0^t f^k_s dW^k_s + \int_0^t \int_{|\xi| \leq \delta} g_s(\xi) \tilde{N}(ds, dz),
\]
\[
f_t^0 = f_0^0 + \int_0^{t \wedge T} f^0_s ds + \int_0^t f^0_s dW^0_s + \int_0^t \int_{|\xi| \leq \delta} g_s^0(\xi) \tilde{N}(ds, dz),
\]
where \( \delta \in (0, 1] \) and \( T \) is a stopping time. For any \( \varepsilon, T \in (0, 1], \) there exist positive random variables \( \zeta_1 \) and \( \zeta_2 \) with \( \mathbb{E}\zeta_1 \leq 1, \mathbb{E}\zeta_2 \leq 1 \) such that
\[
c_1 \int_0^T \left\{ |f_t|^2 + \int_{|\xi| \leq \delta} |g_s(\xi)|^2 \nu(d\xi) \right\} dt \leq (\delta^{-1} + \varepsilon^{-1}) \int_0^T |f_t|^2 dt + \kappa \delta \log \zeta_1 + \kappa(\varepsilon + T\delta) \quad (4.4)
\]
and
\[
c_2 \int_0^T |f_t^0|^2 dt \leq (\delta^{-\frac{1}{2}} + \varepsilon^{-\frac{1}{2}}) \int_0^T |f_t^0|^2 dt + \kappa \delta^{\frac{1}{2}} \log \zeta_2 + \kappa(\varepsilon^{-\frac{1}{2}} + \varepsilon^{\frac{1}{2}} + T\delta^{\frac{1}{2}}), \quad (4.5)
\]
where \( c_1, c_2 \in (0, 1) \) only depends on \( \int_{|x| \leq 1} |x|^2 \nu(dz) \).

**Proof.** First of all, if \( \varepsilon \gg T \), then in view of \( f_t \in \mathcal{S}_m^\kappa \),
\[
\int_0^T \left\{ |f_t|^2 + \int_{|\xi| \leq \delta} |g_s(\xi)|^2 \nu(d\xi) \right\} dt \leq \kappa T \left( 1 + \int_{|x| \leq 1} |x|^2 \nu(dz) \right) \leq C\kappa T,
\]
and
\[
\int_0^T |f_t^0|^2 dt \leq \kappa T \leq C\kappa T,
\]
and hence, (4.4) and (4.5) hold with \( \zeta_1 = \zeta_2 \equiv 1 \). Below, by replacing \((f_t, f_t^k, g_s(\xi))\) and \((f_t^0, g_s^0(\xi))\) with \((f_t, f_t^k, g_s(\xi))/\sqrt{\kappa} \) and \((f_t^0, g_s^0(\xi))/\sqrt{\kappa} \), we shall assume
\[
\kappa = 1, \ \varepsilon \in (0, T).
\]
In particular, we have
\[
|f_t|^2 \vee |f_t^0|^2 \vee |f_t^k|^2 \vee |f_t^0k|^2 \vee |f_t^0k|^2 \vee \sup_{z \in \mathbb{R}^d} \frac{|g_s(\xi)|^2}{1 \wedge |x|^2} \vee \sup_{z \in \mathbb{R}^d} \frac{|g_s^0(\xi)|^2}{1 \wedge |x|^2} \leq 1. \quad (4.6)
\]
We divide the proof into two steps.

(1) For \( t \in [0, T - \varepsilon] \), by Itô’s formula, we have
\[
|f_{t+\varepsilon}|^2 = |f_t|^2 + 2 \int_t^{t+\varepsilon} \langle f_s, 1_{s \leq T} f_s^0 \rangle ds + 2 \int_t^{t+\varepsilon} \langle f_s, f_s^k \rangle dW_s^k + \int_t^{t+\varepsilon} |f_s^k|^2 ds
\]
\[
+ \int_t^{t+\varepsilon} \int_{|\xi| \leq \delta} \left( |f_s^- + g_s(\xi)|^2 - |f_s^-|^2 \right) \tilde{N}(ds, dz) + \int_t^{t+\varepsilon} \int_{|\xi| \leq \delta} |g_s(\xi)|^2 \nu(dz) ds.
\]
Integrating both sides from 0 to \( T - \varepsilon \) and using Fubini’s theorem, we obtain
\[
\int_0^{T-\varepsilon} |f_{t+\varepsilon}|^2 dt \geq 2 \int_0^T \varepsilon_s \langle f_s, 1_{s \leq T} f_s^0 \rangle ds + \int_0^T \varepsilon_s \left( |f_s^k|^2 + \int_{|\xi| \leq \delta} |g_s(\xi)|^2 \nu(dz) \right) ds.
\]
For Here and below, the constant \( \varepsilon_s := (T - \varepsilon) \land s - 0 \lor (s - \varepsilon) \)
and
\[ \xi^{e,k}_s := 2\varepsilon_s(f_s, f_s^k), \quad \eta^e_s(z) := \varepsilon_s(|f_{s^0} + g_s(z)|^2 - |f_{s^0}|^2). \]
For \( I_1^e \), we have
\[ |I_1^e| \leq 2\varepsilon \int_0^T |f_s||f_s^0| ds \leq \int_0^T |f_s|^2 ds + \varepsilon^2 \int_0^T |f_s^0|^2 ds \leq \int_0^T |f_s|^2 ds + T\varepsilon^2. \] (4.7)

For \( I_2^e \), noticing that
\[ |\varepsilon_s - \varepsilon| \leq \varepsilon(1_{(0,\varepsilon)}(s) + 1_{(T - \varepsilon, T)}(s)), \quad s \in [0, T], \]
one easily sees that
\[ I_2^e \overset{4.6}{\geq} \varepsilon \int_0^T \left\{ |f_s^k|^2 + \int_{|z| \leq \delta} |g_s(z)|^2 \nu(dz) \right\} ds - 2 \left( 1 + \int_{|z| \leq 1} |z|^2 \nu(dz) \right) \varepsilon^2. \] (4.9)

For \( I_3^e \), notice that
\[ |\eta^e_s(z)|_{1_{|z| \leq \delta}} \leq \varepsilon(2|f_{s^0}||g_s(z)| + |g_s(z)|^2)_{1_{|z| \leq \delta}} \overset{4.6}{\leq} 3\varepsilon \delta. \]
Thus, by (4.3) with \( R = \frac{1}{3\varepsilon \delta} \), there exists a positive random variable \( M_1^{e,\delta} \) with \( \mathbb{E}M_1^{e,\delta} \leq 1 \) such that
\[ -I_3^e \leq 3\varepsilon \delta \log M_1^{e,\delta} + C(\varepsilon \delta)^{-1} \left( \int_0^T |\xi^{e,k}_s|^2 ds + \int_0^T |\eta^e_s(z)|^2 \nu(dz) ds \right) \]
\[ \overset{4.6}{\leq} 3\varepsilon \delta \log M_1^{e,\delta} + C\varepsilon \delta^{-1} \left( \int_0^T |f_s|^2 ds + \int_0^T \int_{|z| \leq \delta} (|f_s|^2 + |z|^4)^2 \nu(dz) ds \right) \]
\[ \leq 3\varepsilon \delta \log M_1^{e,\delta} + C\varepsilon \delta^{-1} \int_0^T |f_s|^2 ds + CT\varepsilon \delta. \] (4.10)

Here and below, the constant \( C \geq 1 \) only depends on \( \int_{|z| \leq 1} |z|^2 \nu(dz) \). Combining (4.7)--(4.10), we obtain
\[ \int_0^T \left\{ |f_s^k|^2 + \int_{|z| \leq \delta} |g_s(z)|^2 \nu(dz) \right\} ds \]
\[ \leq 3\delta \log M_1^{e,\delta} + C_1(\delta^{-1} + \varepsilon^{-1}) \int_0^T |f_s|^2 ds + C_1(\varepsilon + T\delta), \] (4.11)
which gives (4.4) with \( \kappa = 1 \) by setting \( c_1 := \frac{1}{3C_1} \) and \( \zeta_1 := (M_1^{e,\delta})^{\frac{1}{\kappa+1}}. \)

(2) As above, by Itô’s product formula, we have
\[
\langle f_{t+\varepsilon}, f_{t+\varepsilon}^0 \rangle = \langle f_t, f_t^0 \rangle + \int_t^{t+\varepsilon} \langle f_s, f_s^0 \rangle ds + \int_t^{t+\varepsilon} \langle f_s, f_s^0 \rangle dW_s^k
\]
\[ + \int_t^{t+\varepsilon} \int_{|z| \leq \delta} \langle f_{s^{-}}, g_s^0(z) \rangle \tilde{N}(ds, dz) + \int_t^{t+\varepsilon} 1_{s \leq |f_s|^2} ds \]
\[ + \int_t^{t+\varepsilon} \langle f_s^k, f_s^0 \rangle dW_s^k + \int_t^{t+\varepsilon} \langle g_s(z), f_s^0 \rangle \tilde{N}(ds, dz) \]
Integrating both sides from $0$ to $T - \varepsilon$ and using Fubini’s theorem again, we get
\[
\int_{0}^{T-\varepsilon} (\langle f_{i+\varepsilon}, f_{i+\varepsilon}^{0} \rangle - \langle f_{i}, f_{i}^{0} \rangle) dt = \left\{ \int_{0}^{T} \hat{\xi}_{s}^{\varepsilon,k} dW_{s}^{k} + \int_{0}^{T} \hat{\eta}_{s}^{\varepsilon}(z) d\mathcal{N}(ds, dz) \right\} \\
+ \int_{0}^{T} \varepsilon_{s} \left\{ \langle f_{s}, f_{s}^{0} \rangle + \langle f_{s}^{k}, f_{s}^{0} \rangle + \int_{|s| \leq \delta} (g_{s}(z), g_{s}^{0}(z)) \nu(dz) \right\} ds \\
+ \int_{0}^{T} \varepsilon_{s} 1_{s \leq \varepsilon} |f_{s}^{0}|^2 ds =: J_{1}^{\varepsilon} + J_{2}^{\varepsilon} + J_{3}^{\varepsilon}, \tag{4.12}
\]
where
\[
\hat{\xi}_{s}^{\varepsilon,k} := \varepsilon_{s} \left\{ \langle f_{s}, f_{s}^{0} \rangle + \langle f_{s}^{k}, f_{s}^{0} \rangle \right\}, \quad \hat{\eta}_{s}^{\varepsilon}(z) := \varepsilon_{s} \left\{ \langle f_{s}, g_{s}^{0}(z) \rangle + \langle g_{s}(z), f_{s}^{0} + g_{s}^{0}(z) \rangle \right\}.
\]
For $J_{1}^{\varepsilon}$, notice that
\[
|\hat{\eta}_{s}^{\varepsilon}(z)|_{|s| \leq \delta} \leq \varepsilon_{|f_{s} - \nu g_{s}(z) + |g_{s}(z)(|f_{s}^{0}| + |g_{s}^{0}(z)|)|_{|s| \leq \delta}} \overset{\text{(4.6)}}{\leq} 3\varepsilon \delta \leq 3\varepsilon \delta^{\frac{1}{2}}.
\]
Thus, by (4.3) with $R = \frac{1}{3\varepsilon \delta^{\frac{1}{2}}}$, there exists a positive random variable $M_{2}^{\varepsilon, \delta}$ with $\mathbb{E}M_{2}^{\varepsilon, \delta} \leq 1$ such that
\[
\begin{align*}
-J_{1}^{\varepsilon} \leq 3\varepsilon \delta \frac{1}{2} \log M_{2}^{\varepsilon, \delta} + C(\varepsilon \delta \frac{1}{2})^{-1} \left\{ \int_{0}^{T} |\hat{\xi}_{s}^{\varepsilon,k}|^2 ds + \int_{0}^{T} |\hat{\eta}_{s}^{\varepsilon}(z)|^2 \nu(dz) ds \right\} \\
\overset{\text{(4.6)}}{\leq} 3\varepsilon \delta \frac{1}{2} \log M_{2}^{\varepsilon, \delta} + C\varepsilon \delta^{-\frac{1}{2}} \int_{0}^{T} \left\{ |f_{s}|^2 + |f_{s}^{k}|^2 + \int_{|s| \leq \delta} |g_{s}(z)|^2 \nu(dz) ds \right\} ds \tag{4.13}
\end{align*}
\]
For $J_{2}^{\varepsilon}$, by Young’s inequality and (4.6), we have
\[
|J_{2}^{\varepsilon}| \leq \varepsilon \int_{0}^{T} \left\{ |f_{s}| |f_{s}^{0}| + |f_{s}^{k}| |f_{s}^{0}| \right\} ds + \varepsilon \int_{0}^{T} \int_{|s| \leq \delta} |g_{s}(z)||g_{s}^{0}(z)| \nu(dz) ds \\
\leq \varepsilon \delta^{-\frac{1}{2}} \int_{0}^{T} \left\{ |f_{s}|^2 + |f_{s}^{k}|^2 + \int_{|s| \leq \delta} |g_{s}(z)|^2 \nu(dz) ds \right\} ds + C\varepsilon \delta^{-\frac{1}{2}}
\]
as well as
\[
\left| \int_{0}^{T-\varepsilon} \{ \langle f_{i+\varepsilon}, f_{i+\varepsilon}^{0} \rangle - \langle f_{i}, f_{i}^{0} \rangle \} dt \right| \leq \left( \int_{0}^{T-\varepsilon} + \int_{0}^{\varepsilon} \right) |f_{i} f_{i}^{0}| dt \\
\leq \varepsilon^{-\frac{1}{2}} \int_{0}^{T} |f_{s}|^2 dt + C\varepsilon^{-\frac{1}{2}} \left( \int_{0}^{T-\varepsilon} + \int_{0}^{\varepsilon} \right) |f_{i} f_{i}^{0}|^2 dt \leq \varepsilon^{-\frac{1}{2}} \int_{0}^{T} |f_{s}|^2 dt + C\varepsilon^{-\frac{1}{2}}. \tag{4.14}
\]
For $J_{3}^{\varepsilon}$, it is similar to (4.9) that
\[
J_{3}^{\varepsilon} \geq \varepsilon \int_{0}^{T} 1_{s \leq \varepsilon} |f_{s}^{0}|^2 ds - 2\varepsilon^2 = \varepsilon \int_{0}^{T} |f_{s}^{0}|^2 ds - 2\varepsilon^2. \tag{4.15}
\]
Combining (4.12)–(4.14), we obtain
\[
\int_{0}^{T} |f_{s}^{0}|^2 ds \leq 3\delta \frac{1}{2} \log M_{2}^{\varepsilon, \delta} + C(\delta^{-\frac{1}{2}} + \varepsilon^{-\frac{1}{2}}) \int_{0}^{T} |f_{s}|^2 ds + 2\varepsilon + C\delta \frac{1}{2} + C\varepsilon^{\frac{1}{2}} + C_{0} \delta^{-\frac{1}{2}} \int_{0}^{T} \left\{ |f_{s}^{k}|^2 + \int_{|s| \leq \delta} |g_{s}(z)|^2 \nu(dz) \right\} ds
\]
Remark 4.5. This type of Norris’ estimate was proven by Cass in [10] with \( \kappa = 1 \) by suitable choices of \( c_2 \) and \( \zeta_2 \) as in (1). The proof is thus complete. \( \square \)

**Remark 4.3.** The factor \( T \) before \( \delta^2 \) in (4.5) is quite important for the proof of Lemma 5.1 below. Meanwhile, the freedom of \( \varepsilon \) allows it to be any positive power of \( \delta \) so that for any \( m \geq 1 \) and \( T \in (0, 1) \),

\[
\frac{c_0}{\varepsilon^m} \int_0^T |f_t|^2 \, dt \leq \delta^{-\frac{2m}{m+1}} \left( \int_0^T |f_t|^2 \, dt + \delta^\frac{2}{2} \log \zeta_2 + (\delta^\frac{2}{2} + T \delta^2) \right).
\]

Moreover, the role of stopping time \( \tau \) will become clear in the following corollary.

We have the following corollary.

**Corollary 4.4.** Let \( (f_t)_{t \geq 0} \) and \( (f_t^0)_{t \geq 0} \) be two \( m \)-dimensional semimartingales in \( \mathcal{M}_m \) as given in Theorem 4.2. For \( \delta \in (0, 1) \) and \( \gamma \in (0, \frac{1}{\delta}) \), define a stopping time:

\[
\tau := \inf \left\{ t \geq 0 : |f_t|^2 \lor |f_t^0|^2 \lor |f_t^1|^2 \lor |f_t^{0,1}|^2 \lor |f_t^{1,0}|^2 \lor \sup_{z \in \mathbb{R}^d} \frac{|g_t(z)|^2}{1 \lor |z|^2} \lor \sup_{z \in \mathbb{R}^d} \frac{|g_t^0(z)|^2}{1 \lor |z|^2} \geq \delta^{-\gamma} \right\}.
\]

For any \( \delta, T \in (0, 1) \) and \( \gamma \in (0, \frac{1}{\delta}) \), we have

\[
\mathbb{P} \left\{ \int_0^T |f_t|^2 \, dt \leq \delta^{\frac{1}{2} - \gamma}, \int_0^T |f_t^0|^2 \, dt \geq \delta^{\frac{1}{2} + \gamma}; \tau \geq T \right\} \leq e^{5 - c_0 \delta^\frac{1}{2}}.
\]

**Proof.** In Theorem 4.2 if we replace \( f_t, f_t^1, g_t(z) \) and \( f_t^0, f_t^{0,1}, g_t^0(z) \) by \( f_{1,t}, f_{1,t}^1, f_{1,t}^0, g_{1,t} \) and \( f_{1,t}^0, f_{1,t}^{0,1}, f_{1,t}^{1,0} \) respectively, then by (4.5) with \( \varepsilon = \delta \) and \( \kappa = \delta^{-\gamma} \), we obtain

\[
c_0 \int_0^T |f_t^0|^2 \, dt \leq 2 \delta^{-\frac{2}{2}} \int_0^T |f_t^{0,1}|^2 \, dt + \delta^{\frac{2}{2} - \gamma} \log \zeta_2 + 3 \delta^{\frac{1}{2} - \gamma}.
\]

In particular,

\[
\left\{ \int_0^T |f_t|^2 \, dt \leq \delta^{\frac{1}{2} - \gamma}, \int_0^T |f_t^0|^2 \, dt \geq \delta^{\frac{1}{2} + \gamma}; \tau \geq T \right\} \subset \left\{ \zeta_2 \geq e^{6 \delta^{-\frac{1}{2}} - 5} \right\},
\]

which gives the desired estimate by Chebyshev’s inequality and \( \mathbb{E} \zeta_2 \leq 1 \). \( \square \)

**Remark 4.5.** This type of Norris’ estimate was proven by Cass in [10]. Our proof is simpler, and the dependence of constants are more explicit.

### 5. Proof of Theorem 1.3

In this section we assume that

\[
\sigma(x) = \sigma \text{ is constant and } b \in C^\infty_b(\mathbb{R}^d).
\]

Let us consider the following SDE:

\[
X_t(x) = x + \int_0^t b(X_s(x)) \, ds + \sigma W_s.
\]

Let \( J_t(x) = \nabla X_t(x) \) be the Jacobian matrix of \( X_t(x) \). It is easy to see that

\[
J_t(x) = I + \int_0^t \nabla b(X_s(x)) J_s(x) \, ds.
\]
Let $K_t(x)$ be the inverse matrix of $J_t(x)$. Then
\[
K_t(x) = I - \int_0^t K_s(x) \nabla b(X_s(x)) ds.
\]
In this case, since $J_s(x) = J_t(x)K_s(x)$, the Malliavin covariance matrix $\Sigma_t(x)$ given by (2.9) can be written as:
\[
\Sigma_t(x) = J_t(x) \left( \int_0^t K_s(x) \sigma \sigma^T K_s(x) ds \right) J_t^T(x).
\] (5.1)

The core task of this section is to prove the $L^p$-integrability of the inverse of $\Sigma_t(x)$.

Below, for simplicity of notations, we shall drop the variable “$x$” unless necessary. We first prove the following key lemma.

**Lemma 5.1.** Let $V : \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^d$ be a bounded smooth function with bounded derivatives of all orders. Under (AS$^p$), for any $\beta \in (0 \vee (4\alpha - 7), 1)$, there exist two constants $C_1 \geq 1$ and $c_1 \in (0, 1)$ only depending on $b$, $V$ and $\alpha, \beta, \nu_L$ such that for all $\delta, t \in (0, 1)$,
\[
\sup_{|x|=1} \mathbb{P} \left( \int_0^t |uK_s[b, V](X_s)|^2 ds \geq t^{1-\beta}, \quad \int_0^t |uK_s V(X_s)|^2 ds \leq t^{\delta - 2} \right) \leq C_1 \exp(-c_1 t^\delta),
\] (5.2)
where $[b, V] := b \cdot \nabla V - V \cdot \nabla b$.

**Proof.** We divide the proof into three steps.

(1) Let $h$ be a càdlàg purely discontinuous $\mathbb{R}^d$-valued function with finitely many jumps and $h_0 = 0$. Let $\ell$ be another càdlàg purely discontinuous $\mathbb{R}^d$-valued function with $\ell_0 = 0$. Let $X_t(x; \ell, h)$ solve the following equation:
\[
X_t(x; \ell, h) = x + \int_0^t b(X_s(x; \ell, h)) ds + \sigma \ell_t + \sigma h_t.
\]
Let $K_t(x; \ell, h)$ solve the following matrix-valued ODE:
\[
K_t(x; \ell, h) = I - \int_0^t K_s(x; \ell, h) \nabla b(X_s(x; \ell, h)) ds.
\] (5.3)
By Gronwall’s inequality, it is clear that
\[
|K_t(x; \ell, h)| \leq e^{\|Vh\|_\infty}, \quad \forall t \in [0, 1].
\] (5.4)

For $\delta \in (0, 1)$, let $L_\delta^\circ$ and $\tilde{L}_\delta^\circ$ be defined as in (5.3). Write
\[
X_t^\circ(x) := X_t(x; L_\delta^\circ, 0), \quad K_t^\circ(x) := K_t(x; L_\delta^\circ, 0).
\]
Then $X_t^\circ(x)$ solves the following SDE:
\[
X_t^\circ(x) = x + \int_0^t b(X_s^\circ(x)) ds + \int_{|z| \leq \delta} \sigma z \tilde{N}((0, t], dz).
\] (5.5)
Define
\[
V_0(x) := [b, V](x) + \int_{|z| \leq \delta} (V(x + \sigma z) - V(x) - \sigma z \cdot \nabla V(x)) \nu_L(dz),
\]
\[
V_1(x) := [b, V_0](x) + \int_{|z| \leq \delta} (V_0(x + \sigma z) - V_0(x) - \sigma z \cdot \nabla V_0(x)) \nu_L(dz).
\]
For any $u \in \mathbb{R}^d$ with $|u| \leq e^{\|Vh\|_\infty}$, set
\[
f_t := uK_t^\circ V(X_t^\circ), \quad g_t(z) := uK_t^\circ(V(X_t^\circ + \sigma z) - V(X_t^\circ))
\]
and
\[
f_t^0 := uK_t^\circ V_0(X_t^\circ), \quad f_t^0 := uK_t^\circ V_1(X_t^\circ), \quad g_t^0(z) := uK_t^\circ(V_0(X_t^\circ + \sigma z) - V_0(X_t^\circ)).
\]
By equations (5.3), (5.5) and using Itô’s formula (see (3.15)), we have

\[ f_t = uV(x) + \int_0^t uK^0_s(b, V)(X^0_s)ds + \int_0^t \int_{|z| \leq \delta} g_{s-}(z)\tilde{N}(ds, dz) \]

\[ + \int_0^t \int_{|z| > \delta} uK^0_s(V(X^0_s + \sigma z) - V(X^0_s)) - \sigma z \cdot \nabla V(X^0_s))\nu_L(dz)ds \]

\[ = uV(x) + \int_0^t f^0_t ds + \int_0^t \int_{|z| \leq \delta} g_{s-}(z)\tilde{N}(ds, dz) \]

and

\[ f^0_t = uV_0(x) + \int_0^t f^{00}_s ds + \int_0^t \int_{|z| \leq \delta} g_{s-}(z)\tilde{N}(ds, dz). \]

Since \( V \) has bounded derivatives of all orders, by (5.4) and \( |u| \leq e^{\|V\|_{L^\infty}} \), there exists a constant \( \kappa \geq 1 \) only depending on \( b \) and \( V \) such that for all \( t \in [0, 1] \) and \( z \in \mathbb{R}^d \),

\[ |f_t^2|, |f_t^{02}|, |f_t^{002}| \leq \kappa, \quad |g_t(z)^2|, |g_t(z)|^2 \leq \kappa(1 + |z|^2). \]

By (4.16) with \( m = 5 \), for any \( t \in (0, 1) \), there exists a positive random variable \( \zeta \) with \( \mathbb{E}\zeta \leq 1 \) such that

\[ 2c_0 \int_0^t |f_s^0|^2 ds \leq \delta^{-\frac{\kappa}{2}} \int_0^t |f_s^1|^2 ds + \delta^{\frac{\kappa}{2}} \log \zeta + (\delta^{\frac{\kappa}{2}} + t\delta^{\frac{\kappa}{2}}) \quad \text{a.s.,} \]

where \( c_0 \in (0, 1) \) only depends on \( \kappa \) and \( \int_{|z| \leq 1} |z|^2 \nu_L(dz) \). From this, dividing both sides by \( \delta^{\frac{\kappa}{2}} \) and taking exponential and expectations, we derive that

\[ \mathbb{E} \left\{ \exp \left( 2c_0\delta^{-\frac{\kappa}{2}} \int_0^t |f_s^0|^2 ds - \delta^{-8} \int_0^t |f_s^1|^2 ds \right) \right\} \leq \exp(\delta^2 + t). \quad (5.6) \]

Recalling the definition of \( f^0_t \) and by \( |x + y|^2 \geq \frac{|x|^2}{2} - |y|^2 \), we have

\[ |f^0_t|^2 \geq \frac{|uK^0_s[b, V](X^0_s)|^2}{2} - e^{4\|V\|_{L^\infty}} \|\nabla^2 V\|_{L^\infty} |\sigma|^4 \left( \int_{|z| \leq \delta} |z|^2 \nu_L(dz) \right)^2 \geq \frac{2c_0}{2} \frac{|uK^0_s[b, V](X^0_s)|^2}{2} - C_2\delta^{4-2\sigma}. \]

Thus, by (5.6), we obtain that for all \( \delta, t \in (0, 1) \),

\[ \sup_{|u| \leq \ell} \mathbb{E} \left\{ \exp \left( c_0\delta^{-\frac{\kappa}{2}} \int_0^t |uK^0_s[b, V](X^0_s)|^2 ds - \delta^{-8} \int_0^t |uK^0_s|V(X^0_s)|^2 ds \right) \right\} \leq \exp \left( t(C_3\delta^{\frac{\kappa}{2}-2\sigma} + 1) \right), \quad (5.7) \]

where \( C_3 := 2C_2c_0 \).

(2) Let \( h \) be a càdlàg purely discontinuous \( \mathbb{R}^d \)-valued function with finitely many jumps and \( h_0 = 0 \). Write for \( \delta \in (0, 1) \),

\[ X^0_t(x; h) := X_t(x; L^\delta, h), \quad K^0_t(x; h) := K_t(x; L^\delta, h) \]

and for \( u \in \mathbb{R}^d \),

\[ J^0_t(u, x; h) := \exp \left\{ c_0\delta^{-\frac{\kappa}{2}} \int_0^t |uK^0_s(x; h)[b, V](X^0_s(x; h))|^2 ds \right\} \]

\[ - \delta^{-8} \int_0^t |uK^0_s(x; h)V(X^0_s(x; h))|^2 ds \right\}. \quad (5.8) \]
By (5.7), we have

$$
\sup_{x \in \mathbb{R}^d} \sup_{|u| \leq |v|} \mathbb{E} \mathcal{J}_t^\delta (u, x; 0) \leq \exp \left\{ \delta^2 + t (C_3 \delta^{-2 \alpha - 2} + 1) \right\}.
$$

(5.9)

For $t \in (0, 1)$, let $n$ be the jump number of $\delta h$ before time $t$. Let $0 = t_0 < t_1 < \cdots < t_n \leq t = t_{n+1}$ be the jump time of $\delta h$. It is easy to see that for any $s \in [0, t_{j+1} - t_j)$,

$$
X_{s+t_j}^\delta (x; \delta h) = X_s (X_t^\delta (x; \delta h); \partial_t L_s^\delta, 0)
$$

(5.10)

and

$$
K_{s+t_j}^\delta (x; \delta h) = K_s^\delta (x; \delta h) \cdot K_s (X_t^\delta (x; \delta h); \partial_t L_s^\delta, 0),
$$

(5.11)

where

$$
(\partial_t L_s^\delta := L_{s+t_j}^\delta - L_t^\delta)_{s \geq 0} \overset{(d)}{=} (L_s^\delta)_{s \geq 0}.
$$

(5.12)

Noticing that for $j = 1, \cdots, n$,

$$
\mathcal{J}_t^\delta (u, x; \delta h), K_s^\delta (x; \delta h), X_s^\delta (x; \delta h)
$$

are independent of $X (\cdot; \partial_t L_s^\delta, 0)$ and $K (\cdot; \partial_t L_s^\delta, 0)$

and

$$
X (\cdot; \partial_t L_s^\delta, 0) \overset{(d)}{=} X_s^\delta (\cdot; 0), \quad K (\cdot; \partial_t L_s^\delta, 0) \overset{(d)}{=} K_s^\delta (\cdot; 0),
$$

by (5.10) and (5.11), we obtain that for $|u| = 1$,

$$
\mathbb{E} \mathcal{J}_{t_{j+1}}^\delta (u, x; \delta h) = \mathbb{E} \left( \mathcal{J}_t^\delta (u, x; \delta h) \cdot \left( \mathbb{E} \mathcal{J}_{t_{j+1} - t_j}^\delta (u', y; 0) \right) \right)_{u' = u K_{t_n}^\delta (x; \delta h), y = X_{t_n}^\delta (x; \delta h)}
$$

\begin{align*}
& \leq \exp \left\{ \delta^2 + (t_{n+1} - t_n) (C_3 \delta^{-2 \alpha - 2} + 1) \right\} \mathbb{E} \mathcal{J}_{t_n}^\delta (u, x; \delta h) \\
& \leq \cdots \cdots \\
& \leq \Pi_{j=0}^n \exp \left\{ \delta^2 + (t_{j+1} - t_j) (C_3 \delta^{-2 \alpha - 2} + 1) \right\} \\
& = \exp \left\{ \delta^2 (n + 1) + t_{n+1} (C_3 \delta^{-2 \alpha - 2} + 1) \right\}.
\end{align*}

(5.13)

Let $N_t^\delta$ be the jump number of $\hat{L}_s^\delta$ before time $t$, i.e.,

$$
N_t^\delta = \sum_{x \in (0, t]} 1_{|\Delta L_s^\delta| > 0} = \int_{[0, t]} N((0, t], dz) = \sum_{x \in (0, t]} 1_{|\Delta L_s^\delta| > \delta},
$$

which is a Poisson process with intensity $\int_{[0, t]} \nu_L (dz) =: \lambda_\delta$. Observing that

$$
X (x) = X_t^\delta (x; \hat{L}_s^\delta), \quad K (x) = K_t^\delta (x; \hat{L}_s^\delta),
$$

and recalling (5.8) and the independence of $L_s^\delta$ and $\hat{L}_s^\delta$, we have for any $|u| = 1$,

$$
\mathbb{E} \left( \exp \left\{ c_0 \delta^{-\frac{1}{2}} \int_0^t |u K_s[b, V] (X_s)|^2 ds - \delta^{-8} \int_0^t |u K_s, V (X_s)|^2 ds \right\} \right)
$$

\begin{align*}
& \leq \sum_{n=0}^\infty \exp \left\{ \delta^2 (n + 1) + t (C_3 \delta^{-2 \alpha - 2} + 1) \right\} \mathbb{P}(N_t^\delta = n) \\
& = \exp \left\{ \delta^2 + t (C_3 \delta^{-2 \alpha - 2} + 1) \right\} \sum_{n=0}^\infty \frac{e^{\delta n} (t \lambda_\delta)^n}{n!} e^{-t \lambda_\delta} \\
& = \exp \left\{ \delta^2 + t (C_3 \delta^{-2 \alpha - 2} + 1) + (\delta^2 - 1) t \lambda_\delta \right\}
\end{align*}
\[2.18\]
\[\leq \exp\left\{2 + C_4 t \delta^{2-2\alpha} + C_5 t \delta^{2-\alpha}\right\}, \quad \forall t, \delta \in (0, 1), \]

where in the last step we have used that \(e^x - 1 \leq 3x\) for \(x \in (0, 1)\).

(3) By Chebyshev’s inequality, we have for any \(\beta \in (0, 1)\),
\[
\sup_{|d|=1} \mathbb{P}\left\{ c_0 \delta^{\frac{3}{2}} \int_0^{\bar{\lambda}} |uK_s[b, V](X_s)|^2 ds - \delta^{-8} \int_0^{\bar{\lambda}} |uK_sV(X_s)|^2 ds \geq t \delta^{-\frac{\beta}{2}} \right\}
\leq \exp\left\{2 + C_4 t \delta^{2-2\alpha} + C_5 t \delta^{2-\alpha} - t \delta^{-\frac{\beta}{2}} \right\}.
\]

In particular, if \(\beta \in (0 \lor (4\alpha - 7), 1)\), then there exists a constant \(\delta_0\) such that for all \(\delta \in (0, \delta_0)\) and \(t \in (0, 1)\),
\[
\sup_{|d|=1} \mathbb{P}\left\{ \int_0^{\bar{\lambda}} |uK_s[b, V](X_s)|^2 ds \geq \frac{2t \delta^{\frac{3}{2}}}{c_0}, \quad \int_0^{\bar{\lambda}} |uK_sV(X_s)|^2 ds \leq t \delta^{-8-\frac{\beta}{2}} \right\} \leq \exp(3 - t \delta^{-\frac{\beta}{2}}),
\]

which then gives the desired estimate by adjusting the constants and rescaling \(\delta\).

\[\square\]

**Lemma 5.2.** Let \(t_0 := 1 \land \left(\frac{e^{\frac{1}{2}\|V\|_\infty}}{2\|b\|_\infty}\right)\). Under (1.10) and (AS\(_R^\alpha\)), there exist \(\gamma = \gamma(\alpha) \in (0, 1), C_2 \geq 1\) and \(c_2 \in (0, 1)\) such that for all \(t \in (0, t_0)\) and \(\lambda \geq 1\),
\[\sup_{|d|=1} \mathbb{E} \exp\left\{ -\lambda \int_0^\infty |uK_s(x)|^2 ds \right\} \leq C_2 \exp(-c_2 t \lambda^\gamma). \tag{5.13}\]

**Proof.** For fixed \(\beta \in (0 \lor (4\alpha - 7), 1)\), set \(a := \frac{1-\beta}{18-\beta}\) and define for \(j = 1, \cdots, j_0\),
\[
E_j := \left\{ \int_0^{\bar{\lambda}} |uK_sB_j(X_s)|^2 ds \leq \delta^{\frac{3}{2} \alpha^j(9-\frac{\beta}{2})} \right\}.
\]

Since \(a^{j+1}(9 - \frac{\beta}{2}) = a^{j+1}(\frac{1-\beta}{2})\) and \(B_{j+1} = [b, B_j]\), by (5.2) with \(\delta\) replaced by \(\delta^\alpha\), we have
\[\mathbb{P}(E_j \cap E_{j+1}^c) \leq C_1 \exp(-c_1 t \delta^{\frac{\alpha^j(9-\frac{\beta}{2})}{2}}). \tag{5.14}\]

Noticing that
\[E_1 \subset \bigcap_{j=1}^{j_0} E_j \cup \bigcup_{j=1}^{j_0-1} \left( E_j \cap E_{j+1}^c \right),\]
we have
\[\mathbb{P}(E_1) \leq \mathbb{P}\left( \bigcap_{j=1}^{j_0} E_j \right) + \sum_{j=1}^{j_0-1} \mathbb{P}(E_j \cap E_{j+1}^c). \tag{5.15}\]

On the other hand, noticing that for \(t \leq 1 \land \left(\frac{e^{\frac{1}{2}\|V\|_\infty}}{2\|b\|_\infty}\right) =: t_0\),
\[\inf_{|d|=1} |uK_i| \geq 1 - |K_i - \mathbb{I}| \geq 1 - \|\nabla b\|_\infty \mathbb{E}\|\nabla b\|_\infty t \geq \frac{1}{2}, \]
by (1.10) we have
\[
\bigcap_{j=1}^{j_0} E_j \subset \left\{ \sum_{j=1}^{j_0} \int_0^{\bar{\lambda}} |uK_sB_j(X_s)|^2 ds \leq t \sum_{j=1}^{j_0} \delta^{\frac{\alpha^j(9-\frac{\beta}{2})}{2}} \right\}
\subset \left\{ \kappa_1 \int_0^{\bar{\lambda}} |uK_s|^2 ds \leq t \sum_{j=1}^{j_0} \delta^{\frac{\alpha^j(9-\frac{\beta}{2})}{2}} \right\}
\subset \left\{ \frac{t \kappa_1}{2} \leq t j_0 \delta^{\frac{\alpha^j(9-\frac{\beta}{2})}{2}} \right\} = \emptyset, \tag{5.16}\]
provided \( t \leq t_0 \) and \( \delta \leq \delta_1 \) for some \( \delta_1 \) small enough. Therefore, combining (5.14)-(5.16), we obtain that for all \( t \in (0, t_0) \) and \( \delta \in (0, \delta_1) \),
\[
\mathbb{P} \left\{ \int_0^t |uK_\sigma \sigma'|^2 ds \leq t \delta_0^d \right\} \leq C_1 j_0 \exp \left\{ -c_1 t \delta^{d_0} \right\},
\]
which in turn implies that for \( \theta = \frac{\rho u_0}{1 - p} \) and all \( \varepsilon \in (0, 1) \), \( t \in (0, t_0) \),
\[
\mathbb{P} \left\{ \int_0^t |uK_\sigma \sigma'|^2 ds \leq t \varepsilon \right\} \leq C \exp \left\{ -c_1 t \varepsilon^{-\theta} \right\}.
\]
For \( \lambda \geq t \), setting \( r := (\lambda/t)^{\frac{\theta}{\theta + 1}} \) and \( \xi := \frac{1}{t} \int_0^t |uK_\sigma \sigma'|^2 ds \), we have
\[
\mathbb{E} e^{-\lambda \xi} = \int_0^\infty e^{-\lambda \xi} \mathbb{P}(\xi \leq \varepsilon) d\varepsilon
\leq \int_r^\infty e^{-\lambda \xi} d\xi + C \int_0^r e^{-\lambda \xi - c_1 \lambda \xi^\theta} d\xi
= e^{-\lambda r} + C \int_0^{\lambda r} e^{-s - c_1 \lambda s^\theta} ds
\leq e^{-\alpha r} + Ce^{-c_1 \lambda r^\theta} \int_0^{\lambda r} e^{-s} ds
\leq e^{-\alpha (\lambda r)^{\frac{\theta}{\theta + 1}}} + Ce^{-c_1 (\lambda r)^{\frac{\theta}{\theta + 1}}}.
\]
By replacing \( \lambda \) with \( \lambda t \), we obtain the desired estimate with \( \gamma = \frac{\theta}{1 + \theta} \).

\[\square\]

**Lemma 5.3.** Let \( t_0 := 1 \wedge \left( \frac{\rho u_0}{2 \|\nabla h_0\|} \right) \). Under (1.10) and (AS\( \sigma \)), if \( S_t \) has finite moments of all orders, then for any \( p \geq 1 \), there exists a constant \( C \geq 1 \) such that for all \( t \in (0, t_0) \),
\[
\sup_{x \in \mathbb{R}^d} \mathbb{E} \left( (\det \Sigma_t(x))^{-p} \right) \leq Ct^{-\frac{2(p+1)}{\alpha p}} ,
\]
where \( \gamma \) is the same as in (5.13).

**Proof.** Write
\[
\tilde{\Sigma}_t := \int_0^t K_\sigma \sigma' K_\sigma^* dS_s .
\]
By Lemma [2.4] and (5.13), we have for all \( t \in (0, t_0) \) and \( \lambda \geq 1 \),
\[
\mathbb{E} \exp\{-\lambda (u \tilde{\Sigma}_t u^*)\} = \mathbb{E} \exp\left\{-\lambda \int_0^t |uK_\sigma \sigma'|^2 dS_s \right\}
\leq \left( \mathbb{E} \exp\left\{-c_0 \lambda^{\frac{\theta}{2}} \int_0^t |uK_\sigma \sigma'|^2 dS_s \right\} \right)^{\frac{1}{2}}
\leq C_2^{\frac{1}{2}} \exp\left\{-c_2 t(c_0 \lambda^{\frac{\theta}{2}})^{\gamma}/2 \right\}.
\]

(5.18)

Set \( F(u; \lambda) := \exp\{-\lambda (u \tilde{\Sigma}_t u^*)\} \).

For any \( u_0, u_1 \in \mathbb{S}^{d-1} \), let \( (u_r) \in (0, 1) \) be the shortest arc connecting \( u_0 \) and \( u_1 \). By the mean value theorem, one sees that for some \( r_0 \in (0, 1) \),
\[
|F(u_0; \lambda) - F(u_1; \lambda)| \leq 2 \rho(u_0, u_1) F(u_0; \lambda^\gamma) |\tilde{\Sigma}||,\]
where \( \rho(u_0, u_1) \) denotes the distance of sphere. Since
\[
|\tilde{\Sigma}_t| \leq C|S_t|,
\]

(27)
by Hölder’s inequality and \[5.18\], there exist constants \(C_3 \geq 1, c_3 \in (0, 1)\) such that for all \(t \in (0, t_0), \lambda \geq 1\) and \(u_0, u_1 \in \mathbb{S}^{d-1}\),
\[
\mathbb{E}[F(u_0; \lambda) - F(u_1; \lambda)]^d \leq C_3 p(u_0, u_1)^d \exp(-c_3 t \lambda^\gamma)\lambda^d.
\]
Since \(\mathbb{S}^{d-1}\) is compact and has dimension \(d - 1\), by Kolmogorov’s continuity theorem, we have
\[
\mathbb{E}\left(\sup_{|u|=1} F(u; \lambda)\right) \leq C_4 \exp(-c_4 t \lambda^\gamma)\lambda, \quad t \in (0, t_0), \lambda \geq 1.
\]
Noticing that for a \(d \times d\)-symmetric matrix \(A\),
\[(\det A)\frac{1}{2} \geq \inf_{|u|=1} u Au^*,
\]
we have
\[
\mathbb{E}\left(\left(\det \Sigma_t\right)^{-\frac{1}{2}}\right) = \frac{1}{\Gamma(p)} \int_0^\infty \lambda^{p-1} \mathbb{E} \left\{ -\lambda \inf_{|u|=1} uAu^* \right\} d\lambda
\]
\[
\leq \frac{1}{\Gamma(p)} \int_0^\infty \lambda^{p-1} \mathbb{E} \left\{ -\lambda \inf_{|u|=1} uAu^* \right\} d\lambda
\]
\[
= \frac{1}{\Gamma(p)} \int_0^\infty \lambda^{p-1} \left(\sup_{|u|=1} F(u; \lambda)\right) d\lambda
\]
\[
\leq \frac{1}{\Gamma(p)} \left(\int_0^t \lambda^{p-1} d\lambda + C_4 \int_1^\infty \lambda^p \exp(-c_4 t \lambda^\gamma) d\lambda\right)
\]
\[
= \frac{1}{\Gamma(p)} \left(\frac{1}{p} + C_4 t^{-\frac{2(p+1)}{\gamma}} \int_2^{(p+1)/\gamma} \lambda^p \exp(-c_4 t\gamma) d\lambda\right)
\]
\[
\leq \frac{1}{\Gamma(p)} \left(\frac{1}{p} + C_5 t^{-\frac{2(p+1)}{\gamma}}\right),
\]
which gives
\[
\mathbb{E}\left(\left(\det \Sigma_t\right)^{-\frac{1}{2}}\right) \leq C_6 t^{-\frac{2(p+1)}{\gamma}}, \quad t \in (0, t_0).
\] (5.19)
On the other hand, by \((5.11)\) we have
\[
\Sigma_t = J_t \Sigma_t J_t^*.
\]
Since
\[
\det J_t = \exp\left\{\int_0^t (\text{div} b)(X_s) ds\right\} \geq \exp(-t||\text{div} b||_\infty),
\]
which together with \((5.19)\) yields \((5.17)\). \(\square\)

**Theorem 5.4.** Let \(t_0 := 1 \wedge \left(\frac{e^{c_{\text{lim}}}}{2||b||_\infty}\right)\). Under \((1.10)\) and \((\text{AS}_0)\), for any \(k, m \in \mathbb{N}_0\) with \(k + m \geq 1\), there are \(\gamma_{k,m} > 0\) and \(C = C(k,m) > 0\) such that for all \(f \in C^0_b(\mathbb{R}^d)\) and \(t \in (0, t_0)\),
\[
\sup_{x \in \mathbb{R}^d} \left|\nabla^k \left(\nabla^m f\right)(X_t(x))\right| \leq C||f||_{\infty} t^{-\gamma_{k,m}}.
\] (5.20)

**Proof.** First of all, we assume that
\[
S_t \text{ has finite moments of all orders. (5.21)}
\]
Under this assumption, it is standard to prove that for any \(m, k \in \mathbb{N}_0\) with \(m + k \geq 1\) and \(p \geq 1\) (cf. \([40, \text{Lemma 3.6}]\)),
\[
\sup_{x \in \mathbb{R}^d} \sup_{r \in [0,1]} \mathbb{E}\left(||D^m \nabla^k X_r(x)||_{L^p}^p\right) < +\infty.
\] (5.22)
By the chain rule, we have
\[
\nabla^k \mathbb{E} \left( (\nabla^m f)(X_t(x)) \right) = \sum_{j=1}^k \mathbb{E} \left( (\nabla^{m+j} f)(X_t(x)) G_j(\nabla X_t(x), \cdots, \nabla^k X_t(x)) \right),
\]
where \( G_j, j = 1, \cdots, k \) are real polynomial functions. By Theorem 2.2 and (5.22), through cumbersome calculations, one finds that there exist integer \( n \) and \( p > 1, C > 0 \) such that for all \( t \in (0, 1) \) and \( x \in \mathbb{R}^d \),
\[
\left| \nabla^k \mathbb{E} \left( (\nabla^m f)(X_t(x)) \right) \right| \leq C \| f \|_{\infty} \mathbb{E} \left( \left| (\text{det} \Sigma_t(x))^{-1} \right|_p^p \right) \leq C \| f \|_{\infty} \| (\text{det} \Sigma_t(x))^{-1} \|_{\text{op}}^p.
\]
Estimate (5.20) now follows by Lemma 5.3.

Now, without assumption (5.21), we can use the same argument as in [40] Section 3.3 to prove (5.20). Since it is completely the same, we omit the details. \( \square \)

Now, we are in a position to give

**Proof of Theorem 1.3** Notice that for \( \nu_S(du) = u^{-1-\frac{\alpha}{2}} du \), (AS\( ^\alpha \)) holds. By estimate (5.20) and Sobolev’s embedding theorem, one has (1.11) (see [28], pp.102-103)). Equation (1.12) follows by Itô’s formula. The smoothness of \( \rho_t(x, y) \) with respect to the time variable \( t \) follows by equation (1.12) and a standard bootstrap argument.

**Acknowledgements:**

The author is very grateful to Hua Chen, Zhen-Qing Chen, Zhao Dong, Xuhui Peng and Feng-Yu Wang for their quite useful conversations. This work is supported by NSFs of China (Nos. 11271294, 11325105) and Program for New Century Excellent Talents in University (NCET-10-0654).

**References**

[1] Alexandre R.: Fractional order kinetic equations and hypoellipticity. Anal. Appl., Vol. 10, No. 03 : pp. 237-247 (2012).
[2] Alexandre R., Morimoto Y., Ukai S., Xu C.J. and Yang T.: Uncertainty principle and kinetic equations. J. Func. Anal., 255, 2013-2066(2008).
[3] Applebaum D.: Lévy processes and stochastic calculus. Cambridge Studies in Advanced Math. Cambridge Univ. Press, Cambridge, 2004.
[4] Bally V. and Caramellino L.: On the distance between probability density functions. arXiv:1311.7555v1.
[5] Bichteler K., Gravereaux J.B. and Jacod J.: Malliavin calculus for processes with jumps. Gordan and Breach Science Publishers, 1987.
[6] Bismut J.M.: Martingales, the Malliavin calculus and hypoellipticity under general Hörmander’s conditions. Z. Wahrsch. Verw. Gebiete, 56, 469-505(1981).
[7] Bismut J.M.: Calcul des variations stochastiques et processus de sauts. Z. Wahrsch. Verw. Gebiete, 63, 147-235(1983).
[8] Bogachev V.I.: Differential measures and the Malliavin calculus. AMS, 2010.
[9] Carmona P.: Existence and uniqueness of an invariant measure for a chain of oscillators in contact with two heat baths, Stochastic Process. Appl. 117 (8) (2007) 1076-1092.
[10] Cass T.: Smooth densities for stochastic differential equations with jumps. Stoch. Proc. Appl., 119, no.5, 1416-1435(2009).
[11] Chen H., Li W.X. and Xu C.J.: Gevrey hypoellipticity for a class of kinetic equations. Comm. in Partial Diff. Equa., 36, 693-728(2011).
[12] Dong Z., Peng X., Song Y. and Zhang X.: Strong Feller properties for degenerate SDEs with jumps. arXiv:1312.7380.
[13] Eckmann J.P. and Hairer M.: Non-equilibrium statistical mechanics of strongly an- harmonic chains of oscillators, Comm. Math. Phys. 212 (2000), no. 1, 105-164.
[14] Hörmander L.: Hypoelliptic second order differential equations. Acta Math. 119, 147-171(1967).
[15] 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).
[16] Komatsu T. and Takeuchi A.: Simplified probabilistic approach to the Hörmander theorem. Osaka J. Math., 38(2001), 681-691.
[17] Komatsu T. and Takeuchi A.: On the smoothness of pdf of solutions to SDE with jumps. Int. J. Diff. Eq. Appl., 2, 141-197 (2001).
[18] Kulik A.: Conditions for existence and smoothness of the distribution density for Ornstein-Uhlenbeck processes with Lévy noises. Theory Probab. Math. Statist. no.79, 23-38(2009).
[19] Kunita H: Smooth density of canonical stochastic differential equation with jumps. Astérisque No. 327, 69-91 (2010).
[20] Kunita H.: Nondegenerate SDEs with jumps and their hypoelliptic properties. J. Math. Soc. Japan, Volume 65, Number 3, 687-1035 (2013).
[21] Kusuoka S.: Malliavin calculus for stochastic differential equations driven by subordinated Brownian motions. Kyoto J. of Math. Vol.50, no.3, 491-520(2009).
[22] Kusuoka S. and Stroock D.: Applications of the Malliavin Calculus, Part I. Stochastic Analysis, Proceedings of the Taniguchi International Symposium on Stochastic Analysis North-Holland Mathematical Library Volume 32, Pages 271-306(1984).
[23] Kusuoka S. and Stroock D.: Applications of the Malliavin Calculus, Part II. J. Fac. of Sci., the Univ. of Tokyo, Vol. 32, No.1, 1-76(1985).
[24] Malicet D. and Poly G.: Properties of convergence in Dirichlet structures. J. Funct. Anal. 264, 2077-2096(2013).
[25] Malliavin P.: Stochastic calculus of variations and hypoelliptic operators. In: Proc Inter. Symp. on Stoch. Diff. Equations, Kyoto, 195-263(1976).
[26] Morimoto Y. and Xu C.J.: Hypoellipticity for a class of kinetic equations. J. Math. Kyoto Univ. 47-1, 129-152(2007).
[27] Norris J.: Simplified Malliavin calculus. In: Seminaire de Probabilités XX, Lecture Notes in Math. 1204 , Springer, Berlin, 101-130(1986).
[28] Nualart D.: The Malliavin calculus and related topics. Springer-Verlag, New York, 2006.
[29] Picard J.: On the existence of smooth densities for jump processes. Prob. Theory Rela. Fields, Vol. 105, 481–511(1996).
[30] Priola E. and Zabczyk J.: Densities for Ornstein-Uhlenbeck processes with jumps. Bull. London Math. Soc., Vol 41, 41-50(2009).
[31] Protter, P. E. Stochastic integration and differential equations. Second Edition, Springer-Verlag, Berlin, (2004).
[32] Rey-Bellet L. and Thomas L.E.: Asymptotic behavior of thermal nonequilibrium steady states for a driven chain of anharmonic oscillators, Comm. Math. Phys. 215 (2000), no. 1, 1-24.
[33] Sato K.: Lévy processes and infinite divisible distributions. Cambridge Univ. Press, Cambridge, 1999.
[34] Simon T.: On the absolute continuity of multidimensional Ornstein-Uhlenbeck processes. Probability Theory and Related Fields, Volume 151, Issue 1-2, pp 173-190(2011).
[35] Takeuchi A.: The Malliavin calculus for SDE with jumps and the partially hypoelliptic problem. Osaka J. Math. 39, 523-559(2002).
[36] Watanabe S.: Lectures on stochastic differential equations and Malliavin calculus. Tata Institute of Fundamental Research, Springer-Verlag, 1984.
[37] Wang F.Y. and Wang J.: Harnack inequalities for stochastic equations driven by Lévy noise. J. Math. Anal. Appl., 410, no. 1, 513-523(2014).
[38] Zaslavsky G.M.: The physics of chaos in hamiltonian systems. World Scientific Pub., 2007.
[39] Zhang X.: Derivative formula and gradient estimate for SDEs driven by α-stable processes. Stoch. Proc. Appl. Vol. 123, Issue 4, pp.1213-1228(2013).
[40] Zhang X.: Densities for SDEs driven by degenerate α-stable processes. to appear in Annals of Prob., arXiv:1207.3565
[41] Zhang X.: Fundamental solution of kinetic Fokker-Planck operator with anisotropic nonlocal dissipativity. arXiv:1301.0147
[42] Zhang X.: Fundamental solutions of nonlocal Hörmander’s operators. In preparation.

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