Electronic Journal of Probability

Electron. J. Probab. **20** (2015), no. 21, 1–27.
ISSN: 1083-6489 DOI: 10.1214/EJP.v20-3287

Regularity of density for SDEs driven by degenerate Lévy noises*

Yulin Song† Xicheng Zhang‡

Abstract

By using Bismut's approach to the Malliavin calculus with jumps, we study the regularity of the distributional density for SDEs driven by degenerate additive Lévy noises. Under full Hörmander's conditions, we prove the existence of distributional density and the weak continuity in the first variable of the distributional density. Moreover, under a uniform first order Lie's bracket condition, we also prove the smoothness of the density.

Keywords: Distributional density; Hörmander's condition; Malliavin calculus; Girsanov's theorem.
AMS MSC 2010: 60H15 ; 60H10.
Submitted to EJP on January 27, 2014, final version accepted on March 1, 2015.

1 Introduction

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

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + A_1\mathrm{d}W_t + A_2\mathrm{d}L_t, \quad X_0 = x, \tag{1.1}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ is a smooth vector field, A_1 and A_2 are two constant $d \times d$-matrices, W_t is a d-dimensional standard Brownian motion and L_t is a purely jump d-dimensional Lévy process with Lévy measure $\nu(\mathrm{d}z)$. Let $\Gamma_0 := \{z \in \mathbb{R}^d : 0 < |z| < 1\}$. Throughout this work, we assume that $\frac{\nu(\mathrm{d}z)}{\mathrm{d}z}|_{\Gamma_0} = \kappa(z)$ satisfies the following conditions: for some $\alpha \in (0,2)$ and $m \in \mathbb{N}$,

($\mathbf{H}^\alpha_m$) $\kappa \in C^m(\Gamma_0; (0,\infty))$ is symmetric (i.e. $\kappa(-z) = \kappa(z)$) and satisfies the following Orey's order condition (cf. [19, Proposition 28.3]):

$$\lim_{\varepsilon \downarrow 0} \varepsilon^{\alpha-2} \int_{|z|\leqslant\varepsilon} |z|^2 \kappa(z)\mathrm{d}z =: c_1 > 0, \tag{1.2}$$

and bounded condition: for $j = 1, \cdots, m$ and some $C_j > 0$,

$$|\nabla^j \log \kappa(z)| \leqslant C_j |z|^{-j}, \quad z \in \Gamma_0. \tag{1.3}$$

*Supported by NNSFs of China (Nos. 11271294, 11325105).

†Department of Mathematics, Nanjing University, Nanjing, Jiangsu 210093, P.R.China; School of Mathematical Sciences, Beijing Normal University, Beijing, 100875, P.R.China. E-mail: songyl@amss.ac.cn

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

For example, if $\kappa(z) = a(z)|z|^{-d-\alpha}$ with

$$a(z) = a(-z), \ \ 0 < a_0 \leqslant a(z) \leqslant a_1, \ \ |\nabla^j a(z)| \leqslant a_2, \ j = 1, \cdots, m,$$

then **($\mathbf{H}^\alpha_m$)** holds. Notice that the generator of SDE (1.1) is given by

$$\mathscr{A} f(x) := \frac{1}{2}\nabla^2_{A_1 A_1^*} f(x) + \text{p.v.} \int_{\mathbb{R}^d} (f(x + A_2 z) - f(x))\nu(\mathrm{d}z) + b(x) \cdot \nabla f(x),$$

where A_1^* stands for the transpose of A_1, and

$$b \cdot \nabla f := \sum_{i=1}^d b^i \partial_i f, \ \ \nabla^2_{A_1 A_1^*} f := \sum_{i,j=1}^d (A_1 A_1^*)_{ij} \partial^2_{ij} f, \tag{1.4}$$

and p.v. stands for the Cauchy principle value.

It is well-known that when b is Lipschitz continuous, SDE (1.1) has a unique solution $X_t = X_t(x)$. The aim of this work is to investigate the regularity of the distributional density of $X_t(x)$ under Hörmander's conditions. Let $B_0 := \mathbb{I}_{d\times d}$ be the identity matrix and define $d \times d$-matrix-valued function $B_n(x)$ recursively by

$$B_n(x) := b(x) \cdot \nabla B_{n-1}(x) - \nabla b(x) \cdot B_{n-1}(x) + \tfrac{1}{2}\nabla^2_{A_1 A_1^*} B_{n-1}(x), \ n \in \mathbb{N}.$$

Here and below, $(\nabla b(x))_{ij} := \partial_j b^i(x)$ is the Jacobian matrix, and $b\cdot\nabla B_{n-1}$ and $\nabla^2_{A_1 A_1^*} B_{n-1}$ are defined as in (1.4). Our first main result is about the existence and weak continuity of the distribution density for SDE (1.1) under full Hörmander's condition.

Theorem 1.1. *Assume that* **($\mathbf{H}^\alpha_1$)** *holds and b is smooth and has bounded derivatives of all orders, and for any $x \in \mathbb{R}^d$ and some $n = n(x) \in \mathbb{N}$,*

$$\mathrm{Rank}[A_1, B_1(x)A_1, \cdots, B_n(x)A_1, A_2, B_1(x)A_2, \cdots, B_n(x)A_2] = d. \tag{1.5}$$

Then $X_t(x)$ admits a density $\rho_t(x, y)$ with respect to the Lebesgue measure so that for any bounded measurable function f,

$$x \mapsto \mathcal{P}_t f(x) := \mathbb{E} f(X_t(x)) = \int_{\mathbb{R}^d} f(y)\rho_t(x, y)\mathrm{d}y \text{ is continuous.} \tag{1.6}$$

In particular, the semigroup $(\mathcal{P}_t)_{t\geqslant 0}$ has the strong Feller property.

Remark 1.2. When $A_1 = 0$ and $b(x) = Bx$ is linear, condition (1.5) is called Kalman's rank condition. In this case, the smoothness of the density of the corresponding Ornstein-Uhlenbeck process has been studied in [16, 9].

About the smoothness of the density, we have the following partial result.

Theorem 1.3. *Assume that* **($\mathbf{H}^\alpha_m$)** *holds for some $m \in \mathbb{N}$, and b is smooth and has bounded derivatives of all orders, and*

$$\inf_{x\in\mathbb{R}^d} \inf_{|u|=1} \Big(|uA_1|^2 + |uB_1(x)A_1|^2 + |uA_2|^2 + |uB_1(x)A_2|^2\Big) =: c_2 > 0. \tag{1.7}$$

Then for any $k, n \in \{0\} \cup \mathbb{N}$ with $1 \leqslant k + n \leqslant m$, there are $\gamma_{k,n} > 0$ and $C = C(k, n) > 0$ such that for all $f \in C_b^\infty(\mathbb{R}^d)$ and $t \in (0, 1)$,

$$\sup_{x\in\mathbb{R}^d} |\nabla^k \mathbb{E}((\nabla^n f)(X_t(x)))| \leqslant C\|f\|_\infty t^{-\gamma_{k,n}}, \tag{1.8}$$

where ∇^k denotes the k-order gradient operator. In particular, if $m = \infty$, then $X_t(x)$ admits a smooth density $\rho_t(x, y)$ so that

$$(x, y) \mapsto \rho_t(x, y) \in C_b^\infty(\mathbb{R}^d \times \mathbb{R}^d), \ \ t > 0. \tag{1.9}$$

In the continuous diffusion case (i.e. $A_2 = 0$ and $A_1 = A_1(x)$), under Hörmander's conditions, Malliavin [13] proved that SDE (1.1) has a smooth density by using the stochastic calculus of variations (nowadays, it is also called the Malliavin calculus, and a systematic introduction about the Malliavin calculus is referred to the book [14]). Since the pioneering work of [13], there are many works devoting to extend Malliavin's theory to the jump case (cf. [5, 4, 15, 8] etc.). However, unlike the case of continuous Brownian functionals, there does not exist a unified treatment for Poisson functionals since the canonical Poisson space has a nonlinear structure. We mention that Bismut's approach is based on Girsanov's transformation (cf. [5]), while Picard's approach is to use the difference operator to establish an integration by parts formula (cf. [15]).

When $A_1 = 0$ and $\kappa(z) = c|z|^{-d-\alpha}$, Theorems 1.1 and 1.3 have been proved in [22] and [7] by using the Malliavin calculus for subordinate Brownian motions (cf. [11]). Meanwhile, if $A_1 = 0$ and $\kappa(z) = a(z)|z|^{-d-\alpha}$, Theorem 1.3 is also contained in [21, Theorem 2.2]. About the smoothness of distributional density for degenerate SDEs driven by purely jump noises, Takeuchi [20], Cass [6] and Kunita [10] have already studied this problem under different Hörmander's conditions. However, their results do not cover the present general case (see also [23, 24] for some related works). Compared with [22] and [7], in this work we shall use Bismut's approach to prove Theorems 1.1 and 1.3, and need to assume that the Lévy measure is absolutely continuous with respect to the Lebesgue measure. It is noticed that in [7], the Lévy measure can be singular and the drift is allowed to have arbitrary growth, which cannot be dealt with in the current settings.

In the proof of our main theorems, one of the difficulties we are facing is the infinity of the moments of L_t. To overcome this difficulty, we consider two independent Lévy processes L^0_t and L^1_t with Lévy measures $\nu_0(\mathrm{d}z) := 1_{|z|<1}\kappa(z)\mathrm{d}z$ and $\nu_1(\mathrm{d}z) := 1_{|z|\geqslant 1}\nu(\mathrm{d}z)$ respectively. Clearly,

$$L_t \text{ has the same law as } L^0_t + L^1_t.$$

Notice that L^1_t is a compound Poisson process. Let $0 =: \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_n < \cdots$ be the jump time of L^1_t. It is well-known that

$$\mathscr{E} := \{\tau_n - \tau_{n-1}, n \in \mathbb{N}\},\ \ \mathscr{G} := \{\Delta L^1_{\tau_n} := L_{\tau_n} - L_{\tau_n -}, n \in \mathbb{N}\}$$

are two independent families of i.i.d random variables. Let $\hbar$ be a cádlág purely discontinuous $\mathbb{R}^d$-valued function with finitely many jumps and $\hbar_0 = 0$. Following the argument of [22, Subsection 3.3], we consider the following SDE:

$$\tilde{X}_t(x;\hbar) = x + \int_0^t b(\tilde{X}_s(x;\hbar))\mathrm{d}s + A_1 W_t + A_2 L^0_t + A_2 \hbar_t.$$

Clearly,

$$X_t(x) \stackrel{(d)}{=} \tilde{X}_t(x; L^1_\cdot).$$

If we write

$$\mathcal{P}_t f(x) := \mathbb{E} f(X_t(x)),\ \ \tilde{\mathcal{P}}_t f(x) := \mathbb{E} f(\tilde{X}_t(x;0)),$$

then we have (see [22, (3.19)])

$$\mathcal{P}_t f(x) = \sum_{n=0}^{\infty} \mathbb{E}\left(\tilde{\mathcal{P}}_{\tau_1} \cdots \vartheta_{A_2 \Delta L^1_{\tau_{n-1}}} \tilde{\mathcal{P}}_{\tau_n - \tau_{n-1}} \vartheta_{A_2 \Delta L^1_{\tau_n}} \tilde{\mathcal{P}}_{t-\tau_n} f(x); \tau_n < t \leqslant \tau_{n+1}\right), \quad (1.10)$$

where for a function $g(x)$ and $y \in \mathbb{R}^d$,

$$\vartheta_y g(x) := g(x+y).$$

Basing on (1.10) and as in [22, Subsection 3.3], it suffices to prove Theorems 1.1 and 1.3 for $\tilde{X}_t(x;0)$, that is, we only need to consider the SDE (1.1) driven by W_t and L^0_t.

This paper is organised as follows: in Section 2, we recall Bismut's approach to the Malliavin calculus with jumps. In [4], Bichteler, Gravereaux and Jacod have already systematically introduced it, however, the α-stable like noise does not fall into their framework. Thus, we have to extend the integration by parts formula to the more general class of Lévy measures. Moreover, we also prove a Kusuoka-Stroock's formula for Poisson stochastic integrals. In Section 3, we introduce the reduced Malliavin matrix for SDE (1.1) used in Bismut's approach (cf. [4]), and also give some necessary estimates. In Sections 4 and 5, we shall prove Theorems 1.1 and 1.3.

Convention: The letter C or c with or without subscripts will denote an unimportant constant, whose value may be different in different places.

2 Revisit of Bismut's approach to the Malliavin calculus with jumps

Let $\Gamma \subset \mathbb{R}^d$ be an open set containing the origin. Let us 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. 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
$$\kappa \in C^1(\Gamma_0; (0,\infty)), \ \int_{\Gamma_0} (1 \wedge |z|^2)\kappa(z)\mathrm{d}z < +\infty, \ \ |\nabla \log \kappa(z)| \leqslant C\varrho(z), \tag{2.2}$$
where $\varrho(z)$ is defined by (2.1);
- W and N are independent.

In the following we write

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

2.1 Function spaces

Let $p \geqslant 1$. We introduce the following spaces for later use.

- $\mathbb{L}^1_p$: The space of all predictable processes: $\xi : \Omega \times [0,1] \times \Gamma_0 \to \mathbb{R}^k$ with finite norm:
$$\|\xi\|_{\mathbb{L}^1_p} := \left[\mathbb{E}\left(\int_0^1\!\!\int_{\Gamma_0} |\xi(s,z)|\nu(\mathrm{d}z)\mathrm{d}s\right)^p\right]^{\frac{1}{p}} + \left[\mathbb{E}\int_0^1\!\!\int_{\Gamma_0} |\xi(s,z)|^p\nu(\mathrm{d}z)\mathrm{d}s\right]^{\frac{1}{p}} < \infty.$$

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

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

- $\mathbb{H}_p$: 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 $\mathbf{v} : \Omega \times [0,1] \times \Gamma_0 \to \mathbb{R}^d$ with finite norm:

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

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

$$\mathbb{H}_{\infty-} := \cap_{p\geqslant 1}\mathbb{H}_p, \quad \mathbb{V}_{\infty-} := \cap_{p\geqslant 1}\mathbb{V}_p.$$

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

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

Remark 2.1. For $\xi \in \mathbb{L}_p^1$, if there is a compact subset $U \subset \Gamma_0$ such that $\xi(s,z) = 0$ for all $z \notin U$, then in view of $\kappa \in C^1(\Gamma_0;(0,\infty))$,

$$\|\xi\|_{\mathbb{L}_p^1}^p \asymp \mathbb{E}\left(\int_0^1\int_U |\xi(s,z)|^p\mathrm{d}z\mathrm{d}s\right) = \mathbb{E}\left(\int_0^1\int_{\mathbb{R}^d} |\xi(s,z)|^p\mathrm{d}z\mathrm{d}s\right),$$

where $\asymp$ means that both sides are comparable up to a constant (depending only on U, κ, p, d).

Lemma 2.2. *(i) For any $p \geqslant 1$, the spaces $(\mathbb{H}_p, \|\cdot\|_{\mathbb{H}_p})$ and $(\mathbb{V}_p, \|\cdot\|_{\mathbb{V}_p})$ are Banach spaces.*

(ii) For any $p_2 > p_1 \geqslant 1$, $\mathbb{H}_{p_2} \subset \mathbb{H}_{p_1}$ and $\mathbb{V}_{p_2} \subset \mathbb{V}_{p_1}$.

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

Proof. (i) and (ii) are obvious.

(iii) We only prove the density of $\mathbb{V}_0$ in $\mathbb{V}_p$, i.e., for each $\mathbf{v} \in \mathbb{V}_p$, there exists a sequence $\mathbf{v}_n \in \mathbb{V}_0$ such that

$$\lim_{n\to\infty} \|\mathbf{v}_n - \mathbf{v}\|_{\mathbb{V}_p} = 0.$$

We shall construct the approximation by three steps.

(1) For $\varepsilon \in (0,1)$, define

$$\Gamma_\varepsilon := \Big\{z \in \mathbb{R}^d : \mathbf{d}(z,\Gamma_0^c) > \varepsilon\Big\}.$$

Let $\chi_\varepsilon : \mathbb{R}^d \to [0,1]$ be a smooth function with

$$\chi_\varepsilon|_{\Gamma_{2\varepsilon}} = 1, \quad \chi_\varepsilon|_{\Gamma_\varepsilon^c} = 0, \quad \|\nabla\chi_\varepsilon\|_\infty \leqslant C/\varepsilon. \tag{2.3}$$

For $R>1$, let $\chi_R:\mathbb{R}^d\to[0,1]$ be a smooth function with

$$\chi_R(z)=1,\ \ |z|\leqslant R,\ \ \chi_R(z)=0,\ |z|>2R,\ \ \|\nabla\chi_R\|_\infty\leqslant C/R. \tag{2.4}$$

Let us define

$$\mathbf{v}_{\varepsilon,R}(s,z)=\mathbf{v}(s,z)\chi_\varepsilon(z)\chi_R(z). \tag{2.5}$$

Notice that for $\varepsilon\in(0,1)$ and $R>1$,

$$\begin{aligned}
&|\nabla\mathbf{v}_{\varepsilon,R}(s,z)-\nabla\mathbf{v}(s,z)|\\
&\leqslant C\Big(\varepsilon^{-1}1_{z\in\Gamma_\varepsilon\setminus\Gamma_{2\varepsilon}}+R^{-1}1_{R<|z|<2R}\Big)|\mathbf{v}(s,z)|+\Big(1_{z\in\Gamma^c_{2\varepsilon}}+1_{|z|>R}\Big)|\nabla\mathbf{v}(s,z)|\\
&\leqslant C\varrho(z)\Big(1_{z\in\Gamma^c_{2\varepsilon}}+1_{|z|>R}\Big)|\mathbf{v}(s,z)|+\Big(1_{z\in\Gamma^c_{2\varepsilon}}+1_{|z|>R}\Big)|\nabla\mathbf{v}(s,z)|,
\end{aligned} \tag{2.6}$$

where $\varrho(z)$ is defined by (2.1). By the dominated convergence theorem, we have

$$\lim_{\varepsilon\downarrow0,R\uparrow\infty}\|\mathbf{v}_{\varepsilon,R}-\mathbf{v}\|_{\mathbb{V}_p}=0.$$

(2) Next we can assume that for some compact set $U\subset\Gamma_0$,

$$\mathbf{v}(s,z)=0,\ \ z\notin U. \tag{2.7}$$

Let $\varphi:\mathbb{R}^d\to[0,1]$ be a smooth function with

$$\varphi(x)=1,\ \ |x|\leqslant1,\ \ \varphi(x)=0,\ \ |x|\geqslant2,\ \ \int_{\mathbb{R}^d}\varphi(x)\mathrm{d}x=1.$$

For $\delta\in(0,1)$, set $\varphi_\delta(x):=\delta^{-d}\varphi(\delta^{-1}x)$ and

$$\mathbf{v}_\delta(s,z):=\int_{\mathbb{R}^d}\mathbf{v}(s,z)\varphi_\delta(x-z)\mathrm{d}z. \tag{2.8}$$

By (2.7) and Remark 2.1, it is easy to see that

$$\|\mathbf{v}\|^p_{\mathbb{V}_p}\asymp\mathbb{E}\left(\int_0^1\!\!\int_U(|\mathbf{v}|+|\nabla_z\mathbf{v}|)^p(s,z)\mathrm{d}z\mathrm{d}s\right)=\mathbb{E}\left(\int_0^1\!\!\int_{\mathbb{R}^d}(|\mathbf{v}|+|\nabla_z\mathbf{v}|)^p(s,z)\mathrm{d}z\mathrm{d}s\right). \tag{2.9}$$

Thus,

$$\lim_{\delta\downarrow0}\|\mathbf{v}_\delta-\mathbf{v}\|_{\mathbb{V}_p}=0.$$

(3) Lastly we assume that $\mathbf{v}$ is smooth in z and satisfies (2.7). For $R>1$, we construct $\mathbf{v}_R(s,z)$ as follows:

$$\mathbf{v}_R(\omega,s,z):=\mathbf{v}(\omega,s,z)\cdot1_{\|\mathbf{v}(\omega,s,\cdot)\|_\infty\leqslant R}\cdot1_{\|\nabla\mathbf{v}(\omega,s,\cdot)\|_\infty\leqslant R}.$$

Clearly,

$$\mathbf{v}_R\in\mathbb{V}_0.$$

By (2.9) and the dominated convergence theorem, we have

$$\lim_{R\to\infty}\|\mathbf{v}_R-\mathbf{v}\|_{\mathbb{V}_p}=0.$$

The proof is complete. □

2.2 Girsanov's theorem

We need the following Burkholder's inequality.

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

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

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

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

Proof. (i) Let us write

$$M_t:=\int_0^t\!\!\int_{\Gamma_0}\xi(s,z)\tilde N(\mathrm{d}s,\mathrm{d}z)=\int_0^t\!\!\int_{\Gamma_0}\xi(s,z)N(\mathrm{d}s,\mathrm{d}z)-\int_0^t\!\!\int_{\Gamma_0}\xi(s,z)\nu(\mathrm{d}z)\mathrm{d}s. \tag{2.12}$$

For $p>1$, by Itô's formula and a stopping time technique, we have

$$\mathbb{E}|M_t|^p\leqslant\mathbb{E}\left(\int_0^t\!\!\int_{\Gamma_0}(|M_{s-}+\xi(s,z)|^p-|M_{s-}|^p-p\xi(s,z)\mathrm{sgn}(M_{s-})|M_{s-}|^{p-1})\nu(\mathrm{d}z)\mathrm{d}s\right).$$

By Doob's maximal inequality and Young's inequality, we further have

$$\begin{aligned}\mathbb{E}\left(\sup_{t\in[0,1]}|M_t|^p\right)&\leqslant C_p\mathbb{E}|M_1|^p\leqslant C_p\mathbb{E}\left(\int_0^1\!\!\int_{\Gamma_0}|\xi(s,z)|(|M_{s-}|+|\xi(s,z)|)^{p-1}\nu(\mathrm{d}z)\mathrm{d}s\right)\\&\leqslant C_p\mathbb{E}\left(\sup_{s\in[0,1]}|M_s|^{p-1}\int_0^1\!\!\int_{\Gamma_0}|\xi(s,z)|\nu(\mathrm{d}z)\mathrm{d}s\right)+C_p\|\xi\|^p_{\mathbb{L}^1_p}\\&\leqslant\frac12\mathbb{E}\left(\sup_{s\in[0,1]}|M_s|^p\right)+C_p\|\xi\|^p_{\mathbb{L}^1_p},\end{aligned}$$

which together with (2.12) gives (2.10).

(ii) As above, for $p\geqslant 2$, by Taylor's expansion, we have

$$\begin{aligned}\mathbb{E}\left(\sup_{t\in[0,1]}|M_t|^p\right)&\leqslant C_p\mathbb{E}\left(\int_0^1\!\!\int_{\Gamma_0}|\xi(s,z)|^2(|M_{s-}|+|\xi(s,z)|)^{p-2}\nu(\mathrm{d}z)\mathrm{d}s\right)\\&\leqslant C_p\mathbb{E}\left(\sup_{s\in[0,1]}|M_s|^{p-2}\int_0^1\!\!\int_{\Gamma_0}|\xi(s,z)|^2\nu(\mathrm{d}z)\mathrm{d}s\right)+C_p\|\xi\|^p_{\mathbb{L}^2_p}\\&\leqslant\frac12\mathbb{E}\left(\sup_{s\in[0,1]}|M_s|^p\right)+C_p\|\xi\|^p_{\mathbb{L}^2_p},\end{aligned}$$

which in turn gives (2.11). □

For $\mathbf{v}\in\mathbb{V}_0$ and $\varepsilon>0$, define

$$\gamma_\varepsilon(t,z):=\det\left(I+\varepsilon\nabla_z\mathbf{v}(t,z)\right)\frac{\kappa(z+\varepsilon\mathbf{v}(t,z))}{\kappa(z)}.$$

The following lemma is easy.

Lemma 2.4. *For any $\mathbf{v}\in\mathbb{V}_0$ with compact support $U\subset\Gamma_0$ with respect to z, there exist an $\varepsilon_0>0$ and a constant $C>0$ such that for any $\varepsilon\in(0,\varepsilon_0)$ and all t,z,*

$$|\gamma_\varepsilon(t,z)-1|\leqslant C\varepsilon 1_U(z). \tag{2.13}$$

Moreover, we have

$$\frac{\mathrm{d}\gamma_\varepsilon(t,z)}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0}=\mathrm{div}\mathbf{v}(t,z)+\langle\nabla\log\kappa(z),\mathbf{v}(t,z)\rangle_{\mathbb{R}^d}=\frac{\mathrm{div}(\kappa\mathbf{v})(t,z)}{\kappa(z)}. \tag{2.14}$$

Proof. Since $\mathbf{v}(t,z)=0$ for $z\notin U$, we have

$$\gamma_\varepsilon(t,z)=1,\quad\forall z\notin U.$$

For any $z\in U$, since $\mathbf{v}$ and $\nabla_z\mathbf{v}$ are bounded, we have

$$\begin{aligned}|\gamma_\varepsilon(t,z)-1|&\leqslant|\det(I+\varepsilon\nabla_z\mathbf{v}(t,z))|\left|\frac{\kappa(z+\varepsilon\mathbf{v}(t,z))}{\kappa(z)}-1\right|+|\det(I+\varepsilon\nabla_z\mathbf{v}(t,z))-1|\\&\leqslant\frac{C}{\inf_{z\in U}\kappa(z)}|\kappa(z+\varepsilon\mathbf{v}(t,z))-\kappa(z)|+C\varepsilon,\end{aligned}$$

which gives the desired estimate (2.13) by the compactness of U and $\kappa\in C^1(\Gamma_0;(0,\infty))$. As for (2.14), it follows by a direct calculation. □

For $p\geqslant 1$ and $\Theta:=(h,\mathbf{v})\in\mathbb{H}_p\times\mathbb{V}_p$, we write

$$\mathrm{div}\Theta:=-\int_0^1\langle h(s),\mathrm{d}W_s\rangle_{\mathbb{R}^d}+\int_0^1\!\!\int_{\Gamma_0}\frac{\mathrm{div}(\kappa\mathbf{v})(s,z)}{\kappa(z)}\tilde N(\mathrm{d}s,\mathrm{d}z). \tag{2.15}$$

By Burkholder's inequality and (2.2), we have

$$\mathbb{E}|\mathrm{div}\Theta|^p\leqslant C\Big(\|h\|^p_{\mathbb{H}_p}+\|\mathbf{v}\|^p_{\mathbb{V}_p}\Big). \tag{2.16}$$

Let Q^ε_t solve the following SDE:

$$Q^\varepsilon_t=1-\varepsilon\int_0^t Q^\varepsilon_s\langle h_s,\mathrm{d}W_s\rangle_{\mathbb{R}^d}+\int_0^t\!\!\int_{\Gamma_0}Q^\varepsilon_{s-}(\gamma_\varepsilon(s,z)-1)\tilde N(\mathrm{d}s,\mathrm{d}z), \tag{2.17}$$

whose solution is explicitly given by the Doleans-Dade formula:

$$\begin{aligned}Q^\varepsilon_t=\exp\bigg\{&\int_0^t\!\!\int_{\Gamma_0}\log\gamma_\varepsilon(s,z)N(\mathrm{d}s,\mathrm{d}z)-\int_0^t\!\!\int_{\Gamma_0}(\gamma_\varepsilon(s,z)-1)\nu(\mathrm{d}z)\mathrm{d}s\\&-\varepsilon\int_0^t\langle h_s,\mathrm{d}W_s\rangle_{\mathbb{R}^d}-\frac{\varepsilon^2}{2}\int_0^t|h_s|^2\mathrm{d}s\bigg\}.\end{aligned}$$

Lemma 2.5. *If $\Theta=(h,\mathbf{v})\in\mathbb{H}_0\times\mathbb{V}_0$, then Q^ε_t is a nonnegative martingale and for any $p\geqslant 2$,*

$$\lim_{\varepsilon\downarrow 0}\mathbb{E}\left|\frac{Q^\varepsilon_1-1}{\varepsilon}-\mathrm{div}\Theta\right|^p=0. \tag{2.18}$$

Proof. For any $p\geqslant 2$, by (2.17), (2.13) and (2.10), we have

$$\begin{aligned}\mathbb{E}|Q^\varepsilon_t|^p&\leqslant C+C\varepsilon^p\int_0^t\mathbb{E}(|Q^\varepsilon_s|^p|h_s|^p)\mathrm{d}s+\int_0^t\!\!\int_{\Gamma_0}\mathbb{E}(|Q^\varepsilon_{s-}(\gamma_\varepsilon(s,z)-1)|^p)\nu(\mathrm{d}z)\mathrm{d}s\\&\leqslant C+C\varepsilon^p\left(\|h\|^p_\infty+\nu(U)\right)\int_0^t\mathbb{E}|Q^\varepsilon_s|^p\mathrm{d}s,\end{aligned}$$

which gives

$$\sup_{\varepsilon\in(0,1)}\sup_{t\in[0,1]}\mathbb{E}|Q_t^\varepsilon|^p<+\infty. \tag{2.19}$$

From this and (2.17), one sees that Q_t^ε is a nonnegative martingale and $\mathbb{E}Q_t^\varepsilon=1$.

For (2.18), by equation (2.17) and (2.19), we have

$$\lim_{\varepsilon\downarrow 0}\sup_{t\in[0,1]}\mathbb{E}|Q_t^\varepsilon-1|^p=0,$$

and

$$\begin{aligned}\frac{Q_t^\varepsilon-1}{\varepsilon}-\mathrm{div}\Theta=&\int_0^1(Q_s^\varepsilon-1)\langle h_s,\mathrm{d}W_s\rangle_{\mathbb{R}^d}+\frac{1}{\varepsilon}\int_0^1\!\!\int_{\Gamma_0}(Q_{s-}^\varepsilon-1)(\gamma_\varepsilon(s,z)-1)\tilde N(\mathrm{d}s,\mathrm{d}z)\\&+\int_0^1\!\!\int_{\Gamma_0}\left(\frac{\gamma_\varepsilon(s,z)-1}{\varepsilon}-\frac{\mathrm{div}(\kappa\mathbf{v})(s,z)}{\kappa(z)}\right)\tilde N(\mathrm{d}s,\mathrm{d}z).\end{aligned}$$

Thus, by Burkholder's inequality and Lemma 2.4, we obtain (2.18). □

For $\Theta=(h,\mathbf{v})\in\mathbb{H}_0\times\mathbb{V}_0$ and $\varepsilon>0$, define

$$W_t^\varepsilon:=W_t+\varepsilon\int_0^t h(s)\mathrm{d}s,\quad N^\varepsilon((0,t]\times E):=\int_0^t\!\!\int_{\Gamma_0}1_E(z+\varepsilon\mathbf{v}(s,z))N(\mathrm{d}s,\mathrm{d}z).$$

Then the map

$$\Theta^\varepsilon:(W,N)\mapsto(W^\varepsilon,N^\varepsilon) \tag{2.20}$$

defines a transformation from Ω to Ω. We have (cf. [4, p.64, Theorem 6-16] or [3, p. 185])

Theorem 2.6. *(Girsanov's theorem) For $\Theta=(h,\mathbf{v})\in\mathbb{H}_0\times\mathbb{V}_0$, there exists an $\varepsilon_0>0$ such that for all $\varepsilon\in(0,\varepsilon_0)$, the law of $(W^\varepsilon,N^\varepsilon)$ under $Q_1^\varepsilon\mathbb{P}$ is the same as $\mathbb{P}$, i.e.,*

$$\mathbb{P}=(Q_1^\varepsilon\mathbb{P})\circ(\Theta^\varepsilon)^{-1}.$$

2.3 Malliavin derivative operator

Let $C_{\mathrm{p}}^\infty(\mathbb{R}^m)$ be the class of all smooth functions on $\mathbb{R}^m$ which together with all the derivatives have at most polynomial growth. Let $\mathcal{F}C_{\mathrm{p}}^\infty$ 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 $f\in C_{\mathrm{p}}^\infty(\mathbb{R}^{m_1+m_2})$, $h_1,\cdots,h_{m_1}\in\mathbb{H}_0$ 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

$$\mathcal{F}C_{\mathrm{p}}^\infty\subset\cap_{p\geqslant1}L^p(\Omega,\mathscr{F},\mathbb{P}).$$

For $\Theta=(h,\mathbf{v})\in\mathbb{H}_{\infty-}\times\mathbb{V}_{\infty-}$, let us define

$$\begin{aligned}D_\Theta F:=&\sum_{i=1}^{m_1}(\partial_i f)(\cdot)\int_0^1\langle h(s),h_i(s)\rangle_{\mathbb{R}^d}\mathrm{d}s\\&+\sum_{j=1}^{m_2}(\partial_{j+m_1}f)(\cdot)\int_0^1\!\!\int_{\Gamma_0}\mathbf{v}(s,z)\cdot\nabla_z g_j(s,z)\mu(\mathrm{d}s,\mathrm{d}z),\end{aligned} \tag{2.21}$$

where "$(\cdot)$" stands for $w(h_1), \cdots, w(h_{m_1}), \mu(g_1), \cdots, \mu(g_{m_2})$. By Hölder's inequality and (2.11), it is easy to see that for any $p \geqslant 1$,

$$D_\Theta F \in L^p \text{ and } D_\Theta F = \lim_{\varepsilon \to 0} \frac{F \circ \Theta^\varepsilon - F}{\varepsilon} \text{ in } L^p, \tag{2.22}$$

where Θ^ε is defined by (2.20). Thus, $D_\Theta F$ is well defined, i.e., it does not depend on the representation of F.

We have

Lemma 2.7. *Let $\Theta = (h, \mathbf{v}) \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$ and $\mathrm{div}\Theta$ be defined by (2.15).*

(i) (Density) $\mathcal{F}C_{\mathrm{p}}^\infty$ is dense in $L^p := L^p(\Omega, \mathscr{F}, \mathbb{P})$ for any $p \geqslant 1$.

(ii) (Integration by parts formula) For any $F \in \mathcal{F}C_{\mathrm{p}}^\infty$, we have

$$\mathbb{E}(D_\Theta F) = -\mathbb{E}(F \mathrm{div}\Theta). \tag{2.23}$$

(iii) (Closability) The linear operator $(D_\Theta, \mathcal{F}C_{\mathrm{p}}^\infty)$ is closable in L^p for any $p > 1$.

Proof. (i) is standard by a monotonic argument.

(ii) We first assume $\Theta = (h, \mathbf{v}) \in \mathbb{H}_0 \times \mathbb{V}_0$. By (2.22) and Theorem 2.6 , we have

$$\mathbb{E} D_\Theta F = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \mathbb{E}(F \circ \Theta^\varepsilon - F) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \mathbb{E}((1 - Q_1^\varepsilon) F \circ \Theta^\varepsilon) = -\mathbb{E}(F \mathrm{div}\Theta),$$

where we have used (2.18) in the last step. For general $\Theta = (h, \mathbf{v}) \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$ and $p > 2$, by Lemma 2.4 there exists a sequence of $\Theta_n = (h_n, \mathbf{v}_n) \in \mathbb{H}_0 \times \mathbb{V}_0$ such that

$$\lim_{n \to \infty} (\|h_n - h\|_{\mathbb{H}_p} + \|\mathbf{v}_n - \mathbf{v}\|_{\mathbb{V}_p}) = 0.$$

By the definition of $D_\Theta F$, it is easy to see that

$$\lim_{n \to \infty} \mathbb{E}|D_{\Theta_n} F - D_\Theta F|^2 = 0.$$

Moreover, by (2.16) we also have

$$\lim_{n \to \infty} \mathbb{E}|\mathrm{div}(\Theta_n - \Theta)|^2 \leqslant \lim_{n \to \infty} (\|h_n - h\|_{\mathbb{H}_2}^2 + \|\mathbf{v}_n - \mathbf{v}\|_{\mathbb{V}_2}^2) = 0.$$

By taking limits for $\mathbb{E}(D_{\Theta_n} F) = -\mathbb{E}(F \mathrm{div}\Theta_n)$, we obtain (2.23).

(iii) Fix $p > 1$. Let F_n be a sequence in $\mathcal{F}C_{\mathrm{p}}^\infty$ converging to zero in L^p. Suppose that $D_\Theta F_n$ converges to some ξ in L^p. We want to show $\xi = 0$. For any $G \in \mathcal{F}C_{\mathrm{p}}^\infty$, noticing that $F_n G \in \mathcal{F}C_{\mathrm{p}}^\infty$, by Hölder's inequality, we have

$$\begin{aligned}\mathbb{E}(G\xi) = \lim_{n \to \infty} \mathbb{E}(G D_\Theta F_n) &\overset{(2.22)}{=} \lim_{n \to \infty} \mathbb{E}(D_\Theta(F_n G)) - \lim_{n \to \infty} \mathbb{E}(F_n D_\Theta G) \\ &\overset{(2.23)}{=} -\lim_{n \to \infty} \mathbb{E}(F_n G \mathrm{div}\Theta) = 0.\end{aligned}$$

By (i), we obtain $\xi = 0$. The proof is complete. ☐

Definition 2.8. *For given $\Theta = (h, \mathbf{v}) \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$ and $p > 1$, we define the first order Sobolev space $\mathbb{W}_\Theta^{1,p}$ being the completion of $\mathcal{F}C_{\mathrm{p}}^\infty$ in $L^p(\Omega, \mathscr{F}, \mathbb{P})$ with respect to the norm:*

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

Clearly, $\mathbb{W}_\Theta^{1,p_2} \subset \mathbb{W}_\Theta^{1,p_1}$ for $p_2 > p_1 > 1$. We shall write

$$\mathbb{W}_\Theta^{1,\infty-} := \cap_{p>1} \mathbb{W}_\Theta^{1,p}.$$

We have the following integration by parts formula.

Theorem 2.9. *Let* $\Theta = (h, \mathbf{v}) \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$ *and* $p > 1$. *For any* $F \in \mathbb{W}^{1,p}_{\Theta}$, *we have*

$$\mathbb{E}(D_\Theta F) = -\mathbb{E}(F \mathrm{div}\Theta), \tag{2.24}$$

where $\mathrm{div}\Theta$ *is defined by (2.15).*

Proof. Let $F_n \in \mathcal{F}C^\infty_{\mathrm{p}}$ converge to F in $\mathbb{W}^{1,p}_{\Theta}$. By (2.23) we have

$$\mathbb{E}(D_\Theta F_n) = -\mathbb{E}(F_n \mathrm{div}\Theta).$$

By taking limits, we obtain (2.24). □

Moreover, we also have the following chain rule.

Proposition 2.10. *(Chain rule) Let* $\Theta = (h, \mathbf{v}) \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$. *For* $m, k \in \mathbb{N}$, *let* $F = (F_1, \cdots, F_m) \in (\mathbb{W}^{1,\infty-}_{\Theta})^m$ *and* $\varphi \in C^\infty_{\mathrm{p}}(\mathbb{R}^m; \mathbb{R}^k)$. *Then the composition* $\varphi(F) \in (\mathbb{W}^{1,\infty-}_{\Theta})^k$ *and*

$$D_\Theta \varphi(F) = D_\Theta F \cdot \nabla\varphi(F).$$

Proof. Since $\varphi \in C^\infty_{\mathrm{p}}(\mathbb{R}^m; \mathbb{R}^k)$, we can assume that for some $r \in \mathbb{N}$,

$$|\nabla\varphi(x)| \leqslant C(1 + |x|^r). \tag{2.25}$$

For any fixed $p > r + 1$, let $F_n \in (\mathcal{F}C^\infty_{\mathrm{p}})^m$ converge to F in $(\mathbb{W}^{1,p}_{\Theta})^m$. Since $\varphi(F_n) \in (\mathcal{F}C^\infty_{\mathrm{p}})^k$, by (2.22) it is easy to see that

$$D_\Theta \varphi(F_n) = D_\Theta F_n \cdot \nabla\varphi(F_n).$$

For any $q \in (1, \frac{p}{r+1})$, by Hölder's inequality and (2.25), we have

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}|D_\Theta F_n \cdot \nabla\varphi(F_n) - D_\Theta F \cdot \nabla\varphi(F)|^q \\
&\quad \leqslant C \lim_{n\to\infty} \left(\mathbb{E}|D_\Theta F_n|^p\right)^{\frac{q}{p}} \left(\mathbb{E}|\nabla\varphi(F_n) - \nabla\varphi(F)|^{\frac{qp}{p-q}}\right)^{\frac{p-q}{p}} \\
&\quad + C \lim_{n\to\infty} \left(\mathbb{E}|D_\Theta F_n - D_\Theta F|^p\right)^{\frac{q}{p}} \left(1 + \mathbb{E}|F|^{\frac{rqp}{p-q}}\right)^{\frac{p-q}{p}} = 0,
\end{aligned}$$

and also

$$\lim_{n\to\infty} \mathbb{E}|\varphi(F_n) - \varphi(F)|^q \leqslant C \lim_{n\to\infty} \left(\mathbb{E}|F_n - F|^p\right)^{\frac{q}{p}} \left(1 + \mathbb{E}|F_n|^{\frac{rqp}{p-q}} + \mathbb{E}|F|^{\frac{rqp}{p-q}}\right)^{\frac{p-q}{p}} = 0.$$

Thus, by definition we have $\varphi(F) \in (\mathbb{W}^{1,q}_{\Theta})^k$ and

$$D_\Theta \varphi(F) = D_\Theta F \cdot \nabla\varphi(F).$$

Since $p > r + 1$ is arbitrary and $q \in (1, \frac{p}{r+1})$, we obtain $\varphi(F) \in (\mathbb{W}^{1,\infty-}_{\Theta})^k$. □

2.4 Kusuoka-Stroock's formula

In this subsection we are about to establish a commutation formula between the gradient and Poisson stochastic integrals. On Wiener space this formula is given by Kusuoka and Stroock [12]. On configuration space similar formula is proven in [18].

Proposition 2.11. *Fix* $\Theta = (h, \mathbf{v}) \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$. *Let* $\eta(\omega, s, z) : \Omega \times [0,1] \times \Gamma_0 \to \mathbb{R}$ *be a measurable map and satisfy that for each* $(s, z) \in [0,1] \times \Gamma_0$,

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

and

$$s \mapsto \eta(s,z), D_\Theta \eta(s,z), \nabla_z \eta(s,z) \text{ are left-continuous and } \mathscr{F}_s\text{-adapted}, \tag{2.26}$$

and for any $p > 1$,

$$\mathbb{E}\left[\sup_{s\in[0,1]} \sup_{z\in\Gamma_0} \left(\frac{|\eta(s,z)|^p + |D_\Theta \eta(s,z)|^p}{(1 \wedge |z|)^p} + |\nabla_z \eta(s,z)|^p \right)\right] < +\infty. \tag{2.27}$$

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

$$D_\Theta \mathscr{I}(\eta) = \int_0^1 \int_{\Gamma_0} D_\Theta \eta(s,z) \tilde{N}(\mathrm{d}s, \mathrm{d}z) + \int_0^1 \int_{\Gamma_0} \mathbf{v}(s,z) \cdot \nabla \eta(s,z) N(\mathrm{d}s, \mathrm{d}z). \tag{2.28}$$

Proof. (i) First of all, we assume that $\eta(s,z) = 1_{(t_0,t_1]}(s)\eta(z)$, where $\eta(z)$ is $\mathscr{F}_{t_0}$-measurable, and satisfies (2.27) and

$$z \mapsto \eta(z) \text{ has compact support } U \subset \Gamma_0. \tag{2.29}$$

For $n \in \mathbb{N}$, let $\mathbb{D}_n$ be the grid of $\mathbb{R}^d$ with step 2^{-n}. For a point $z \in \mathbb{R}^d$, let $\phi_n(z)$ be the left-lower corner point in $\mathbb{D}_n$ which is closest to z. For $\varepsilon \in (0,1)$ and $R > 1$, let χ_ε and χ_R be defined by (2.3) and (2.4). For $\delta \in (0,1)$, let $\eta_\delta(z)$ be defined as in (2.8), and let us define

$$\eta^{\delta,n}_{\varepsilon,R}(\omega, y) := \chi_\varepsilon(y)\chi_R(y) \int_0^{y_1} \cdots \int_0^{y_d} (\partial_{z_1} \cdots \partial_{z_d} \eta_\delta)(\omega, \phi_n(z)) \mathrm{d}z_1 \cdots \mathrm{d}z_d.$$

From this definition, we can write

$$\eta^{\delta,n}_{\varepsilon,R}(\omega, z) = \sum_{k=1}^m \xi_j(\omega) g_j(z),$$

where $\xi_j \in \mathbb{W}^{1,\infty-}_{\Theta}$ is $\mathscr{F}_{t_0}$-measurable and g_j is smooth and has support

$$U_{\varepsilon,R} := \Gamma_\varepsilon \cap \{z : |z| \leqslant 2R\} \subset \Gamma_0.$$

By definition (2.21), it is easy to check that $\mathscr{I}(\eta^{\delta,n}_{\varepsilon,R}) := \int_U \eta^{\delta,n}_{\varepsilon,R}(z) \tilde{N}((t_0,t_1], \mathrm{d}z) \in \mathbb{W}^{1,\infty-}_{\Theta}$ and

$$D_\Theta \mathscr{I}(\eta^{\delta,n}_{\varepsilon,R}) = \int_{U_{\varepsilon,R}} D_\Theta \eta^{\delta,n}_{\varepsilon,R}(z) \tilde{N}((t_0,t_1], \mathrm{d}z) + \int_{t_0}^{t_1} \int_{U_{\varepsilon,R}} \mathbf{v}(s,z) \cdot \nabla \eta^{\delta,n}_{\varepsilon,R}(z) N(\mathrm{d}s, \mathrm{d}z).$$

Thus, for proving (2.28), by Lemma 2.3 it suffices to prove that for any $p > 1$,

$$\lim_{R\to\infty} \lim_{\varepsilon\to 0} \lim_{\delta\to\infty} \lim_{n\to\infty} \left(\|\eta^{\delta,n}_{\varepsilon,R} - \eta\|_{\mathbb{L}^1_p} + \|D_\Theta(\eta^{\delta,n}_{\varepsilon,R} - \eta)\|_{\mathbb{L}^1_p} \right) = 0,$$

and

$$\lim_{R\to\infty} \lim_{\varepsilon\to 0} \lim_{\delta\to\infty} \lim_{n\to\infty} \|\mathbf{v} \cdot \nabla(\eta^{\delta,n}_{\varepsilon,R} - \eta)\|_{\mathbb{L}^1_p} = 0. \tag{2.30}$$

We only prove the second limit. The first limit is similar. For fixed ε, R, set $\eta_{\varepsilon,R} := \chi_\varepsilon \chi_R \eta$. Since for $z \notin U_{\varepsilon,R}$,

$$\eta^{\delta,n}_{\varepsilon,R}(z) = \eta_{\varepsilon,R}(z) = 0,$$

by Remark 2.1 and Hölder's inequality, we have

$$\begin{aligned}
&\lim_{\delta\to 0}\lim_{n\to\infty} \|\mathbf{v}\cdot\nabla(\eta^{\delta,n}_{\varepsilon,R} - \eta_{\varepsilon,R})\|^p_{\mathbb{L}^1_p} \\
&\leqslant C \lim_{\delta\to 0}\lim_{n\to\infty} \mathbb{E}\int_{t_0}^{t_1}\int_{U_{\varepsilon,R}} |\mathbf{v}(s,z)\cdot\nabla(\eta^{\delta,n}_{\varepsilon,R} - \eta_{\varepsilon,R})(s,z)|^p \mathrm{d}z\mathrm{d}s \\
&\leqslant C \lim_{\delta\to 0}\lim_{n\to\infty} \left(\mathbb{E}\int_{t_0}^{t_1}\int_{U_{\varepsilon,R}} |\nabla(\eta^{\delta,n}_{\varepsilon,R} - \eta_{\varepsilon,R})(s,z)|^{2p} \mathrm{d}z\mathrm{d}s\right)^{\frac{1}{2}} = 0. \qquad (2.31)
\end{aligned}$$

On the other hand, since η has compact support U, by (2.27) and the dominated convergence theorem, we have

$$\lim_{R\to\infty}\lim_{\varepsilon\to 0} \|\mathbf{v}\cdot\nabla(\eta_{\varepsilon,R} - \eta)\|^p_{\mathbb{L}^1_p} = \lim_{R\to\infty}\lim_{\varepsilon\to 0} \|(1-\chi_\varepsilon\chi_R)\mathbf{v}\cdot\nabla\eta\|^p_{\mathbb{L}^1_p} = 0. \qquad (2.32)$$

Combining (2.31) and (2.32), we obtain (2.30).

(ii) Next we assume that for some compact set $U \subset \Gamma_0$,

$$\eta(s,z) = 0, \quad z \notin U. \qquad (2.33)$$

For $n \in \mathbb{N}$, let $s_k := k/n$ and define

$$\eta_n(s,z) := \sum_{k=1}^n 1_{(s_{k-1},s_k]}(s)\eta(s_{k-1},z).$$

In this case, we have

$$\mathscr{I}(\eta_n) = \sum_{k=1}^n \left(\int_{\Gamma_0} \eta(s_{k-1},z)N((s_{k-1},s_k],\mathrm{d}z) - \frac{1}{n}\int_{\Gamma_0} \eta(s_{k-1},z)\nu(\mathrm{d}z)\right)$$

By (i), we have

$$\mathscr{I}(\eta_n) \in \mathbb{W}^{1,\infty-}_\Theta$$

and

$$D_\Theta \mathscr{I}(\eta_n) = \int_0^1\int_{\Gamma_0} D_\Theta\eta_n(s,z)\tilde{N}(\mathrm{d}s,\mathrm{d}z) + \int_0^1\int_{\Gamma_0} \mathbf{v}(s,z)\cdot\nabla\eta_n(s,z)N(\mathrm{d}s,\mathrm{d}z).$$

By Lemma 2.3 and (2.33), for any $p \geqslant 2$, we have

$$\begin{aligned}
&\mathbb{E}\left|\int_0^1\int_{\Gamma_0} (D_\Theta\eta_n(s,z) - D_\Theta\eta(s,z))\tilde{N}(\mathrm{d}s,\mathrm{d}z)\right|^p \\
&\qquad \leqslant C\mathbb{E}\left(\int_0^1\int_U |D_\Theta\eta_n(s,z) - D_\Theta\eta(s,z)|^p \mathrm{d}z\mathrm{d}s\right)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\left|\int_0^1\int_{\Gamma_0} \mathbf{v}(s,z)\cdot\nabla(\eta_n - \eta)(s,z)N(\mathrm{d}s,\mathrm{d}z)\right|^p \\
&\qquad \leqslant C\mathbb{E}\left(\int_0^1\int_U |\nabla\eta_n(s,z) - \nabla\eta(s,z)|^p\,|\mathbf{v}(s,z)|^p \mathrm{d}z\mathrm{d}s\right).
\end{aligned}$$

By the assumptions and the dominated convergence theorem, both of them converge to zero as $n\to\infty$, and we obtain (2.28).

(iii) We now drop the assumption (2.33). Define

$$\eta_{\varepsilon,R}(s,z):=\chi_\varepsilon(z)\chi_R(z)\eta(s,z),$$

where χ_ε and χ_R are the same as in (2.3) and (2.4). By (ii), we have

$$D_\Theta\mathscr{I}(\eta_{\varepsilon,R})=\int_0^1\!\!\int_{\Gamma_0}D_\Theta\eta_{\varepsilon,R}(s,z)\tilde N(\mathrm{d}s,\mathrm{d}z)+\int_0^1\!\!\int_{\Gamma_0}\mathbf{v}(s,z)\cdot\nabla\eta_{\varepsilon,R}(s,z)N(\mathrm{d}s,\mathrm{d}z).$$

For proving (2.28), it suffices to prove that for any $p\geqslant 2$,

$$I^{(1)}_{\varepsilon,R}:=\mathbb{E}\left|\int_0^1\!\!\int_{\Gamma_0}(1-\chi_\varepsilon(z)\chi_R(z))D_\Theta\eta(s,z)\tilde N(\mathrm{d}s,\mathrm{d}z)\right|^p\to 0,$$
$$I^{(2)}_{\varepsilon,R}:=\mathbb{E}\left|\int_0^1\!\!\int_{\Gamma_0}\mathbf{v}(s,z)\cdot\nabla(\eta_{\varepsilon,R}-\eta)(s,z)N(\mathrm{d}s,\mathrm{d}z)\right|^p\to 0,$$

as $R\to\infty$ and $\varepsilon\to 0$. The first limit follows by (2.11), (2.27) and the dominated convergence theorem. For the second limit, noticing that as in (2.6),

$$|\nabla\eta_{\varepsilon,R}(s,z)-\nabla\eta(s,z)|\leqslant C\varrho(z)\Big(1_{z\in\Gamma^c_{2\varepsilon}}+1_{|z|>R}\Big)|\eta(s,z)|+\Big(1_{z\in\Gamma^c_{2\varepsilon}}+1_{|z|>R}\Big)|\nabla\eta(s,z)|,$$

by (2.10) we have

$$\begin{aligned}&\mathbb{E}\left|\int_0^1\!\!\int_{\Gamma_0}\mathbf{v}(s,z)\cdot\nabla(\eta_{\varepsilon,R}-\eta)(s,z)N(\mathrm{d}s,\mathrm{d}z)\right|^p\\
&\quad\leqslant C\mathbb{E}\left(\sup_{s,z}(|\eta(s,z)|+|\nabla\eta(s,z)|)\int_0^1\!\!\int_{\Gamma^c_{2\varepsilon}\cup\{|z|>R\}}|\varrho(z)\mathbf{v}(s,z)|\nu(\mathrm{d}z)\mathrm{d}s\right)^p\\
&\quad\quad+C\mathbb{E}\left(\sup_{s,z}(|\eta(s,z)|+|\nabla\eta(s,z)|)^p\int_0^1\!\!\int_{\Gamma^c_{2\varepsilon}\cup\{|z|>R\}}|\varrho(z)\mathbf{v}(s,z)|^p\nu(\mathrm{d}z)\mathrm{d}s\right),\end{aligned}$$

which converges to zero as $\varepsilon\to 0$ and $R\to\infty$. The proof is complete. □

3 Reduced Malliavin matrix for SDEs driven by Lévy noises

As discussed in the introduction, in the remainder of this paper, we shall assume

$$\Gamma_0=\{z\in\mathbb{R}^d:0<|z|<1\},$$

and

$$\tfrac{\nu(\mathrm{d}z)}{\mathrm{d}z}|_{\Gamma_0}=\kappa(z)\text{ with }\kappa\text{ satisfying }\mathbf{(H}^\alpha_1\mathbf{)}.$$

Let $N(\mathrm{d}t,\mathrm{d}z)$ be the Poisson random measure associated with L^0_t, i.e.,

$$N((0,t]\times E):=\sum_{s\leqslant t}1_E(L^0_s-L^0_{s-}),\quad E\in\mathscr{B}(\Gamma_0).$$

Since $\nu(\mathrm{d}z)$ is symmetric, by Lévy-Itô's decomposition, we can write

$$L^0_t=\int_0^t\!\!\int_{\Gamma_0}z\tilde N(\mathrm{d}s,\mathrm{d}z)=\int_0^t\!\!\int_{\Gamma_0}z(N(\mathrm{d}s,\mathrm{d}z)-\mathrm{d}s\nu(\mathrm{d}z)).$$

By Proposition 2.11, for any $\Theta = (h, \mathbf{v}) \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$, we have $W_t, L_t^0 \in \mathbb{W}_\Theta^{1,\infty-}$ and

$$D_\Theta W_t = \int_0^t h(s)\mathrm{d}s, \quad D_\Theta L_t^0 = \int_0^t\!\!\int_{\Gamma_0} \mathbf{v}(s,z)N(\mathrm{d}s,\mathrm{d}z). \tag{3.1}$$

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

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + A_1\mathrm{d}W_t + A_2\mathrm{d}L_t^0, \quad X_0 = x.$$

Proposition 3.1. *Assume that* $b \in C^1$ *has bounded derivative. For fixed* $\Theta = (h, \mathbf{v}) \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$*, we have* $X_t \in \mathbb{W}_\Theta^{1,\infty-}$ *and*

$$D_\Theta X_t = \int_0^t \nabla b(X_s) D_\Theta X_s \mathrm{d}s + A_1 \int_0^t h(s)\mathrm{d}s + A_2 \int_0^t\!\!\int_{\Gamma_0} \mathbf{v}(s,z)N(\mathrm{d}s,\mathrm{d}z). \tag{3.2}$$

Proof. Consider the following Picard's iteration: $X_t^0 = x$ and for $n \in \mathbb{N}$,

$$X_t^n = x + \int_0^t b(X_s^{n-1})\mathrm{d}s + A_1 W_t + A_2 L_t^0.$$

It is by now standard to prove that for any $t \geqslant 0$ and $p \geqslant 1$,

$$\lim_{n\to\infty} \mathbb{E}|X_t^n - X_t|^p = 0. \tag{3.3}$$

Since $\Theta \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$, by (3.1) and the induction, we have $X_t^n \in \mathbb{W}_\Theta^{1,\infty-}$ and

$$D_\Theta X_t^n = \int_0^t \nabla b(X_s^{n-1}) D_\Theta X_t^{n-1} \mathrm{d}s + A_1 \int_0^t h(s)\mathrm{d}s + A_2 \int_0^t\!\!\int_{\Gamma_0} \mathbf{v}(s,z)N(\mathrm{d}s,\mathrm{d}z).$$

By Gronwall's inequality, it is easy to prove that for any $T > 0$ and $p \geqslant 1$,

$$\sup_{n\in\mathbb{N}} \sup_{t\in[0,T]} \mathbb{E}|D_\Theta X_t^n|^p < +\infty.$$

Let Y_t solve the following SDE:

$$Y_t = \int_0^t \nabla b(X_s) Y_s \mathrm{d}s + A_1 \int_0^t h(s)\mathrm{d}s + A_2 \int_0^t\!\!\int_{\Gamma_0} \mathbf{v}(s,z)N(\mathrm{d}s,\mathrm{d}z).$$

By Fatou's lemma and (3.3), we have

$$\varlimsup_{n\to\infty} \mathbb{E}|D_\Theta X_t^n - Y_t|^p \leqslant \|\nabla b\|_\infty^p \int_0^t \varlimsup_{n\to\infty} \mathbb{E}|D_\Theta X_s^{n-1} - Y_s|^p \mathrm{d}s,$$

which then gives

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

Thus, by (3.3) we have $X_t \in \mathbb{W}_\Theta^{1,p}$ and $D_\Theta X_t = Y_t$. The proof is complete. □

Let $J_t := J_t(x) := \nabla X_t(x)$ be the Jacobian matrix and $K_t := K_t(x) := J_t^{-1}(x)$. Then J_t and K_t solve the following ODEs

$$J_t = \mathbb{I} + \int_0^t \nabla b(X_s) J_s \mathrm{d}s, \quad K_t = \mathbb{I} - \int_0^t K_s \nabla b(X_s)\mathrm{d}s, \tag{3.4}$$

and it is easy to see that

$$\sup_{t\in[0,1]} \sup_{x\in\mathbb{R}^d} |J_t(x)| \leqslant \mathrm{e}^{\|\nabla b\|_\infty}, \quad \sup_{t\in[0,1]} \sup_{x\in\mathbb{R}^d} |K_t(x)| \leqslant \mathrm{e}^{\|\nabla b\|_\infty}. \tag{3.5}$$

By (3.2) and the formula of constant variation, we have

$$D_\Theta X_t = J_t \int_0^t K_s A_1 h(s)\mathrm{d}s + J_t \int_0^t\!\!\int_{\Gamma_0} K_s A_2 \mathbf{v}(s,z) N(\mathrm{d}s,\mathrm{d}z). \tag{3.6}$$

Below, let $\zeta(z)$ be a nonnegative smooth function with

$$\zeta(z) = |z|^3, \quad |z| \leqslant 1/4, \quad \zeta(z) = 0, \quad |z| > 1/2. \tag{3.7}$$

Let us choose

$$\Theta_j(x) = (h_j(x;\cdot), \mathbf{v}_j(x;\cdot))$$

with

$$h_j(x;s) = (K_s(x)A_1)^*_{\cdot j}, \quad \mathbf{v}_j(x;s,z) = (K_s(x)A_2)^*_{\cdot j}\zeta(z).$$

Lemma 3.2. *Under* $(\mathbf{H}^\alpha_1)$*, for each* $j = 1, \cdots, d$ *and* $x \in \mathbb{R}^d$*, we have* $\Theta_j(x) \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$ *and*

$$\mathrm{div}\Theta_j(x) = -\sum_l \int_0^1 (K_s(x)A_1)_{lj} \mathrm{d}W^l_s + \sum_l \int_0^1\!\!\int_{\Gamma_0} (K_s(x)A_2)_{lj}\eta_l(z)\tilde{N}(\mathrm{d}z,\mathrm{d}s), \tag{3.8}$$

where $\eta_l(z) := \partial_l\zeta(z) + \zeta(z)\partial_l \log\kappa(z)$*. In particular, for any* $p \geqslant 2$*,*

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}|\mathrm{div}\Theta_j(x)|^p < +\infty. \tag{3.9}$$

Proof. Since $\mathbf{d}(z,\Gamma^c_0) \geqslant |z| \wedge (1-|z|)$, by (3.5) and (3.7), it is easy to check that $\Theta_j(x) \in \mathbb{H}_{\infty-} \times \mathbb{V}_{\infty-}$. Moreover, by definition (2.15) we immediately have (3.8). As for (3.9), it follows by (2.16). □

Write

$$\mathbf{\Theta} := (\Theta_1, \cdots, \Theta_d), \quad (D_{\mathbf{\Theta}} X_t)_{ij} := D_{\Theta_j} X^i_t$$

and

$$\Sigma_t(x) := \int_0^t K_s(x)A_1A_1^*K_s^*(x)\mathrm{d}s + \int_0^t\!\!\int_{\Gamma_0} K_s(x)A_2A_2^*K_s^*(x)\zeta(z)N(\mathrm{d}s,\mathrm{d}z), \tag{3.10}$$

then by (3.6),

$$D_{\mathbf{\Theta}} X_t(x) = J_t(x)\Sigma_t(x). \tag{3.11}$$

The matrix $\Sigma_t(x)$ is called the reduced Malliavin matrix (cf. [4, p. 89, (7-20)] and [21, (2.12)]).

Lemma 3.3. *Assume that* $b \in C^\infty$ *has bounded derivatives of all orders. For any* $k, n \in \mathbb{N}\cup\{0\}$ *with* $k+n \geqslant 1$*,* $j_1, \cdots, j_n \in \{1, \cdots, d\}$ *and* $p \geqslant 2$*, we have*

$$\sup_{t\in[0,1]}\sup_{x\in\mathbb{R}^d} \mathbb{E}|D_{\Theta_{j_1}}\cdots D_{\Theta_{j_n}}\nabla^k X_t(x)|^p < \infty, \tag{3.12}$$

$$\sup_{t\in[0,1]}\sup_{x\in\mathbb{R}^d} \mathbb{E}|D_{\Theta_{j_1}}\cdots D_{\Theta_{j_n}} J_t(x)|^p < \infty, \tag{3.13}$$

$$\sup_{t\in[0,1]}\sup_{x\in\mathbb{R}^d} \mathbb{E}|D_{\Theta_{j_1}}\cdots D_{\Theta_{j_n}} K_t(x)|^p < \infty, \tag{3.14}$$

$$\sup_{t\in[0,1]}\sup_{x\in\mathbb{R}^d} \mathbb{E}|D_{\Theta_{j_1}}\cdots D_{\Theta_{j_n}} \Sigma_t(x)|^p < \infty. \tag{3.15}$$

Moreover, under $(\mathbf{H}^\alpha_m)$ *with* $m \geqslant 2$*, for any* $n \leqslant m-1$*, we also have*

$$\sup_{t\in[0,1]}\sup_{x\in\mathbb{R}^d} \mathbb{E}|D_{\Theta_{j_1}}\cdots D_{\Theta_{j_n}} \mathrm{div}\Theta_i(x)|^p < \infty. \tag{3.16}$$

Proof. First of all, by equation (3.4) and induction, it is easy to prove that for any $k \in \mathbb{N}$,

$$\sup_{t\in[0,1]} \sup_{x\in\mathbb{R}^d} \sup_{\omega} |\nabla^k X_t(x,\omega)| < +\infty. \tag{3.17}$$

By (3.10), (3.5) and inequality (2.10), we have for any $p \geqslant 1$,

$$\sup_{t\in[0,1]} \sup_{x\in\mathbb{R}^d} \mathbb{E}|\Sigma_t(x)|^p < \infty,$$

which together with (3.11) and (3.5) yields (3.12) with $n = 1$ and $k = 0$. By induction, the higher order derivatives for (3.12)-(3.15) follow by (3.2), (3.10), Proposition 2.11 and (2.10).

We now look at (3.16). By (3.8) and Proposition 2.11, we have

$$\begin{aligned} D_{\Theta_j} \mathrm{div}\Theta_i = &- \sum_l \int_0^1 (D_{\Theta_j} K_s A_1)_{li} \mathrm{d}W_s^l - \sum_l \int_0^1 (K_s A_1)_{li}(K_s A_1)_{lj} \mathrm{d}s \\ &+ \sum_{l,l'} \int_0^1 \int_{\Gamma_0} (K_s A_2)_{li}(K_s A_2)_{l'j} \partial_{l'} \eta_l(z)\zeta(z) N(\mathrm{d}z, \mathrm{d}s) \\ &+ \sum_l \int_0^1 \int_{\Gamma_0} (D_{\Theta_j} K_s A_2)_{li} \eta_l(z) \tilde{N}(\mathrm{d}z, \mathrm{d}s). \end{aligned}$$

Recalling $\eta_l(z) := \partial_l \zeta(z) + \zeta(z)\partial_l \log \kappa(z)$ and $\zeta(z)$ given by (3.7), by **($\mathbf{H}_m^\alpha$)** we have

$$|\eta_l(z)| \leqslant C|z|^2, \quad |\partial_{l'}\eta_l(z)\zeta(z)| \leqslant C|z|^4, \quad z \in \Gamma_0.$$

By (2.10) and (3.14), we obtain (3.16) for $n = 1$. The higher order derivative estimates follow by induction. □

4 Proof of Theorem 1.1

4.1 Invertibility of Σ_t

We need the following easy fact (c.f. [22, Lemma 2.1]). For the readers' convenience, a short proof is provided here.

Lemma 4.1. *Set $\Delta L_s^0 := L_s^0 - L_{s-}^0$ and define*

$$\Omega_0 := \Big\{\omega : \{s : |\Delta L_s^0(\omega)| \neq 0\} \text{ is dense in } [0,\infty)\Big\}.$$

Under **($\mathbf{H}_1^\alpha$)**, *we have* $\mathbb{P}(\Omega_0) = 1$.

Proof. Define a stopping time $\tau := \inf\{t > 0 : |L_t^0| = 0\}$. As in the proof of [22, Lemma 2.1], it suffices to prove that

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

For any $\varepsilon \in (0,1)$, we have

$$\begin{aligned} \varepsilon^2 \geqslant \mathbb{E}\left(|\Delta L_\tau^0|^2 1_{|\Delta L_\tau^0| \leqslant \varepsilon}\right) &= \mathbb{E}\left(\sum_{0<s\leqslant\tau} |\Delta L_s^0|^2 1_{|\Delta L_s^0|\leqslant\varepsilon}\right) = \mathbb{E}\left(\int_0^\tau \int_{|z|\leqslant\varepsilon} |z|^2 N(\mathrm{d}s, \mathrm{d}z)\right) \\ &= \mathbb{E}\left(\int_0^\tau \int_{|z|\leqslant\varepsilon} |z|^2 \nu(\mathrm{d}z)\mathrm{d}s\right) = \int_{|z|\leqslant\varepsilon} |z|^2 \kappa(z) \mathrm{d}z \mathbb{E}\tau, \end{aligned}$$

which, together with **($\mathbf{H}_\kappa^\alpha$)** and letting $\varepsilon \to 0$, implies that

$$\mathbb{E}\tau = 0 \Rightarrow \mathbb{P}(\tau = 0) = 1.$$

The proof is complete. □

Let $V:\mathbb{R}^d\to\mathbb{R}^d\times\mathbb{R}^d$ be a matrix-valued C_b^2-function. Below we set

$$M_t^V:=\sum_{l=1}^d\int_0^t K_s\partial_l V(X_{s-})\mathrm{d}W_s^l,\quad G_t^V:=\int_0^t\!\!\int_{\Gamma_0}K_s(A_2z\cdot\nabla)V(X_{s-})\tilde{N}(\mathrm{d}s,\mathrm{d}z),$$

and

$$H_t^V:=\sum_{0<s\leqslant t}K_s\Big(V(X_s)-V(X_{s-})-(\Delta X_s\cdot\nabla)V(X_{s-})\Big).$$

We have

Lemma 4.2. *There exists a subsequence $n_m\to\infty$ such that $\mathbb{P}(\Omega_1^V)=1$, where*

$$\Omega_1^V:=\Bigg\{\omega:H_t^V=\lim_{n_m\to\infty}\sum_{0<s\leqslant t}K_s\Big(V(X_s)-V(X_{s-})-(\Delta X_s\cdot\nabla V)(X_{s-})\Big)1_{|\Delta L_s^0|>\frac{1}{n_m}}$$

$$\text{and } G_t^V=\lim_{n_m\to\infty}\sum_{s\in(0,t]}K_s(A_2\Delta L_s^0\cdot\nabla)V(X_{s-})1_{|\Delta L_s^0|>\frac{1}{n_m}}\text{ uniformly in } t\in[0,1]\Bigg\}.$$

Proof. By $\nu(\mathrm{d}z)=\nu(-\mathrm{d}z)$ and Doob's maximal inequality, we have

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\in[0,1]}\left|G_t^V-\sum_{s\in(0,t]}K_s(A_2\Delta L_s^0\cdot\nabla)V(X_{s-})1_{|\Delta L_s^0|>\frac{1}{n}}\right|^2\right)\\
&\quad=\mathbb{E}\left(\sup_{t\in[0,1]}\left|\int_0^t\!\!\int_{|z|\leqslant\frac{1}{n}}K_s(A_2z\cdot\nabla)V(X_{s-})\tilde{N}(\mathrm{d}s,\mathrm{d}z)\right|^2\right)\\
&\quad\leqslant 4\mathbb{E}\left|\int_0^1\!\!\int_{|z|\leqslant\frac{1}{n}}K_s(A_2z\cdot\nabla)V(X_{s-})\tilde{N}(\mathrm{d}s,\mathrm{d}z)\right|^2\\
&\quad\leqslant 4\mathbb{E}\left(\int_0^1\!\!\int_{|z|\leqslant\frac{1}{n}}|K_s(A_2z\cdot\nabla)V(X_{s-})|^2\nu(\mathrm{d}z)\mathrm{d}s\right)\\
&\quad\leqslant C\int_{|z|\leqslant\frac{1}{n}}|z|^2\nu(\mathrm{d}z)\to0,\quad n\to\infty.
\end{aligned}$$

Similarly, we also have

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\in[0,1]}\left|\sum_{0<s\leqslant t}K_s\Big(V(X_s)-V(X_{s-})-(\Delta X_s\cdot\nabla)V(X_{s-})\Big)1_{|\Delta L_s^0|\leqslant\frac{1}{n}}\right|\right)\\
&\quad=\mathbb{E}\left(\sup_{t\in[0,1]}\left|\int_0^t\!\!\int_{|z|\leqslant\frac{1}{n}}K_s\Big(V(X_{s-}+z)-V(X_{s-})-(A_2z\cdot\nabla V)(X_{s-})\Big)N(\mathrm{d}s,\mathrm{d}z)\right|\right)\\
&\quad\leqslant\mathbb{E}\left(\int_0^1\!\!\int_{|z|\leqslant\frac{1}{n}}|K_s(V(X_{s-}+z)-V(X_{s-})-(A_2z\cdot\nabla)V(X_{s-}))|\nu(\mathrm{d}z)\mathrm{d}s\right)\\
&\quad\leqslant C\int_{|z|\leqslant\frac{1}{n}}|z|^2\nu(\mathrm{d}z)\to0,\quad n\to\infty.
\end{aligned}$$

The proof is complete. □

By [17, p.64, Theorem 21 and p.68, Theorem 23], we have

Lemma 4.3. *For $n \in \mathbb{N}$, let $t_k := k/n$. There exists a subsequence $n_m \to \infty$ such that*

$$\mathbb{P}(\Omega_2^V \cap \Omega_3^V) = 1,$$

where

$$\Omega_2^V := \left\{ \omega : \lim_{n_m \to \infty} \sum_{k=0}^{n_m-1} \sum_l K_{t_k} \partial_l V(X_{t_k})(W^l_{t_{k+1}\wedge t} - W^l_{t_k \wedge t}) = M_t^V \text{ uniformly in } t \in [0,1] \right\},$$

$$\Omega_3^V := \Bigg\{ \omega : \ \forall i,j,i',j' = 1,\cdots,d, \lim_{n_m \to \infty} \sum_{k=0}^{n_m-1} (M^V_{t_{k+1}\wedge t} - M^V_{t_k\wedge t})_{ij} (M^V_{t_{k+1}\wedge t} - M^V_{t_k\wedge t})_{i'j'}$$
$$= \sum_{k,l,k'} \int_0^t (K_s)_{ik} \partial_l V_{jk}(X_s)(K_s)_{i'k'} \partial_l V_{j'k'}(X_s) \mathrm{d}s \text{ uniformly in } t \in [0,1] \Bigg\}.$$

By Itô's formula, we also have

Lemma 4.4. *Let $[b,V] := b \cdot \nabla V - \nabla b \cdot V + \frac{1}{2} \nabla^2_{A_1 A_1^*} V$ and define*

$$\Omega_4^V := \left\{ \omega : K_t V(X_t) = V(x) + \int_0^t K_s [b,V](X_s) \mathrm{d}s + H_t^V + M_t^V + G_t^V, \forall t \geqslant 0 \right\}.$$

Then $\mathbb{P}(\Omega_4^V) = 1$.

Now we can prove the following key lemma.

Lemma 4.5. *Fix $x \in \mathbb{R}^d$. Let $B_0 := \mathbb{I}$ and for $n \in \mathbb{N}$,*

$$B_n(x) := [b, B_{n-1}](x) := b(x) \cdot \nabla B_{n-1}(x) - \nabla b(x) \cdot B_{n-1}(x) + \tfrac{1}{2} \nabla^2_{A_1 A_1^*} B_{n-1}(x).$$

Assume that for some $n = n(x) \in \mathbb{N}$,

$$\mathrm{Rank}[A_1, B_1(x)A_1, \cdots, B_n(x)A_1, A_2, B_1(x)A_2, \cdots, B_n(x)A_2] = d. \tag{4.1}$$

Then under $(\mathbf{H}_1^\alpha)$*, for any $t > 0$, $\Sigma_t(x)$ is almost surely invertible.*

Proof. Set

$$\tilde{\Omega} := \cap_{n=1}^\infty (\Omega_1^{B_n} \cap \Omega_2^{B_n} \cap \Omega_3^{B_n} \cap \Omega_4^{B_n}) \cap \Omega_0.$$

Then by Lemmas 4.1-4.4, we have

$$\mathbb{P}(\tilde{\Omega}) = 1.$$

We want to prove that under (4.1), for each $t > 0$, the reduced Malliavin matrix $\Sigma_t(x,\omega)$ is invertible for each $\omega \in \tilde{\Omega}$. Without loss of generality, we assume $t = 1$ and fix an $\omega \in \tilde{\Omega}$. For simplicity of notation, we shall drop (x,ω) below. By (3.10), for a row vector $u \in \mathbb{R}^d$ we have

$$\begin{aligned} u\Sigma_1 u^* &= \int_0^1 |uK_s A_1|^2 \mathrm{d}s + \int_0^1 \int_{\Gamma_0} |uK_s A_2|^2 \zeta(z) N(\mathrm{d}s, \mathrm{d}z) \\ &= \int_0^1 |uK_s A_1|^2 \mathrm{d}s + \sum_{s \leqslant 1} |uK_s A_2|^2 \zeta(\Delta L_s^0) 1_{|\Delta L_s^0| \neq 0}. \end{aligned}$$

Suppose that for some $u \in \mathbb{S}^{d-1}$,

$$u\Sigma_1 u^* = 0.$$

Since $s \mapsto K_s$ is continuous and $\omega \in \Omega_0$, we have

$$|uK_s A_1|^2 = |uK_s A_2|^2 = 0, \quad \forall s \in [0,1].$$

Hence, by (3.4) we have

$$0 = uK_tA_i = uA_i - \int_0^t uK_s\nabla b(X_s)A_i\mathrm{d}s, \forall t \in [0,1],\ i = 1,2,$$

which implies that

$$uA_i = 0, \quad i = 1,2,$$

and

$$uK_t\nabla b(X_t)A_i = uK_tB_1(X_t)A_i = 0, \quad t \in [0,1], \quad i = 1,2.$$

Now we use induction to prove

$$uK_tB_n(X_t)A_i = 0, \quad t \in [0,1], \quad i = 1,2. \tag{4.2}$$

Suppose that (4.2) holds for some $n \in \mathbb{N}$. In view of $\omega \in \Omega_4^{B_n}$, we have for all $t \in [0,1]$,

$$uK_tB_n(X_t)A_i = uB_n(x)A_i + \int_0^t uK_sB_{n+1}(X_s)A_i\mathrm{d}s + uH_t^{B_n}A_i + uM_t^{B_n}A_i + uG_t^{B_n}A_i.$$

By the induction hypothesis and the definition of $H_t^{B_n}$, we further have

$$\begin{aligned} 0 &= \int_0^t uK_sB_{n+1}A_i(X_s)\mathrm{d}s - \sum_{0<s\leqslant t} uK_s(\Delta X_s\cdot\nabla)B_n(X_{s-})A_i + uM_t^{B_n}A_i + uG_t^{B_n}A_i \\ &= \int_0^t uK_sB_{n+1}A_i(X_s)\mathrm{d}s + uM_t^{B_n}A_i,\ \forall t \in [0,1] \text{ (since } \omega \in \Omega_1^{B_n}), \end{aligned} \tag{4.3}$$

which together with $\omega \in \Omega_3^{B_n}$ implies that

$$\begin{aligned} 0 &= \lim_{n_m\to\infty} \sum_{k=0}^{n_m-1} \langle uM_{t_{k+1}}^{B_n}A_i - uM_{t_k}^{B_n}A_i, uM_{t_{k+1}}^{B_n}A_i - uM_{t_k}^{B_n}A_i\rangle_{\mathbb{R}^d} \\ &= \sum_l \int_0^1 |uK_s\partial_lB_n(X_s)A_i|^2\mathrm{d}s. \end{aligned}$$

In particular,

$$uK_s\partial_lB_n(X_s)A_i = 0, \quad \forall s \in [0,1].$$

Since $\omega \in \Omega_2^{B_n}$, we also have

$$uM_t^{B_n}A_i = 0, \quad \forall t \in [0,1],$$

which together with (4.3) implies that

$$uK_sB_{n+1}A_i(X_s) = 0, \quad \forall s \in [0,1].$$

Thus, we obtain

$$uA_i = uB_1A_i = \cdots uB_nA_i = 0, \quad i = 1,2,$$

which is contradict with (4.1). The proof is complete. □

4.2 Proof of Theorem 1.1

Now we can finish the proof of Theorem 1.1 by the same argument as in [7]. We divide the proof into two steps.

(1) Let $\mathrm{GL}(d) \simeq \mathbb{R}^d \times \mathbb{R}^d$ be the set of all $d \times d$-matrix. Define

$$\mathbb{M}_n := \Big\{\Sigma \in \mathrm{GL}(d) : |\Sigma| \leqslant n, \ \ \det(\Sigma) \geqslant 1/n\Big\}.$$

Then $\mathbb{M}_n$ is a compact subset of $\mathrm{GL}(d)$. Let $\Phi_n \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ be a smooth function with

$$\Phi_n|_{\mathbb{M}_n} = 1, \ \ \Phi_n|_{\mathbb{M}^c_{n+1}} = 0, \ \ 0 \leqslant \Phi_n \leqslant 1.$$

Below we fix $t > 0$ and $x \in \mathbb{R}^d$. For each $n \in \mathbb{N}$, let us define a finite measure μ_n by

$$\mu_n(A) := \mathbb{E}\Big[1_A(X_t)\Phi_n(\Sigma_t)\Big], \ \ A \in \mathscr{B}(\mathbb{R}^d).$$

For each $\varphi \in C_b^\infty(\mathbb{R}^d)$, by the chain rule and (3.11), we have

$$D_{\mathbf{\Theta}}(\varphi(X_t)) = \nabla\varphi(X_t) D_{\mathbf{\Theta}} X_t = \nabla\varphi(X_t) J_t \Sigma_t,$$

where $\nabla = (\partial_1, \cdots, \partial_d)$. So,

$$\nabla\varphi(X_t) = D_{\mathbf{\Theta}}(\varphi(X_t))\Sigma_t^{-1} K_t. \tag{4.4}$$

Thus, by the integration by parts formula (2.24), we have for $i = 1, \cdots, d$,

$$\begin{aligned}\int_{\mathbb{R}^d} \partial_i\varphi(y)\mu_n(\mathrm{d}y) &= \mathbb{E}[\partial_i\varphi(X_t)\Phi_n(\Sigma_t)] \\ &= \sum_j \mathbb{E}\Big[D_{\Theta_j}(\varphi(X_t))(\Sigma_t^{-1}K_t)_{ij}\Phi_n(\Sigma_t)\Big] = \mathbb{E}[\varphi(X_t)H_t^i],\end{aligned} \tag{4.5}$$

where

$$H_t^i := -\sum_j \left((\Sigma_t^{-1}K_t)_{ij}\Phi_n(\Sigma_t)\mathrm{div}(\Theta_j) + D_{\Theta_j}((\Sigma_t^{-1}K_t)_{ij}\Phi_n(\Sigma_t))\right).$$

From this and using Lemma 3.3, by cumbersome calculations, we derive that

$$\left|\int_{\mathbb{R}^d} \partial_i\varphi(y)\mu_n(\mathrm{d}y)\right| \leqslant C_n\|\varphi\|_\infty, \ \ i = 1, \cdots, d,$$

where C_n is independent of t, x. Hence, μ_n is absolutely continuous with respect to the Lebesgue measure (cf. [14]), and by the Sobolev embedding theorem (cf. [1]), the density $p_n(y)$ satisfies that for any $q \in [1, d/(d-1))$,

$$\int_{\mathbb{R}^d} p_n(y)^q \mathrm{d}y \leqslant C_{d,q,n},$$

where the constant $C_{d,q,n}$ is independent of t, x. Therefore, for any Borel set $F \subset \mathbb{R}^d$ and $R > 0$, we have

$$\mu_n(F) = \int_F p_n(y)\mathrm{d}y \leqslant m(F)R + \int_{F\cap\{p_n > R\}} p_n(y)\mathrm{d}y \leqslant m(F)R + \frac{C_{d,q,n}}{R^{q-1}}, \tag{4.6}$$

where m is the Lebesgue measure and $q > 1$. In particular, for any Lebesgue zero measure set $A \subset \mathbb{R}^d$,

$$\mathbb{E}\Big[1_A(X_t)\Phi_n(\Sigma_t)\Big] = 0.$$

By Lemma 4.5 and the dominated convergence theorem, we obtain that for any Lebesgue zero measure set $A \subset \mathbb{R}^d$,

$$\mathbb{E}[1_A(X_t)] = 0,$$

which means that the law of X_t is absolutely continuous with respect to the Lebesgue measure.

(2) Let $\chi_n \in C^\infty(\mathbb{R}^d)$ be a smooth function with

$$\chi_n|_{\{|x|\leqslant n\}} = 1, \quad \chi_n|_{\{|x|>n+1\}} = 0, \quad 0 \leqslant \chi_n \leqslant 1. \tag{4.7}$$

Let f be a bounded nonnegative measurable function. By Lusin's theorem, for any $\varepsilon > 0$, there exist a set $F_\varepsilon \subset \{x \in \mathbb{R}^d : |x| < n+1\}$ and a nonnegative continuous function $g \in C_c(\mathbb{R}^d)$ such that

$$f\chi_n|_{F_\varepsilon^c} = g|_{F_\varepsilon^c}, \quad \|g\|_\infty \leqslant \|f\|_\infty, \quad m(F_\varepsilon) < \varepsilon.$$

Let $\mu_{t,x;n}$ be defined by

$$\mu_{t,x;n}(A) := \mathbb{E}\Big[1_A(X_t(x))\Phi_n(\Sigma_t(x))\Big], \quad A \in \mathscr{B}(\mathbb{R}^d).$$

By the dominated convergence theorem and (4.6), we have for any $R > 0$,

$$\begin{aligned}
&\varlimsup_{x\to x_0} \mathbb{E}\Big[(f\chi_n)(X_t(x))\Phi_n(\Sigma_t(x))\Big] \\
&\quad \leqslant \varlimsup_{x\to x_0} \mathbb{E}\Big[g(X_t(x))\Phi_n(\Sigma_t(x))\Big] + \varlimsup_{x\to x_0} \mathbb{E}\Big[|f\chi_n - g|(X_t(x))\Phi_n(\Sigma_t(x))\Big] \\
&\quad \leqslant \mathbb{E}\Big[g(X_t(x_0))\Phi_n(\Sigma_t(x_0))\Big] + 2\|f\|_\infty \varlimsup_{x\to x_0} \mu_{t,x;n}(F_\varepsilon) \\
&\quad \leqslant \mathbb{E}\Big[(f\chi_n)(X_t(x_0))\Phi_n(\Sigma_t(x_0))\Big] + \mathbb{E}\Big[(g - f\chi_n)(X_t(x_0))\Phi_n(\Sigma_t(x_0))\Big] \\
&\qquad + 2\|f\|_\infty \left(m(F_\varepsilon)R + \frac{C_{d,q,n}}{R^{q-1}}\right) \\
&\quad \leqslant \mathbb{E}f(X_t(x_0)) + 4\|f\|_\infty \left(m(F_\varepsilon)R + \frac{C_{d,q,n}}{R^{q-1}}\right).
\end{aligned}$$

First letting $\varepsilon \to 0$ and then $R \to \infty$, we obtain for $n \in \mathbb{N}$,

$$\varlimsup_{x\to x_0} \mathbb{E}\Big[(f\chi_n)(X_t(x))\Phi_n(\Sigma_t(x))\Big] \leqslant \mathbb{E}f(X_t(x_0)). \tag{4.8}$$

On the other hand, by the definition (3.10) of $\Sigma_t(x)$, it is easy to see that

$$x \mapsto X_t(x),\ \Sigma_t(x) \text{ are continuous in probability.}$$

Thus, by the dominated convergence theorem and (4.8), we have

$$\begin{aligned}
&\varlimsup_{x\to x_0} \mathbb{E}f(X_t(x)) \\
&\leqslant \varlimsup_{x\to x_0} \mathbb{E}\Big[(f\chi_n)(X_t(x))\Phi_n(\Sigma_t(x))\Big] + \|f\|_\infty \varlimsup_{x\to x_0} \mathbb{E}\Big[1 - \chi_n(X_t(x))\Phi_n(\Sigma_t(x))\Big] \\
&= \varlimsup_{x\to x_0} \mathbb{E}\Big[(f\chi_n)(X_t(x))\Phi_n(\Sigma_t(x))\Big] + \|f\|_\infty \mathbb{E}\Big[1 - \chi_n(X_t(x_0))\Phi_n(\Sigma_t(x_0))\Big] \\
&\leqslant \mathbb{E}f(X_t(x_0)) + \|f\|_\infty \mathbb{P}\Big(\{\Sigma_t(x_0) \notin \mathbb{M}_n\} \cup \{|X_t(x_0)| > n\}\Big),
\end{aligned} \tag{4.9}$$

which, by Lemma 4.5 and letting $n \to \infty$, implies

$$\varlimsup_{x\to x_0} \mathbb{E}f(X_t(x)) \leqslant \mathbb{E}f(X_t(x_0)).$$

Applying the above limit to the nonnegative function $\|f\|_\infty - f(x)$, we also have

$$\varlimsup_{x\to x_0} \mathbb{E}(\|f\|_\infty - f(X_t(x))) \leqslant \|f\|_\infty - \mathbb{E}f(X_t(x_0)) \Rightarrow \varliminf_{x\to x_0} \mathbb{E}f(X_t(x)) \geqslant \mathbb{E}f(X_t(x_0)).$$

Thus, we obtain the desired continuity (1.6).

5 Proof of Theorem 1.3

5.1 Norris' type estimate

We first recall the following Norris' type estimate (cf. [23]).

Lemma 5.1. *Let $Y_t = y + \int_0^t \beta_s \mathrm{d}s$ be an $\mathbb{R}^d$-valued process, where β_t takes the following form:*

$$\beta_t = \beta_0 + \int_0^t \gamma_s \mathrm{d}s + \int_0^t Q_s \mathrm{d}W_s + \int_0^t\int_{\Gamma_0} g_s(z)\tilde{N}(\mathrm{d}s, \mathrm{d}z),$$

where $\gamma_t : \mathbb{R}_+ \to \mathbb{R}^d$, $Q_t : \mathbb{R}_+ \to \mathbb{R}^d \times \mathbb{R}^d$ and $g_t(z) : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ are three left continuous $\mathscr{F}_t$-adapted processes. Suppose that for some $R > 0$,

$$|\beta_t|, |Q_t|, |\gamma_t| \leqslant R, \quad |g_t(z)| \leqslant R(1 \wedge |z|). \tag{5.1}$$

Then there exists a constant $C \geqslant 1$ such that for any $t \in (0,1)$, $\delta \in (0, \frac{1}{3})$ and $\varepsilon \in (0, t^3)$,

$$P\left\{\int_0^t |Y_s|^2 \mathrm{d}s < \varepsilon, \int_0^t |\beta_s|^2 \mathrm{d}s \geqslant 9R^2\varepsilon^\delta\right\} \leqslant 4\exp\left\{-\frac{\varepsilon^{\delta-\frac{1}{3}}}{CR^4}\right\}. \tag{5.2}$$

The following lemma is simple.

Lemma 5.2. *Assume that for some $\alpha \in (0,2)$,*

$$\lim_{\varepsilon\to 0} \varepsilon^{\alpha-2} \int_{|z|\leqslant\varepsilon} |z|^2 \nu(\mathrm{d}z) =: c_1 > 0. \tag{5.3}$$

Then for any $p \geqslant 2$, there exist constants $\varepsilon_0, c_2 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$,

$$\int_{|z|\leqslant\varepsilon} |z|^p \nu(\mathrm{d}z) \geqslant c_2 \varepsilon^{p-\alpha}. \tag{5.4}$$

Proof. For any $\delta \in (0,1)$, by (5.3), there is an $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$,

$$(1-\delta)c_1\varepsilon^{2-\alpha} \leqslant \int_{|z|\leqslant\varepsilon} |z|^2\nu(\mathrm{d}z) \leqslant (1+\delta)c_1\varepsilon^{2-\alpha}.$$

Hence,

$$\begin{aligned}
\int_{|z|\leqslant\varepsilon} |z|^p\nu(\mathrm{d}z) &= \sum_{n=0}^\infty \int_{2^{-(n+1)}\varepsilon<|z|\leqslant 2^{-n}\varepsilon} |z|^p\nu(\mathrm{d}z) \\
&\geqslant \sum_{n=0}^\infty (2^{-(n+1)}\varepsilon)^{p-2} \int_{2^{-(n+1)}\varepsilon<|z|\leqslant 2^{-n}\varepsilon} |z|^2\nu(\mathrm{d}z) \\
&\geqslant \sum_{n=0}^\infty (2^{-(n+1)}\varepsilon)^{p-2} \left((1-\delta)c_1(2^{-n}\varepsilon)^{2-\alpha} - (1+\delta)c_1(2^{-(n+1)}\varepsilon)^{2-\alpha}\right) \\
&= \varepsilon^{p-\alpha}c_1 2^{\alpha-p} \sum_{n=0}^\infty 2^{-n(p-\alpha)} \left(2^{2-\alpha}(1-\delta) - (1+\delta)\right),
\end{aligned}$$

which gives (5.4) by letting δ small enough. □

We also need the following estimate.

Lemma 5.3. *Let g_t be a nonnegative, bounded and predictable processes. Under* **($\mathbf{H}_1^\alpha$)***, there exist constants $\lambda_0, c_0 \geqslant 1$ depending on the bound of g_t such that for all $\varepsilon \in (0, \frac{1}{\lambda_0})$,*

$$\mathbb{P}\left\{\int_0^t\!\!\int_{\Gamma_0} g_s\zeta(z)N(\mathrm{d}s,\mathrm{d}z) \leqslant \varepsilon; \int_0^t g_s\mathrm{d}s > \varepsilon^{\frac{\alpha}{6}}\right\} \leqslant \exp\left\{1 - c_0\varepsilon^{-\frac{\alpha}{6}}\right\}, \tag{5.5}$$

where $\zeta(z)$ is defined by (3.7).

Proof. Define

$$\beta_t^\lambda := \int_{\Gamma_0}(1 - \mathrm{e}^{-\lambda g_t\zeta(z)})\nu(\mathrm{d}z)$$

and

$$M_t^\lambda := -\lambda\int_0^t\!\!\int_{\Gamma_0} g_s\zeta(z)N(\mathrm{d}s,\mathrm{d}z) + \int_0^t \beta_s^\lambda\mathrm{d}s.$$

By Itô's formula, we have

$$\mathrm{e}^{M_t^\lambda} = 1 + \int_0^t\!\!\int_{\Gamma_0}\mathrm{e}^{M_{s-}^\lambda}(\mathrm{e}^{-\lambda g_s\zeta(z)} - 1)\tilde{N}(\mathrm{d}s,\mathrm{d}z).$$

Since $1 - \mathrm{e}^{-x} \leqslant 1 \wedge x$ for any $x \geqslant 0$, we have

$$M_t^\lambda \leqslant \int_0^t \beta_s^\lambda\mathrm{d}s \leqslant \int_0^t\!\!\int_{\Gamma_0}(1 \wedge (\lambda g_s\zeta(z)))\nu(\mathrm{d}z)\mathrm{d}s.$$

Hence,

$$\mathbb{E}\mathrm{e}^{M_t^\lambda} = 1. \tag{5.6}$$

Since for any $\kappa \in (0,1)$ and $x \leqslant -\log\kappa$, $1 - \mathrm{e}^{-x} \geqslant \kappa x$, letting $\kappa = \frac{1}{\mathrm{e}}$ and by (5.4), there exist $\lambda_0, c_0 > 1$ such that for all $\lambda \geqslant \lambda_0$,

$$\begin{aligned}\beta_s^\lambda &\geqslant \int_{|z|\leqslant((\|g\|_\infty+1)\lambda)^{-1/3}}(1 - \mathrm{e}^{-\lambda g_s\zeta(z)})\nu(\mathrm{d}z)\\ &\geqslant \frac{\lambda g_s}{\mathrm{e}}\int_{|z|\leqslant((\|g\|_\infty+1)\lambda)^{-1/3}}|z|^3\nu(\mathrm{d}z) \geqslant c_0\lambda^{\frac{\alpha}{3}}g_s.\end{aligned}$$

Thus,

$$\begin{aligned}&\left\{\int_0^t\!\!\int_{\Gamma_0} g_s\zeta(z)N(\mathrm{d}s,\mathrm{d}z) \leqslant \varepsilon; \int_0^t g_s\mathrm{d}s > \varepsilon^{\frac{\alpha}{6}}\right\}\\ &\quad= \left\{\mathrm{e}^{M_t^\lambda} \geqslant \mathrm{e}^{-\lambda\varepsilon + \int_0^t\beta_s^\lambda\mathrm{d}s}; \int_0^t g_s\mathrm{d}s > \varepsilon^{\frac{\alpha}{6}}\right\} \subset \left\{\mathrm{e}^{M_t^\lambda} \geqslant \mathrm{e}^{-\lambda\varepsilon + c_0\lambda^{\frac{\alpha}{3}}\varepsilon^{\frac{\alpha}{6}}}\right\},\end{aligned}$$

which, by Chebyschev's inequality, (5.6) and letting $\lambda = \frac{1}{\varepsilon}$, gives the desired estimate. □

5.2 Proof of Theorem 1.3

Lemma 5.4. *Under (1.7), there exist constants $C_1, C_2 \in (0,1)$ independent of the starting point x and $t_0 \in (0,1)$ such that for all $t \in (0,t_0)$ and $\varepsilon \in (0, C_1t^4)$,*

$$\sup_{|u|=1}\mathbb{P}\left(\int_0^t(|uK_sA_1|^2 + |uK_sA_2|^2)\mathrm{d}s \leqslant \varepsilon\right) \leqslant 8\exp\left\{-C_2\varepsilon^{-\frac{1}{12}}\right\}. \tag{5.7}$$

Proof. Fix $u \in \mathbb{S}^{d-1}$ and set for $i = 1, 2$ and $j = 1, \cdots, d$,

$$
\begin{aligned}
Y_t^i &:= uK_tA_i, \quad \beta_t^i := uK_t\nabla b(X_t)A_i, \quad Q_t^{ij} := \sum_k uK_t(A_1)_{kj}\partial_j\nabla b(X_t)A_i,\\
\gamma_t^i &:= uK_t\Bigg[\Big((b\cdot\nabla)\nabla b - (\nabla b)^2 + \tfrac{1}{2}\nabla^2_{A_1A_1^*}\nabla b\Big)(X_t)\\
&\quad + \int_{\Gamma_0}\Big(\nabla b(X_t + A_2z) - \nabla b(X_t) - 1_{|z|\leqslant 1}(A_2z\cdot\nabla)\nabla b(X_t)\Big)\nu(\mathrm{d}z)\Bigg]A_i,\\
g_t^i(z) &:= uK_t(\nabla b(X_{t-} + A_2z) - \nabla b(X_{t-}))A_i.
\end{aligned}
$$

By equations (3.4) and Itô's formula, one sees that

$$Y_t^i = uA_i + \int_0^t \beta_s^i \mathrm{d}s,$$

and

$$\beta_t^i = u\nabla b(x)A_i + \int_0^t \gamma_s^i\mathrm{d}s + \sum_j\int_0^t Q_s^{ij}\mathrm{d}W_s^j + \int_0^t\!\!\int_{\Gamma_0} g_s^i(y)\tilde{N}(\mathrm{d}s, \mathrm{d}y).$$

By the assumptions, it is easy to see that for some $R > 0$,

$$|g_t^i(z)| \leqslant R(1\wedge|z|), \quad |\beta_t^i| + |\gamma_t^i| + |Q_t^{ij}| \leqslant R.$$

Notice that

$$
\begin{aligned}
&\left\{\int_0^t(|Y_s^1|^2 + |Y_s^2|^2)\mathrm{d}s \leqslant \varepsilon, \int_0^t(|\beta_s^1|^2 + |\beta_s^2|^2)\mathrm{d}s > 18R^2\varepsilon^{\frac{1}{4}}\right\}\\
&\subset \left\{\int_0^t |Y_s^1|^2\mathrm{d}s \leqslant \varepsilon, \int_0^t|\beta_s^1|^2 > 9R^2\varepsilon^{\frac{1}{4}}\right\} \cup \left\{\int_0^t |Y_s^2|^2\mathrm{d}s \leqslant \varepsilon, \int_0^t|\beta_s^2|^2 > 9R^2\varepsilon^{\frac{1}{4}}\right\}.
\end{aligned}
$$

By Lemma 5.1, we have for some $C_2 \in (0, 1)$,

$$\mathbb{P}\left\{\int_0^t(|Y_s^1|^2 + |Y_s^2|^2)\mathrm{d}s \leqslant \varepsilon, \int_0^t(|\beta_s^1|^2 + |\beta_s^2|^2)\mathrm{d}s > 18R^2\varepsilon^{\frac{1}{4}}\right\} \leqslant 8\exp\left\{-C_2\varepsilon^{-\frac{1}{12}}\right\}. \tag{5.8}$$

On the other hand, noticing that

$$|uK_t| \geqslant 1 - \int_0^t|uK_s|\cdot|\nabla b(X_s)|\mathrm{d}s \overset{(3.5)}{\geqslant} 1 - t\|\nabla b\|_\infty \mathrm{e}^{\|\nabla b\|_\infty t} \geqslant \tfrac{1}{2},$$

provided $t < 1\wedge(2\|\nabla b\|_\infty\mathrm{e}^{\|\nabla b\|_\infty})^{-1}$, we have

$$
\begin{aligned}
&\mathbb{P}\left\{\int_0^t(|Y_s^1|^2 + |Y_s^2|^2)\mathrm{d}s \leqslant \varepsilon, \int_0^t(|\beta_s^1|^2 + |\beta_s^2|^2)\mathrm{d}s \leqslant 18R^2\varepsilon^{\frac{1}{4}}\right\}\\
&\quad\leqslant \mathbb{P}\left\{\int_0^t(|Y_s^1|^2 + |Y_s^2|^2 + |\beta_s^1|^2 + |\beta_s^2|^2)\mathrm{d}s \leqslant \varepsilon + 18R^2\varepsilon^{\frac{1}{4}}\right\}\\
&\quad\overset{(1.7)}{\leqslant} \mathbb{P}\left\{c_2\int_0^t|uK_s|^2\mathrm{d}s \leqslant \varepsilon + 18R^2\varepsilon^{\frac{1}{4}}\right\} \leqslant \mathbb{P}\left\{\frac{c_2t}{4} \leqslant (1 + 18R^2)\varepsilon^{\frac{1}{4}}\right\},
\end{aligned} \tag{5.9}
$$

which equals zero provided $\varepsilon < (\frac{c_2t}{4(1+18R^2)})^4$. If we choose

$$C_1 := \left(\frac{c_2}{4(1 + 18R^2)}\right)^4, \quad t_0 := 1\wedge(2\|\nabla b\|_\infty\mathrm{e}^{\|\nabla b\|_\infty})^{-1},$$

then combining this with (5.8) and (5.9), we obtain (5.7). □

Lemma 5.5. *Under* $(\mathbf{H}_1^\alpha)$ *and (1.7), there exist constants* $C_1, C_2 \in (0,1), C_3 > 1$ *independent of the starting point* x *and* $t_0 \in (0,1)$ *such that for all* $t \in (0, t_0)$ *and* $\varepsilon \in (0, C_1 t^{24/\alpha})$,

$$\sup_{|u|=1} \mathbb{P}(u\Sigma_t u^* \leqslant \varepsilon) \leqslant C_3 \exp\left\{-C_2 \varepsilon^{-\frac{\alpha}{72}}\right\}, \tag{5.10}$$

where Σ_t *is defined by (3.10).*

Proof. Noticing that

$$u\Sigma_t u^* := \int_0^t |uK_s A_1|^2 \mathrm{d}s + \int_0^t\!\!\int_{\Gamma_0} |uK_s A_2|^2 \zeta(z) N(\mathrm{d}s, \mathrm{d}z),$$

we have

$$\begin{aligned}
\mathbb{P}(u\Sigma_t u^* \leqslant \varepsilon) &\leqslant \mathbb{P}\left(u\Sigma_t u^* \leqslant \varepsilon; \int_0^t (|uK_sA_1|^2 + |uK_sA_2|^2)\mathrm{d}s > \varepsilon^{\frac{\alpha}{6}}\right)\\
&\quad + \mathbb{P}\left(\int_0^t (|uK_sA_1|^2 + |uK_sA_2|^2)\mathrm{d}s \leqslant \varepsilon^{\frac{\alpha}{6}}\right)\\
&\leqslant \mathbb{P}\left(u\Sigma_t u^* \leqslant \varepsilon; \int_0^t |uK_sA_1|^2\mathrm{d}s > \frac{\varepsilon^{\frac{\alpha}{6}}}{2}\right)\\
&\quad + \mathbb{P}\left(u\Sigma_t u^* \leqslant \varepsilon; \int_0^t |uK_sA_2|^2\mathrm{d}s > \frac{\varepsilon^{\frac{\alpha}{6}}}{2}\right)\\
&\quad + \mathbb{P}\left(\int_0^t (|uK_sA_1|^2 + |uK_sA_2|^2)\mathrm{d}s \leqslant \varepsilon^{\frac{\alpha}{6}}\right),
\end{aligned}$$

which gives the desired estimate by Lemmas 5.3 and 5.4. □

Now we are in a position to give:

Proof of Theorem 1.3. By Lemma 5.5 and a standard compactness argument (cf. [14, p.133 Lemma 2.31] or [22]), for any $p \geqslant 1$, there exist constants $C_p > 0$ and $\gamma(p) > 0$ such that for all $t \in (0,1)$,

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}(\det \Sigma_t(x))^{-p} \leqslant C_p t^{-\gamma(p)}. \tag{5.11}$$

Now, by the chain rule, we have

$$\nabla^k \mathbb{E}\Big((\nabla^n f)(X_t(x))\Big) = \sum_{j=1}^k \mathbb{E}\Big((\nabla^{n+j} f)(X_t(x)) G_j(\nabla X_t(x), \cdots, \nabla^k X_t(x))\Big),$$

where $\{G_j, j = 1, \cdots, k\}$ are real polynomial functions. Using (4.4) and as in (4.5) (cf. [14, p.100, Proposition 2.1.4]), by Lemma 3.3 and Hölder's inequality, there exist $p_1, p_2 > 1$, $C > 0$ independent of x such that for all $t \in (0,1)$,

$$|\nabla^k \mathbb{E}((\nabla^n f)(X_t(x)))| \leqslant C\|f\|_\infty (\mathbb{E}(\det \Sigma_t(x))^{-p_1})^{1/p_2} \leqslant Ct^{-\gamma(p_1)/p_2}.$$

The proof is complete. □

References

[1] Adams R. and Fournier J.: Sobolev spaces. Second Edition, Academic Press, Amsterdam Boston, 2003. MR-2424078

[2] Applebaum D.: Lévy processes and stochastic calculus. Cambridge Studies in Advanced Math. Cambrdige Univ. Press, Cambridge, 2004. MR-2072890

[3] Bichteler K.: Stochastic integration with jumps. Encyclopedia of Mathematics and its Applications (No. 89), Cambridge University Press, 2002. MR-1906715

[4] Bichteler K., Gravereaux J.B. and Jacod J.: Malliavin calculus for processes with jumps. Gordan and Breach Science Publishers, 1987. MR-1008471

[5] Bismut J.M.: Calcul des variations stochastiques et processus de sauts. *Z. Wahrsch. Verw. Gebiete* **63**, (1983), 147–235. MR-0701527

[6] Cass T.: Smooth densities for stochastic differential equations with jumps. *Stoch. Proc. Appl.* **119**, (2009), 1416-1435. MR-2513114

[7] Dong Z., Peng X., Song Y. and Zhang X.: Strong Feller properties for degenerate SDEs with jumps. arXiv:1312.7380

[8] Ishikawa Y. and Kunita H.: Malliavin calculus on the Wiener-Poisson space and its application to canonical SDE with jumps. *Stoch. Proc. Appl.* **116**, (2006), 1743-1769. MR-2307057

[9] Kulik A.: Conditions for existence and smoothness of the distribution density for Ornstein-Uhlenbeck processes with Lévy noises. *Theory Probab. Math. Statist.* **79**, (2009), 23–38. MR-2494533

[10] Kunita H.: Nondegenerate SDEs with jumps and their hypoelliptic properties. *J. Math. Soc. Japan* **65**, (2013), 687–1035. MR-3084987

[11] Kusuoka S.: Malliavin calculus for stochastic differential equations driven by subordinated Brownian motions. *Kyoto J. of Math.* **50**, (2009), no.3, 491–520. MR-2723861

[12] 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 32, (1984), 271–306. MR-0780762

[13] Malliavin P.: Stochastic calculus of variations and hypoelliptic operators. *In: Proc Inter. Symp. on Stoch. Diff. Equations*, Kyoto, (1976), 195–263. MR-0536013

[14] Nualart D.: The Malliavin calculus and related topics. Springer-Verlag, New York, 2006. MR-2200233

[15] Picard J.: On the existence of smooth densities for jump processes. *Prob. Theory Rela. Fields* **105**, (1996), 481–511. MR-1402654

[16] Priola E. and Zabczyk J.: Densities for Ornstein-Uhlenbeck processes with jumps. *Bull. London Math. Soc.* **41**, (2009), 41–50. MR-2481987

[17] Protter, P. E. Stochastic integration and differential equations. Second Edition, Springer-Verlag, Berlin, (2004). MR-2020294

[18] Ren, J., Röckner, M., Zhang, X.: Kusuoka-Stroock formula on configuration space and regularities of local times with jumps. *Potential Analysis* **26**, (2007), no.4, 363–396. MR-2300338

[19] Sato K.: Lévy processes and infinite divisible distributions. Cambridge Univ. Press, Cambridge, 1999. MR-1739520

[20] Takeuchi A.: The Malliavin calculus for SDE with jumps and the partially hypoelliptic problem. *Osaka J. Math.* **39**, (2002), 523–559. MR-1932281

[21] Xu L.: Smooth densities of stochastic differential equations forced by degenerate stable like noise. arXiv:1308.1124

[22] Zhang X.: Densities for SDEs driven by degenerate α-stable processes. *Ann. Probab.* **42**, (2014), 1885–1910. MR-3262494

[23] Zhang X.: Fundamental solution of kinetic Fokker-Planck operator with anisotropic nonlocal dissipativity. SIAM J. Math. Anal. **46**, (2014), 2254ĺC2280. MR-3225504

[24] Zhang X.: Nonlocal Hörmander's hypoellipticity theorem. arXiv:1306.5016

Acknowledgments. The authors would like to thank Professors Zhao Dong, Feng-Yu Wang and Xuhui Peng, Lihu Xu for their interests and stimulating discussions. Special thanks go to Lihu Xu for sending us his preprint paper [21].