The Annals of Probability
2014, Vol. 42, No. 5, 1885–1910
DOI: 10.1214/13-AOP900

DENSITIES FOR SDES DRIVEN BY DEGENERATE α-STABLE PROCESSES

BY XICHENG ZHANG[1]

Wuhan University

In this work, by using the Malliavin calculus, under Hörmander's condition, we prove the existence of distributional densities for the solutions of stochastic differential equations driven by degenerate subordinated Brownian motions. Moreover, in a special degenerate case, we also obtain the smoothness of the density. In particular, we obtain the existence of smooth heat kernels for the following fractional kinetic Fokker–Planck (nonlocal) operator:

$$\mathscr{L}_b^{(\alpha)} := \Delta_{\mathrm{v}}^{\alpha/2} + \mathrm{v} \cdot \nabla_x + b(x, \mathrm{v}) \cdot \nabla_{\mathrm{v}}, \qquad x, \mathrm{v} \in \mathbb{R}^d,$$

where $\alpha \in (0, 2)$ and $b : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ is smooth and has bounded derivatives of all orders.

1. Introduction and main results. Consider the following stochastic differential equation (abbreviated as SDE) in $\mathbb{R}^d$:

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t + \int_{\mathbb{R}^d - \{0\}} g(X_t, z)\tilde{N}(\mathrm{d}t, \mathrm{d}z), \tag{1.1}$$
$$X_0 = x,$$

where $b : \mathbb{R}^d \to \mathbb{R}^d, \sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ and $g : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ are smooth functions, $(W_t)_{t \geq 0}$ is a standard d-dimensional Brownian motion, and $\tilde{N}(\mathrm{d}t, \mathrm{d}z)$ is an independent compensated Poisson random measure on $\mathbb{R}^d - \{0\}$ with intensity measure $\mathrm{d}t\nu(\mathrm{d}z)$. Below, we always assume that b, σ and g have bounded derivatives of all orders. Let us define the vector fields

$$V_0 := (b_j - \tfrac{1}{2}\partial_l \sigma_{jk} \sigma_{lk})\partial_j \quad \text{and} \quad V_i := \sigma_{ij}\partial_j, \qquad i = 1, \dots, d,$$

where we have used the convention: a repeated index in a product will be summed automatically. Set $\mathscr{V}_0 := \{V_1, \dots, V_d\}$ and define recursively

$$\mathscr{V}_k := \{[V_0, V], [V_1, V], \dots, [V_d, V], V \in \mathscr{V}_{k-1}\}, \qquad k \in \mathbb{N},$$

where $[V_i, V] := V_i V - V V_i$ denotes the Lie bracket. It is well known that when $g \equiv 0$ (i.e., no jump part) and if $\bigcup_{k \in \mathbb{N}} \mathscr{V}_k$ spans $\mathbb{R}^d$ at all points x (called Hörmander's condition), then the solution $X_t(x)$ of SDE (1.1) admits a smooth density

---

Received July 2012; revised November 2013.

[1] Supported by NNSFs of China (Nos. 11271294, 11325105) and Program for New Century Excellent Talents in University (NCET-10-0654).

Primary 60H07, 60H10; secondary 35Q84.

Key words and phrases. Malliavin calculus, Hörmander's condition, α-stable process, distributional density, SDE.

$p_t(x, y)$, which was originally initiated by Malliavin [14] (see [15] for a systematic introduction). Moreover, by Itô's formula, $p_t(x, y)$ satisfies the following Fokker–Planck equation:

$$\partial_t p_t(x, y) = \tfrac{1}{2}\sigma_{ik}(x)\sigma_{jk}(x)\,\partial_{y_i}\,\partial_{y_j} p_t(x, y) + \partial_{y_i}\big(b_i(y)p_t(x, y)\big)$$

with $p_0(x, y) = \delta_x(y)$.

Malliavin's probabilistic proof about Hörmander's theorem is based on the stochastic calculus of variations on the Wiener space invented by himself [14]. Since then, there are many works devoted to extending the Malliavin calculus to the Poisson space case (see, e.g., [2, 4, 5, 13, 16], etc.). In these works, the existence and smoothness of the distributional densities for SDEs with jumps were obtained, where various nondegeneracy conditions about $g(x, z)\nu(\mathrm{d}z)$ are imposed. We particularly mention that Kusuoka in [12] developed the Malliavin calculus for subordinated Brownian motions, and obtained the existence of smooth densities for SDEs driven by nondegenerate subordinated Brownian motions. His argument will be discussed later.

On the other hand, assuming that $\nu(\mathrm{d}z) = \mathrm{d}z/|z|^{d+\alpha}$, where $\alpha \in (0, 2)$, and $g(x, z)$ satisfies some boundedness and smoothness conditions, Takeuchi in [20], Corollary 1, proved that the solution $X_t(x)$ of SDE (1.1) has a smooth density with respect to the Lebesgue measure under some uniform Hörmander's conditions. Notice that Takeuchi's conditions allow pure-jump degenerate noises. In [7], Cass obtained a similar result. It is remarkable that recently, Kunita in [11] proved the analytic property of distributional density to SDE (1.1) under weaker Hörmander's conditions. His proofs are based on the Malliavin calculus on the Wiener–Poisson spaces developed in [8] and [10]. Moreover, an estimate for discontinuous semimartingales due to Komatsu and Takeuchi [9] plays a crucial role in Takeuchi and Kunita's proofs. It is emphasized that all these results assume that g is *bounded* or the Lévy measure ν has *finite* moments of all orders. Thus, the interesting α-stable noise is ruled out.

In this work, we consider the following simple SDE:

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + A\,\mathrm{d}L_t, \qquad X_0 = x \in \mathbb{R}^d, \tag{1.2}$$

where $A = (a_{ij})$ is a $d \times d$-matrix, and $(L_t)_{t \geq 0}$ is a rotationally invariant d-dimensional α-stable process, that is, its characteristic function is given by

$$\mathbb{E}\mathrm{e}^{\mathrm{i}z \cdot L_t} = \mathrm{e}^{-t|z|^\alpha}, \qquad \alpha \in (0, 2). \tag{1.3}$$

We are interested in the problem that under what degenerate conditions on A together with b, $X_t(x)$ admits a smooth density with respect to the Lebesgue measure. Let us first look at the linear case of Ornstein–Uhlenbeck processes, that is,

$$\mathrm{d}X_t = BX_t\,\mathrm{d}t + A\,\mathrm{d}L_t, \qquad X_0 = x, \tag{1.4}$$

where B is a $d \times d$-matrix. The generator of this SDE is given by $\mathcal{L}_A^{(\alpha)} + Bx \cdot \nabla$, where the nonlocal operator $\mathcal{L}_A^{(\alpha)}$ is defined by

$$\mathcal{L}_A^{(\alpha)} f(x) := \text{P.V.} \int_{\mathbb{R}^d} [f(x + Ay) - f(x)] \frac{\mathrm{d}y}{|y|^{d+\alpha}}, \tag{1.5}$$

where P.V. stands for the Cauchy principal value. Recently, Priola and Zabczyk [17] proved that X_t has a smooth density under the following Kalman's condition (see also [6] for further discussions on this condition):

$$\text{Rank}[A, BA, \ldots, B^{d-1}A] = d. \tag{1.6}$$

In fact, the solution of (1.4) is explicitly given by

$$X_t = \mathrm{e}^{tB} x + \int_0^t \mathrm{e}^{(t-s)B} A \, \mathrm{d}L_s =: \mathrm{e}^{tB} x + Z_t.$$

Using the approximation of step functions, by (1.3) it is easy to see that

$$\mathbb{E}\mathrm{e}^{\mathrm{i}z \cdot Z_t} = \mathbb{E} \exp\left\{\mathrm{i}z \cdot \int_0^t \mathrm{e}^{(t-s)B} A \, \mathrm{d}L_s\right\} = \exp\left\{-\int_0^t |z^* \mathrm{e}^{(t-s)B} A|^\alpha \, \mathrm{d}s\right\},$$

where $*$ stands for the transpose of a column vector. Hence, for any $m \in \mathbb{N}$,

$$\begin{aligned} \int_{\mathbb{R}^d} |z|^m \mathbb{E}\mathrm{e}^{\mathrm{i}z \cdot Z_t} \, \mathrm{d}z &= \int_{\mathbb{R}^d} |z|^m \exp\left\{-\int_0^t |z^* \mathrm{e}^{(t-s)B} A|^\alpha \, \mathrm{d}s\right\} \mathrm{d}z \\ &\leq \int_{\mathbb{R}^d} |z|^m \exp\left\{-|z|^\alpha \inf_{|a|=1} \int_0^t |a\mathrm{e}^{sB} A|^\alpha \, \mathrm{d}s\right\} \mathrm{d}z. \end{aligned}$$

Here and below, "a" denotes a row vector in $\mathbb{R}^d$. By (1.6), one has

$$\inf_{|a|=1} \int_0^t |a\mathrm{e}^{sB} A|^\alpha \, \mathrm{d}s > 0$$

and so,

$$\int_{\mathbb{R}^d} |z|^m \mathbb{E}\mathrm{e}^{\mathrm{i}z \cdot Z_t} \, \mathrm{d}z < +\infty \qquad \forall m \in \mathbb{N}.$$

Thus, Z_t admits a smooth density by [19], Proposition 28.1, and so does X_t.

We now turn to the nonlinear case. Before stating our main results, we first recall some notions about the subordinated Brownian motions. Let $(S_t)_{t \geq 0}$ be a subordinator (an increasing one-dimensional Lévy process) on $\mathbb{R}_+$ with Laplace transform:

$$\mathbb{E}\mathrm{e}^{-sS_t} = \exp\left\{t \int_0^\infty (\mathrm{e}^{-su} - 1) \nu_S(\mathrm{d}u)\right\},$$

where ν_S (called the Lévy measure of S_t) satisfies $\nu_S(\{0\}) = 0$ and

$$\int_0^\infty (1 \wedge u) \nu_S(\mathrm{d}u) < +\infty.$$

Below, we assume that $(S_t)_{t\geq 0}$ is independent of $(W_t)_{t\geq 0}$ and

$$P\{\omega:\exists t>0 \text{ such that } S_t(\omega)=0\}=0, \tag{1.7}$$

which means that for almost all ω, $t\mapsto S_t(\omega)$ is *strictly* increasing (see Lemma 2.1 below). Notice that the Poisson process does not satisfy such an assumption, but the α-stable subordinator satisfies this assumption (see [3], p. 88, Theorem 11). Essentially, condition (1.7) is a nondegenerate assumption, and says that the subordinator has infinitely many jumps on any interval. In particular, the process defined by

$$L_t := W_{S_t}, \qquad t\geq 0, \tag{1.8}$$

is a Lévy process (called subordinated Brownian motion) with characteristic function:

$$\mathbb{E}e^{iz\cdot L_t}=\exp\left\{t\int_{\mathbb{R}^d}(e^{iz\cdot y}-1-iz\cdot y1_{|y|\leq 1})\nu_L(dy)\right\},$$

where ν_L is the Lévy measure given by

$$\nu_L(\Gamma)=\int_0^\infty(2\pi s)^{-d/2}\left(\int_\Gamma e^{-|y|^2/2s}\,dy\right)\nu_S(ds). \tag{1.9}$$

Obviously, ν_L is a symmetric measure.

The first aim of this paper is to prove the following existence result of distributional density to SDE (1.2) under Hörmander's condition as in [20] and [11].

THEOREM 1.1. *Let $b:\mathbb{R}^d\to\mathbb{R}^d$ be a C^∞-function with bounded partial derivatives of first order. For $x\in\mathbb{R}^d$, let $X_t(x)$ solve SDE* (1.2) *with subordinated Brownian motion L_t. Assume that for some $n=n(x)\in\mathbb{N}$,*

$$\mathrm{Rank}[A, B_1(x)A, B_2(x)A,\dots,B_n(x)A]=d, \tag{$\mathscr{H}_n$}$$

where $B_1(x):=(\nabla b)_{ij}(x)=(\partial_j b^i(x))_{ij}$, and for $n\geq 2$,

$$B_n(x):=(b^i\,\partial_i B_{n-1})(x)-(\nabla b\cdot B_{n-1})(x). \tag{1.10}$$

Then the law of $X_t(x)$ is absolutely continuous with respect to the Lebesgue measure. In particular, the density $p_t(x,y)$ solves the following nonlocal Fokker–Plack equation in the weak or distributional sense:

$$\partial_t p_t(x,y)=\mathcal{L}_A p_t(x,\cdot)(y)+\partial_{y_i}(b_i(y)p_t(x,y)) \tag{1.11}$$

with $p_0(x,y)=\delta_x(y)$, where

$$\mathcal{L}_A f(y):=\mathrm{P.V.}\int_{\mathbb{R}^d}[f(y+Az)-f(y)]\nu_L(dz).$$

REMARK 1.2. If we assume that L_t has finite moments of all orders, then this result is contained in [11], Theorem 5.1. In fact, Kunita also obtained the smoothness of the density. Nevertheless, our proof is simpler in this case. Notice that if $b(x) = Bx$, then condition $(\mathscr{H}_n)$ reduces to (1.6).

For the smoothness of $p_t(x, y)$, we shall assume the following uniform Hörmander's condition:

$$(U\mathscr{H}_1) \qquad \inf_{x\in\mathbb{R}^d} \inf_{|a|=1} \big(|aA|^2 + |a\nabla b(x)A|^2\big) =: c_1 > 0$$

and prove the following partial result.

THEOREM 1.3. *Let $b:\mathbb{R}^d \to \mathbb{R}^d$ be a C^∞-function with bounded partial derivatives of all orders. In addition to $(U\mathscr{H}_1)$, we assume that the Lévy measure ν_S satisfies for some $\theta \in (0, \frac{1}{2})$,*

$$\lim_{\varepsilon\downarrow 0} \frac{1}{\varepsilon^{1-2\theta}} \int_0^\varepsilon u\nu_S(\mathrm{d}u) =: c_\theta > 0. \tag{1.12}$$

Then the density $p_t(x, y)$ is a smooth function on $(0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d$, and for each $t > 0$,

$$(x, y) \mapsto p_t(x, y) \in C_b^\infty\big(\mathbb{R}^d \times \mathbb{R}^d\big).$$

In particular, for all $(t, x, y) \in (0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d$,

$$\partial_t p_t(x, y) = \mathcal{L}_A p_t(\cdot, y)(x) + b(x) \cdot \nabla_x p_t(x, y). \tag{1.13}$$

REMARK 1.4. Condition $(U\mathscr{H}_1)$, compared with $(\mathscr{H}_n)$, is much stronger, and will be used to prove the L^p-integrability of the inverse of the Mallavin covariance matrix defined by (2.6) and (3.6) below, where the key point is to prove a Norris' type lemma (see Lemma 3.4 below). We conjecture that a similar $(U\mathscr{H}_n)$ as in [11] should imply the smoothness of $p_t(x, y)$. Nevertheless, the following stochastic Hamilton system driven by a subordinated Brownian motion satisfies $(U\mathscr{H}_1)$:

$$\begin{cases} \mathrm{d}X_t = \nabla_y H(X_t, Y_t)\,\mathrm{d}t, & X_0 = x \in \mathbb{R}^d, \\ \mathrm{d}Y_t = -\nabla_x H(X_t, Y_t)\,\mathrm{d}t + A\,\mathrm{d}L_t, & Y_0 = y \in \mathbb{R}^d, \end{cases} \tag{1.14}$$

where A is a $d \times d$-invertible matrix, and $H:\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is a C^2-Hamiltonian function so that $y \mapsto H(x, y)$ is strictly convex or concave.

REMARK 1.5. Let $\nu_S(\mathrm{d}u) = u^{-(1+\alpha)}\,\mathrm{d}u$ be the Lévy measure of an α-stable subordinator. It is easy to see that (1.12) holds for $\theta = \alpha/2$.

The argument for proving Theorems 1.1 and 1.3 is different from Takeuchi and Kunita's works. We shall follow Kusuoka's method [12]. The advantage of which is that it is not necessary to develop a *new* Malliavin calculus for jump processes, and moreover, one can obtain some quantitive estimates about the semigroup (see Theorem 3.8 below); while the drawback of which is of course the loss of generality. It is noticed that in [12], Kusuoka considered the SDE driven by multiplicative noises. However, it seems that there is a gap in the calculations about the Malliavin covariance matrix (see [12], Theorem 3.3) since the solution of SDE (1.1) usually does not form a stochastic diffeomorphism flow if there is no further restriction on the jump size (cf. [18], p. 328). This is also why we have to confine ourself to the additive noise.

Let us now describe the argument (see also [21]). Let $(\mathbb{W}, \mathbb{H}, \mu_{\mathbb{W}})$ be the classical Wiener space, that is, $\mathbb{W}$ is the space of all continuous functions from $\mathbb{R}^+$ to $\mathbb{R}^d$ with vanishing values at starting point 0, $\mathbb{H} \subset \mathbb{W}$ is the Cameron–Martin space consisting of all absolutely continuous functions with square integrable derivatives, and $\mu_{\mathbb{W}}$ is the Wiener measure so that the coordinate process

$$W_t(w) := w_t$$

is a standard d-dimensional Brownian motion.

Let $\mathbb{S}$ be the space of all increasing, purely discontinuous and càdlàg functions from $\mathbb{R}_+$ to $\mathbb{R}_+$ with $\ell_0 = 0$, which is endowed with the Skorohod metric and the probability measure $\mu_{\mathbb{S}}$ so that the coordinate process

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

has the same law as the given subordinator. Consider the following product probability space:

$$(\Omega, \mathscr{F}, P) := \big(\mathbb{W} \times \mathbb{S}, \mathscr{B}(\mathbb{W}) \times \mathscr{B}(\mathbb{S}), \mu_{\mathbb{W}} \times \mu_{\mathbb{S}}\big)$$

and define

$$L_t(w, \ell) := w_{\ell_t}.$$

Then $(L_t)_{t\geq 0}$ has the same law as the given subordinated Brownian motion. In particular, the solution $X_t(x)$ of SDE (1.2) can be regarded as a functional of w and ℓ and

$$\mathbb{E} f\big(X_t(x)\big) = \int_{\mathbb{S}} \int_{\mathbb{W}} f\big(X_t(x, w_\ell)\big) \mu_{\mathbb{W}}(\mathrm{d}w) \mu_{\mathbb{S}}(\mathrm{d}\ell). \tag{1.15}$$

The advantage of this viewpoint is that we can use the classical Malliavin calculus to study the Brownian functional $w \to X_t(x, w_\ell)$ (see [12]). Thus, in order to prove Theorem 1.1, it is enough to prove that for each $\ell \in \mathbb{S}$, the law of $w \mapsto X_t(x, w_\ell)$ under $\mu_{\mathbb{W}}$ is absolutely continuous with respect to the Lebesgue measure. In order to prove Theorem 1.3, the key point is to prove the L^p-integrability of the inverse of the Malliavin covariance matrix so that we can use the integration by parts formula to derive some gradient estimates (see Theorem 3.8 below), which then implies the smoothness of the density by Sobolev's embedding theorem.

This paper is organized as follows: in Section 2, we prove Theorem 1.1 by using the Malliavin calculus, where the main point is to prove the invertibility of the Malliavin covariance matrix $(\Sigma_t^{\ell})_{\ell=S}$ in (2.7) below. In Section 3, we prove Theorem 1.3 by establishing a Norris' type lemma as in [7]. In order to overcome the nonintegrability of α-stable processes, we shall separately consider the small jumps and the large jumps of the subordinator. In particular, the asymptotic estimate of small times about the semigroup plays a crucial role.

2. Proof of Theorem 1.1.

We need the following simple lemma about the density of the jump number of the subordinator.

LEMMA 2.1. *For $s > 0$, set $\Delta\ell_s := \ell_s - \ell_{s-}$ and*

$$\mathbb{S}_0 := \{\ell \in \mathbb{S} : \{s : \Delta\ell_s > 0\} \text{ is dense in } [0, \infty)\}.$$

Under (1.7), *we have* $\mu_{\mathbb{S}}(\mathbb{S}_0) = 1$.

PROOF. Let $\mathscr{I}$ be the total of all rational intervals in $[0, \infty)$, that is,

$$\mathscr{I} := \{I = (a, b) : 0 \le a < b \text{ are rational numbers}\}.$$

For $I \in \mathscr{I}$, let us write

$$\mathbb{S}_I := \{\ell \in \mathbb{S} : I \subset \{s : \Delta\ell_s = 0\}\}.$$

It is easy to see that

$$\mathbb{S} - \mathbb{S}_0 = \bigcup_{I \in \mathscr{I}} \mathbb{S}_I.$$

Thus, for proving $\mu_{\mathbb{S}}(\mathbb{S}_0) = 1$, it is enough to prove that for each $I = (a, b) \in \mathscr{I}$,

$$\mu_{\mathbb{S}}(\mathbb{S}_I) = \mu_{\mathbb{S}}(\{\ell \in \mathbb{S} : (a, b) \subset \{s : \Delta\ell_s = 0\}\}) = 0,$$

which, by the stationarity of the subordinator, is equivalent to

$$\mu_{\mathbb{S}}(\{\ell \in \mathbb{S} : (0, b - a) \subset \{s : \Delta\ell_s = 0\}\}) = 0. \tag{2.1}$$

Since

$$\{\ell \in \mathbb{S} : (0, b - a) \subset \{s : \Delta\ell_s = 0\}\} = \{\ell \in \mathbb{S} : \ell_s = 0,\ \forall s \in (0, b - a)\}$$

by (1.7), we obtain (2.1), and complete the proof. □

For a functional F on $\mathbb{W}$, the Malliavin derivative of F along the direction $h \in \mathbb{H}$ is defined as

$$D_h F(w) := \lim_{\varepsilon \to 0} \frac{F(w + \varepsilon h) - F(w)}{\varepsilon} \qquad \text{in } L^2(\mathbb{W}, \mu_{\mathbb{W}}). \tag{2.2}$$

If $h \mapsto D_h F$ is bounded, then there exists a unique $DF \in L^2(\mathbb{W}, \mu_{\mathbb{W}}; \mathbb{H})$ such that

$$\langle DF, h\rangle_{\mathbb{H}} = D_h F \qquad \forall h \in \mathbb{H}.$$

In this case, we shall write $F \in \mathscr{D}(D)$ and call DF the Malliavin gradient of F (cf. [15]).

For $\ell \in \mathbb{S}_0$ and $x \in \mathbb{R}^d$, let $X_t^\ell(x) = X_t^\ell$ solve the following SDE:

$$X_t^\ell = x + \int_0^t b(X_s^\ell)\,\mathrm{d}s + AW_{\ell_t}. \tag{2.3}$$

Let $J_t^\ell := J_t^\ell(x) := \nabla X_t^\ell(x)$ be the derivative matrix of $X_t^\ell(x)$ with respect to the initial value x. It is easy to see that

$$J_t^\ell = I + \int_0^t \nabla b(X_s^\ell) \cdot J_s^\ell\,\mathrm{d}s. \tag{2.4}$$

Let K_t^ℓ be the inverse matrix of J_t^ℓ. Then K_t^ℓ satisfies

$$K_t^\ell = I - \int_0^t K_s^\ell \cdot \nabla b(X_s^\ell)\,\mathrm{d}s. \tag{2.5}$$

Moreover, by definition (2.2) and equation (2.3), it is easy to see that $X_t^\ell(x) \in \mathscr{D}(D)$ and for any $h \in \mathbb{H}$,

$$D_h X_t^\ell = \int_0^t \nabla b(X_s^\ell) D_h X_s^\ell\,\mathrm{d}s + Ah_{\ell_t}.$$

The Malliavin covariance matrix is defined by

$$(\Sigma_t^\ell)_{ij} := \langle D(X_t^\ell)^i, D(X_t^\ell)^j\rangle_{\mathbb{H}}. \tag{2.6}$$

The following lemma provides an explicit expression of Σ_t^ℓ in terms of J_t^ℓ (cf. [12]), which is crucial in the Malliavin's proof of Hörmander's hypoellipticity theorem.

LEMMA 2.2. *We have*

$$\Sigma_t^\ell = J_t^\ell \left(\int_0^t K_s^\ell AA^* (K_s^\ell)^*\,\mathrm{d}\ell_s \right) (J_t^\ell)^*, \tag{2.7}$$

where $*$ *denotes the transpose of a matrix.*

PROOF. For $\varepsilon \in (0,1)$, we define

$$\ell_t^\varepsilon := \frac{1}{\varepsilon} \int_t^{t+\varepsilon} \ell_s\,\mathrm{d}s = \int_0^1 \ell_{\varepsilon s + t}\,\mathrm{d}s. \tag{2.8}$$

Since $t \mapsto \ell_t$ is strictly increasing and right continuous, it follows that for each $t \geq 0$,

$$\ell_t^\varepsilon \downarrow \ell_t \qquad \text{as } \varepsilon \downarrow 0. \tag{2.9}$$

Moreover, $t \mapsto \ell_t^\varepsilon$ is absolutely continuous and strictly increasing. Let γ^ε be the inverse function of ℓ^ε, that is,

$$\ell^\varepsilon_{\gamma_t^\varepsilon} = t, \qquad t \geq \ell_0^\varepsilon \quad \text{and} \quad \gamma^\varepsilon_{\ell_t^\varepsilon} = t, \qquad t \geq 0.$$

By definition, γ_t^ε is also absolutely continuous on $[\ell_0^\varepsilon, \infty)$. Let $X_t^{\ell^\varepsilon}$ solve the following SDE:

$$X_t^{\ell^\varepsilon} = x + \int_0^t b(X_s^{\ell^\varepsilon})\,\mathrm{d}s + A(W_{\ell_t^\varepsilon} - W_{\ell_0^\varepsilon}).$$

Let us now define

$$Y_t^{\ell^\varepsilon}(x) := X^{\ell^\varepsilon}_{\gamma_t^\varepsilon}(x), \qquad t \geq \ell_0^\varepsilon.$$

By the change of variables, one sees that

$$Y_t^{\ell^\varepsilon} = x + \int_{\ell_0^\varepsilon}^t b(Y_s^{\ell^\varepsilon})\dot{\gamma}_s^\varepsilon\,\mathrm{d}s + A(W_t - W_{\ell_0^\varepsilon}).$$

It is well known that [cf. [15], p. 127, (2.60)]

$$\langle DY_t^{\ell^\varepsilon}, DY_t^{\ell^\varepsilon}\rangle_{\mathbb{H}} = \nabla Y_t^{\ell^\varepsilon}\left(\int_{\ell_0^\varepsilon}^t (\nabla Y_s^{\ell^\varepsilon})^{-1}AA^*((\nabla Y_s^{\ell^\varepsilon})^{-1})^*\,\mathrm{d}s\right)(\nabla Y_t^{\ell^\varepsilon})^*.$$

By the change of variables again, we obtain

$$\begin{aligned}\langle DX_t^{\ell^\varepsilon}, DX_t^{\ell^\varepsilon}\rangle_{\mathbb{H}} &= \nabla X_t^{\ell^\varepsilon}\left(\int_{\ell_0^\varepsilon}^{\ell_t^\varepsilon} (\nabla Y_s^{\ell^\varepsilon})^{-1}AA^*((\nabla Y_s^{\ell^\varepsilon})^{-1})^*\,\mathrm{d}s\right)(\nabla X_t^{\ell^\varepsilon})^* \\ &= \nabla X_t^{\ell^\varepsilon}\left(\int_0^t (\nabla X_s^{\ell^\varepsilon})^{-1}AA^*((\nabla X_s^{\ell^\varepsilon})^{-1})^*\,\mathrm{d}\ell_s^\varepsilon\right)(\nabla X_t^{\ell^\varepsilon})^* \qquad (2.10) \\ &= J_t^{\ell^\varepsilon}\left(\int_0^t K_s^{\ell^\varepsilon}AA^*(K_s^{\ell^\varepsilon})^*\,\mathrm{d}\ell_s^\varepsilon\right)(J_t^{\ell^\varepsilon})^*.\end{aligned}$$

From equation (2.3), it is easy to see that for each $t \geq 0$ and $w \in \mathbb{W}$,

$$\lim_{\varepsilon\downarrow 0}|X_t^{\ell^\varepsilon}(w) - X_t^\ell(w)| \leq C\lim_{\varepsilon\downarrow 0}|W_{\ell_t^\varepsilon}(w) - W_{\ell_t}(w)| = 0.$$

Thus, by equations (2.4) and (2.5), we also have

$$\lim_{\varepsilon\downarrow 0}\sup_{s\in[0,t]}|J_s^{\ell^\varepsilon}(w) - J_s^\ell(w)| = 0$$

and

$$\lim_{\varepsilon\downarrow 0}\sup_{s\in[0,t]}|K_s^{\ell^\varepsilon}(w) - K_s^\ell(w)| = 0.$$

Taking limits for both sides of (2.10), we obtain (2.7) (see [21]). □

The following lemma is a direct application of Itô's formula (cf. [18], p. 81, Theorem 33).

LEMMA 2.3. *Let $V:\mathbb{R}^d\to\mathbb{M}^d$ be a $d\times d$-matrix valued smooth function. We have*

$$\begin{aligned}K^\ell_t V(X^\ell_t) = V(x) &+ \int_0^t K^\ell_s(b\cdot\nabla V - \nabla b\cdot V)(X^\ell_s)\,\mathrm{d}s\\ &+ \sum_{0<s\le t} K^\ell_s\big(V(X^\ell_s) - V(X^\ell_{s-}) - \nabla V(X^\ell_{s-})\cdot\Delta X^\ell_s\big)\\ &+ \int_0^t K^\ell_s\cdot(\nabla V)(X^\ell_{s-})\cdot A\,\mathrm{d}W_{\ell_s},\end{aligned}$$

where $\Delta X^\ell_s := X^\ell_s - X^\ell_{s-} = A(W_{\ell_s} - W_{\ell_{s-}})$.

We are now in a position to give the following.

PROOF OF THEOREM 1.1. By Lemma 2.1 and (1.15), it is enough to prove that for each $\ell\in\mathbb{S}_0$, the law of X^ℓ_t under $\mu_{\mathbb{W}}$ is absolutely continuous with respect to the Lebesgue measure. By [15], page 97, Theorem 2.1.2, it suffices to prove that Σ^ℓ_t is invertible. Since J^ℓ_t is invertible, by (2.7) we only need to show that for any row vector $a\neq 0\in\mathbb{R}^d$,

$$\int_0^t |aK^\ell_s A|^2\,\mathrm{d}\ell_s > 0. \tag{2.11}$$

Suppose that

$$\int_0^t |aK^\ell_s A|^2\,\mathrm{d}\ell_s = \sum_{s\in(0,t]} |aK^\ell_s A|^2\Delta\ell_s = 0,$$

then by Lemma 2.1 and the continuity of $s\mapsto|aK^\ell_s A|$, we have

$$aK^\ell_s A = 0 \qquad \forall s\in[0,t].$$

Thus, by (2.5) we get

$$0 = aK^\ell_{t'}A = aA - \int_0^{t'} aK^\ell_s(\nabla b)(X^\ell_s)A\,\mathrm{d}s \qquad \forall t'\in[0,t],$$

which in turn implies that

$$aA = 0 \tag{2.12}$$

and by the right continuity of $s\mapsto X^\ell_s$,

$$aK^\ell_s B_1(X^\ell_s)A = aK^\ell_s(\nabla b)(X^\ell_s)A = 0 \qquad \forall s\in[0,t]. \tag{2.13}$$

Now we use the induction to prove that for each $n\in\mathbb{N}$,

$$aK^\ell_s B_n(X^\ell_s)A = 0 \qquad \forall s\in[0,t]. \tag{2.14}$$

Suppose that (2.14) is true for some n. By Lemma 2.3, we have

$$K_t^\ell B_n(X_t^\ell) = B_n(x) + \int_0^t K_s^\ell B_{n+1}(X_s^\ell)\,\mathrm{d}s + M_t + V_t,$$

where

$$M_t := \int_0^t K_s^\ell \cdot (\nabla B_n)(X_{s-}^\ell) \cdot A\,\mathrm{d}W_{\ell_s}$$

and

$$V_t := \sum_{0<s\le t} K_s^\ell\big(B_n(X_s^\ell) - B_n(X_{s-}^\ell) - (\nabla B_n)(X_{s-}^\ell)\cdot \Delta(AW_{\ell_s})\big).$$

Thus, by (2.14) we have

$$\int_0^{t'} aK_s^\ell B_{n+1}(X_s^\ell)A\,\mathrm{d}s + aM_{t'}A + aV_{t'}A = 0 \qquad \forall t' \in [0,t]. \tag{2.15}$$

By the inductive assumption (2.14), we have

$$aK_s^\ell B_n(X_s^\ell)A = aK_s^\ell B_n(X_{s-}^\ell)A = 0.$$

Hence,

$$\begin{aligned} aV_{t'}A &= -\sum_{0<s\le t'} aK_s^\ell \cdot (\nabla B_n)(X_{s-}^\ell)\cdot \Delta(AW_{\ell_s})\cdot A \\ &= -\int_0^{t'} aK_s^\ell \cdot (\nabla B_n)(X_{s-}^\ell)\cdot A\,\mathrm{d}W_{\ell_s}\cdot A = -aM_{t'}A, \end{aligned}$$

which together with (2.15) implies that

$$aK_s^\ell B_{n+1}(X_s^\ell)A = 0 \qquad \forall s \in [0,t].$$

The assertion (2.14) is thus proved. Combining (2.12) and (2.14) and by letting $s \to 0$, we obtain

$$aA = aB_1(x)A = \cdots = aB_n(x)A = 0,$$

which is contrary to ($\mathscr{H}_n$). The proof is thus complete. □

3. Proof of Theorem 1.3.

3.1. *Norris' type lemma.* In this section, we use the following filtration:

$$\mathscr{F}_t := \sigma\{W_{S_s}, S_s : s \le t\}.$$

Clearly, for $t > s$, $W_{S_t} - W_{S_s}$ and $S_t - S_s$ are independent of $\mathscr{F}_s$.

Let us first prove the following estimate of exponential type about the subordinator S_t.

LEMMA 3.1. *Let* $f_t:\mathbb{R}_+\to\mathbb{R}_+$ *be a bounded continuous nonnegative* $\mathscr{F}_t$*-adapted process. For any* $\varepsilon,\delta>0$, *we have*

$$P\left\{\int_0^t f_s\,\mathrm{d}S_s\le\varepsilon;\int_0^t f_s\,\mathrm{d}s>\delta\right\}\le \mathrm{e}^{1-\phi(1/\varepsilon)\delta},$$

where

$$\phi(\lambda):=\frac{\lambda}{2}\int_0^{(\log 2)/(\lambda\|f\|_\infty)}u\nu_S(\mathrm{d}u),\qquad \lambda>0,$$

and ν_S *is the Lévy measure of the subordinator* S_t.

PROOF. For $\lambda>0$, set

$$g_s^\lambda:=\int_0^\infty(1-\mathrm{e}^{-\lambda f_s u})\nu_S(\mathrm{d}u)$$

and

$$M_t^\lambda:=-\lambda\int_0^t f_s\,\mathrm{d}S_s+\int_0^t g_s^\lambda\,\mathrm{d}s.$$

Let $\mu(t,\mathrm{d}u)$ be the Poisson random measure associated with S_t, that is,

$$\mu(t,U):=\sum_{s\le t}1_U(\Delta S_s),\qquad U\in\mathscr{B}(\mathbb{R}_+).$$

Let $\tilde\mu(t,\mathrm{d}u)$ be the compensated Poisson random measure of $\mu(t,\mathrm{d}u)$, that is,

$$\tilde\mu(t,\mathrm{d}u)=\mu(t,\mathrm{d}u)-t\nu_S(\mathrm{d}u).$$

Then we can write

$$\int_0^t f_s\,\mathrm{d}S_s=\int_0^t\int_0^\infty f_s u\mu(\mathrm{d}s,\mathrm{d}u).$$

By Itô's formula, we have

$$\mathrm{e}^{M_t^\lambda}=1+\int_0^t\int_0^\infty \mathrm{e}^{M_{s-}^\lambda}[\mathrm{e}^{-\lambda f_s u}-1]\tilde\mu(\mathrm{d}s,\mathrm{d}u).$$

Since for $x>0$,

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

we have

$$g_s^\lambda\le\int_0^\infty(1\wedge(\lambda\|f\|_\infty u))\nu_S(\mathrm{d}u)$$

and

$$M_t^\lambda\le\int_0^t g_s^\lambda\,\mathrm{d}s\le t\int_0^\infty(1\wedge(\lambda\|f\|_\infty u))\nu_S(\mathrm{d}u).$$

Hence, for any $\lambda > 0$ and $t > 0$,

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

On the other hand, since for any $\kappa \in (0, 1)$ and $0 \le x \le -\log k$,

$$1 - \mathrm{e}^{-x} \ge \kappa x,$$

we have

$$\begin{aligned} g_s^\lambda &\ge \int_0^{(\log 2)/(\lambda\|f\|_\infty)} (1 - \mathrm{e}^{-\lambda f_s u})\nu_S(\mathrm{d}u) \\ &\ge \frac{\lambda f_s}{2} \int_0^{(\log 2)/(\lambda\|f\|_\infty)} u\nu_S(\mathrm{d}u) = \phi(\lambda) f_s. \end{aligned}$$

Thus,

$$\begin{aligned} \left\{\int_0^t f_s\,\mathrm{d}S_s \le \varepsilon; \int_0^t f_s\,\mathrm{d}s > \delta\right\} &\subset \left\{\mathrm{e}^{M_t^\lambda} \ge \mathrm{e}^{-\lambda\varepsilon + \int_0^t g_s^\lambda\,\mathrm{d}s}; \int_0^t g_s^\lambda\,\mathrm{d}s > \phi(\lambda)\delta\right\} \\ &\subset \left\{\mathrm{e}^{M_t^\lambda} \ge \mathrm{e}^{-\lambda\varepsilon + \phi(\lambda)\delta}\right\}, \end{aligned}$$

which then implies the result by Chebyshev's inequality and letting $\lambda = \frac{1}{\varepsilon}$. □

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

$$N(t, \Gamma) = \sum_{s \le t} 1_\Gamma(L_s - L_{s-}), \qquad \Gamma \in \mathscr{B}(\mathbb{R}^d).$$

Let $\widetilde{N}(t, \mathrm{d}y)$ be the compensated Poisson random measure of $N(t, \mathrm{d}y)$, that is,

$$\widetilde{N}(t, \mathrm{d}y) = N(t, \mathrm{d}y) - t\nu_L(\mathrm{d}y),$$

where ν_L is the Lévy measure of L_t given by (1.9). By Lévy–Itô's decomposition (cf. [1]), we have

$$L_t = W_{S_t} = \int_{|y| \le 1} y\widetilde{N}(t, \mathrm{d}y) + \int_{|y| > 1} yN(t, \mathrm{d}y). \tag{3.1}$$

We recall the following result about the exponential estimate of discontinuous martingales (cf. [7], Lemma 1).

LEMMA 3.2. *Let $f_t(y)$ be a bounded $\mathscr{F}_t$-predictable process with bound A. Then for any $\delta, \rho > 0$, we have*

$$\begin{aligned} &P\left\{\sup_{t \in [0,T]} \left|\int_0^t \int_{\mathbb{R}^d} f_s(y)\widetilde{N}(\mathrm{d}s, \mathrm{d}y)\right| \ge \delta, \int_0^T \int_{\mathbb{R}^d} |f_s(y)|^2 \nu_L(\mathrm{d}y)\,\mathrm{d}s < \rho\right\} \\ &\quad \le 2\exp\left(-\frac{\delta^2}{2(A\delta + \rho)}\right). \end{aligned}$$

The following lemma is contained in the proof of Norris' lemma (cf. [15], p. 137).

LEMMA 3.3. *For $T > 0$, let f be a bounded measurable $\mathbb{R}^d$-valued function on $[0, T]$. Assume that for some $\varepsilon < T$ and $x \in \mathbb{R}^d$,*

$$\int_0^T \left| x + \int_0^t f_s \, \mathrm{d}s \right|^2 \mathrm{d}t \le \varepsilon^3. \tag{3.2}$$

Then we have

$$\sup_{t \in [0,T]} \left| \int_0^t f_s \, \mathrm{d}s \right| \le 2(1 + \|f\|_\infty)\varepsilon.$$

PROOF. By (3.2) and Chebyshev's inequality, we have

$$\mathrm{Leb}\left\{ t \in [0, T] : \left| x + \int_0^t f_s \, \mathrm{d}s \right| \ge \varepsilon \right\} \le \varepsilon < T.$$

Thus, for each $t \in [0, T]$, there exits an $s \in [0, T]$ such that

$$|s - t| \le \varepsilon \quad \text{and} \quad \left| x + \int_0^s f_r \, \mathrm{d}r \right| < \varepsilon.$$

Consequently, for such t, s,

$$\left| x + \int_0^t f_r \, \mathrm{d}r \right| \le \left| x + \int_0^s f_r \, \mathrm{d}r \right| + \left| \int_s^t f_r \, \mathrm{d}r \right| \le \varepsilon + \varepsilon \|f\|_\infty.$$

In particular,

$$|x| \le \varepsilon + \varepsilon \|f\|_\infty,$$

hence,

$$\left| \int_0^t f_s \, \mathrm{d}s \right| \le |x| + \left| x + \int_0^t f_s \, \mathrm{d}s \right| \le 2(\varepsilon + \varepsilon \|f\|_\infty).$$

The proof is finished. □

We now prove the following Norris' type lemma (cf. [7, 15]).

LEMMA 3.4. *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:*

$$\beta_t = \beta_0 + \int_0^t \gamma_s \, \mathrm{d}s + \int_0^t \int_{\mathbb{R}^d} g_s(y) \tilde{N}(\mathrm{d}s, \mathrm{d}y),$$

where γ_t and $g_t(y)$ are two $\mathscr{F}_t$-predictable $\mathbb{R}^d$-valued processes. Suppose that for some nonrandom constants $C_1, C_2 \ge 1$ and all $s \ge 0$, $y \in \mathbb{R}^d$,

$$|\beta_s| + |\gamma_s| \le C_1, \qquad |g_s(y)| \le C_2(1 \wedge |y|). \tag{3.3}$$

Then for any $\delta \in (0, \frac{1}{3})$, *there exists* $\varepsilon_0 = \varepsilon_0(C_1, C_2, \nu_L, \delta) \in (0, 1)$ *such that for all* $T \in (0, 1)$ *and* $\varepsilon \in (0, T^3 \wedge \varepsilon_0)$,

$$P\left\{\int_0^T |Y_s|^2 \,\mathrm{d}s < \varepsilon, \int_0^T |\beta_s|^2 \,\mathrm{d}s \geq 9C_1^2 \varepsilon^\delta\right\} \leq 2\exp\left\{-\frac{\varepsilon^{\delta-(1/3)}}{9C_1}\right\}. \tag{3.4}$$

PROOF. Let us define

$$h_t := \int_0^t \beta_s \,\mathrm{d}s, \qquad M_t := \int_0^t \int_{\mathbb{R}^d} \langle h_s, g_s(y)\rangle_{\mathbb{R}^d} \widetilde{N}(\mathrm{d}s, \mathrm{d}y)$$

and

$$E_1 := \left\{\int_0^T |Y_s|^2 \,\mathrm{d}s < \varepsilon\right\}, \qquad E_2 := \left\{\sup_{t\in[0,T]} |h_t| \leq 2(1 + C_1)\varepsilon^{1/3}\right\},$$

$$E_3 := \{\langle M\rangle_T \leq C_3 \varepsilon^{2/3}\}, \qquad E_4 := \left\{\sup_{t\in[0,T]} |M_t| \leq \varepsilon^\delta\right\},$$

$$E_5 := \left\{\int_0^T |\beta_s|^2 \,\mathrm{d}s < 9C_1^2 \varepsilon^\delta\right\},$$

where C_3 is determined below.

First of all, by Lemma 3.3, one sees that for $\varepsilon < T^3$,

$$E_1 \subset E_2 \subset E_3, \tag{3.5}$$

where the second inclusion is due to

$$\begin{aligned}\langle M\rangle_T &= \int_0^T \int_{\mathbb{R}^d} \left|\langle h_s, g_s(y)\rangle_{\mathbb{R}^d}\right|^2 \nu_L(\mathrm{d}y) \,\mathrm{d}s \\ &\leq 4(1 + C_1)^2 C_2^2 \left(\int_{\mathbb{R}^d} 1 \wedge |y|^2 \nu_L(\mathrm{d}y)\right) \varepsilon^{2/3} =: C_3 \varepsilon^{2/3}.\end{aligned}$$

On the other hand, by the integration by parts formula, we have

$$\int_0^T |\beta_t|^2 \,\mathrm{d}t = \int_0^T \langle \beta_t, \mathrm{d}h_t\rangle_{\mathbb{R}^d} = \langle \beta_T, h_T\rangle_{\mathbb{R}^d} - \int_0^T \langle h_t, \gamma_t\rangle_{\mathbb{R}^d} \,\mathrm{d}t - M_T.$$

From this, one sees that on $E_2 \cap E_4$,

$$\begin{aligned}\int_0^T |\beta_t|^2 \,\mathrm{d}t &\leq 2C_1(1 + C_1)\varepsilon^{1/3}(1 + T) + \varepsilon^\delta \\ &\leq \big(4C_1(1 + C_1) + 1\big)\varepsilon^\delta \leq 9C_1^2 \varepsilon^\delta.\end{aligned}$$

This means that

$$E_2 \cap E_4 \subset E_5,$$

which together with (3.5) gives

$$E_1 \cap E_5^c \subset E_1 \cap E_4^c \subset E_2 \cap E_3 \cap E_4^c.$$

Thus, by Lemma 3.2 we have

$$P(E_1 \cap E_5^c) \leq 2\exp\left(-\frac{\varepsilon^{2\delta}}{2(2(1+C_1)\varepsilon^{(1/3)+\delta} + C_3\varepsilon^{2/3})}\right)$$

and (3.4) follows by choosing ε_0 with $C_3\varepsilon_0^{(1/3)-\delta} = 1$. $\square$

Below we set

$$\Sigma_t := \Sigma_t^\ell|_{\ell=S}, \qquad K_t := K_t^\ell|_{\ell=S}, \qquad J_t := J_t^\ell|_{\ell=S}. \tag{3.6}$$

The following lemma is a key step for proving the smoothness of $p_t(x, y)$.

LEMMA 3.5. *Let $\theta \in (0, \frac{1}{2})$ be given in* (1.12). *Under* ($U\mathscr{H}_1$) *and* (1.12), *for any $p > 1$, there exist $C_0 = C_0(p, \theta) > 0$ and $C_1 = C_1(p, \theta) > 0$ such that for all $t \in (0, 1)$ and $\varepsilon \in (0, C_0 t^{8/\theta})$*,

$$\sup_{|a|=1} P\left\{\int_0^t |aK_sA|^2\, \mathrm{d}S_s \leq \varepsilon\right\} \leq C_1\varepsilon^p. \tag{3.7}$$

PROOF. By Lemma 3.1 and (1.12), for the given θ in (1.12), there exists an $\varepsilon_0 = \varepsilon_0(\theta) > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ and $t \in (0, 1)$,

$$\begin{aligned}
&P\left\{\int_0^t |aK_sA|^2\, \mathrm{d}S_s \leq \varepsilon\right\} \\
&\quad \leq P\left\{\int_0^t |aK_sA|^2\, \mathrm{d}S_s \leq \varepsilon, \int_0^t |aK_sA|^2\, \mathrm{d}s \geq \varepsilon^\theta\right\} \\
&\qquad + P\left\{\int_0^t |aK_sA|^2\, \mathrm{d}s < \varepsilon^\theta\right\} \\
&\quad \leq \exp\left\{1 - \frac{1}{2\varepsilon^{1-\theta}} \int_0^{C\varepsilon} u\nu_S(\mathrm{d}u)\right\} + P\left\{\int_0^t |aK_sA|^2\, \mathrm{d}s < \varepsilon^\theta\right\} \\
&\quad \leq \exp\{1 - \varepsilon^{-\theta/2}\} + P\left\{\int_0^t |aK_sA|^2\, \mathrm{d}s < \varepsilon^\theta\right\}.
\end{aligned} \tag{3.8}$$

Notice that by (3.1),

$$X_t = x + \int_0^t b(X_s)\, \mathrm{d}s + \int_{|y|\leq 1} Ay\widetilde{N}(t, \mathrm{d}y) + \int_{|y|>1} AyN(t, \mathrm{d}y).$$

If we set $Y_t := aK_tA$ and

$$\begin{aligned}
&\beta_t := aK_t\nabla b(X_t)A, \qquad g_t(y) := aK_t\big(\nabla b(X_{t-} + Ay) - \nabla b(X_{t-})\big)A, \\
&\gamma_t := \int_{\mathbb{R}^d} aK_t\big(\nabla b(X_t + Ay) - \nabla b(X_t) - 1_{|y|\leq 1}Ay \cdot \nabla^2 b(X_t)\big)A\nu_L(\mathrm{d}y) \\
&\qquad + aK_tB_2(X_t)A,
\end{aligned}$$

then by equation (2.5) and Itô's formula, one sees that $Y_t = aA + \int_0^t \beta_s \, \mathrm{d}s$ and

$$\beta_t = a\nabla b(x)A + \int_0^t \gamma_s \, \mathrm{d}s + \int_0^t \int_{\mathbb{R}^d} g_s(y)\widetilde{N}(\mathrm{d}s, \mathrm{d}y).$$

By the assumptions, it is easy to see that

$$|\beta_t| + |\gamma_t| \le C_1\bigl(\|\nabla b\|_\infty, \|\nabla^3 b\|_\infty, \|A\|\bigr)$$

and

$$\bigl|g_t(y)\bigr| \le C_2\bigl(\|\nabla b\|_\infty, \|\nabla^2 b\|_\infty, \|A\|\bigr)\bigl(1 \wedge |y|\bigr).$$

Fix $\delta \in (0, \frac{1}{3})$. Define now

$$E_t^\varepsilon := \left\{\int_0^t |aK_sA|^2 \, \mathrm{d}s < \varepsilon^\theta\right\}, \qquad F_t^\varepsilon := \left\{\int_0^t \bigl|aK_s\nabla b(X_s)A\bigr|^2 \, \mathrm{d}s < 9C_1^2\varepsilon^{\theta\delta}\right\}.$$

Then, by Lemma 3.4, there is an $\varepsilon_0 \in (0, 1)$ such that for all $t \in (0, 1)$ and $\varepsilon \in (0, t^3 \wedge \varepsilon_0)$,

$$\begin{aligned} P\bigl(E_t^\varepsilon\bigr) &= P\bigl(E_t^\varepsilon \cap \bigl(F_t^\varepsilon\bigr)^c\bigr) + P\bigl(E_t^\varepsilon \cap F_t^\varepsilon\bigr) \\ &\le 2\exp\bigl\{-\varepsilon^{\theta(\delta-(1/3))}/(9C_1)\bigr\} + P\bigl(E_t^\varepsilon \cap F_t^\varepsilon\bigr). \end{aligned}$$

Define

$$\tau := \inf\bigl\{s \ge 0 : |K_s - I| \ge \tfrac{1}{2}\bigr\} \wedge t.$$

Then

$$P\bigl(E_t^\varepsilon \cap F_t^\varepsilon\bigr) \le P\bigl(E_t^\varepsilon \cap F_t^\varepsilon \cap \bigl\{\tau \ge \varepsilon^{\delta\theta/2}\bigr\}\bigr) + P\bigl(\tau < \varepsilon^{\delta\theta/2}\bigr).$$

By Chebyshev's inequality, we have for any $p > 1$,

$$\begin{aligned} P\bigl(\tau < \varepsilon^{\delta\theta/2}\bigr) &\le P\left\{\sup_{s \in (0, \varepsilon^{\delta\theta/2} \wedge t)} |K_s - I| \ge \frac{1}{2}\right\} \\ &\le 2^p\mathbb{E}\Bigl(\sup_{s \in (0, \varepsilon^{\delta\theta/2} \wedge t)} |K_s - I|^p\Bigr) \\ &\le C\bigl(\varepsilon^{\delta\theta/2} \wedge t\bigr)^p \end{aligned}$$

and by $(U\mathscr{H}_1)$,

$$\begin{aligned} E_t^\varepsilon \cap F_t^\varepsilon &\subset \left\{\int_0^t \bigl(|aK_sA|^2 + \bigl|aK_s\nabla b(X_s)A\bigr|^2\bigr) \, \mathrm{d}s < \varepsilon^\theta + 9C_1^2\varepsilon^{\delta\theta}\right\} \\ &\subset \left\{\int_0^t \frac{|aK_sA|^2 + |aK_s\nabla b(X_s)A|^2}{|aK_s|^2} |aK_s|^2 \, \mathrm{d}s < \bigl(1 + 9C_1^2\bigr)\varepsilon^{\delta\theta}\right\} \\ &\subset \left\{c_1 \int_0^t |aK_s|^2 \, \mathrm{d}s < \bigl(1 + 9C_1^2\bigr)\varepsilon^{\delta\theta}\right\}. \end{aligned}$$

Since on $\{\tau \geq \varepsilon^{\delta\theta/2}\}$,

$$|aK_s| \geq 1 - |K_s - I| \geq \tfrac{1}{2}, \qquad |a| = 1, \qquad s \in [0, \varepsilon^{\delta\theta/2} \wedge t],$$

it is easy to see that for any $\varepsilon < t^{2/\delta\theta} \wedge (\frac{c_1}{4(1+9C_1^2)})^{2/(\delta\theta)}$,

$$E_t^\varepsilon \cap F_t^\varepsilon \cap \{\tau \geq \varepsilon^{\delta\theta/2}\} \subset \{c_1(\varepsilon^{\delta\theta/2} \wedge t)/4 < (1 + 9C_1^2)\varepsilon^{\delta\theta}\} = \varnothing.$$

Hence, for any $p > 1$, if one takes $\delta = \frac{1}{4}$ and $C_0 = C_0(\varepsilon_0, p, \theta, c_1)$ being small enough, then for all $t \in (0, 1)$ and $\varepsilon \in (0, C_0 t^{8/\theta})$,

$$P(E_t^\varepsilon) \leq C\varepsilon^{\theta p/8},$$

which together with (3.8) yields (3.7) by resetting $p = \frac{8p'}{\theta}$. □

3.2. *S_t has finite moments of all orders*. In this subsection, we suppose that S_t has finite moments of all orders and $b \in C^\infty(\mathbb{R}^d)$ has bounded derivatives of all orders. The following lemma is standard.

LEMMA 3.6. *For any $m, k \in \{0\} \cup \mathbb{N}$ with $m + k \geq 1$ and $p \geq 1$, we have*

$$\sup_{x \in \mathbb{R}^d} \sup_{t \in [0,1]} \mathbb{E}(\|D^m \nabla^k X_t^\ell(x)\|_{\mathbb{H}^{\otimes m}}^p |_{\ell = S}) < +\infty. \tag{3.9}$$

PROOF. Noticing that

$$DX_t^\ell(x) = \int_0^t \nabla b(X_s^\ell(x)) DX_s^\ell(x)\, \mathrm{d}s + \cdot \wedge \ell_t$$

and

$$\nabla X_t^\ell(x) = I + \int_0^t \nabla b(X_s^\ell(x)) \nabla X_s^\ell(x)\, \mathrm{d}s,$$

we have

$$\|DX_t^\ell(x)\|_{\mathbb{H}} \leq \|\nabla b\|_\infty \int_0^t \|DX_s^\ell(x)\|_{\mathbb{H}}\, \mathrm{d}s + \ell_t^{1/2}$$

and

$$|\nabla X_t^\ell(x)| \leq 1 + \|\nabla b\|_\infty \int_0^t |\nabla X_s^\ell(x)|\, \mathrm{d}s.$$

By Gronwall's inequality, we obtain

$$\|DX_t^\ell(x)\|_{\mathbb{H}} \leq \ell_t^{1/2} + \mathrm{e}^{\|\nabla b\|_\infty t} \int_0^t \ell_s^{1/2}\, \mathrm{d}s$$

and

$$\|\nabla X_t^\ell(x)\|_{\mathbb{H}} \leq \mathrm{e}^{\|\nabla b\|_\infty t}.$$

Hence, for any $p \geq 1$,

$$\mathbb{E}(\|DX_t^\ell\|_{\mathbb{H}}^p|_{\ell=s}) \leq C\mathbb{E}|S_t|^{p/2} + C\int_0^t \mathbb{E}|S_s|^{p/2}\,\mathrm{d}s < +\infty.$$

Thus, we obtain (3.9) for $m + k = 1$. For the general m and k, it follows by similar calculations and the induction. □

We recall the following main criterion in the Malliavin calculus that a random vector admits a smooth density (cf. [15], pp. 100–103).

PROPOSITION 3.7. *Let $F = (F^1, \dots, F^d)$ be a smooth Wiener functional and $(\Sigma_F)_{ij} := \langle DF^i, DF^j\rangle_{\mathbb{H}}$ be the Malliavin covariance matrix. We assume that for all $p \geq 2$,*

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

Let G be another smooth Wiener functional and $\varphi \in C_b^\infty(\mathbb{R}^d)$. Then for any multi-index $\alpha = (\alpha_1, \dots, \alpha_m) \in \{1, 2, \dots, d\}^m$,

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

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

$$H_{(i)}(F, G) := \sum_j D^*\big(G(\Sigma_F^{-1})_{ij}DF^j\big),$$

$$H_\alpha(F, G) := H_{(\alpha_m)}\big(F, H_{(\alpha_1,\dots,\alpha_{m-1})}(F, G)\big).$$

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

$$\big\|H_\alpha(F, G)\big\|_p \leq C\big\|(\det \Sigma_F)^{-1}\big\|_{p_1}^{n_1}\|DF\|_{m,p_2}^{n_2}\|G\|_{m,p_3}.$$

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

Now we can prove the following gradient estimate.

THEOREM 3.8. *Under $(U\mathscr{H}_1)$ and (1.12), for any $k, m \in \{0\} \cup \mathbb{N}$ with $k + m \geq 1$, there are $\gamma_{k,m} > 0$ and $C = C(k, m) > 0$ such that for all $f \in C_b^\infty(\mathbb{R}^d)$ and $t \in (0, 1)$,*

$$\sup_{x\in\mathbb{R}^d}\big|\nabla^k\mathbb{E}\big((\nabla^m f)(X_t(x))\big)\big| \leq C\|f\|_\infty t^{-\gamma_{k,m}}. \tag{3.10}$$

PROOF. We first prove that there exists a constant $\gamma > 0$ such that for any $p \geq 1$, some $C = C(p) > 0$ and all $t \in (0, 1)$,

$$\|(\det \Sigma_t)^{-1}\|_p \leq C t^{-\gamma}, \tag{3.11}$$

which, by (2.7), is equivalent to prove that

$$\left\| \det\left(\int_0^t K_s A A^* K_s^* \,\mathrm{d}S_s\right)^{-1} \right\|_p \leq C t^{-\gamma}.$$

Since the determinant of a matrix is greater than d-times its smallest eigenvalue, that is,

$$\left(\inf_{|a|=1} \int_0^t |aK_s A|^2 \,\mathrm{d}S_s\right)^d \leq \det\left(\int_0^t K_s A A^* K_s^* \,\mathrm{d}S_s\right),$$

it suffices to prove that for some $\gamma' > 0$,

$$\left\| \left(\inf_{|a|=1} \int_0^t |aK_s A|^2 \,\mathrm{d}S_s\right)^{-1} \right\|_p \leq C t^{-\gamma'},$$

which will follow by showing that for all $p \geq 1$ and $\varepsilon \in (0, C_p t^{\gamma'})$,

$$P\left\{\inf_{|a|=1} \int_0^t |aK_s A|^2 \,\mathrm{d}S_s \leq \varepsilon\right\} \leq C\varepsilon^p.$$

Since S_t has finite moments of all orders, this estimate follows by (3.7) and a compact argument (see [15], p. 133, Lemma 2.3.1, for more details).

Next, by the chain rule, we have

$$\begin{aligned}
&\nabla^k \mathbb{E}((\nabla^m f)(X_t(x))) \\
&\quad = \sum_{j=1}^k \mathbb{E}((\nabla^{m+j} f)(X_t(x)) G_j(\nabla X_t(x), \ldots, \nabla^k X_t(x))) \\
&\quad = \sum_{j=1}^k \mathbb{E}(\mathbb{E}((\nabla^{m+j} f)(X_t^\ell(x)) G_j(\nabla X_t^\ell(x), \ldots, \nabla^k X_t^\ell(x)))|_{\ell=S}),
\end{aligned}$$

where $\{G_j, j = 1, \ldots, k\}$ are real polynomial functions. By Proposition 3.7, Lemma 3.6 and Hölder's inequality, there exist integer n and $p > 1$, $C > 0$ such that for all $t \in (0, 1)$,

$$\begin{aligned}
|\nabla^k \mathbb{E}((\nabla^m f)(X_t(x)))| &\leq C\|f\|_\infty \mathbb{E}(\|(\det \Sigma_t^\ell)^{-1}\|_p^n|_{\ell=S}) \\
&\leq C\|f\|_\infty \|(\det \Sigma_t)^{-1}\|_{np}^n.
\end{aligned}$$

Estimate (3.10) now follows by (3.11). □

3.3. *Without the finiteness assumption of moments.* Let S'_t be a subordinator with Lévy measure $1_{(0,1)}(u)\nu_S(\mathrm{d}u)$ and independent of $(W_t)_{t\geq 0}$. Let $p'_t(x,y)$ be the distributional density of $X'_t(x)$, where $X'_t(x)$ solves the following SDE:

$$X'_t(x) = x + \int_0^t b\big(X'_s(x)\big)\,\mathrm{d}s + AW_{S'_t}.$$

Let us write

$$\mathcal{P}'_t f(x) := \mathbb{E} f\big(X'_t(x)\big) = \int_{\mathbb{R}^d} f(y)p'_t(x,y)\,\mathrm{d}y.$$

We first prepare two simple lemmas for later use.

LEMMA 3.9. *Let $f \in C_b^\infty(\mathbb{R}^d)$. For any $m \in \mathbb{N}$, there exists a constant $C_{m,b} \geq 1$ such that for all $x \in \mathbb{R}^d$ and $t \in [0,1]$,*

$$\big|\nabla^m \mathcal{P}'_t f(x)\big| \leq C_{m,b} \sum_{k=1}^m \mathcal{P}'_t\big|\nabla^k f\big|(x). \tag{3.12}$$

PROOF. By the chain rule, (3.12) follows by the following estimate:

$$\sup_{t\in[0,1]} \sup_{x\in\mathbb{R}^d} \big|\nabla^m X'_t(x)\big| \leq C_{m,b}, \tag{3.13}$$

which has been proved in estimating (3.9). □

LEMMA 3.10. *Let $J'_t(x) := \nabla X'_t(x)$ and $K'_t(x)$ be the inverse matrix of $J'_t(x)$. Let $f = (f_{kl}) \in C_b^\infty(\mathbb{R}^d)$ be an $\mathbb{R}^m \times \mathbb{R}^m$ valued function. Then for any $j = 1,\dots,d$ and $k,l = 1,\dots,m$, we have the following formula*:

$$\mathcal{P}'_t(\partial_j f_{kl})(x) = \operatorname{div} Q_{kl}^{\cdot j}(t,x;f) - G_{kl}^j(t,x;f), \tag{3.14}$$

where

$$Q_{kl}^{ij}(t,x;f) := \mathbb{E}\big(f_{kl}\big(X'_t(x)\big)\big(K'_t(x)\big)_{ij}\big), \tag{3.15}$$

$$G_{kl}^j(t,x;f) := \mathbb{E}\big(f_{kl}\big(X'_t(x)\big)\operatorname{div}\big(K'_t\big)_{\cdot j}(x)\big). \tag{3.16}$$

Moreover, for any $m \in \{0\}\cup\mathbb{N}$, we have

$$\sup_{t\in[0,1]} \sup_{x\in\mathbb{R}^d} \big|\nabla^m K'_t(x)\big| \leq \widetilde{C}_{m,b}, \tag{3.17}$$

where $\widetilde{C}_{m,b} \geq 1$.

PROOF. Noticing that

$$\nabla\big(f\big(X'_t(x)\big)\big) = (\nabla f)\big(X'_t(x)\big)\nabla X'_t(x) = (\nabla f)\big(X'_t(x)\big)J'_t(x),$$

we have

$$(\nabla f)\big(X'_t(x)\big) = \nabla\big(f\big(X'_t(x)\big)\big)K'_t(x) = \operatorname{div}\big(f\big(X'_t\big)K'_t\big)(x) - f\big(X'_t(x)\big)\operatorname{div} K'_t(x),$$

which in turn gives (3.14) by taking expectations. As for (3.17), it follows by equation

$$K'_t(x) = I - \int_0^t K'_s(x)\cdot\nabla b\big(X'_s(x)\big)\,\mathrm{d}s$$

and estimate (3.13). □

Below, let $\mathscr{C} := \{\tau_1, \tau_2, \dots, \tau_n, \dots\}$ and $\mathscr{G} := \{\xi_1, \xi_2, \dots, \xi_n, \dots\}$ be two independent families of i.i.d. random variables in $\mathbb{R}^+$ and $\mathbb{R}^d$, respectively, which are also independent of $(W_t, S'_t)_{t\geq 0}$. We assume that τ_1 obeys the exponential distribution of parameter

$$\lambda := \nu_S\big([1,\infty)\big)$$

and ξ_1 has the distributional density

$$\frac{1}{\nu_S([1,\infty))}\int_1^\infty (2\pi s)^{-d/2}\mathrm{e}^{-|x|^2/2s}\nu_S(\mathrm{d}s).$$

Set $\tau_0 := 0$ and $\xi_0 := 0$, and define

$$N_t := \max\{n : \tau_0 + \tau_1 + \cdots + \tau_n \leq t\} = \sum_{n=0}^\infty 1_{\{\tau_0+\cdots+\tau_n\leq t\}}$$

and

$$H_t := \xi_0 + \xi_1 + \cdots + \xi_{N_t} = \sum_{j=0}^{N_t}\xi_j.$$

Then H_t is a compound Poisson process with Lévy measure

$$\nu_H(\Gamma) = \int_1^\infty (2\pi s)^{-d/2}\left(\int_\Gamma \mathrm{e}^{-|y|^2/2s}\,\mathrm{d}y\right)\nu_S(\mathrm{d}s).$$

Moreover, it is easy to see that H_t is independent of $W_{S'_t}$, and

$$(AW_{S_t})_{t\geq 0} \overset{(d)}{=} (AW_{S'_t} + AH_t)_{t\geq 0}. \tag{3.18}$$

Let $\hbar_t$ be a càdlàg purely discontinuous $\mathbb{R}^d$-valued function with finite many jumps and $\hbar_0 = 0$. Let $X^\hbar_t(x)$ solve the following SDE:

$$X^\hbar_t(x) = x + \int_0^t b\big(X^\hbar_s(x)\big)\,\mathrm{d}s + AW_{S'_t} + \hbar_t.$$

Let n be the jump number of $\hbar$ before time t. Let $0 = t_0 < t_1 < t_2 < \cdots < t_n < t$ be the jump time of $\hbar$. By the Markovian property of $X^\hbar_t(x)$, we have the following formula:

$$
\begin{aligned}
&\mathbb{E}f\big(X^{\hbar}_t(x)\big)\\
&\quad=\int_{\mathbb{R}^d}\bigg(\int_{\mathbb{R}^d}\bigg(\int_{\mathbb{R}^d}\cdots\bigg(\int_{\mathbb{R}^d}p'_{t_1}(x,y_1)p'_{t_2-t_1}(y_1+\Delta\hbar_{t_1},y_2)\,\mathrm{d}y_1\bigg)\\
&\qquad\qquad\qquad\cdots p'_{t_n-t_{n-1}}(y_{n-1}+\Delta\hbar_{t_{n-1}},y_n)\,\mathrm{d}y_{n-1}\bigg)\\
&\qquad\qquad\qquad\qquad\times p'_{t-t_n}(y_n+\Delta\hbar_{t_n},z)\,\mathrm{d}y_n\bigg)f(z)\,\mathrm{d}z\\
&\quad=\mathcal{P}'_{t_1}\cdots\vartheta_{\Delta\hbar_{t_{n-1}}}\mathcal{P}'_{t_n-t_{n-1}}\vartheta_{\Delta\hbar_{t_n}}\mathcal{P}'_{t-t_n}f(x),
\end{aligned}
$$

where

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

Now, by (3.18) we have

$$X_t(x)\overset{(d)}{=}X^{\hbar}_t(x)|_{\hbar=AH.}$$

and so,

$$
\begin{aligned}
\mathcal{P}_t f(x)&=\mathbb{E}f\big(X_t(x)\big)=\mathbb{E}\big(\mathbb{E}f\big(X^{\hbar}_t(x)\big)|_{\hbar=AH.}\big)\\
&=\sum_{n=0}^{\infty}\mathbb{E}\big(\mathcal{P}'_{\tau_1}\cdots\vartheta_{A\xi_{n-1}}\mathcal{P}'_{\tau_n}\vartheta_{A\xi_n}\mathcal{P}'_{t-(\tau_0+\tau_1+\cdots+\tau_n)}f(x);N_t=n\big).
\end{aligned}
$$

In view of

$$\{N_t=n\}=\{\tau_0+\cdots+\tau_n\le t<\tau_0+\cdots+\tau_{n+1}\}$$

and that $\mathscr{C}$ is independent of $\mathscr{G}$, we further have

$$
\begin{aligned}
\mathcal{P}_t f(x)&=\sum_{n=1}^{\infty}\bigg\{\int_{t_1+\cdots+t_n<t<t_1+\cdots+t_{n+1}}\lambda^{n+1}\mathrm{e}^{-\lambda(t_1+\cdots+t_n+t_{n+1})}\\
&\qquad\times\mathbb{E}\big(\mathcal{P}'_{t_1}\cdots\vartheta_{A\xi_{n-1}}\mathcal{P}'_{t_n}\vartheta_{A\xi_n}\mathcal{P}'_{t-(t_1+\cdots+t_n)}f(x)\big)\,\mathrm{d}t_1\cdots\mathrm{d}t_{n+1}\bigg\}\\
&\quad+\mathcal{P}'_t f(x)P(N_t=0) \qquad (3.19)\\
&=\sum_{n=1}^{\infty}\bigg\{\lambda^n\mathrm{e}^{-\lambda t}\int_{t_1+\cdots+t_n<t}\mathbb{E}I^{A\boldsymbol{\xi}}_f(t_1,\ldots,t_n,t,x)\,\mathrm{d}t_1\cdots\mathrm{d}t_n\bigg\}\\
&\quad+\mathcal{P}'_t f(x)\mathrm{e}^{-\lambda t},
\end{aligned}
$$

where $\boldsymbol{\xi}:=(\xi_1,\ldots,\xi_n)$, and

$$I^{\mathbf{y}}_f(t_1,\ldots,t_n,t,x):=\mathcal{P}'_{t_1}\cdots\vartheta_{y_{n-1}}\mathcal{P}'_{t_n}\vartheta_{y_n}\mathcal{P}'_{t-(t_1+\cdots+t_n)}f(x)$$

with $\mathbf{y} := (y_1, \dots, y_n)$.

Now we can complete the proof of Theorem 1.3.

PROOF OF THEOREM 1.3. We first establish the same gradient estimate as in (3.10).

If we let $t_{n+1} := t - (t_1 + \cdots + t_n) > 0$, then there is at least one $j \in \{1, 2, \dots, n+1\}$ such that

$$t_j \geq \frac{t}{n+1}. \tag{3.20}$$

Thus, we have

$$\begin{aligned}
\left|\nabla_x I_f^{\mathbf{y}}(t_1, \dots, t_n, t, x)\right| &\overset{(3.12)}{\leq} C_{1,b}^{j-1} \|\nabla_x \mathcal{P}'_{t_j} \cdots \vartheta_{y_{n-1}} \mathcal{P}'_{t_n} \vartheta_{y_n} \mathcal{P}'_{t_{n+1}} f\|_\infty \\
&\overset{(3.10)}{\leq} C C_{1,b}^{j-1} t_j^{-\gamma_{1,0}} \|\mathcal{P}'_{t_{j+1}} \cdots \vartheta_{y_{n-1}} \mathcal{P}'_{t_n} \vartheta_{y_n} \mathcal{P}'_{t_{n+1}} f\|_\infty \\
&\leq C C_{1,b}^{n} \big(t/(n+1)\big)^{-\gamma_{1,0}} \|f\|_\infty.
\end{aligned}$$

Here and below, the various constant C is independent of t and n. Hence, by (3.19) we have

$$\begin{aligned}
\left|\nabla \mathcal{P}_t f(x)\right| &\leq C \|f\|_\infty t^{-\gamma_{1,0}} \mathrm{e}^{-\lambda t} \\
&\quad \times \left(1 + \sum_{n=1}^\infty \lambda^n C_{1,b}^n (n+1)^{\gamma_{1,0}} \int_{t_1 + \cdots + t_n < t} \mathrm{d}t_1 \cdots \mathrm{d}t_n\right) \\
&= C \|f\|_\infty t^{-\gamma_{1,0}} \mathrm{e}^{-\lambda t} \left(\sum_{n=0}^\infty \lambda^n C_{1,b}^n (n+1)^{\gamma_{1,0}} \frac{t^n}{n!}\right) \\
&\leq C \|f\|_\infty t^{-\gamma_{1,0}}.
\end{aligned} \tag{3.21}$$

Thus, we obtain (3.10) with $k = 1$ and $m = 0$.

For $k, l = 1, \dots, d$, set $F_{kl}^{(0)}(x) := 1_{k=l} f(x)$ and $R_l^{(0)}(x) := 0$. Let us recursively define for $m = 0, 1, \dots, n$,

$$F_{kl}^{(m+1)}(x) := \sum_{i=1}^d Q_{il}^{ki}\big(t_{n+1-m}, x; \vartheta_{y_{n+1-m}} F^{(m)}\big),$$

$$R_l^{(m+1)}(x) := \sum_{i=1}^d G_{il}^{i}\big(t_{n+1-m}, x; \vartheta_{y_{n+1-m}} F^{(m)}\big),$$

where $y_{n+1} := 0$, Q_{il}^{ki} and G_{il}^i are defined by (3.15) and (3.16). From these definitions and by (3.17), it is easy to see that

$$\begin{aligned}
\|F_{kl}^{(m+1)}\|_\infty &\leq d \|F_{kl}^{(m)}\|_\infty \mathbb{E}\|K'_t(x)\| \leq \widetilde{C}_{0,b} \|F_{kl}^{(m)}\|_\infty \\
&\leq \widetilde{C}_{0,b}^{m+1} \|F_{kl}^{(0)}\|_\infty \leq \widetilde{C}_{0,b}^{m+1} \|f\|_\infty
\end{aligned}$$

and

$$\|R_l^{(m+1)}\|_\infty \le \|F_{kl}^{(m)}\|_\infty \mathbb{E}\big|\operatorname{div} K_t'(x)\big| \le \widetilde{C}_{0,b}^m \widetilde{C}_{1,b} \|f\|_\infty.$$

By repeatedly using Lemma 3.10, we have

$$\begin{aligned}
&\big|I_{\partial_l f}^{\mathbf{y}}(t_1,\ldots,t_n,t,x)\big| \\
&\quad = \left|\mathcal{P}_{t_1}' \cdots \vartheta_{y_{j-1}} \mathcal{P}_{t_j}' \operatorname{div} F_{\cdot l}^{(n+1-j)}(x) - \sum_{m=1}^{n+1-j} \mathcal{P}_{t_1}' \cdots \vartheta_{y_{n+1-m}} \mathcal{P}_{t_{n+1-m}}' R_l^{(m)}(x)\right| \\
&\quad \overset{(3.10)}{\le} C t_j^{-\gamma_{0,1}} \|F_{\cdot l}^{(n+1-j)}\|_\infty + \sum_{m=1}^{n+1-j} \|R_l^{(m)}\|_\infty \\
&\quad \overset{(3.20)}{\le} C\big(t/(n+1)\big)^{-\gamma_{0,1}} \widetilde{C}_{0,b}^n \|f\|_\infty + C\widetilde{C}_{0,b}^n \|f\|_\infty.
\end{aligned}$$

As in estimating (3.21), we obtain (3.10) with $k = 0$ and $m = 1$. For the general m and k, the gradient estimate (3.10) follows by similar calculations and the induction.

Lastly, by estimate (3.10) and Sobolev's embedding theorem (see [15], pp. 102–103), one has that for each $t > 0$,

$$(x, y) \mapsto p_t(x, y) \in C_b^\infty\big(\mathbb{R}^d \times \mathbb{R}^d\big).$$

The smoothness of $p_t(x, y)$ with respect to the time variable t follows by equation (1.11) and the standard bootstrap argument. As for equation (1.13), it follows by

$$\frac{\mathrm{d}\mathcal{P}_t f(x)}{\mathrm{d}t} = \mathcal{L}_A \mathcal{P}_t f(x) + b(x) \cdot \nabla_x \mathcal{P}_t f(x),$$

where $f \in C_b^\infty(\mathbb{R}^d)$. $\square$

Acknowledgements. The author is very grateful to Professors Hua Chen, Zhen-Qing Chen and Feng-Yu Wang for their quite useful conversations. The referees' very useful suggestions are also deeply acknowledged.

REFERENCES

[1] APPLEBAUM, D. (2009). *Lévy Processes and Stochastic Calculus*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **116**. Cambridge Univ. Press, Cambridge. MR2512800

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

[3] BERTOIN, J. (1996). *Lévy Processes*. *Cambridge Tracts in Mathematics* **121**. Cambridge Univ. Press, Cambridge. MR1406564

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

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

[6] BONDARCHUK, S. V. and KULIK, O. M. (2008). Conditions for the existence and smoothness of a density of distribution for Ornstein–Uhlenbeck processes with Lévy noise. *Theory Probab. Math. Statist.* **79** 20–33. MR2494533

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

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

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

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

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

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

[13] LÉANDRE, R. (1988). Régularité de processus de sauts dégénérés. II. *Ann. Inst. Henri Poincaré Probab. Stat.* **24** 209–236. MR0953118

[14] MALLIAVIN, P. (1978). Stochastic calculus of variation and hypoelliptic operators. In *Proceedings of the International Symposium on Stochastic Differential Equations* (*Res. Inst. Math. Sci., Kyoto Univ., Kyoto*, 1976) 195–263. Wiley, New York. MR0536013

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

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

[17] PRIOLA, E. and ZABCZYK, J. (2009). Densities for Ornstein–Uhlenbeck processes with jumps. *Bull. Lond. Math. Soc.* **41** 41–50. MR2481987

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

[19] SATO, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge Univ. Press, Cambridge. MR1739520

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

[21] ZHANG, X. (2013). Derivative formulas and gradient estimates for SDEs driven by α-stable processes. *Stochastic Process. Appl.* **123** 1213–1228. MR3016221

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