Smoothness of the distribution of the supremum of a multi-dimensional diffusion process

Masafumi Hayashi *†‡ Arturo Kohatsu-Higa §¶

October 17, 2011

Abstract

In this article we deal with a multi-dimensional diffusion whose corresponding diffusion vector fields are commutative, and study its joint distribution at the time when a component attains its maximum on finite time interval. Under regularity and ellipticity conditions we shall show the smoothness of this joint distribution.

Keywords: Malliavin Calculus, maximum process, stochastic differential equations

AMS Subject Classification Number (2010): 60H07, 60H10

1 Introduction

In this paper we shall consider the following stochastic differential equation:

$$\begin{cases} dX_t^i &= \sum_{j=1}^d \sigma_j^i(X_t) \circ dW_t^j + b^i(X_t)dt \\ X_0 &= x^0, \end{cases} \tag{1}$$

where $b : \mathbf{R}^d \to \mathbf{R}^d$ and $\sigma_j^i : \mathbf{R}^d \to \mathbf{R}$ $(i, j = 1, \ldots, d)$ are smooth functions with bounded derivatives, $(W^1, \ldots, W^d)$ is a d-dimensional Brownian motion and $x^0 = (x_1^0, \ldots, x_d^0) \in \mathbf{R}^d$. The main goal of the present study is to determine the smoothness of the distribution of $F = (X_{\theta^1}^1, \ldots, X_{\theta^1}^d)$ where θ_1 is the time when X_t^1 attains its maximum on the finite time interval $[0, T]$. This type of random variables are important in a wide variety of applied fields such as in path dependent options in mathematical finance. Furthermore, the distribution of the maximal process $F^1 = X_{\theta^1}^1 = \sup_{0 \le t \le T} X_t^1$ is closely related to that of the first hitting time $T_{a_0} := \inf\{t > 0\,;\, X_t^1 \ge a_0\}$ by which we can define the default of risky assets in structural models. Application in the calculation of

*Center for the Study of Finance and Insurance, Osaka University

†mhayashi@sigmath.es.osaka-u.ac.jp

‡Japan Science and Technology Agency

§Department of Mathematical Sciences, Ritsumeikan University

¶Japan Science and Technology Agency

Greeks can be found in Gobet-Kohatsu [6]. Other applications of the present results in finance are currently under study.

The first mathematical problem that appears in this study is the uniqueness of the time at which the supremum of X is achieved in the time interval $[0,T]$. To obtain this uniqueness, it is clear that some conditions are needed in order to avoid the constancy of the process at some time intervals. These conditions are expressed in Proposition 1.

To study the distribution of F, we shall use techniques from Malliavin calculus. Malliavin calculus is a tool to study the smoothness of the distribution of Wiener functionals. However, because of the singularity of the random variable F as a functional of the underlying Wiener process, it is difficult to apply Malliavin calculus for F directly. Although the supremum of a smooth continuous process belongs to $\mathbf{D}^{1,2}$ (see [11], [2]), we cannot expect the smoothness of the supremum process. In [8] the authors have studied the density for the supremum of a fractional Brownian motion and proved the smoothness of its density. The density for the supremum of Wiener sheet is also studied in [5] (see also Section 2.1.7 of Nualart [10]). To avoid the singularity, the authors established the integration by parts formula without using the Malliavin covariance matrix. We shall also establish this sort of integration by parts formula (Theorem 12) however, contrary to the cases of the Wiener sheet or the fractional Brownian motion, the derivative of F has a somewhat more complicated expression in general.

In this paper, to avoid such complexity, we assume that the vector fields $V_j = \sum_{i=1}^d \sigma_j^i(x)\frac{\partial}{\partial x^i}$ $(j = 1,\ldots,d)$ are commutative. Diffusions with commutative vector fields are studied in [4]. In their paper, the authors gave some formulas related to the Malliavin derivative of diffusions with commutative vector fields. As we shall mention in Remark 8, under the assumption of the commutativity of V_j, we can find a relationship between the Malliavin derivative D and the vector field V_j in the chain rule:

$$D^l\psi(F) = (V_l\psi)(F)1_{[0,\theta_1]}. \tag{2}$$

Therefore to obtain an integration by parts formula we need to find a function ψ so that $V_l\psi = f$ for a given function f and estimate its sup-norm. These properties are obtained in Theorem 17.

This paper is organized as follows: in the following subsection, we shall mention our assumptions, and exhibit the main result. In Section 2, we prove the differentiability of F and give the formula of its derivative from which we can deduce (2). We shall establish the integration by parts formula with respect to the vector fields V_j in Section 3. The discussion of this section is a modification of Section 2.1.7 in Nualart [10]. In Section 4, we study the properties of the function ψ in function of f. Finally, we shall give the proof of the main result in Section 5. In Section 6, we shall generalize our result for an SDE which has a general drift, as it is the case in most financial applications. In Section 7 (Appendix) we shall prove the uniqueness of θ_1, that is, X_t^1 attains its maximum on $[0,T]$ on a unique point θ_1.

We denote by $C_b^\infty(\mathbf{R}^d : \mathbf{R}^m)$ the space of smooth functions with bounded derivatives defined in $\mathbf{R}^d$ taking values in $\mathbf{R}^m$. Similarly, we denote by $C_c^\infty(\mathbf{R}^d : \mathbf{R}^m)$ the space of smooth function with compact support and $C_p^\infty(\mathbf{R}^d : \mathbf{R}^m)$ the space of smooth function such that f and all of its partial derivatives have polynomial growth.

For fundamental results in Malliavin calculus and notations, we refer to Nualart [10]. In particular we will denote by D^i the stochastic derivative with respect to the i-th noise W^i and its adjoint on the time interval by δ^i.

As it is usual, constants may change from one line to the next without further comment.

1.1 Assumptions and main result

Let $(\Omega, \mathcal{F}, \mathbf{P}), \{\mathcal{F}_t\}_{t\in[0,T]}, W_t = (W_t^1, \ldots, W_t^d),\ t \in [0,T]$ be a d-dimensional standard Brownian motion. We shall consider the following d-dimensional stochastic differential equation:

$$\begin{cases} dX_t^i &= \sum_{j=1}^d \sigma_j^i(X_t) \circ dW_t^j \\ X_0 &= x^0 = (x_1^0, \ldots, x_d^0), \end{cases} \tag{3}$$

where $\sigma : \mathbf{R}^d \to \mathbf{R}^d \times \mathbf{R}^d$ is a matrix valued function and $\circ dW_s$ denotes the Stratonovich integral. Put

$$a(x) = \sigma\, \sigma^T(x),$$

where σ^T is the transpose matrix of σ.

Throughout this paper, we assume the following assumption.

Assumption A

(A1). There exists a constant $\Lambda \in (0,1)$ such that for any $\xi,\ x \in \mathbf{R}^d$

$$\Lambda|\xi|^2 \le \langle \xi\, ,\, a(x)\xi \rangle \le \Lambda^{-1}|\xi|^2,$$

holds.

(A2). $\sigma \in C_b^\infty(\mathbf{R}^d : \mathbf{R}^d \times \mathbf{R}^d)$. Note that (A1) implies that σ is also bounded.

(A3). Set for any $f \in C^1(\mathbf{R}^d : \mathbf{R})$

$$V_j f(x) = \sum_{i=1}^d \sigma_j^i(x) \frac{\partial f}{\partial x^i}(x).$$

The vector fields $V_1, \ldots, V_d$ are commutative in the sense that $V_i V_j = V_j V_i$ for all $1 \le i, j \le d$.

Set for $i = 1, \ldots, d$

$$M_t^i = \sup_{0 \le s \le t} X_s^i,$$

and for simplicity we denote by

$$M = M_T = (M_T^1, \dots, M_T^d).$$

Without further mentioning, we remark that under Assumption (A), $M_T \in \cap_{p\geq 1} L^p(\Omega)$.

Proposition 1 *Suppose that condition (A) is satisfied and that*

$$Leb \left\{ x\,;\, \sum_{j=1}^{d} \sigma_j^1(x) = 0 \right\} = 0. \tag{4}$$

Then, with probability one, $\{X_t^1;\ t \in [0,T]\}$ attains its maximum on $[0,T]$ on a unique random point θ^1.

This Proposition is a generalization of that for 1-dim Brownian motion see Remark 8.16 in Chapter 2 of [7]. Therefore the natural idea is to reduce to that case. In fact, in view of Girsanov theorem, X^1 is a martingale under a measure that is equivalent to $\mathbf{P}$, and therefore a time changed Brownian motion.

We define

$$F^i = X_{\theta^1}^i.$$

The following theorem is our main result.

Theorem 2 *Suppose that Assumption A and (4) hold. Further we suppose that*

$$\inf_{x\in\mathbf{R}^d} \sigma_1^1(x) > 0. \tag{5}$$

Then $F = (F^1, \dots, F^d)$ has a smooth density on $(X_0^1, \infty) \times \mathbf{R}^{d-1}$.

Remark 3 *a) Condition (5) is used in (28) which is a condition to assure that the mapping that defines X^1 as a function of the Wiener process is an increasing function with respect to W^1.*

b) A similar result for $(X_{\theta^i}^1, \dots, X_{\theta^i}^d)$ can be achieved if we replace the index 1 by $i \neq 1$.

c) The theorem is also valid for the following SDE:

$$X_t^i = x^0 + \sum_{j=1}^{d} \int_0^t \sigma_j^i(X_s) \circ dW_t^j + \int_0^t b^i(X_s)\, ds,$$

where b is a bounded smooth function. The proof for this case will be given in Section 6.

Example 4 *Let us consider a one dimensional diffusion X_t:*

$$X_t = x^0 + \int_0^t \sigma(X_s) \circ dW_s.$$

Suppose that σ is a bounded smooth function and that $\inf_{x\in\mathbf{R}} \sigma(x) > 0$ then Assumption A holds. Define

$$\begin{cases} \frac{d\varphi(x)}{dx} & = \sigma(\varphi(x)) \\ \varphi(0) & = x^0. \end{cases}$$

One can check that $X_t = \varphi(W_t)$ (see [3]). Since $x \to \varphi(x)$ is an increasing function, we have $\sup_{0<t\leq T} X_t = \varphi(\sup_{0<t\leq T} W_t)$. Since $\sup_{0<s\leq T} W_s$ has the smooth density on $(0, \infty)$, so does X_t. Because W_t attains its maximum on $[0, T]$ on a unique point, so does X_t. In multi-dimensional cases, the uniqueness of the time when X^1 attains its maximum on $[0, T]$ is not clear in general. Here we are able to prove this uniqueness under the assumption (4).

In mathematical finance, Assumption (A3) is not binding as the following multi-dimensional Black-Scholes type example shows.

Example 5 *Suppose that $\sigma : \mathbf{R} \to \mathbf{R}^d \times \mathbf{R}^d$ satisfies (A1) and (A2). Further assume that*

$$\sigma^i_j(x_1, \ldots, x_d) = \sigma^i_j(x_j).$$

Then σ satisfies Assumption (A).

2 The differentiability of F

In this section we shall show the differentiability of F. Note that under Assumption A we have

Proposition 6 *Under Assumption A, there exists a diffeomorphism $\varphi = (\varphi_1, \ldots, \varphi_d) : \mathbf{R}^d \to \mathbf{R}^d$ satisfying*

$$\begin{cases} \frac{\partial(\varphi_1,\ldots,\varphi_d)}{\partial(x_1,\ldots,x_d)} = \sigma \circ \varphi \\ \varphi_i(0) = x^0_i \qquad\qquad (1 \leq i \leq d), \end{cases} \tag{6}$$

where we denote by $\frac{\partial(\varphi_1,\ldots,\varphi_d)}{\partial(x_1,\ldots,x_d)}$ the Jacobi matrix of φ.

Proof. Let $\phi_j(u)$ be the exponential map, that is, the one parameter group generated by the operator V_j:

$$\phi_j(u) = \mathrm{Exp}(uV_j),$$

and put

$$\varphi(x_1, \ldots, x_d) = \phi_1(x_1) \circ \cdots \circ \phi_d(x_d)(x^0).$$

Then because $V_1, \ldots, V_j$ are commutative, φ satisfies (6). From the assumption (A1) σ is bounded. Hence the existence and differentiability of φ^{-1} follows from Corollary 4.1 in [12]. ■

By using φ, we can solve the stochastic differential equation (3) as follows:

$$X_t^i = \varphi_i(W_t^1, \ldots, W_t^d). \tag{7}$$

In [4], the authors studied this sort of formula for time inhomogeneous SDE (see Theorem 3.1 in [4]).

We denote by D the H-derivative operator. Then from (7) we can evaluate the derivative $D_r^l X_t^i$ as

$$D_r^l X_t^i = \sum_{k=1}^d \frac{\partial \varphi_i}{\partial x^k}(W_t)\, D_r^l W_t^k = \sigma_l^i(X_t) 1_{[0,t]}(r). \tag{8}$$

Theorem 7 *For each i, F^i is differentiable, and it holds that*

$$D_r^l F^i = \sigma_l^i(F) 1_{[0,\theta^1]}(r). \tag{9}$$

Remark 8 *The relationship in* (2) *can be deduced from this theorem.*

Proof. We shall prove the theorem in the case $i \neq 1$ only, since the case $i = 1$ is easier than the case $i \neq 1$. The proof follows along the lines of the proof of Proposition 2.1.10 in page 109 of [10]. We give the details here for easy reference.

Let $S_0 = \{t_1, t_2, \ldots\}$ be a dense subset of $[0, T]$. Put

$$M_n^1 = \max\{X_{t_1}^1, \ldots, X_{t_n}^1\}.$$

Then, we have $M_n^1 \to M^1$ **P**-a.s.. Define

$$\begin{aligned} A_1 &= \{X_{t_1}^1 = M_n^1\}, \\ A_2 &= \{X_{t_1}^1 \neq M_n^1,\, X_{t_2}^1 = M_n^1\}, \\ &\vdots \\ A_k &= \{X_{t_1}^1 \neq M_n^1, \ldots, X_{t_{k-1}}^1 \neq M_n^1, X_{t_k}^1 = M_n^1\}. \end{aligned}$$

Then one can check that

$$A_k \cap A_m = \emptyset \qquad \text{if } k \neq m, \tag{10}$$

and

$$\bigcap_{p=1}^k A_p^c = \bigcap_{p=1}^k \{X_{t_p}^1 \neq M_n^1\}. \tag{11}$$

Define

$$G_n^i = \sum_{p=1}^n 1_{A_p} X_{t_p}^i.$$

Then $G_n^i \to F^i$ **P**-a.s. We shall prove that G_n^i is differentiable and

$$D_\cdot^l G_n^i = \sum_{p=1}^n 1_{A_p} D_\cdot^l X_{t_p}^i, \tag{12}$$

holds. Let $\phi(x)$ be a non-negative smooth function on $\mathbf{R}$ such that $\phi(0) = 1$, and the support of ϕ is contained in $[-1, 1]$. For $\varepsilon > 0$, we set $\phi_\varepsilon(x) = \phi(\frac{x}{\varepsilon})$. Inductively, we shall introduce a sequence of random variables

$$\Phi_1^\varepsilon = \phi_\varepsilon(M_n^1 - X_{t_1}^1),$$
$$\Phi_k^\varepsilon = (1 - \sum_{p=1}^{k-1} \Phi_p^\varepsilon)\phi_\varepsilon(M_n^1 - X_{t_k}^1) \qquad \text{(for } k = 2, \dots, n).$$

By (10), (11) and induction, one can show that

$$\lim_{\varepsilon \downarrow 0} \Phi_k^\varepsilon = 1_{A_k} \qquad \mathbf{P}\text{-}a.s.$$

Thus we have

$$G_n^i = \lim_{\varepsilon \downarrow 0} \sum_{k=1}^n \Phi_k^\varepsilon X_{t_k}^i \qquad \mathbf{P}\text{-}a.s.$$

Note that Φ_k^ε is differentiable, since M_n^1 is differentiable (see Proposition 2.1.20 in Nualart [10]). Further we have for any $k = 1, \dots, n$

$$\begin{aligned}
\lim_{\varepsilon \to 0} D^l \phi_\varepsilon(M_n^1 - X_{t_k}^1) &= \lim_{\varepsilon \to 0} \phi_\varepsilon'(M_n^1 - X_{t_k}^1) D^l(M_n^1 - X_{t_k}^1) \\
&= 1_{\{M_n^1 = X_{t_k}^1\}} D^l(M_n^1 - X_{t_k}^1) \\
&= 0,
\end{aligned}$$

where the last equality follows from local property (see Proposition 1.3.16 in Nualart [10]). Hence one can prove by induction that

$$\lim_{\varepsilon \to 0} D^l \Phi_k^\varepsilon = 0 \qquad \mathbf{P}\text{-}a.s.$$

Therefore (12) holds true, since D^l is closable.

Further, it follows from (10) that

$$\mathbf{E}\left[\left\| \sum_{p=1}^n 1_{A_p} D_\cdot^l X_{t_p}^i \right\|_{\mathbf{L}^2([0,T];\mathbf{R})} \right] \le \mathbf{E}[\sup_{t \in [0,T]} \|D_\cdot X_t^i\|_{\mathbf{L}^2([0,t];\mathbf{R})}] < \infty.$$

Hence $\lim_n \sum_{p=1}^n 1_{A_p} D_r^l X_{t_p}^i$ exists. Therefore F_1^i is differentiable and (9) holds, since D^l is closable. ∎

3 Integration by parts formula

In this section, we shall establish the integration by parts formula with respect to the operator $V_j = \sum_{i=1}^d \sigma_j^i(x)\frac{\partial}{\partial x^i}$. Most of the arguments are slight variations of similar ones in [10], we write them here so that the article is self-contained. We start recalling the Garsia, Rodemich, and Rumsey's lemma (see [10]):

Lemma 9 *Let p, $\Psi : \mathbf{R}_+ \to \mathbf{R}_+$ be continuous and strictly increasing functions such that $\Psi(0) = 0$ and $\lim_{t\to\infty}\Psi(t) = \infty$. Let $\phi : \mathbf{R}^d \to E$ be a continuous function which values in separable Banach space $(E\,,\, \|\cdot\|)$. Suppose that*

$$\Gamma = \int_{B_\rho}\int_{B_\rho} \Psi\left(\frac{\|\phi(t)-\phi(s)\|}{p(|t-s|)}\right) ds\, dt < \infty,$$

where $B_\rho = \{x\, : |x_0 - x| < \rho\}$. Then it holds that

$$\|\phi(t)-\phi(s)\| \le 8\int_0^{2|t-s|} \Psi^{-1}\left(\frac{4^{d+1}\Gamma}{\lambda u^{2d}}\right) p(du),$$

where λ is a constant depending only on d.

Lemma 10 *Suppose that p and r satisfies that $2d < 1 + 2pr < p$, and define*

$$Y_t = \int_{[0,t]^2} \frac{|W_s - W_{s'}|^{2p}}{|s-s'|^{1+2pr}}\, dsds'. \tag{13}$$

Then for any $a_0 > X_0^1$ there exists $R = R(p, r, a_0) \in (0, 1]$ such that

$$Y_t \le R \qquad \Rightarrow \qquad M_t^1 \le a_0. \tag{14}$$

Proof. We use Lemma 9 with

$$\rho = t, \qquad p(u) = u^{r+\frac{1}{2p}}, \qquad \Psi(u) = u^{2p}.$$

Note that since $r < \frac{p-1}{2p}$, Y_t is finite a.s.. If $Y_t \le R$, then Lemma 9 shows that

$$\begin{aligned} |W_s - W_{s'}| &\le CR^{\frac{1}{2p}} \int_0^{2|s-s'|} u^{-\frac{d}{p}}\, u^{r+\frac{1}{2p}-1}\, du \\ &= CR^{\frac{1}{2p}}|s-s'|^{r-\frac{2d-1}{2p}}, \end{aligned}$$

for any $s, s' \in (0, t]$. Since $W_0 = 0$ and $\frac{2d-1}{2p} < r$, we have

$$Y_t \le R \qquad \Rightarrow \qquad \sup_{0<s\le t} |W_s| \le CR^{\frac{1}{2p}}.$$

Recall that we can write $X_t = \varphi(W_t)$ where φ is introduced in (6). Hence if $Y_t \le R$ we have

$$\begin{aligned} |\, X_s^1 - X_0^1 \,| \le |\varphi_1(W_s) - \varphi_1(0)| &\le \sup_{|y|\le CR^{\frac{1}{2p}}} |\nabla\varphi_1(y)|\, |W_s| \\ &\le C \sup_{|y|\le C} |\nabla\varphi_1(y)|\, R^{\frac{1}{2p}}. \end{aligned}$$

The assumption (A1) implies that σ is bounded. Hence, $\sup_{|y|\le C}|\nabla\varphi_1(y)| = \sup_{|y|\le C}|\sigma(\varphi(y))| < \infty$. Therefore we can choose $R\in(0,1]$ small enough so that for any $s\in[0,t]$

$$X_s^1 \le X_0^1 + C\sup_{|y|\le C}|\nabla\varphi_1(y)|\,R^{\frac{1}{2p}} \le a_0.$$

This implies (14). ■

Note that Y_t is continuous and increasing. Hence there exists the inverse function $Y_t^{-1} := Y^{-1}(t)$.

Proposition 11 *With the notation of Lemma 10, we have that if* $\varepsilon>0$

$$\mathbf{P}(\,Y^{-1}(R)<\varepsilon\,)\le C\varepsilon^{2p}, \tag{15}$$

where C *depends on* R*,* p *and* d*.*

Proof. Let $q>1$ be the constant that satisfies $\frac{1}{p}+\frac{1}{q}=1$. By the Hölder's inequality, we have

$$\begin{aligned}\mathbf{P}(\,Y^{-1}(R)<\varepsilon\,) &= \mathbf{P}(\,R<Y_\varepsilon\,)\\ &\le R^{-p}\mathbf{E}\left[\left|\int_{[0,\varepsilon]^2}\frac{|W_s-W_{s'}|^{2p}}{|s-s'|^{1+2pr}}\,ds\,ds'\right|^p\right]\\ &\le R^{-p}\left[\int_{[0,\varepsilon]^2}ds\,ds'\right]^{\frac{p}{q}}\left[\int_{[0,\varepsilon]^2}\mathbf{E}\left[\frac{|W_s-W_{s'}|^{2p^2}}{|s-s'|^{(1+2pr)p}}\right]\,ds\,ds\right].\end{aligned}$$

The Hölder continuity of Brownian motion yields

$$\begin{aligned}\int_{[0,\varepsilon]^2}\mathbf{E}\left[\frac{|W_s-W_{s'}|^{2p^2}}{|s-s'|^{(1+2pr)p}}\right]\,ds\,ds' &\le C\varepsilon^2\sup_{s,s'\in[0,1]}|s-s'|^{p(p-(1+2pr))}\\ &\le C'\varepsilon^2.\end{aligned}$$

Therefore we obtain (15). ■

The following theorem is our main result in this section.

Theorem 12 *Let* $a_0>X_0^1$ *be fixed. There exists a process* v_t *such that*

(i) $v_t\in D^\infty$*;*

(ii) For any $f\in C_b^\infty(\mathbf{R}^d:\mathbf{R})$*, on the event* $\{M^1>a_0\}$ *it holds that*

$$(V_jf)(F)=\int_0^T D_r^j f(F)\,v_r\,dr.$$

Proof. We define

$$T_{a_0} = \inf\{t > 0\,;\, M_t^1 > a_0\},$$

and choose $R \in (0,1)$ such that (14) holds. Let $\rho : \mathbf{R}_+ \to [0,1]$ be a smooth function such that

$$\rho(x) = \begin{cases} 0 & \text{if } x > R \\ 1 & \text{if } 0 < x < \frac{R}{2}. \end{cases}$$

Set

$$u_t = \rho(Y_t),$$

and

$$H = \int_0^T u_s\, ds.$$

On the event $\{M^1 > a_0\}$ we have $T_{a_0} \leq \theta^1$, and further Lemma 10 implies $Y_t > R$ if $t > T_{a_0}$. Hence from the definition of ρ we have $u_t = 0$ if $t > T_{a_0}$. Thus we have

$$\begin{aligned} \int_0^T D_r^j f(F)\, u_r\, dr &= \sum_{k=1}^d \int_0^T \frac{\partial f}{\partial x_k}(F) D_r^j F_k\, u_r dr \\ &= (V_j f)(F) \int_0^T 1_{[0,\theta^1]}(r)\, u_r\, dr \\ &= (V_j f)(F)\, H, \end{aligned}$$

where we have used Theorem 7. By the definition of ρ we have

$$H \geq \int_0^T 1_{\{Y_s < \frac{R}{2}\}}\, ds = Y^{-1}(\frac{R}{2}).$$

Hence, Proposition 11 implies that $H^{-1} \in L^p(\Omega)$. Thus, on the event $\{M^1 > a_0\}$, we have

$$(V_j f)(F) = \int_0^T D_r^j f(F)\, v_r\, dr,$$

for $v_r = \frac{u_r}{H}$. ∎

4 Integration of the operator V_j and uniform estimates

Let $(\alpha_1, \cdots, \alpha_d) \in \mathbf{Z}_{\geq 0}^d$ be a multi-index, we denote by $|\alpha| = \alpha_1 + \cdots + \alpha_d$ and

$$\partial^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}.$$

Recall that φ is the solution to (6). For $\psi \in C_c^\infty(\mathbf{R}^d : \mathbf{R})$, since φ^{-1} is continuous, there exists $K > 0$ such that for any multi-index α

$$\mathrm{supp}((\partial^\alpha \psi) \circ \varphi) \subset \prod_{j=1}^d [-K, K].$$

Throughout this section, $\psi \in C_c^\infty(\mathbf{R}^d : \mathbf{R})$ and K will be fixed.

Lemma 13 *Under the notation above, for $j = 1, \ldots, d$, we define $h_j(\psi) : \mathbf{R^d} \to \mathbf{R}$ to be*

$$h_j(\psi)(y_1, \ldots, y_d) = \int_{-K}^{y_j} \psi \circ \varphi(y_1, \ldots, y_{j-1}, \theta, y_{j+1}, \ldots, y_d)\, d\theta, \tag{16}$$

where φ is the solution to (6). Then we have

$$V_j(h_j(\psi) \circ \varphi^{-1}) = \psi, \tag{17}$$

where V_j is the vector field defined in (A3).

Proof. Put $k_j = h_j(\psi) \circ \varphi^{-1}$. Since φ is a solution to (6) we have

$$\psi \circ \varphi(y) = \frac{\partial}{\partial y_j} h_j(\psi)(y) = \frac{\partial}{\partial y_j}(k_j \circ \varphi(y)) = \sum_{i=1}^d \frac{\partial k_j}{\partial x_i} \circ \varphi(y) \frac{\partial \varphi_i}{\partial y_j}(y) = (V_j k_j)(\varphi(y)).$$

This means (17). ■

For $g \in C_p^\infty(\mathbf{R}^d : \mathbf{R})$ and for a multi-index α we define

$$\begin{aligned} H_0(\alpha, g)(y) &= ((\partial^\alpha \psi)\ \cdot g) \circ \varphi(y), \\ H_k(\alpha, g)(y) &= h_d(h_{d-1}(\cdots(h_1(H_{k-1}(\alpha, g)))\cdots))(y) \\ &= \int_{-K}^{y_d} \cdots \int_{-K}^{y_1} H_{k-1}(\alpha, g)(\theta_1, \ldots, \theta_d)\, d\theta_1 \cdots d\theta_d, \end{aligned}$$

for $k = 1, 2, \ldots$.

Lemma 14 *Let α be a multi-index with $|\alpha| \geq 1$. Suppose that for some i, $\alpha_i \geq 1$, and put $\alpha' = (\alpha_1, \ldots, \alpha_{i-1}, \alpha_i - 1, \alpha_{i+1}, \ldots, \alpha_d)$. Then for any $g \in C_p^\infty(\mathbf{R}^d : \mathbf{R})$, there exist $g_{ij} \in C_p^\infty(\mathbf{R}^d : \mathbf{R})$ $(j = 1, 2, \ldots, d)$ such that*

$$\begin{aligned} &H_1(\alpha, g)(y) \\ &= \sum_{j=1}^d \int_{-K}^{y_d} \cdots \int_{-K}^{y_{j+1}} \int_{-K}^{y_{j-1}} \cdots \int_{-K}^{y_1} H_0(\alpha', g_{ij})(\bar{\theta}_j(y_j)) d\theta_1 \cdots d\theta_{j-1} d\theta_{j+1} \cdots d\theta_d \\ &\quad - \sum_{j=1}^d H_1(\alpha', g_{ij})(y), \end{aligned} \tag{18}$$

where we put $\bar{\theta}_j(y_j) = (\theta_1, \ldots, \theta_{j-1}, y_j, \theta_{j+1}, \ldots, \theta_d)$.

Proof. We shall prove by induction with respect to $n := |\alpha|$. For $n = 1$, we suppose without loss of generality that $(\alpha_1, \ldots, \alpha_d) = (1, 0, \ldots, 0)$. Since the Jacobi matrix of φ is $\sigma \circ \varphi$ we have

$$\frac{\partial \psi}{\partial x_1} \circ \varphi(y) = \sum_{j=1}^{d} \left(\frac{\partial \psi(\varphi(y))}{\partial y_j} \right) (\sigma^{-1})_1^j \circ \varphi(y).$$

Applying this formula and the integration by parts formula, we have

$$\begin{aligned} &H_1(\alpha, g)(y) \\ &= \int_{-K}^{y_d} \cdots \int_{-K}^{y_1} \left(\frac{\partial \psi}{\partial x_1} \cdot g \right) \circ \varphi(\theta_1, \ldots, \theta_d) d\theta_1 \cdots d\theta_d \\ &= \sum_{j=1}^{d} \int_{-K}^{y_d} \cdots \int_{-K}^{y_{j+1}} \int_{-K}^{y_{j-1}} \cdots \int_{-K}^{y_1} (\psi \cdot g_{1,j}) \circ \varphi(\bar{\theta}_j(y_j)) d\theta_1 \cdots d\theta_{j-1} d\theta_{j+1} \cdots d\theta_d \\ &- \sum_{j=1}^{d} \int_{-K}^{y_d} \cdots \int_{-K}^{y_1} (\psi \cdot g_{1j}) \circ \varphi(\theta) d\theta_1 \cdots d\theta_d, \end{aligned} \tag{19}$$

where we have put

$$g_{1j}(x) = \frac{\partial ((g \cdot (\sigma^{-1})_1^j) \circ \varphi)}{\partial y_j} (\varphi^{-1}(x)).$$

It follows from (A1) that σ^{-1} is bounded. Hence we have $g_{1j} \in C_p^\infty(\mathbf{R}^d : \mathbf{R})$. In view of the definition of H_0 and H_1, (19) is nothing but (18) for $n = 1$.

Let $n > 1$ be fixed and suppose that for any multi-index α with $|\alpha| \le n$ and for any $g \in C_p^\infty(\mathbf{R}^d : \mathbf{R})$ the equality (18) holds. Let $\alpha = (\alpha_1, \ldots, \alpha_d)$ with $|\alpha| = n + 1$. We may assume without loss of generality that $\alpha_1 \ge 1$. Put $\alpha' = (\alpha_1 - 1, \alpha_2, \ldots, \alpha_d)$. Then integration by parts formula shows

$$\begin{aligned} &H_1(\alpha, g)(y) \\ &= \int_{-K}^{y_d} \cdots \int_{-K}^{y_1} (\partial^\alpha \psi \cdot g) \circ \varphi(\theta) \, d\theta_1 \cdots d\theta_d \\ &= \sum_{j=1}^{d} \int_{-K}^{y_d} \cdots \int_{-K}^{y_{j+1}} \int_{-K}^{y_{j-1}} \cdots \int_{-K}^{y_1} (\partial^{\alpha'} \psi \cdot g_{1j}) \circ \varphi(\bar{\theta}_j(y_j)) d\theta_1 \cdots d\theta_{j-1} d\theta_{j+1} \cdots d\theta_d \\ &- \sum_{j=1}^{d} \int_{-K}^{y_d} \cdots \int_{-K}^{y_1} (\partial^{\alpha'} \psi \cdot g_{1j}) \circ \varphi(\theta) \, d\theta. \end{aligned}$$

Noting the definition of H_1 we obtain (18) for any multi-index. ■

Lemma 15 *Let α be a multi-index with $|\alpha| \ge 1$. Suppose that $\alpha_i \ge 1$, and put $\alpha' = (\alpha_1, \ldots, \alpha_{i-1}, \alpha_i - 1, \alpha_{i+1}, \ldots, \alpha_d)$. Then for any $g \in C_p^\infty(\mathbf{R}^d : \mathbf{R})$, there* *exist* $g_{ij} \in C_p^\infty(\mathbf{R}^d : \mathbf{R})$ $(j = 1, 2, \ldots, d)$ *such that for any* $k \geq 1$

$$\begin{aligned} &H_k(\alpha, g)(y) \\ &= \int_{-K}^{y_d} \cdots \int_{-K}^{y_{j+1}} \int_{-K}^{y_{j-1}} \cdots \int_{-K}^{y_1} H_{k-1}(\alpha', g_{ij})(\bar{\theta}_j(y_j)) d\theta_1 \cdots d\theta_{j-1} d\theta_{j+1} \cdots d\theta_d \\ &\quad - \sum_{j=1}^{d} H_k(\alpha', g_{ij})(\theta), \end{aligned} \tag{20}$$

where we have put $\bar{\theta}_j(y_j) = (\theta_1, \ldots, \theta_{j-1}, y_j, \theta_{j+1}, \ldots, \theta_d)$.

Proof. We prove the lemma by induction with respect to k. For the case $k = 1$ we have already proved in Lemma 14. Let $k > 1$ be fixed and suppose that (20) holds for k. Recall that H_{k+1} is defined by

$$H_{k+1}(\alpha, g)(y) := \int_{-K}^{y_d} \cdots \int_{-K}^{y_1} H_k(\alpha, g)(\theta_1, \ldots, \theta_d)\, d\theta_1 \cdots d\theta_d. \tag{21}$$

By the assumption of the induction we can replace the integrand in (21) by the right hand side in (20). Thus, using the Fubini's theorem and noting the definition of H_k, we obtain (20) for $k + 1$. ■

Lemma 16 *Let* n *be fixed. For any multi-index* α *with* $|\alpha| \leq n$ *and for any* $g \in C_p^\infty(\mathbf{R}^d : \mathbf{R})$*, we have*

$$|H_n(\alpha, g)(y)| \leq |\psi|_\infty \, |Q_{\alpha,g}(y)|, \tag{22}$$

where $Q_{\alpha,g}$ *is a polynomial function.*

Proof. We prove this by induction with respect to n. Note that $\partial^\alpha \psi = \psi$ if $|\alpha| = 0$. Thus, by the definition of $H_0(\alpha, g)$ we have

$$|H_0(\alpha, g)(y)| \leq |\psi|_\infty \, |g \circ \varphi(y)| \leq C|\psi|_\infty (1 + |\varphi(y)|^m),$$

for some $m \in \mathbf{N}$ since $g \in C_p^\infty(\mathbf{R}^d : \mathbf{R})$. Note that

$$|\varphi(y)| \leq |\varphi(0)| + |\varphi(y) - \varphi(0)| \leq |\varphi(0)| + \int_0^1 |\nabla \varphi(\rho\, y) \cdot y| \, d\rho \leq C(1 + |y|),$$

where in the second inequality we have used the mean value theorem and in the last inequality we have used the fact that φ is a solution to (6) and that σ is bounded. Thus we have (22) for $n = 0$.

Suppose that (22) holds for any multi-index α with $|\alpha| \leq n$. Let $\alpha = (\alpha_1, \ldots, \alpha_d)$ be a multi-index with $|\alpha| = n+1$. We assume that $\alpha_i \geq 1$ for some

i. Put $\alpha' = (\alpha_1, \ldots, \alpha_{i-1}, \alpha_i - 1, \alpha_{i+1}, \ldots, \alpha_d)$. By using Lemma 15, we have

$$\begin{aligned}
&|H_{n+1}(\alpha, g)(y)| \\
&\leq \int_{-K}^{y_d} \cdots \int_{-K}^{y_{j+1}} \int_{-K}^{y_{j-1}} \cdots \int_{-K}^{y_1} \left| H_n(\alpha', g_{ij})(\bar{\theta}_j(y_j)) \right| d\theta_1 \cdots d\theta_{j-1} d\theta_{j+1} \cdots d\theta_d \\
&\quad + \sum_{j=1}^{d} |H_{n+1}(\alpha', g_{ij})(\theta)|,
\end{aligned} \tag{23}$$

where $g_{ij} \in C_p^\infty(\mathbf{R}^d : \mathbf{R})$ $j = 1, 2, \ldots, d$. By the assumption of the induction the first term in the right hand side is dominated by

$$\int_{-K}^{y_d} \cdots \int_{-K}^{y_{j+1}} \int_{-K}^{y_{j-1}} \cdots \int_{-K}^{y_1} |\psi|_\infty \, |Q_{\alpha', g}(\theta)| \, d\theta_1 \cdots d\theta_{j-1} d\theta_{j+1} \cdots d\theta_d.$$

It follows from the definition of $H_n(\alpha, g)$ the second term in (23) can be estimated as

$$\begin{aligned}
\sum_{j=1}^{d} |H_{n+1}(\alpha', g_{ij})(y)| &\leq \sum_{j=1}^{d} \int_{-K}^{y_d} \cdots \int_{-K}^{y_1} |H_n(\alpha', g_{ij})(\theta_1, \ldots, \theta_d)| \, d\theta_1 \cdots d\theta_d \\
&\leq \sum_{j=1}^{d} \int_{-K}^{y_d} \cdots \int_{-K}^{y_1} |\psi|_\infty \left| Q_{\alpha', g_{ij}}(\theta) \right| d\theta_1 \cdots d\theta_d,
\end{aligned}$$

where in the last inequality we have used the assumption of the induction. Thus we have (22). ■

Define

$$f_n(x) = H_n(\alpha, 1) \circ \varphi^{-1}(x). \tag{24}$$

Then we have

Theorem 17 *Let ψ be a smooth function with compact support. Suppose that*

$$\mathrm{supp}(\psi) \subset [a_0, \infty) \times \mathbf{R}^{d-1},$$

and (5) *holds. For any multi-index $(\alpha_1, \ldots, \alpha_d)$, there exist functions $f_1, \ldots, f_{|\alpha|}$ such that*

1. $$f_0(x) = \partial^\alpha \psi(x),$$

   and

   $$(V_1 \cdots V_d) f_n = f_{n-1} \tag{25}$$

2. There exists a polynomial Q such that

$$|f_{|\alpha|}(x)| \leq |\psi|_\infty | \, Q \circ \varphi^{-1}(x)|. \tag{26}$$

3. For any $m \leq n$

$$f_m(x) = 0 \qquad \text{if } x_1 \leq a_0. \tag{27}$$

Proof. Let f_m $(m = 0, 1, 2 \ldots n)$ be the functions defined in (24). Recall that φ is a diffeomorphism. Note that it follows from the definition of H_n, (16) and (17) that

$$(V_1 \cdots V_d)^n f_n(x) = \partial^\alpha \psi(x).$$

Hence, from Lemma 13 and Lemma 16, the functions $\{f_m\}$ satisfy (25) and (26). Thus all we have to do is to show that the functions f_m satisfy (27). Define

$$a(y_2, \ldots, y_d) = \inf\{\theta \, ; \, \varphi_1(\theta, y_2, \ldots, y_d) > a_0\}.$$

Fix $x = (x_1, \ldots, x_d)$ with $x_1 < a_0$, we can find $y = (y_1, \ldots, y_d)$ such that $x = \varphi(y)$ since $\varphi(y)$ is a diffeomorphism. From (5), we have

$$\frac{\partial \varphi_1}{\partial \theta}(\theta, y_2, \ldots, y_d) = \sigma_1^1(\varphi(\theta, y_2, \ldots, y_d)) > 0, \tag{28}$$

that is, $\theta \to \varphi_1(\theta, y_2, \ldots, y_d)$ is strictly increasing for fixed $(y_2, \ldots, y_d)$. Hence, we have $y_1 \leq a(y_2, \ldots, y_d)$. Recall that f_n can be written as follows

$$f_n(x) = H_n(\alpha, 1)(y) = \int_{-K}^{y_d} \cdots \int_{-K}^{y_1} H_{n-1}(\alpha, 1)(\theta) d\theta.$$

By induction one can show that the integrand in the above multiple integral vanishes since $y_1 \leq a(y_2, \ldots, y_d)$. ■

5 Proof of Theorem 2

In this section, we shall prove Theorem 2. Let $\psi \in C_c^\infty(\mathbf{R}^d : \mathbf{R})$ be fixed and suppose that supp $(\psi) \subset [a_0, \infty) \times \mathbf{R}^{d-1}$ where $a_0 > X_0^1$. Let $(\alpha_1, \ldots, \alpha_d)$ be a multi-index with $\alpha_1 + \cdots + \alpha_d = n$. All we have to do is to show that for any $G \in D^\infty$

$$\left| \mathbf{E} \left[\frac{\partial^n \psi}{\partial x^{\alpha_1} \cdots \partial x^{\alpha_d}}(F) G \right] \right| \leq C |\psi|_\infty, \tag{29}$$

which will imply the smoothness of the law of $F = (F_1, \ldots, F_d)$ on $(a_0, \infty) \times \mathbf{R}^{d-1}$ (see Lemma 2.1.5 in Nualart [9]).

In view of Theorem 17, we can find $f_0, f_1, \ldots, f_n$ that satisfy (25), (26) and (27). Since $\operatorname{supp}(\psi) \subset [a_0, \infty)$ it follows from (25), Theorem 12 and duality that for any $G \in D^\infty$

$$\begin{aligned}\mathbf{E}\left[\frac{\partial^n \psi}{\partial x^{\alpha_1} \cdots \partial x^{\alpha_d}}(F) G\right] &= \mathbf{E}[(V_1 \cdots V_d) f_1(F) G] \\ &= \mathbf{E}\left[\int_0^T D_r^1 (V_2 \cdots V_d) f_1(F)\, v_r\, dr\, G\right] \\ &= \mathbf{E}[(V_2 \cdots V_d) f_1(F) \delta^1(v.G)].\end{aligned}$$

By iterating the same argument we have

$$\mathbf{E}\left[\frac{\partial^n \psi}{\partial x^{\alpha_1} \cdots \partial x^{\alpha_d}}(F) G\right] = \mathbf{E}\left[f_1(F) \delta^d(v.\delta^{d-1}(v. \cdots \delta^1(v.G))))\right] = \mathbf{E}[f_1(F) \delta(v \cdot G)].$$

Here $\delta(v \cdot G) = \delta^d(v....\delta^1(v.G)...)$. From Theorem 17 we have $f_2(x) = 0$ provided $x_1 < a_0$ and $f_1 = (V_1 \cdots V_d) f_2$, hence we can apply this argument recursively:

$$\begin{aligned}\mathbf{E}\left[\frac{\partial^n \psi}{\partial x^{\alpha_1} \cdots \partial x^{\alpha_d}}(F) G\right] &= \mathbf{E}\left[f_1(F) \delta(v \cdot G)\right] \\ &= \mathbf{E}[(V_1 \cdots V_d) f_2(F) \delta(v \cdot G)] \\ &\vdots \\ &= \mathbf{E}[f_n(F) \delta(v \cdot \delta(v \cdot \cdots \delta(v \cdot \delta(v \cdot G))))].\end{aligned}$$

It follows from (26) that

$$\begin{aligned}\left|\mathbf{E}\left[\frac{\partial^n \psi}{\partial x^{\alpha_1} \cdots \partial x^{\alpha_d}}(F) G\right]\right| &\leq \|f_n(F)\|_{\mathbf{L}^2(\Omega)} \, \|\delta(v.\delta(v. \cdots \delta(v.\delta(v.G))))]\|_{\mathbf{L}^2(\Omega)} \\ &\leq C |\psi|_\infty.\end{aligned}$$

Thus we obtain (29).

6 S.D.E. with commutative vector fields and with drift

Let $(\Omega', \mathcal{F}', \mathbf{Q}), \{\mathcal{F}'_t\}_{t \in [0,T]}, B_t = (B_t^1, \ldots, B_t^d)$ be a d-dimensional standard Brownian motion. In this section, we shall consider the following d-dimensional stochastic differential equation that has the drift term:

$$\begin{cases} dZ_t^i &= \sum_{j=1}^d \sigma_j^i(Z_{s-}) \circ dB_s^j + b^i(Z_s)\, ds \\ Z_0^i &= z^i. \end{cases} \tag{30}$$

We assume that Assumption A, (5) and (4) hold. Further in this section we assume

Assumption B $b : \mathbf{R}^d \to \mathbf{R}^d$ is bounded smooth function with bounded derivatives.

Since σ is uniformly non-degenerate, we can define

$$\gamma(x) := \sigma^{-1} b(x).$$

Note that γ is a bounded smooth function, and hence

$$K_t = 1 - \sum_{j=1}^{d} \int_0^t K_s \gamma_j(X_s)\, dB_s^j$$

is a martingale. Define

$$\frac{d\mathbf{Q}'}{d\mathbf{Q}}\bigg|_{\mathcal{F}_T} = K_T,$$

then Girsanov's theorem shows that, under $\mathbf{Q}'$,

$$\hat{B}_t^j = B_t^j - \langle B^j, K\rangle_t = B_t^j + \int_0^t \gamma_j(X_s)\, ds$$

is a Brownian motion. Note that X can be written as

$$dX_t^i = \sum_{j=1}^{d} \sigma_j^i(X_{s-}) \circ dB_s^j + b^i(X_s)\, ds = \sigma_j^i(X_{s-}) \circ d\hat{B}_s^j.$$

Now let $(\Omega, \mathcal{F}, \mathbf{P}), \{\mathcal{F}_t\}_{t\in[0,T]}, W_t = (W_t^1, \dots, W_t^d)$ be another d-dimensional standard Brownian motion and consider the following stochastic differential equation

$$\begin{cases} dX_t^i &= \sum_{j=1}^d \sigma_j^i(X_s) \circ dW_s^j \\ X_0^i &= z^i. \end{cases} \tag{31}$$

By putting

$$G_T = \exp\left\{ \sum_{j=1}^{d} \int_0^T \gamma_j(X_s) dW_s^j - \frac{1}{2}\int_0^T \sum_{j=1}^{d} \gamma_j^2(X_s)\, ds \right\},$$

we have, for any $\psi \in C_c^\infty((X_0^1, \infty) \times \mathbf{R}^{d-1} : \mathbf{R})$ and for any $n \in \mathbf{N}$,

$$\begin{aligned} \left| \mathbf{E_Q}\left[\frac{\partial^n \psi}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} (\sup_{0\le s\le T} Z_s) \right] \right| &= \left| \mathbf{E_{Q'}}\left[\frac{\partial^n \psi}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} (\sup_{0\le s\le T} Z_s)\, K_T^{-1} \right] \right| \\ &= \left| \mathbf{E_P}\left[\frac{\partial^n \psi}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} (\sup_{0\le s\le T} X_s)\, G_T \right] \right| \\ &\le C|\psi|_\infty, \end{aligned}$$

where we denote by $\mathbf{E_Q}$ the mathematical expectation with respect to the probability measure $\mathbf{Q}$, and in the last inequality we have used (29).

7 Appendix: The time when X_t^1 attains its maximum is unique

In this section, we shall prove Proposition 1. We assume that Assumption (A) holds throughout this section. First of all we shall prove that X_t^1 does not attain its maximum on $[0,T]$ at the terminal time T:

Proposition 18 *Suppose that* (4) *holds. Then we have*

$$\mathbf{P}(X_T^1 = \sup_{0\le s\le T} X_s^1) = 0.$$

For the proof of Proposition 18, we need the law of the iterated logarithm:

Lemma 19 *Let X_t be the solution to* (3)*. We have*

$$\mathbf{P}\left(\limsup_{t\to T} \frac{X_T^1 - X_t^1}{\sqrt{(T-t)\log|\log(T-t)|}} = \sum_{j=1}^d \sigma_j^1(X_T)\right) = 1 \tag{32}$$

$$\mathbf{P}\left(\liminf_{t\to T} \frac{X_T^1 - X_t^1}{\sqrt{(T-t)\log|\log(T-t)|}} = -\sum_{j=1}^d \sigma_j^1(X_T)\right) = 1. \tag{33}$$

Remark 20 *On the event $\{\sum_{j=1}^d \sigma_j^1(X_T) < 0\}$, the limit superior is strictly negative, and the limit inferior is strictly positive. This is a contradiction. Hence we have $\sum_{j=1}^d \sigma_j^1(X_T) \ge 0$ $\mathbf{P}$-a.s..*

Proof. By the law of the iterated logarithm for Brownian motion (see Theorem 9.23 in Chapter 2 of [7]) we have

$$\mathbf{P}\left(\limsup_{t\to T} \frac{W_T^j - W_t^j}{\sqrt{(T-t)\log|\log(T-t)|}} = 1, 1 \le {}^\forall j \le d\right) = 1. \tag{34}$$

Since by (7), X_t^1 can be written as $X_t^1 = \varphi_1(W_t)$, the mean value theorem and (6) shows that

$$\begin{aligned}
&\limsup_{t\to T} \frac{X_T^1 - X_t^1}{\sqrt{(T-t)\log|\log(T-t)|}} \\
&= \limsup_{t\to T} \sum_{j=1}^d \int_0^1 \frac{\partial\varphi_1}{\partial x_j}(W_t + \theta(W_T - W_t))\, d\theta \; \frac{W_T^j - W_t^j}{\sqrt{(T-t)\log|\log(T-t)|}} \\
&= \limsup_{t\to T} \sum_{j=1}^d \int_0^1 \sigma_j^1 \circ \varphi(W_t + \theta(W_T - W_t))\, d\theta \; \frac{W_T^j - W_t^j}{\sqrt{(T-t)\log|\log(T-t)|}} \\
&= \sum_{j=1}^d \sigma_j^1(X_T) \qquad \mathbf{P}\text{-a.s.,}
\end{aligned}$$

where in the last equality we have used (34) and the dominated convergence theorem. Similarly one can prove (33). ■

Proof of Proposition 18. As we mentioned in Remark 20, $\sum_{j=1}^d \sigma_j^1(X_T) \geq 0$ **P**-a.s. Hence, for any $\varepsilon > 0$, we have

$$\begin{aligned} \mathbf{P}\left(X_T^1 = \sup_{0\leq s\leq T} X_s^1\right) \leq & \mathbf{P}\left(X_T^1 - \sup_{0\leq s\leq T} X_s^1 = 0, \sum_{j=1}^d \sigma_j^1(X_T) > \epsilon\right) \\ & + \mathbf{P}\left(0 \leq \sum_{j=1}^d \sigma_j^1(X_T) \leq \epsilon\right). \end{aligned} \tag{35}$$

It follows from Lemma 19 that

$$\begin{aligned} & \mathbf{P}\left(X_T^1 - \sup_{0\leq s\leq T} X_s^1 = 0, \sum_{j=1}^d \sigma_j^1(X_T) > \epsilon\right) \\ & \leq \mathbf{P}\left(X_T^1 - X_s^1 \geq 0 \ {}^\forall s \in [0,T], \liminf_{t\to T} \frac{X_T^1 - X_t^1}{\sqrt{(T-t)\log|\log(T-t)|}} \leq -\varepsilon\right) \\ & = 0. \end{aligned}$$

Therefore for any $\varepsilon > 0$ we have

$$\mathbf{P}\left(X_T^1 = \sup_{0\leq s\leq T} X_s^1\right) \leq \mathbf{P}\left(0 \leq \sum_{j=1}^d \sigma_j^1(X_T) \leq \epsilon\right).$$

Letting ε tend to zero, we have

$$\mathbf{P}\left(X_T^1 = \sup_{0\leq s\leq T} X_s^1\right) \leq \mathbf{P}\left(\sum_{j=1}^d \sigma_j^1(X_T) = 0\right) = 0.$$

The last equality follows from the assumption (4) and the fact that the law of X_T has the smooth density. ■

In view of Girsanov's theorem, it is enough to show Proposition 1 for the following stochastic differential equation

$$X_t^1 = x_1^0 + \sum_{j=1}^d \int_0^t \sigma_j^1(X_s)\, dW_s^j. \tag{36}$$

We define

$$\begin{aligned} \bar{\theta}_t &= \sup\{s \leq t\,;\, X_s^1 = \sup_{0\leq r\leq t} X_r^1\} \\ \underline{\theta}_t &= \inf\{s \leq t\,;\, X_s^1 = \sup_{0\leq r\leq t} X_r^1\}, \end{aligned}$$

and we shall show that

$$\mathbf{P}(\underline{\theta}_T < \bar{\theta}_T < T) = 0. \tag{37}$$

Note that Proposition 18 yields $\bar{\theta}_T < T$ **P**-a.s., and obviously $\bar{\theta}_t \geq \underline{\theta}_t$ **P**-a.s.. Hence (37) is equivalent to the assertion of Proposition 1.

From (A1) we have $\inf_x a_1^1(x) > 0$ thus the quadratic variation of X^1

$$\langle X^1 \rangle_t = \int_0^t a_1^1(X_s) ds$$

is a strictly increasing continuous process. Hence,

$$B_t = X^1_{\langle X^1 \rangle_t^{-1}}$$

is a $\mathcal{F}_t$-adapted 1-dimensional Brownian motion. We can write

$$X_t^1 = B_{\langle X^1 \rangle_t},$$

and

$$\begin{aligned} \bar{\theta}_T &= \sup \left\{ s \leq T \, ; \, B_{\langle X^1 \rangle_s} = \sup_{0 \leq u \leq T} B_{\langle X^1 \rangle_u} \right\}, \\ \underline{\theta}_T &= \inf \left\{ s \leq T \, ; \, B_{\langle X^1 \rangle_s} = \sup_{0 \leq u \leq T} B_{\langle X^1 \rangle_u} \right\}. \end{aligned}$$

Define

$$\begin{aligned} \bar{\tau}_t &= \sup \left\{ r \leq t \, ; \, B_r = \sup_{0 \leq s \leq t} B_s \right\}, \\ \underline{\tau}_t &= \inf \left\{ r \leq t \, ; \, B_r = \sup_{0 \leq s \leq t} B_s \right\}. \end{aligned} \tag{38}$$

Then we have

$$\bar{\theta}_T = \langle X^1 \rangle^{-1}_{\bar{\tau}_{\langle X^1 \rangle_T}}, \qquad \underline{\theta}_T = \langle X^1 \rangle^{-1}_{\underline{\tau}_{\langle X^1 \rangle_T}}. \tag{39}$$

Proof of Proposition 1 .
We shall show (37). Since $\langle X^1 \rangle_t^{-1}$ is strictly increasing, (39) shows

$$\begin{aligned} \mathbf{P}(\underline{\theta}_T < \bar{\theta}_T < T) &= \mathbf{P}(\underline{\tau}_{\langle X^1 \rangle_T} < \bar{\tau}_{\langle X^1 \rangle_T} < \langle X^1 \rangle_T) \\ &= \mathbf{P}\Big(\bigcup_{\substack{r_1, r_2 \in \mathbf{Q} \\ \mathbf{r_1 < r_2}}} \{ \underline{\tau}_{\langle X^1 \rangle_T} < r_1 < \bar{\tau}_{\langle X^1 \rangle_T} < r_2 < \langle X^1 \rangle_T \} \Big) \\ &\leq \sum_{\substack{r_1, r_2 \in \mathbf{Q} \\ \mathbf{r_1 < r_2}}} \mathbf{P}(\underline{\tau}_{\langle X^1 \rangle_T} < r_1 < \bar{\tau}_{\langle X^1 \rangle_T} < r_2 < \langle X^1 \rangle_T). \end{aligned} \tag{40}$$

On the event $\{\underline{\tau}_{\langle X^1\rangle_T} < r_1 < \bar{\tau}_{\langle X^1\rangle_T} < r_2 < \langle X^1\rangle_T\}$, the definition of $\bar{\tau}$ shows

$$\bar{\tau}_{\langle X^1\rangle_T} = \sup\{s \le r_2\,;\, B_s = \sup_{0\le r\le r_2} B_r\}. \tag{41}$$

The definition of $\underline{\tau}$ also shows

$$\underline{\tau}_{\langle X^1\rangle_T} = \inf\{s \le r_1\,;\, B_s = \sup_{0\le r\le r_1} B_r\}. \tag{42}$$

Since $B_{\bar{\tau}_{\langle X^1\rangle_T}} = B_{\underline{\tau}_{\langle X^1\rangle_T}}$ holds $\mathbf{P}$-a.s., (41) and (42) implies

$$\sup_{0\le r\le r_2} B_r = \sup_{0\le r\le r_1} B_r.$$

Thus we have

$$\begin{aligned}&\mathbf{P}(\underline{\tau}_{\langle X^1\rangle_T} < r_1 < \bar{\tau}_{\langle X^1\rangle_T} < r_2 < \langle X^1\rangle_T)\\ &\le \mathbf{P}(\sup\{s \le r_2\,;\, B_s = \sup_{0\le r\le r_2} B_r\} \in (r_1, r_2),\ \sup_{0\le r\le r_2} B_r = \sup_{0\le r\le r_1} B_r) = 0,\end{aligned}$$

where the last equality follows from Proposition 4 in Section VI of Bertoin [1]. Therefore the right hand side of (40) equals to zero. ■

At first glance one may think that the above proof may be extended by using the uniqueness of the maximum process in the open interval $[0, < X^1 >_T +\varepsilon)$ and then taking $\varepsilon \to 0$ but this arguments fails because this argument assumes a priori that Proposition 18 is satisfied.

References

[1] Bertoin, Jean, Levy processes. Cambridge Tracts in Mathematics, 121. Cambridge University Press, Cambridge, 1996

[2] Bouleau, Nicolas and Hirsch, Francis, Dirichlet forms and analysis on Wiener space. de Gruyter Studies in Mathematics, 14. Walter de Gruyter & Co., Berlin, 1991.

[3] Doss, Halim, Liens entre équations différentielles stochastiques et ordinaires. Ann. Inst. H. Poincare Sect. B (N.S) 13 (1977), no. 2, 99–125.

[4] Detemple, Jerome, Garcia, Rene and Rindisbacher, Marcel, Representation formulas for Malliavin derivatives of diffusion processes. Finance Stoch. 9 (2005), no. 3, 349–367.

[5] Florit, Carme, and Nualart, David, A local criterion for smoothness of densities and application to the supremum of the Brownian sheet. Statist. Probab. Lett. 22 (1995), no. 1, 25–31.

[6] Gobet, Emmanuel and Kohatsu-Higa, Arturo, Computation of Greeks for barrier and look-back options using Malliavin calculus. Electron. Comm. Probab. 8 (2003), 51?62

[7] Karatzas, Ioannis and Shreve, Steven E. Brownian motion and stochastic calculus. Second edition. Graduate Texts in Mathematics, 113. Springer-Verlag, New York, 1991.

[8] N. Lanjri Zadi and D. Nualart, Smoothness of the law of the supremum of the fractional Brownian motion. Electron. Comm. Probab. 8 (2003), 102–111.

[9] Nualart, David, The Malliavin calculus and related topics. Probability and its Applications (New York). Springer-Verlag, Berlin, 1995.

[10] Nualart, David, The Malliavin calculus and related topics. Second edition. Probability and its Applications (New York). Springer-Verlag, Berlin, 2006.

[11] Nualart, David and Vives, Josep, Continuité absolue de la loi du maximum d'un processus continu. C. R. Acad. Sci. Paris Ser. I Math. 307 (1988), no. 7, 349–354.

[12] Plastock, Roy, Homeomorphisms between Banach spaces. Trans. Amer. Math. Soc. 200 (1974), 169–183.