ELSEVIER
Bull. Sci. math. 134 (2010) 826–856
BULLETIN DES SCIENCES MATHÉMATIQUES
www.elsevier.com/locate/bulsci

# Support theorem for stochastic variational inequalities ☆

Jiagang Ren *, Siyan Xu

*School of Mathematics and Computational Science, Sun Yat-Sen University, Guangzhou, Guangdong 510275, PR China*

Received 26 April 2010
Available online 23 May 2010

**Abstract**

We prove a support theorem of the type of Stroock–Varadhan for solutions of stochastic variational inequalities.


*MSC:* 60H10; 60F99; 47H04

*Keywords:* Stochastic variational inequality; Multivalued maximal monotone operator; Limit theorems; Approximate continuity; Support theorem

## 1. Introduction and main result

The aim of this paper is to describe the support of the law of the solution of the following stochastic variational inequality (SVI in short):

$$\begin{cases} dX(t) \in b\big(X(t)\big)\,dt + \sigma\big(X(t)\big)\circ dw(t) - \partial\varphi\big(X(t)\big)\,dt, & t \in \mathbb{R}_+, \\ X(0) = x \in \overline{D(\partial\varphi)}, \end{cases} \tag{1}$$

where $\varphi$ is a convex function and $\partial\varphi$ is its subdifferential. SVIs of this type have been investigated by many authors in the past two decades (see e.g. [3,4,7–9] and reference therein) and they include as a special case stochastic differential equations (SDE in short) in convex domains reflected at the boundaries. Also it is needless to say that they reduce to usual SDEs if $\varphi$ is dif-

☆ Work supported by NSFC (Grant No. 10871215), Ph.D. Programs Foundation of Ministry of Education of China (Grant No. 20060558051) and China postdoctoral science foundation (Grant No. 20090460824).
* Corresponding author.
*E-mail addresses:* renjg@mail.sysu.edu.cn (J. Ren), xsy_00@hotmail.com (S. Xu).

0007-4497/$ – see front matter 
doi:10.1016/j.bulsci.2010.05.001

ferentiable. For connection of SVIs with parabolic and elliptic Neumann problems and parabolic variational inequalities we refer to [3,15].

The support of the law of a diffusion defined by an SDE was first characterized by Stroock and Varadhan in [18] and this result has undergone various extensions ever since, among which for the most recent ones we mention only [14]. For reflected diffusions in *smooth* ($C^2$) domains this was done by Doss and Priouret in [10]. Their approach, however, involves a heavy localization procedure which does not seem applicable to convex domains with only Lipschitz boundaries.

As was originally done by Stroock and Varadhan and is now a standard approach, it consists of proving two inclusion relations to characterize the support of a diffusion: the direct one and the inverse one. The direct inclusion involves essentially a limit theorem for the equation. Such theorem was proved in [10] for reflected diffusions in *smooth domains* and in [16] for reflected diffusions in convex domains but with *constant diffusion coefficients* and some other extra assumptions which are not easy to verify. In Section 3, we shall prove a limit theorem for general SVIs in the form of (1). Here we would like to point out that compared with our recent work [17], the main difference is that in that paper we approximate SVIs by ODEs while in the present one we approximate SVIs by ordinary variational inequalities. Each one has its advantage and its disadvantage: the former is good for establishing various regularity properties for solutions of SVIs and the latter is adequate for determining the support.

The inverse inclusion is deduced from the Denjoy approximate continuity of solutions of SVIs at sufficiently regular sample paths. For SDEs with *smooth* reflecting boundaries this was proved in [10] and for multivalued SDEs with *bounded* multivalued maximal monotone operators this was proved recently in [19]. This boundedness assumption is, however, so strong that it even excludes the case of reflected diffusions. In the present paper we shall be able to remove this assumption, see Theorem 5.10.

Combining the two inclusions will yield the main result of the paper, Theorem 3.2.

The paper is organized as follows: in Section 2 we prepare necessary preliminary materials and we state our main result in Section 3. Section 4 and Section 5 will be devoted to the proof of the main result.

## 2. Preliminaries

In this section we collect some materials which will be needed below.

$\mathcal{T}$ will be the space of càdlàg real-valued functions of finite variation defined on $\mathbb{R}_+$ with the metric

$$d(f,g) := \int_0^\infty \frac{|f(t)-g(t)|}{1+|f(t)-g(t)|} e^{-t}\,dt.$$

This metric obviously corresponds to convergence in Lebesgue measure and is also equivalent to the convergence at each point of continuity of the target function. It is trivial that $(\mathcal{T},d)$ is separable and it is an easy consequence of the standard diagonalization argument (see e.g. [6, pp. 210–212]) that it is also complete. Hence it is a Polish space. Furthermore, we have (see e.g. [6, Lemma 13.15])

**Lemma 2.1.** *For any increasing positive function $C(t)$, the set*

$$\{f\colon |f|_t \leqslant C(t),\ \forall t\}$$

*is compact in $(\mathcal{T},d)$, where $|f|_t$ stands for the total variation of $f$ on $[0,t]$.*

For $\kappa \in \mathcal{T}$ which is increasing, define its inverse by

$$\kappa^{-1}(t) := \inf\{s\colon \kappa(s) > t\}.$$

Then it is easy to see that

$$\left(\kappa^{-1}\right)^{-1} = \kappa.$$

Let $\mathscr{W}^m := C(\mathbb{R}_+, \mathbb{R}^m)$ be the space of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^m$, endowed with compact uniform convergence topology.

We will need the following lemmas.

**Lemma 2.2.** *Let, for each $n$, $y_n \in \mathscr{W}^m$, $y \in \mathscr{W}^m$ and $\theta_n$ an increasing càdlàg function from $\mathbb{R}_+$ to itself with $\theta_n(0) = 0$ and $\sup_n \theta_n(t) < \infty$ for all $t$. Suppose $y_n \to y$ in $\mathscr{W}^m$. Define $x_n(t) := y_n(\theta_n(t))$. Then there exist a subsequence $\{n_k\}$ and an increasing càdlàg function $\theta$ such that $\theta_{n_k}$ converges to $\theta$ at each continuity point of $\theta$ and $x_{n_k}(t)$ converges to $x(t)$ for all but countably many $t$ where $x(t) := y(\theta(t))$. If furthermore $y(u) = y(v)$ provided that $\theta^{-1}(u) = \theta^{-1}(v)$ (or, equivalently, $y(\theta(t)) = y(\theta(t-))$ for all $t$), then $x \in \mathscr{W}^m$ and $x_{n_k} \to x$ in $\mathscr{W}^m$.*

**Proof.** Since $\sup_n \theta_n(t) < \infty$ for all $t$, it follows from Lemma 2.1 that $\{\theta_n\}$ is compact in $\mathcal{T}$. Consequently the theorem follows directly from (the proof of) [13, Lemma 2.3]. (Notice that in that lemma $y_n$ are càdlàg functions and the condition "strictly increasing" is required, but it is clear from the proof that if we are restricted to the continuous function space $\mathscr{W}^m$, the strictness can be dropped.) □

**Lemma 2.3.** *If $\theta_n(t) \to \theta(t)$ at each continuity point of $\theta$, and $\theta$ and $\theta_n$ are strictly increasing, then $\theta_n^{-1} \to \theta^{-1}$ in $C(\mathbb{R}_+, \mathbb{R}_+)$.*

**Proof.** We first prove $\theta_n^{-1} \to \theta^{-1}$ pointwisely. Let $\theta^{-1}(t) = s$. By the strict increasingness of $\theta$ we have $\theta(u) > t$ for $u > s$ and $\theta(u) < t$ for $u < s$. Now for any $\varepsilon > 0$, choose two points $s - \varepsilon < s_1 < s < s_2 < s + \varepsilon$ of continuity of $\theta$. Then for large $n$ we have

$$\theta_n(s_1) < t < \theta_n(s_2)$$

which implies

$$s - \varepsilon \leqslant \theta_n^{-1}(t) \leqslant s + \varepsilon.$$

Now we prove the convergence is in fact uniform on each finite interval. For simplicity we do this for [0, 1]. For every $\varepsilon > 0$, chose an $m$ such that

$$\theta^{-1}\left(k2^{-m}\right) - \theta^{-1}\left((k-1)2^{-m}\right) < \varepsilon, \quad \forall k = 1, 2, \ldots, m.$$

For large $n$ we have

$$\left|\theta_n^{-1}\left(k2^{-m}\right) - \theta^{-1}\left(k2^{-m}\right)\right| < \varepsilon, \quad \forall k = 0, 1, \ldots, m.$$

Then, since $\theta_n$ is increasing we have for all $t \in [0, 1]$

$$\left|\theta_n^{-1}(t) - \theta^{-1}(t)\right| \leqslant 6\varepsilon.$$

This completes the proof. □

Given a multivalued operator $A$ from $\mathbb{R}^m$ to $2^{\mathbb{R}^m}$, define:

$$D(A) := \{x \in \mathbb{R}^m\colon\ A(x) \neq \emptyset\},$$
$$Im(A) := \bigcup_{x \in D(A)} A(x),$$
$$Gr(A) := \{(x, y) \in \mathbb{R}^{2m}\colon\ x \in \mathbb{R}^m,\ y \in A(x)\}.$$

$A^{-1}$ is defined by: $x \in A^{-1}(y) \Leftrightarrow y \in A(x)$.

A maximal monotone operator $A$ is a multivalued operator satisfying the following conditions:
(i) Monotonicity:

$$\langle y_1 - y_2, x_1 - x_2 \rangle \geqslant 0, \quad \forall (x_1, y_1), (x_2, y_2) \in Gr(A).$$

(ii) Maximality:

$$(x_1, y_1) \in Gr(A) \quad \Leftrightarrow \quad \{\langle y_1 - y_2, x_1 - x_2 \rangle \geqslant 0,\ \forall (x_2, y_2) \in Gr(A)\}.$$

Then we have (see [1] or [5]).

**Proposition 2.4.**

(1) *For each $x \in D(A)$, $A(x)$ is a closed and convex subset of $\mathbb{R}^m$. In particular, there is a unique $y \in A(x)$ such that $|y| = \inf\{|z|\colon\ z \in Ax\}$. $A^\circ(x) := y$ is called the minimal section of $A$, and we have*

$$x \in D(A) \quad \Leftrightarrow \quad |A^\circ(x)| < +\infty.$$

(2) *The resolvent operator $J_n := (1 + \frac{1}{n}A)^{-1}$ is single-valued and Lipschitz continuous with Lipschitz constant* 1. *Moreover, $\lim_{n\uparrow\infty} J_n x = x$ for any $x \in D(A)$.*
(3) *The Yosida approximation $A_n := n(1 - J_n)$ is monotone and Lipschitz continuous with Lipschitz constant $n$. Moreover, as $n \uparrow \infty$*

$$A_n(x) \to A^\circ(x) \quad \text{and} \quad |A_n(x)| \uparrow |A^\circ(x)| \quad \text{if } x \in D(A).$$

The following lemma which will be needed is proved in [17].

**Lemma 2.5.** *If $x \notin D(A)$, $x_n \to x$, then*

$$\liminf_{n\to\infty} |A_n(x_n)| = \infty.$$

We give the following definition for convenience.

**Definition 2.6.** Let $F, G$ be two continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^m$ and suppose furthermore that $F(t) \in \overline{D(A)}$ for all $t$ and $G$ is of finite variation. We say that $dG(t) \in A(F(t))\,dt$ if for every pair of continuous functions $(\alpha, \beta)$ satisfying

$$(\alpha(t), \beta(t)) \in Gr(A),$$

we have

$$\langle F(t) - \alpha(t), dG(t) - \beta(t)dt \rangle \geqslant 0.$$

Then we have the following result due to [8].

**Lemma 2.7.** *There exist $a \in \mathbb{R}^m$, $c_1 > 0$, $c_2 \geqslant 0$, such that if $dG(t) \in A(F(t))\,dt$, then for all $0 \leqslant s \leqslant t < \infty$, we have*

$$\int_s^t \langle F(u) - a, dG(u) \rangle \geqslant c_1\big(|G|_t - |G|_s\big) - c_2 \int_s^t \big|F(u) - a\big|\,du - c_1 c_2 (t - s).$$

Natural and important examples of maximal monotone operators are subdifferentials of convex functions. More precisely, Let $\varphi$ be a proper convex function on $\mathbb{R}^m$, i.e., $\varphi$ is a function from $\mathbb{R}^m$ to $(-\infty, +\infty]$ such that $\varphi \not\equiv +\infty$ and

$$\varphi\big(\lambda x + (1 - \lambda) y\big) \leqslant \lambda \varphi(x) + (1 - \lambda)\varphi(y), \quad \forall \lambda \in (0, 1).$$

We also suppose that $\varphi$ is lower-semicontinuus (l.s.c) and define its effective domain by

$$D(\varphi) := \big\{x \in \mathbb{R}^m\colon\ \varphi(x) < \infty\big\},$$

and its subdifferential by

$$\partial\varphi(x) := \big\{y\colon\ \varphi(x) \leqslant \varphi(z) + (x - z, y)\big\}.$$

We set

$$D(\partial\varphi) := \big\{x\colon\ \partial\varphi(x) \neq \emptyset\big\}.$$

Then it is well known that $\partial\varphi$ is a multivalued maximal monotone operator and $D(\partial\varphi)$ is a dense subset of $D(\varphi)$ and $D(\varphi)^o = D(\partial\varphi)^o$ (see [1,2,5]).

For more examples of multivalued maximal monotone operators and applications of SVIs, we refer to [9].

The following estimate, due to [5], will play a key role in this paper.

**Lemma 2.8.** *Suppose that $f \in L^2([0, T])$ and $u(t)$ be the unique solution to the deterministic differential equation $\dot{u}(t) \in -\partial\varphi(u(t)) + f$ with $u(0) \in D(\varphi)$. Then we have*

(i) *$u$ is absolutely continuous in $[0, T]$;*

(ii) $$\left[\int_0^T \big|\dot{u}(t)\big|^2\,dt\right]^{\frac{1}{2}} \leqslant \left[\int_0^T \big|f(t)\big|^2\,dt\right]^{\frac{1}{2}} + \sqrt{\big|\varphi\big(u(0)\big)\big|}.$$

## 3. Main result

To state our main result, we have to fix some more notations first. Denote by

$$\Omega := C_0\big([0, \infty), \mathbb{R}^d\big)$$

the space of continuous functions from $[0, \infty)$ to $\mathbb{R}^d$ which are null at 0, and denote the generic point of $\Omega$ by $\omega$. For $\omega \in \Omega$, set

$$w_t(\omega) := \omega(t).$$

Endow $\Omega$ with the compact uniform convergence topology. Denote by $\mathcal{F}$ the associated Borel $\sigma$-algebra and set

$$\mathcal{F}_t := \sigma(w_s, s \leqslant t).$$

Let $P$ stand for the canonical Wiener measure on $(\Omega, \mathcal{F})$. Then $w_t$ is a standard Brownian motion on $(\Omega, \mathcal{F}, P)$.

Set

$$\mathscr{S} := \{f \in \Omega:\ f \text{ is smooth}\},$$
$$\mathscr{S}_p := \{f \in \Omega:\ f \text{ is piecewise smooth}\}.$$

From now on, $b$ and $\sigma$ will be $C_b^3$-maps from $\mathbb{R}^m$ to $\mathbb{R}^m$ and $\mathbb{R}^m \times \mathbb{R}^d$ respectively, $\varphi$ is a proper l.s.c convex function on $\mathbb{R}^m$ with $D(\partial\varphi)^o \neq \emptyset$. We suppose further that $D(\varphi)$ is closed and $\varphi$ is bounded on it (which in particular implies that $D(\varphi)$ is closed). Consider the following Stratonovich SVI:

$$\begin{cases} dX(t) \in b\big(X(t)\big)\,dt + \sigma\big(X(t)\big) \circ dw(t) - \partial\varphi\big(X(t)\big)\,dt, \quad t \in \mathbb{R}_+, \\ X(0) = x \in \overline{D(\partial\varphi)}. \end{cases} \tag{2}$$

We have the following definition from [7,8]:

**Definition 3.1.** A pair of continuous and $\mathscr{F}_t$-adapted processes $(X, K)$ is called a solution of (1) if

(i) $X(0) = x$ and $X(t) \in \overline{D(\partial\varphi)}$ $a.s.$;
(ii) $K = \{K(t),\ t \in \mathbb{R}_+\}$ is of finite variation and $K(0) = 0$ a.s.;
(iii) $dX(t) = b(X(t))\,dt + \sigma(X(t)) \circ dw(t) - dK(t)$, $t \in \mathbb{R}_+$, a.s.;
(iv) almost surely, $dK(t) \in \partial\varphi(X(t))\,dt$.

By [7,8], Eq. (2) has a unique solution $(X(t), K(t))$.

For $h \in \mathscr{S}_p$, consider the following deterministic variational inequality

$$\begin{cases} d\xi(t) \in b\big(\xi(t)\big)\,dt + \sigma\big(\xi(t)\big)\dot{h}(t)\,dt - \partial\varphi\big(\xi(t)\big)\,dt, \quad t \in \mathbb{R}_+, \\ \xi(0) = x \in \overline{D(\partial\varphi)}. \end{cases} \tag{3}$$

By a classical result in [5], this inequality admits a unique solution and we shall denote it by $\xi(h, x)$. Set then

$$\eta(h, x, t) := \int_0^t b\big(\xi(h, x, s)\big)\,ds + \int_0^t \sigma\big(\xi(h, x, s)\big)\dot{h}(s)\,ds - \xi(h, x, t),$$
$$\mathscr{S}^x := \big\{\big(\xi(h, x), \eta(h, x)\big):\ h \in \mathscr{S}\big\},$$
$$\mathscr{S}_p^x := \big\{\big(\xi(h, x), \eta(h, x)\big):\ h \in \mathscr{S}_p\big\}.$$

We can now state our main result:

**Theorem 3.2.** *Denote by $P_x$ the law of $(X, K)$, the unique solution of* (1)*, in $\mathscr{W}^{2m}$. Then*

$$\operatorname{supp}(P_x) = \overline{\mathscr{S}_p^x} = \overline{\mathscr{S}^x}.$$

This theorem will be a direct consequence of Theorems 4.9 and 5.10.

## 4. Limit theorem

Now consider the following deterministic variational inequality

$$\begin{cases} \dot{X}_n(t) \in b\big(X_n(t)\big) + \sigma\big(X_n(t)\big)\dot{w}_n(t) - \partial\varphi\big(X_n(t)\big), \\ X_n(0) = x \in \overline{D(\partial\varphi)}, \end{cases} \tag{4}$$

where

$$\dot{w}_n(t) = 2^n\big[w\big(t_n^+\big) - w(t_n)\big], \quad t_n^+ = \frac{[2^n t]+1}{2^n}, \ t_n = \frac{[2^n t]}{2^n}.$$

Here $[a]$ stands for the integer part of $a$.

By [5, Proposition 3.12], Eq. (4) has a unique solution $X_n$. Set

$$K_n(t) := \int_0^t b\big(X_n(s)\big)\,ds + \int_0^t \sigma\big(X_n(s)\big)\dot{w}_n(s)\,ds - X_n(t).$$

Since $t \to X_n(t)$ is continuous and $x \in \overline{D(\partial\varphi)}$, we have $X_n(t) \in D(\partial\varphi) \subset \overline{D(\partial\varphi)} = D(\varphi)$, $\forall t \in \mathbb{R}_+^m$. Then we have by Lemma 2.8

$$\begin{aligned} \int_{u_n}^u \big|\dot{X}_n(t)\big|^2\,dt &\leqslant \int_{u_n}^u \big|b\big(X_n(v)\big) + \sigma\big(X_n(v)\big)\dot{w}_n(v)\big|^2\,dv + C \\ &\leqslant C\big(1 + 2^{-n}\big|\dot{w}_n(u)\big|^2\big), \end{aligned} \tag{5}$$

where $C$ is a constant independent of $n$.

Now we can state the main result of this section.

**Theorem 4.1.** *$(X_n, K_n)$ converges in $\mathscr{W}^{2m}$ to $(X, K)$ in probability.*

The rest of this section is devoted to the proof of this theorem. First we note that it is easy to deduce from Lemma 2.1 the following tightness criterion (see also [7]):

**Lemma 4.2.** *Let $\{\theta_n\}_{n\in\mathbb{N}^*}$ be a family of càdlàg increasing processes with $\theta_n(0) = 0$. If for all $t > 0$ there exists a constant $C(t) > 0$ such that*

$$\sup_{n\in\mathbb{N}^*} \mathbb{E}\big[\theta_n(t)\big] \leqslant C(t),$$

*then $(\theta_n)_{n\in\mathbb{N}^*}$ is tight on $\mathcal{T}$.*

We set

$$\begin{aligned} &\theta_n(t) := |K_n|_t + t, \\ &\tau_n := \theta_n^{-1}, \\ &M_n(t) := \int_0^t \sigma\big(X_n(u)\big)\dot{w}_n(u)\,du, \\ &Y_n(t) := X_n\big(\tau_n(t)\big), \end{aligned}$$

$$
\begin{aligned}
&H_n(t) := K_n\big(\tau_n(t)\big),\\
&w_n(t) := \int_0^t \dot{w}_n(u)\,du,\\
&a^{ij}(x) := \sum_{k=1}^{d} \sigma_k^i(x)\sigma_k^j(x),\\
&\alpha_n(t) := \sigma\big(X_n(t_n)\big)\dot{w}_n(t),\\
&\big(\sigma'\sigma\big)_i^{l,l'}(x) := \sum_{j=1}^{m}\big(\partial_j\sigma^{il}(x)\big)\sigma^{jl'}(x),\\
&(Lf)(x) := \sum_{i=1}^{m}\left[b_i(x) + \sum_{k=1}^{d}\sum_{j=1}^{m}\left(\frac{\partial}{\partial x^j}\sigma_k^i(x)\right)\sigma_k^j(x)\right]\partial_i f(x) + \frac{1}{2}\sum_{i,j=1}^{m} a^{ij}(x)\partial_i\partial_j f(x).
\end{aligned}
$$

We now prove

**Theorem 4.3.** *$(H_n, M_n, Y_n, w_n, \theta_n)_{n\in\mathbb{N}^*}$ is tight in $\mathscr{W}^{3m} \times \Omega \times \mathcal{T}$.*

**Proof.** Since

$$
t - s = \theta_n\big(\tau_n(t)\big) - \theta_n\big(\tau_n(s)\big) = |K_n|_{\tau_n(t)} - |K_n|_{\tau_n(s)} + \tau_n(t) - \tau_n(s),
$$

we have for $s \leqslant t$

$$
\begin{aligned}
\big|H_n(t) - H_n(s)\big| &= \big|K_n\big(\tau_n(t)\big) - K_n\big(\tau_n(s)\big)\big|\\
&\leqslant |K_n|_{\tau_n(t)} - |K_n|_{\tau_n(s)}\\
&= t - s - \big(\tau_n(t) - \tau_n(s)\big)\\
&\leqslant t - s.
\end{aligned}
\tag{6}
$$

Thus the tightness of $\{H_n\}$ follows. That of $\{w_n\}$ is trivial since

$$
\mathbb{E}\big[\big|w_n(t) - w_n(s)\big|^{2p}\big] \leqslant C|t - s|^p, \quad \forall p \geqslant 1. \tag{7}
$$

Next we look at $\theta_n$. Let $a$ be as in Proposition 2.7. We have for all $0 \leqslant s \leqslant t < \infty$,

$$
\begin{aligned}
\big|X_n(s) - a\big|^2 = |x - a|^2
&+ 2\int_0^s \big\langle X_n(u) - a, b\big(X_n(u)\big)\big\rangle\,du\\
&+ 2\int_0^s \big\langle X_n(u) - a, \sigma\big(X_n(u)\big)\dot{w}_n(u)\big\rangle\,du\\
&- 2\int_0^s \big\langle X_n(u) - a, dK_n(u)\big\rangle.
\end{aligned}
$$

By Proposition 2.7

$$\begin{aligned}\left|X_n(s)-a\right|^2 &\leqslant C+Ct+C\int_0^s \left|X_n(u)-a\right|^2 du \\ &\quad + C\int_0^s \left\langle X_n(u)-a, \sigma\big(X_n(u)\big)\dot{w}_n(u)\right\rangle du - C|K_n|_s, \end{aligned} \tag{8}$$

where we have used the boundedness of $b$ and Young's inequality. Let

$$I := \mathbb{E}\left[\sup_{0\leqslant s\leqslant t}\left|\int_0^s f\big(X_n(u)\big)\dot{w}_n(u)\,du\right|\right],$$

where $f(X_n(u)) = (X_n(u)-a)^*\sigma(X_n(u))$. Then

$$\begin{aligned} I &\leqslant \mathbb{E}\left[\sup_{0\leqslant s\leqslant t}\left|\int_0^s f\big(X_n(u_n)\big)\dot{w}_n(u)\,du\right|\right] \\ &\quad + \mathbb{E}\left[\sup_{0\leqslant s\leqslant t}\left|\int_0^s\int_{u_n}^u \frac{\partial f(X_n(v))}{\partial v}\dot{w}_n(u)\,dv\,du\right|\right] \\ &:= I_1 + I_2. \end{aligned} \tag{9}$$

Noticing that

$$\int_0^s f\big(X_n(u_n)\big)\dot{w}_n(u)\,du = \int_0^{s_n^+} \xi_u\,dw(u),$$

where

$$\xi_u = 2^n \int_{u_n}^{u_n^+\wedge s} f\big(X_n(v_n)\big)\,dv = 2^n\big(u_n^+\wedge s - u_n\big) f\big(X_n(u_n)\big).$$

We have by BDG inequality and the boundedness of $\sigma$

$$\begin{aligned} I_1 &\leqslant CE\left(\int_0^{t_n^+} |\xi_u|^2\,du\right)^{\frac{1}{2}} \\ &\leqslant C\mathbb{E}\left(\int_0^t \left|X_n(u_n)-a\right|^2 du\right)^{\frac{1}{2}} \\ &\leqslant C\int_0^t \mathbb{E}\Big[\sup_{0\leqslant u\leqslant s}\left|X_n(u)-a\right|^2\Big]\,ds + C(t). \end{aligned} \tag{10}$$

By $\dot{X}_n(t) = b(X_n(t)) + \sigma(X_n(t))\dot{w}_n(t) - \dot{K}_n(t)$ and the boundedness of $b$ and $\sigma$, we have

$$\begin{aligned}
I_2 \leqslant{}& \mathbb{E}\Bigg[\sup_{0\leqslant s\leqslant t}\Bigg|\int_0^s\int_{u_n}^u \big(\big|X_n(v)-a\big|\big|\nabla\sigma\big(X_n(v)\big)\big|+\big|\sigma\big(X_n(v)\big)\big|\big) \\
&\quad\times\big(\big|b\big(X_n(v)\big)\big|+\big|\sigma\big(X_n(v)\big)\big|\big|\dot{w}_n(v)\big|+\big|\dot{K}_n(v)\big|\big)\big|\dot{w}_n(u)\big|\,dv\,du\Bigg|\Bigg] \\
\leqslant{}& \mathbb{E}\Bigg[\sup_{0\leqslant s\leqslant t}\Bigg|\int_0^s\int_{u_n}^u \big(C\big|X_n(v)-a\big|+C\big)\big(C+C\big|\dot{w}_n(v)\big|+\big|\dot{K}_n(v)\big|\big)\big|\dot{w}_n(u)\big|\,dv\,du\Bigg|\Bigg] \\
\leqslant{}& C\mathbb{E}\Bigg[\int_0^t\int_{u_n}^u\big|\dot{w}_n(u)\big|\,dv\,du\Bigg]+C\mathbb{E}\Bigg[\int_0^t\int_{u_n}^u\big|\dot{w}_n(u)\big|^2\,dv\,du\Bigg] \\
&+C\mathbb{E}\Bigg[\int_0^t\int_{u_n}^u\big|\dot{K}_n(v)\big|\big|\dot{w}_n(u)\big|\,dv\,du\Bigg]+C\mathbb{E}\Bigg[\int_0^t\int_{u_n}^u\big|X_n(v)-a\big|\big|\dot{w}_n(u)\big|\,dv\,du\Bigg] \\
&+C\mathbb{E}\Bigg[\int_0^t\int_{u_n}^u\big|X_n(v)-a\big|\big|\dot{w}_n(u)\big|^2\,dv\,du\Bigg] \\
&+C\mathbb{E}\Bigg[\int_0^t\int_{u_n}^u\big|X_n(v)-a\big|\big|\dot{K}_n(v)\big|\big|\dot{w}_n(u)\big|\,dv\,du\Bigg] \\
\leqslant{}& C\mathbb{E}\Bigg[\int_0^t\int_{u_n}^u\big|\dot{w}_n(u)\big|\,dv\,du\Bigg]+C\mathbb{E}\Bigg[\int_0^t\int_{u_n}^u\big|\dot{w}_n(u)\big|^2\,dv\,du\Bigg] \\
&+C\mathbb{E}\Bigg[\int_0^t\int_{u_n}^u\big|\dot{X}_n(v)\big|\big|\dot{w}_n(u)\big|\,dv\,du\Bigg]+C\mathbb{E}\Bigg[\int_0^t\int_{u_n}^u\big|X_n(v)-a\big|\big|\dot{w}_n(u)\big|\,dv\,du\Bigg] \\
&+C\mathbb{E}\Bigg[\int_0^t\int_{u_n}^u\big|X_n(v)-a\big|\big|\dot{w}_n(u)\big|^2\,dv\,du\Bigg] \\
&+C\mathbb{E}\Bigg[\int_0^t\int_{u_n}^u\big|X_n(v)-a\big|\big|\dot{X}_n(v)\big|\big|\dot{w}_n(u)\big|\,dv\,du\Bigg] \\
:={}& I_{21}+I_{22}+I_{23}+I_{24}+I_{25}+I_{26}. \qquad (11)
\end{aligned}$$

It is easily seen that

$$I_{21}\leqslant C(t)2^{-\frac{n}{2}} \qquad (12)$$

and

$$I_{22}\leqslant C(t). \qquad (13)$$

By Lemma 2.8 and the boundedness of $b$ and $\sigma$,

$$\begin{aligned}
I_{23} &\leqslant C\mathbb{E}\left[\int_0^t \left\{\int_{u_n}^u \left|\dot{X}_n(v)\right|^2 dv\right\}^{\frac{1}{2}} \left\{\int_{u_n}^u \left|\dot{w}_n(u)\right| dv\right\}^{\frac{1}{2}} du\right] \\
&\leqslant C\mathbb{E}\left[\int_0^t \left\{\int_{u_n}^u \left|b\big(X_n(v)\big) + \sigma\big(X_n(v)\big)\dot{w}_n(v)\right|^2 dv\right\}^{\frac{1}{2}} 2^{-\frac{n}{2}} \left|\dot{w}_n(u)\right| du\right] + C(t) \\
&\leqslant C\mathbb{E}\left[\left|\dot{w}_n(t)\right|^2 2^{-n}\right] + C(t) \\
&\leqslant C(t). \qquad (14)
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
I_{24} &= C\int_0^t \int_{u_n}^u \mathbb{E}\left[\left|X_n(v) - a\right|\left|\dot{w}_n(u)\right|\right] dv\, du \\
&\leqslant C\int_0^t \int_{u_n}^u \left(\mathbb{E}\left|X_n(v) - a\right|^2\right)^{\frac{1}{2}} \left(\mathbb{E}\left|\dot{w}_n(u)\right|^2\right)^{\frac{1}{2}} dv\, du \\
&\leqslant C2^{-\frac{n}{2}} \int_0^t \mathbb{E}\Big[\sup_{0\leqslant v\leqslant u} \left|X_n(v) - a\right|^2\Big] du + C(t)2^{-\frac{n}{2}} \qquad (15)
\end{aligned}$$

and

$$\begin{aligned}
I_{25} &= C\int_0^t \int_{u_n}^u \mathbb{E}\left[\left|X_n(v) - a\right|\left|\dot{w}_n(u)\right|^2\right] dv\, du \\
&\leqslant C\int_0^t \int_{u_n}^u \left(\mathbb{E}\left[\left|X_n(v) - a\right|^2\right]\right)^{\frac{1}{2}} \left(\mathbb{E}\left[\left|\dot{w}_n(u)\right|^4\right]\right)^{\frac{1}{2}} dv\, du \\
&\leqslant C\int_0^t \mathbb{E}\Big[\sup_{0\leqslant v\leqslant u} \left|X_n(v) - a\right|^2\Big] du + C(t). \qquad (16)
\end{aligned}$$

Furthermore,

$$\begin{aligned}
I_{26} &\leqslant C\mathbb{E}\left[\int_0^t \left\{\int_{u_n}^u \left|X_n(v) - a\right|^2 dv\right\}^{\frac{1}{2}} \left\{\int_{u_n}^u \left|\dot{X}_n(v)\right|^2 \left|\dot{w}_n(u)\right|^2 dv\right\}^{\frac{1}{2}} du\right] \\
&\leqslant C\mathbb{E}\left[\int_0^t \left\{\int_{u_n}^u \left|X_n(v) - a\right|^2 dv\right\}^{\frac{1}{2}} \left|\dot{w}_n(u)\right| \left\{\int_{u_n}^u \left|\dot{X}_n(v)\right|^2 dv\right\}^{\frac{1}{2}} du\right] \\
&\leqslant C2^{-\frac{n}{2}}\mathbb{E}\left[\int_0^t \left\{\int_{u_n}^u \left|X_n(v) - a\right|^2 dv\right\}^{\frac{1}{2}} \left|\dot{w}_n(u)\right|^2 du\right]
\end{aligned}$$

$$+ C2^{-\frac{n}{2}}\mathbb{E}\Big[\int_0^t \Big\{\int_{u_n}^u \big|X_n(v)-a\big|^2\,dv\Big\}^{\frac{1}{2}}\big|\dot w_n(u)\big|\,du\Big]$$

$$\leqslant C\int_0^t \mathbb{E}\Big[\sup_{0\leqslant v\leqslant u}\big|X_n(v)-a\big|^2\Big]\,du + C(t). \tag{17}$$

Combining (9), (10), (11)–(16), (17) gives

$$I \leqslant C\int_0^t \mathbb{E}\Big[\sup_{0\leqslant v\leqslant u}\big|X_n(v)-a\big|^2\Big]\,du + C(t).$$

Hence

$$\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t}\big|X_n(s)-a\big|^2\Big] \leqslant C(t) + C\int_0^t \mathbb{E}\Big[\sup_{0\leqslant v\leqslant u}\big|X_n(v)-a\big|^2\Big]\,du.$$

Using Gronwall's inequality,

$$\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t}\big|X_n(s)-a\big|^2\Big] \leqslant C(t)e^{Ct}. \tag{18}$$

We obtain by (8)

$$\mathbb{E}\big[|K_n|_t\big] \leqslant C(t)e^{Ct}. \tag{19}$$

That is

$$\sup_{n\in\mathbb{N}^*}\mathbb{E}\big[\theta_n(t)\big] \leqslant C(t) < \infty, \quad 0\leqslant t<\infty. \tag{20}$$

Therefore, in virtue of Lemma 4.2, $\theta_n(t)$ is tight.

$\forall p\geqslant 1$, we also have

$$\begin{aligned}
&\mathbb{E}\big[\big|M_n(t)-M_n(s)\big|^{2p}\big]\\
&\quad = \mathbb{E}\Big[\Big|\int_s^t \sigma\big(X_n(u)\big)\dot w_n(u)\,du\Big|^{2p}\Big]\\
&\quad \leqslant C\mathbb{E}\Big[\Big|\int_s^t \sigma\big(X_n(u_n)\big)\dot w_n(u)\,du\Big|^{2p}\Big]\\
&\qquad + C\mathbb{E}\Big[\Big|\int_s^t \big(\sigma\big(X_n(u)\big)-\sigma\big(X_n(u_n)\big)\big)\dot w_n(u)\,du\Big|^{2p}\Big]\\
&\quad \leqslant C(t-s)^p + C\mathbb{E}\Big[\Big|\int_s^t\int_{u_n}^u \frac{\partial}{\partial x_j}\sigma^{il}\big(X_n(v)\big)b^j\big(X_n(v)\big)\dot w_n^l(u)\,dv\,du\Big|^{2p}\Big]\\
&\qquad + C\mathbb{E}\Big[\Big|\int_s^t\int_{u_n}^u \big(\sigma'\sigma\big)^{l,l'}\big(X_n(v)\big)\dot w_n^l(u)\dot w_n^{l'}(u)\,dv\,du\Big|^{2p}\Big]
\end{aligned}$$

$$
\begin{aligned}
&+C\mathbb{E}\Bigg[\bigg|\int_s^t\int_{u_n}^u \frac{\partial}{\partial x_j}\sigma^{il}\big(X_n(v)\big)\dot{K}_n^j(v)\dot{w}_n^l(u)\,dv\,du\bigg|^{2p}\Bigg]\\
&\leqslant C(t-s)^p+C(t-s)^{2p}\\
&\quad+C(t-s)^{2p-1}\mathbb{E}\Bigg[\int_s^t du\bigg|\int_{u_n}^u \big|\dot{K}_n(v)\big|\big|\dot{w}_n(u)\big|\,dv\bigg|^{2p}\Bigg]\\
&\leqslant C(t-s)^p+C(t-s)^{2p}\\
&\quad+C(t-s)^{2p-1}\mathbb{E}\Bigg[\int_s^t du\bigg|\int_{u_n}^u \big|\dot{w}_n(u)\big|\,dv\bigg|^{2p}\Bigg]\\
&\quad+C(t-s)^{2p-1}\mathbb{E}\Bigg[\int_s^t du\bigg|\int_{u_n}^u \big|\dot{w}_n(u)\big|^2\,dv\bigg|^{2p}\Bigg]\\
&\quad+C(t-s)^{2p-1}\mathbb{E}\Bigg[\int_s^t du\bigg|\int_{u_n}^u \big|\dot{X}_n(v)\big|\big|\dot{w}_n(u)\big|\,dv\bigg|^{2p}\Bigg]\\
&\leqslant C(t-s)^p+C(t-s)^{2p}\\
&\quad+C(t-s)^{2p-1}\mathbb{E}\Bigg[\int_s^t\bigg(\int_{u_n}^u \big|\dot{X}_n(v)\big|^2\,dv\bigg)^p\bigg(\int_{u_n}^u \big|\dot{w}_n(v)\big|^2\,dv\bigg)^p du\Bigg]\\
&\leqslant C(t-s)^p+C(t-s)^{2p}\\
&\quad+Cn^{2p}2^{-2np}(t-s)^{2p-1}\mathbb{E}\Bigg[\int_s^t\bigg(\int_{u_n}^u \big|\dot{X}_n(v)\big|^2\,dv\bigg)^p 2^{-np}\big|\dot{w}_n(v)\big|^{2p}\,du\Bigg]\\
&\leqslant C(t-s)^p+C(t-s)^{2p}.
\end{aligned}
$$

That is

$$
\mathbb{E}\big[\big|M_n(t)-M_n(s)\big|^{2p}\big]\leqslant C(t-s)^p. \tag{21}
$$

Furthermore,

$$
\begin{aligned}
\mathbb{E}\big[\big|Y_n(t)-Y_n(s)\big|^{2p}\big]&=\mathbb{E}\big[\big|X_n\big(\tau_n(t)\big)-X_n\big(\tau_n(s)\big)\big|^{2p}\big]\\
&\leqslant C\mathbb{E}\Bigg[\bigg|\int_{\tau_n(s)}^{\tau_n(t)} b\big(X_n(u)\big)\,du\bigg|^{2p}\Bigg]+C\mathbb{E}\big[\big|M_n\big(\tau_n(t)\big)-M_n\big(\tau_n(s)\big)\big|^{2p}\big]\\
&\quad+C\mathbb{E}\big[\big|H_n(t)-H_n(s)\big|^{2p}\big]\\
&\leqslant C\mathbb{E}\big[\big|\tau_n(t)-\tau_n(s)\big|^{2p}\big]+C(t-s)^{2p}\\
&\quad+C\mathbb{E}\big[\big|M_n\big(\tau_n(t)\big)-M_n\big(\tau_n(s)\big)\big|^{2p}\big].
\end{aligned}
$$

Using $|\tau_n(t)-\tau_n(s)|\leqslant|t-s|$, similarly to (21), we have

$$\begin{aligned}\mathbb{E}\big[\big|M_n\big(\tau_n(t)\big)-M_n\big(\tau_n(s)\big)\big|^{2p}\big]&=\mathbb{E}\Bigg[\Bigg|\int_{\tau_n(s)}^{\tau_n(t)}\sigma\big(X_n(u)\big)\dot{w}_n(u)\,du\Bigg|^{2p}\Bigg]\\&\leqslant C\mathbb{E}\Bigg[\Bigg|\int_{\tau_n(s)}^{\tau_n(t)}\sigma\big(X_n(u_n)\big)\dot{w}_n(u)\,du\Bigg|^{2p}\Bigg]\\&\quad+C\mathbb{E}\Bigg[\Bigg|\int_{\tau_n(s)}^{\tau_n(t)}\big(\sigma\big(X_n(u)\big)-\sigma\big(X_n(u_n)\big)\big)\dot{w}_n(u)\,du\Bigg|^{2p}\Bigg]\\&\leqslant C(t-s)^p.\end{aligned}\tag{22}$$

Hence

$$\mathbb{E}\big[\big|Y_n(t)-Y_n(s)\big|^{2p}\big]\leqslant C(t-s)^p.\tag{23}$$

Combining (6), (7), (20), (21), (41) gives the desired tightness by Aldous's theorem (see [12]). □

Denote by $\{L_n, n\in N^*\}$ the distribution of $(\tau_n, H_n, M_n, Y_n, w_n, \theta_n)$ on $\mathscr{W}^{3m+1}\times\Omega\times\mathcal{T}$. Since $\mathscr{W}^{3m+1}\times\Omega\times\mathcal{T}$ is a Polish space, by Prokhorov's theorem and Proposition 4.3, there exists $L_{n_k}$ and probability $L$ on $\mathscr{W}^{3m+1}\times\Omega\times\mathcal{T}$ such that $L_{n_k}\to L(k\to\infty)$. To simplify the notation, we suppose that $L_n\to L(n\to\infty)$. By Skorohod's representation theorem, there exists a probability space $(\hat{\Omega},\hat{\mathscr{F}},\hat{P})$ on which are defined random variables $(\hat{\tau}_n,\hat{H}_n,\hat{M}_n,\hat{Y}_n,\hat{w}_n,\hat{\theta}_n)$, $(\hat{\tau},\hat{H},\hat{M},\hat{Y},\hat{w},\hat{\theta})$ such that

$$\begin{aligned}&(\hat{\tau}_n,\hat{H}_n,\hat{M}_n,\hat{Y}_n,\hat{w}_n,\hat{\theta}_n)\sim(\tau_n,H_n,M_n,Y_n,w_n,\theta_n),\\&\hat{P}^{(\hat{\tau},\hat{H},\hat{M},\hat{Y},\hat{w},\hat{\theta})}=L,\end{aligned}\tag{24}$$

and as $n\to\infty$,

$$(\hat{\tau}_n,\hat{H}_n,\hat{M}_n,\hat{Y}_n,\hat{w}_n,\hat{\theta}_n)\to(\hat{\tau},\hat{H},\hat{M},\hat{Y},\hat{w},\hat{\theta})\quad a.s.\tag{25}$$

in $\mathscr{W}^{3m+1}\times\Omega\times\mathcal{T}$. Define

$$\begin{aligned}&\hat{X}_n(t):=\hat{Y}_n\big(\hat{\theta}_n(t)\big),\qquad \hat{X}(t):=\hat{Y}\big(\hat{\theta}(t)\big),\\&\hat{K}_n(t):=\hat{H}_n\big(\hat{\theta}_n(t)\big),\qquad \hat{K}(t):=\hat{H}\big(\hat{\theta}(t)\big).\end{aligned}\tag{26}$$

We now pass from the convergence of $\hat{Y}_n$ to that of $\hat{X}_n$. First, note that according to Lemma 2.2, this will be done if we prove the following

**Theorem 4.4.** *There exists $\hat{\Omega}_0\in\hat{\mathscr{F}}$, $\hat{P}(\hat{\Omega}_0)=1$ such that for all $\hat{\omega}\in\hat{\Omega}_0$, if there exists $0\leqslant s\leqslant t<\infty$ satisfying $\hat{\tau}(s)(\hat{\omega})=\hat{\tau}(t)(\hat{\omega})$, then $\hat{Y}(s)(\hat{\omega})=\hat{Y}(t)(\hat{\omega})$.*

**Proof.** The proof is the same as that of [17, Theorem 3.2], except the Step (A) there. But this is even easier here. In fact, since $\hat{X}_n(t,\omega)\in\overline{D(\partial\varphi)}$, there exists an $\hat{N}_1$ with $\hat{P}(\hat{N}_1)=0$ such that $\hat{X}(t,\omega)=\lim_{n\mapsto\infty}\hat{X}_n(t,\omega)\in\overline{D(\partial\varphi)}$, $\forall(t,\omega)\in\mathbb{R}_+\times\hat{N}_1^c$, a.s. and thus Step (A) is done. □

By Lemma 2.2 and Theorem 4.4, there is a subsequence $\{n_k\}$ such that

$$(\hat{X}_{n_k}, \hat{K}_{n_k}) \to (\hat{X}, \hat{K}) \quad \text{in } \mathscr{W}^{2m} \text{a.s.} \tag{27}$$

Consequently, the image measure of $(\hat{X}_{n_k}, \hat{K}_{n_k})$ has a weak limit $\mu$ in $\mathscr{W}^{2m}$.

Denote by $\mu_n$ the law of $(X_n, K_n)$. Since $(\hat{X}_n, \hat{K}_n)$ and $(X_n, K_n)$ are identically distributed, $\mu$ is a weak limit of $\{\mu_n\}$ in $\mathscr{W}^{2m}$.

Starting from the very beginning with an arbitrary subsequence and repeating the above reasoning, we know that any subsequence has a weakly convergent sub-subsequence and we thus arrive at the following result.

**Theorem 4.5.** *$\{\mu_n\}$ is relatively compact in $\mathscr{W}^{2m}$.*

Next we shall prove that the whole sequence $\{\mu_n\}$ converges weakly to a unique limit and we shall identify this limit.

Let

$$\mathscr{V}^m := \big\{V\colon \mathbb{R}_+ \mapsto \mathbb{R}^m,\ V(0) = 0,\ V \text{ is continuous and of finite variation on compacts}\big\}.$$

By Theorem 4.5, $\{\mu_n\}$ has a weak limit. Using the equivalence of weak solution and martingale problem, we shall prove:

**Theorem 4.6.** *Suppose that $\mu$ is a weak limit of $\{\mu_n\}$. Let $(x(\cdot), v(\cdot))$ be coordinate processes on $\mathscr{W}^m \times \mathscr{V}^m$. Then, under $\mu$,*

$$f\big(x(t)\big) - f\big(x(s)\big) + \int_s^t \big\langle \nabla f\big(x(u)\big), dv(u)\big\rangle - \int_s^t Lf\big(x(u)\big)\, du$$

*is a martingale for all $f \in C_b^2$.*

**Proof.** By a density argument, it suffices to prove that

$$\mathbb{E}^{\mu}\big[F \cdot \big(f\big(x(t)\big) - f\big(x(s)\big)\big)\big] = \mathbb{E}^{\mu}\left[F \cdot \int_s^t \big(Lf\big(x(u)\big)\, du - \big\langle \nabla f\big(x(u)\big),\, dv(u)\big\rangle\big)\right]$$

for all $f \in C_0^{\infty}(\mathbb{R}^m)$, $0 \leqslant s < t$, and bounded $\mathscr{B}_s(\mathscr{W}^m) \times \mathscr{B}_s(\mathscr{V}^m)$ measurable $F : \Omega \mapsto \mathbb{R}$. Clearly, it will suffice to do this when $s$ and $t$ have the form $k/2^N$ and $F$ is bounded continuous, and $\mathscr{B}_s(\mathscr{W}^m) \times \mathscr{B}_s(\mathscr{V}^m)$ measurable. Observe that

$$\begin{aligned}
&\mathbb{E}^{\mu_n}\left[F \cdot \left(f\big(x(t)\big) - f\big(x(s)\big) + \int_s^t \big\langle \nabla f\big(x(u)\big),\, dv(u)\big\rangle\right)\right] \\
&\quad = \mathbb{E}^{\mu_n}\left[F \cdot \int_s^t \big\langle \nabla f\big(x(u)\big), b\big(x(u)\big)\big\rangle\, du\right] + \mathbb{E}^{P}\left[F \cdot \int_s^t \big\langle \nabla f\big(X_n(u)\big), \alpha_n(u)\big\rangle\, du\right] \\
&\qquad + \mathbb{E}^{P}\left[F \cdot \int_s^t \big\langle \nabla f\big(X_n(u)\big), \dot{M}_n(u) - \alpha_n(u)\big\rangle\, du\right]
\end{aligned}$$

$$+\mathbb{E}^P\left[\int_s^t \langle \nabla f(X_n(u)), dK_n(u)\rangle\right]$$
$$:= J_{1,n} + J_{2,n} + J_{3,n} + J_{4,n}.$$

Clearly,

$$J_{1,n} \to \mathbb{E}^\mu\left[F\cdot\int_s^t \langle \nabla f(x(u)), b(x(u))\rangle\, du\right], \tag{28}$$

and

$$J_{4,n} \to \mathbb{E}^P\left[\int_s^t \langle \nabla f(X(u)), dK(u)\rangle\right] = \mathbb{E}^\mu\left[F\cdot\int_s^t \langle \nabla f(x(u)), dv(u)\rangle\right]. \tag{29}$$

We have to consider $J_{2,n}$ and $J_{3,n}$. First we prove

$$J_{2,n} \to \mathbb{E}^\mu\left[F\cdot\int_s^t L_u^0 f(x(u))\, du\right], \tag{30}$$

where

$$L_u^0 = \frac{1}{2}a^{ij}(x)\left(\partial^2/\partial x_i\partial x_j\right).$$

Let $H(x)$ denote the Hessian matrix of $f$. Since

$$\mathbb{E}^P\left[\alpha_n(u) \mid \mathscr{B}_{u_n}(\mathscr{W}^m)\times\mathscr{B}_{u_n}(\mathscr{V}^m)\right] = 0,$$

we have

$$\begin{aligned}
J_{2,n} &= \mathbb{E}^P\left[F\cdot\int_s^t \langle \nabla f(X_n(u_n)), \alpha_n(u)\rangle\, du\right] \\
&\quad + \mathbb{E}^P\left[F\cdot\int_s^t \langle \nabla f(X_n(u)) - \nabla f(X_n(u_n)), \alpha_n(u)\rangle\, du\right] \\
&= \mathbb{E}^P\left[F\cdot\int_s^t \langle \nabla f(X_n(u)) - \nabla f(X_n(u_n)), \alpha_n(u)\rangle\, du\right] \\
&= \mathbb{E}^P\left[F\cdot\int_s^t du\int_{u_n}^u dv\langle \dot{M}_n(v), H(X_n(v))\alpha_n(u)\rangle\right] \\
&\quad + \mathbb{E}^P\left[F\cdot\int_s^t du\int_{u_n}^u dv\langle b(X_n(v)), H(X_n(v))\alpha_n(u)\rangle\right] \\
&\quad - \mathbb{E}^P\left[F\cdot\int_s^t du\int_{u_n}^u \langle H(X_n(v))\alpha_n(u), dK_n(v)\rangle\right] \\
&:= K_{1,n} + K_{2,n} + K_{3,n}.
\end{aligned}$$

An elementary calculus gives $|K_{2,n}| \to 0$. Since

$$K_{3,n} \leqslant \mathbb{E}\left[\left|\int_s^t du \int_{u_n}^u \left\langle H\big(X_n(v)\big)\alpha_n(u), dK_n(v)\right\rangle\right|\right],$$

we have to show that $|\int_s^t du \int_{u_n}^u \langle H(X_n(v))\alpha_n(u), dK_n(v)\rangle|$ is uniformly integrable and converges to zero in probability.

Since

$$\begin{aligned}
&\int_{u_n}^u \left\langle H\big(X_n(v)\big)\alpha_n(u), dK_n(v)\right\rangle \\
&\quad = \left\langle H\big(X_n(u)\big)\alpha_n(u), K_n(u) - K_n(u_n)\right\rangle \\
&\qquad - \int_{u_n}^u \left\langle \sum_{i=1}^m \alpha_n^i(u)\left(\sum_{j=1}^m \frac{\partial H_{ki}(X_n(v))}{x_j}\dot{X}_n(v)\right), K_n(v) - K_n(u_n)\right\rangle dv \\
&\quad \leqslant C\big|\dot{w}_n(u)\big|\big|K_n(u) - K_n(u_n)\big| + \int_{u_n}^u \big|\dot{w}_n(u)\big|\big|K_n(v) - K_n(u_n)\big|\big|\dot{X}_n(v)\big|\, dv
\end{aligned}$$

and

$$\begin{aligned}
\big|K_n(u) - K_n(u_n)\big| &\leqslant \int_{u_n}^u \big|\dot{K}_n(v)\big|\, dv \\
&\leqslant \int_{u_n}^u \big|\dot{X}_n(v)\big|\, dv + C2^{-n} + C2^{-n}\big|\dot{w}_n(u)\big| \\
&\leqslant 2^{-\frac{n}{2}}\left\{\int_{u_n}^u \big|\dot{X}_n(v)\big|^2\, dv\right\}^{\frac{1}{2}} + C2^{-n} + C2^{-n}\big|\dot{w}_n(u)\big| \\
&\leqslant C2^{-n} + C2^{-n}\big|\dot{w}_n(u)\big|,
\end{aligned}$$

we have

$$\begin{aligned}
&\left|\int_{u_n}^u \left\langle H\big(X_n(v)\big)\alpha_n(u), dK_n(v)\right\rangle\right| \\
&\quad \leqslant C2^{-n}\big|\dot{w}_n(u)\big| + C2^{-n}\big|\dot{w}_n(u)\big|^2 \\
&\qquad + C2^{-n}\int_{u_n}^u \big|\dot{w}_n(u)\big|\big|\dot{X}_n(v)\big|\, dv + C2^{-n}\int_{u_n}^u \big|\dot{w}_n(u)\big|^2\big|\dot{X}_n(v)\big|\, dv \\
&\quad \leqslant C2^{-n}\big|\dot{w}_n(u)\big| + C2^{-n}\big|\dot{w}_n(u)\big|^2 + C2^{-2n}\big|\dot{w}_n(u)\big| \\
&\qquad + C2^{-2n}\big|\dot{w}_n(u)\big|^2 + C2^{-2n}\big|\dot{w}_n(u)\big|^3.
\end{aligned}$$

For any $p > 1$,

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_s^t du \int_{u_n}^u \langle H(X_n(v))\alpha_n(u), dK_n(v)\rangle\right|^p\right] \\
&\quad \leqslant C(t-s)^{p-1}\mathbb{E}\left[\int_s^t \left|\int_{u_n}^u \langle H(X_n(v))\alpha_n(u), dK_n(v)\rangle\right|^p du\right] \\
&\quad \leqslant C(t-s)^p\left\{\mathbb{E}\left[2^{-np}\left|\dot{w}_n(u)\right|^p\right] + \mathbb{E}\left[2^{-np}\left|\dot{w}_n(u)\right|^{2p}\right] + \mathbb{E}\left[2^{-2np}\left|\dot{w}_n(u)\right|^p\right]\right. \\
&\qquad \left. + \mathbb{E}\left[2^{-2np}\left|\dot{w}_n(u)\right|^{2p}\right] + \mathbb{E}\left[2^{-2np}\left|\dot{w}_n(u)\right|^{3p}\right]\right\} \\
&\quad \leqslant C.
\end{aligned}$$

Therefore, $|\int_s^t du \int_{u_n}^u \langle H(X_n(v))\alpha_n(u), dK_n(v)\rangle|$ is uniformly integrable.

Set

$$J_n(u) := \int_{u_n}^u \langle H(X_n(v)), dK_n(v)\rangle,$$

$$\beta_n(u) := \int_{u_n}^u \alpha_n(v)\, dv,$$

$$\gamma_n(u) := \int_0^u \alpha_n(v)\, dv.$$

As $J_n(u_n) = 0$, we have

$$\begin{aligned}
&\int_s^t du \int_{u_n}^u \langle H(X_n(v))\alpha_n(u), dK_n(v)\rangle \\
&\quad = \int_s^t \alpha_n(u) J_n(u)\, du \\
&\quad = \int_s^t J_n(u)\, d\gamma_n(u) \\
&\quad = \int_s^t J_n(u)\, d\beta_n(u) \\
&\quad = J_n(t)\beta_n(t) - J_n(s)\beta_n(s) - \int_s^t \beta_n(u)\, dJ_n(u).
\end{aligned}$$

Since as $n \to \infty$

$$\begin{aligned}&\left|\beta_n(u)\right| \leqslant C\left|w\left(u_n^+\right)-w(u_n)\right| \to 0,\\&\sup_n E\left[V_t(J_n)\right] \leqslant C \sup_n E\left[V_t(K_n)\right] < \infty,\end{aligned}$$

we have

$$\int_s^t du \int_{u_n}^u \left\langle H\left(X_n(v)\right)\alpha_n(u), dK_n(v)\right\rangle \to 0$$

in probability. Hence $|K_{3,n}| \to 0$.

For $K_{1,n}$, we have

$$\begin{aligned}K_{1,n} &= \mathbb{E}^P\left[F \cdot \int_s^t du \int_{u_n}^u dv \left\langle \alpha_n(v), H\left(X_n(v)\right)\alpha_n(u)\right\rangle\right]\\&\quad + \mathbb{E}^P\left[F \cdot \int_s^t du \int_{u_n}^u dv \int_{v_n}^v dr \left\langle \frac{\partial}{\partial x_j}\sigma^{il}\left(X_n(r)\right)b^j\left(X_n(r)\right)\dot{w}_n^l, H\left(X_n(v)\right)\alpha_n(u)\right\rangle\right]\\&\quad + \mathbb{E}^P\left[F \cdot \int_s^t du \int_{u_n}^u dv \int_{v_n}^v dr \left\langle \left(\sigma'\sigma\right)^{l,l'}\left(X_n(r)\right)\dot{w}_n^l\dot{w}_n^{l'}, H\left(X_n(v)\right)\alpha_n(u)\right\rangle\right]\\&\quad - \mathbb{E}^P\left[F \cdot \int_s^t du \int_{u_n}^u dv \int_{v_n}^v dr \left\langle \frac{\partial}{\partial x_j}\sigma^{il}\left(X_n(r)\right)\dot{K}_n^j(v)\dot{w}_n^l, H\left(X_n(v)\right)\alpha_n(u)\right\rangle\right]\\&:= K_{4,n} + K_{5,n} + K_{6,n} + K_{7,n}.\end{aligned}$$

Again, we have $|K_{5,n}| \to 0$, $|K_{6,n}| \to 0$, $|K_{7,n}| \to 0$ by simple calculus. We only need to consider $K_{4,n}$.

$$\begin{aligned}K_{4,n} &= \mathbb{E}^P\left[F \cdot \int_s^t du \int_{u_n}^u dv \left\langle \alpha_n(v), H\left(X_n(u_n)\right)\alpha_n(u)\right\rangle\right]\\&\quad + \mathbb{E}^P\left[F \cdot \int_s^t du \int_{u_n}^u dv \left\langle \alpha_n(v), \left(H\left(X_n(v)\right) - H\left(X_n(u_n)\right)\right)\alpha_n(u)\right\rangle\right]\\&:= K_{8,n} + K_{9,n}.\end{aligned}$$

Since $|K_{9,n}| \to 0$, it remains to examine $K_{8,n}$.

$$\begin{aligned}K_{8,n} &= 2^n\mathbb{E}^P\left[F \cdot \int_s^t du \int_{u_n}^u dv\, tr\left[\sigma^*\left(X_n(v_n)\right)H\left(X_n(u_n)\right)\sigma\left(X_n(u_n)\right)\right]\right]\\&= 2^n\mathbb{E}^{\mu_n}\left[F \cdot \int_s^t du \int_{u_n}^u dv\, tr\left[\sigma^*\left(x(v_n)\right)H\left(x(u_n)\right)\sigma\left(x(u_n)\right)\right]\right].\end{aligned}$$

Since for bounded measurable functions $\varphi$ and $\psi$ on $[s,t]$,

$$2^n \int_s^t \varphi(u)\,du \int_{u_n}^u \psi(v)\,dv \to \frac{1}{2}\int_s^t \varphi(u)\psi(u)\,du$$

(see [18, Lemma 4.2]), using $\mu_n \to \mu$,

$$K_{8,n} \to \frac{1}{2}\mathbb{E}^\mu\left[F \cdot \int_s^t tr\big[\sigma^*\big(x(u)\big)H\big(x(u)\big)\sigma\big(x(u)\big)\big]\,du\right],$$

(30) is proved.

Finally we treat $J_{3,n}$. We write

$$\begin{aligned}
J_{3,n} &= \mathbb{E}^P\left[F \cdot \int_s^t \int_{u_n}^u dv \big\langle \nabla f\big(X_n(u)\big), \big(\sigma'\sigma\big)^{l,l'}\big(X_n(v)\big)\dot{w}_n^l \dot{w}_n^{l'}\big\rangle\right] \\
&\quad + \mathbb{E}^P\left[F \cdot \int_s^t \int_{u_n}^u dv \left\langle \nabla f\big(X_n(u)\big), \frac{\partial}{\partial x_j}\sigma^{il}\big(X_n(v)\big)b^j\big(X_n(v)\big)\dot{w}_n^l\right\rangle\right] \\
&\quad - \mathbb{E}^P\left[F \cdot \int_s^t \int_{u_n}^u dv \left\langle \nabla f\big(X_n(u)\big), \frac{\partial}{\partial x_j}\sigma^{il}\big(X_n(v)\big)\dot{K}_n^j(v)\dot{w}_n^l\right\rangle\right] \\
&:= L_{1,n} + L_{2,n} + L_{3,n}.
\end{aligned}$$

Clearly, $|L_{2,n}| \to 0$ and $|L_{3,n}| \to 0$. Observe that

$$\begin{aligned}
L_{1,n} &= \mathbb{E}^P\left[F \cdot \int_s^t \big\langle \nabla f\big(X_n(u_n)\big), \big(\sigma'\sigma\big)^{l,l'}\big(X_n(u_n)\big)(u-u_n)\dot{w}_n^l \dot{w}_n^{l'}\big\rangle\right] \\
&\quad + \mathbb{E}^P\bigg[F \cdot \int_s^t du \int_{u_n}^u dv \big\langle \nabla f\big(X_n(u_n)\big), \\
&\qquad\qquad \big[\big(\sigma'\sigma\big)^{l,l'}\big(X_n(v)\big) - \big(\sigma'\sigma\big)^{l,l'}\big(X_n(u_n)\big)\big]\dot{w}_n^l \dot{w}_n^{l'}\big\rangle\bigg] \\
&\quad + \mathbb{E}^P\left[F \cdot \int_s^t du \int_{u_n}^u dv \big\langle \nabla f\big(X_n(u)\big) - \nabla f\big(X_n(u_n)\big), \big(\sigma'\sigma\big)^{l,l'}\big(X_n(v)\big)\dot{w}_n^l \dot{w}_n^{l'}\big\rangle\right] \\
&:= L_{4,n} + L_{5,n} + L_{6,n}.
\end{aligned}$$

By $|L_{5,n}| \to 0$, $|L_{6,n}| \to 0$ and $L_{4,n} \to \frac{1}{2}\mathbb{E}^\mu[F \cdot \int_s^t \langle \nabla f(x(u)), \sigma'\sigma(x(u))\rangle\,du]$, we get

$$J_{3,n} \to \frac{1}{2}\mathbb{E}^\mu\left[F \cdot \int_s^t \big\langle \nabla f\big(x(u)\big), \sigma'\sigma\big(x(u)\big)\big\rangle\,du\right]. \tag{31}$$

The proof is completed. □

Using the same argument as in the proof of [7, Proposition 5.13], we can prove:

**Proposition 4.7.** *If $(\alpha, \beta)$ are continuous functions satisfying*

$$\big(\alpha(t), \beta(t)\big) \in Gr(A), \quad \forall t \in \mathbb{R}_+,$$

*then the measure*

$$\big\langle x(t) - \alpha(t), dv(t) - \beta(t)dt \big\rangle$$

*is positive on $\mathbb{R}^+$, $\mu$-a.s.*

Now, instead of (2) we consider the following system

$$\begin{cases} dX^i(t) \in \sigma^{ij}\big(X(t)\big) \circ dw^j(t) + b^i\big(X(t)\big)\,dt - \big(\partial\varphi\big)^i\big(X(t)\big)\,dt, \quad i = 1, 2, \dots, m, \\ dX^{m+j}(t) = dw^j(t), \\ X^i(0) = x^i, \\ X^{m+j}(0) = 0. \end{cases} \tag{32}$$

Denote by $\nu_n$ the law of $(w_n, X_n, K_n)$ in $\Omega \times \mathscr{W}^m \times \mathscr{V}^m$.

Applying Theorem 4.6 and Proposition 4.7 to the above system and using the uniqueness in distribution of the solution of (2) we obtain

**Theorem 4.8.** *On $(\Omega \times \mathscr{W}^m \times \mathscr{V}^m, \nu)$, $t \mapsto w(t)$ is a Brownian motion and $(X, K)$ is solution of the following multivalued Stratonovich SDE*:

$$\begin{cases} dX(t) \in b\big(X(t)\big)\,dt + \sigma\big(X(t)\big) \circ dw(t) - \partial\varphi\big(X(t)\big)\,dt, \\ X(0) = x \in \overline{D(\partial\varphi)}. \end{cases} \tag{33}$$

*Moreover, $\nu$ is the unique weak limit of $\{\nu_n\}$.*

Finally, using an argument analogous to [17, Theorem 6.2] we can prove

**Theorem 4.9.** *$(X_n, K_n)$ converges in $\mathscr{W}^{2m}$ to $(X, K)$ in probability.*

## 5. Approximate continuity

In this section we further suppose that $D(\varphi)$ is bounded and the notations $a_{ij}(x)$, $(\sigma'\sigma)_i^{l,l'}(x)$ and $(Lf)(x)$ which will be needed are defined as in Section 3.

Set

$$H_0 := \mathbb{W}_0^d \cap C_b^2.$$

Let $h \in H_0$ and denote by $\xi(t)$ the unique solution, whose existence and uniqueness is assured by [5, Proposition 3.12], of the following DVI:

$$\begin{cases} \dot{\xi}(t) \in b\big(\xi(t)\big) + \sigma\big(\xi(t)\big)\dot{h}(t) - \partial\varphi\big(\xi(t)\big), \\ \xi(0) = x \in \overline{D(\partial\varphi)}. \end{cases} \tag{34}$$

Then

$$\dot{\eta}(t) := -\dot{\xi}(t) + b\big(\xi(t)\big) + \sigma\big(\xi(t)\big)\dot{h}(t) \in \partial\varphi\big(\xi(t)\big).$$

The following two lemmas and corollary are taken from [11].

**Lemma 5.1.** *$P(\|w\|_T < \varepsilon) \sim C\exp(-\frac{C}{\varepsilon^2})$ as $\varepsilon \downarrow 0$.*

**Lemma 5.2.** *Set $\kappa^{ij}(t) := \frac{1}{2}\int_0^t [w^i(s)\,dw^j(s) - w^j(s)\,dw^i(s)]$, $i, j = 1, \dots, d$. Then for all $i, j = 1, \dots, d$,*

$$\lim_{M\uparrow\infty} \sup_{0<\delta\leqslant 1} P\big(\big\|\kappa^{ij}\big\|_T > M\delta \bigm| \|w\|_T < \delta\big) = 0.$$

**Corollary 5.3.** *Let $\zeta^{ij}(t) := \int_0^t w^i(s) \circ dw^j(s)$, $i, j = 1, \dots, d$. Then for all $i, j = 1, \dots, d$, we have*

$$\lim_{M\uparrow\infty} \sup_{0<\delta\leqslant 1} P\big(\big\|\zeta^{ij}\big\|_T > M\delta \bigm| \|w\|_T < \delta\big) = 0.$$

*In particular, for every $\varepsilon > 0$ and $\alpha \in (0, 1)$,*

$$P\big(\big\|\zeta^{ij}\big\|_T > \varepsilon\delta^\alpha \bigm| \|w\|_T < \delta\big) \to 0 \quad \text{as } \delta \downarrow 0. \tag{35}$$

We will need two more lemmas.

**Lemma 5.4.** *Let $|K|_T$ be the total variation of $K$ on $[0, T]$. Then there exists a strictly positive constant $\alpha > 0$, such that*

$$\mathbb{E}\big[e^{\alpha|K|_T^2}\big] < \infty. \tag{36}$$

**Proof.** By Lemma 2.7 we have

$$\begin{aligned}
\big|X(t) - a\big|^2 &\leqslant C + Ct + 2\int_0^t \big\langle X(s) - a, \sigma\big(X(s)\big) \circ dw(s)\big\rangle + C\int_0^t \big|X(s) - a\big|\,ds \\
&\quad - 2\int_0^t \big\langle X(s) - a, dK(s)\big\rangle \\
&\leqslant C + Ct + C\int_0^t \big|X(s) - a\big|\,ds - C|K|_t + 2\int_0^t \big\langle X(s) - a, \sigma\big(X(s)\big) \circ dw(s)\big\rangle.
\end{aligned}$$

Hence

$$|K|_T \leqslant CT + 2\sup_{0\leqslant t\leqslant T}\left|\int_0^t \big\langle X(s) - a, \sigma\big(X(s)\big) \circ dw(s)\big\rangle\right|. \tag{37}$$

Let

$$N_t := \int_0^t \big\langle X(s) - a, \sigma\big(X(s)\big) \circ dw(s)\big\rangle.$$

Since $X$ and $\sigma$ are bounded, we have

$$\langle N, N\rangle_t \leqslant Ct.$$

Hence there exists $\alpha_1 > 0$ such that

$$\mathbb{E}\left[e^{\alpha_2 \|N\|_T^2}\right] < \infty. \tag{38}$$

Combining (37) and (38) gives the desired result. □

**Lemma 5.5.** *Let $M_t := \int_0^t \langle X(s) - a, \sigma(X(s)) \circ dw(s)\rangle$. Then for all $\varepsilon > 0$, we have*

$$P\left(\|M\|_T \geqslant \varepsilon\delta^{-\frac{1}{2}} \mid \|w\|_T < \delta\right) \to 0, \quad \delta \downarrow 0, \tag{39}$$

$$P\left(|K|_T \geqslant \varepsilon\delta^{-\frac{1}{2}} \mid \|w\|_T < \delta\right) \to 0, \quad \delta \downarrow 0. \tag{40}$$

**Proof.** Obviously

$$P\left(|K|_T \geqslant \varepsilon\delta^{-\frac{3}{2}} \mid \|w\|_T < \delta\right) \sim \frac{Ce^{-C\varepsilon\delta^{-3}}}{Ce^{-C\delta^{-2}}} \to 0, \quad \delta \downarrow 0. \tag{41}$$

Applying Ito's formula to $M_t$, we have

$$\begin{aligned}
M_t &= \sum_{i=1}^m \int_0^t \left(X^i(s) - a\right)\sigma_k^i\left(X(s)\right) \circ dw^k(s) \\
&= \sum_{i=1}^m \left(X^i(t) - a\right)\sigma_k^i\left(X(t)\right)w^k(t) - \sum_{i=1}^m \int_0^t w^k(s)\left(X^i(s) - a\right) \circ d\sigma_k^i\left(X(s)\right) \\
&\quad - \sum_{i=1}^m \int_0^t w^k(s)\sigma_k^i\left(X(s)\right) \circ d\left(X^i(s) - a\right) \\
&= \sum_{i=1}^m \left(X^i(t) - a\right)\sigma_k^i\left(X(t)\right)w^k(t) - \sum_{i=1}^m \int_0^t w^k(s)\left(X^i(s) - a\right)\sigma_{k,l}^i\left(X(s)\right) \\
&\quad \times \left[\sigma_\alpha^l\left(X(s)\right) \circ dw^\alpha(s) + b^l\left(X(s)\right)ds - dK^l(s)\right] \\
&\quad - \sum_{i=1}^m \int_0^t w^k(s)\sigma_k^i\left(X(s)\right)\left[\sigma_\alpha^i\left(X(s)\right) \circ dw^\alpha(s) + b^i\left(X(s)\right)ds - dK^i(s)\right] \\
&=: \sum_{j=1}^7 J_j.
\end{aligned} \tag{42}$$

We have to prove that $P(\|J_j\|_T \geqslant \varepsilon\delta^{-\frac{1}{2}} | \|w\|_T < \delta) \to 0$ as $\delta \downarrow 0$, for $j = 1, 2, \ldots, 7$.

Obviously, there is no problem for $J_1$, $J_3$ and $J_6$. For $J_2$, we have

$$\begin{aligned}
J_2 &= \sum_{i=1}^m \int_0^t w^k(s)\left(X^i(s) - a\right)\sigma_{k,l}^i\left(X(s)\right)\sigma_\alpha^l\left(X(s)\right) \circ dw^\alpha(s) \\
&= \sum_{i=1}^m \int_0^t w^k(s)\left(X^i(s) - a\right)\sigma_{k,l}^i\left(X(s)\right)\sigma_\alpha^l\left(X(s)\right) dw^\alpha(s)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2}\sum_{i=1}^{m}\int_0^t \big[\sigma_{k,l}^i\sigma_\alpha^l\sigma_\beta^i\big]\big(X(s)\big)w^k(s)\delta^{\alpha\beta}\,ds\\
&+\frac{1}{2}\sum_{i=1}^{m}\int_0^t \big(X^i(s)-a\big)\frac{\partial}{\partial x^n}\big[\sigma_{k,l}^i\sigma_\alpha^l\big]\big(X(s)\big)\sigma_\beta^n\big(X(s)\big)w^k(s)\delta^{\alpha\beta}\,ds\\
:={}& J_{21}+J_{22}+J_{23}.
\end{aligned}
$$

It is obvious that $J_{21}$ is a martingle with

$$
\langle J_{21},J_{21}\rangle_t=\sum_{i=1}^{m}\sum_{i'=1}^{m}\int_0^t \big(X^i(s)-a\big)\big(X^{i'}(s)-a\big)\big[\sigma_{k,l}^i\sigma_{k',l'}^{i'}a^{ll'}\big]\big(X(s)\big)w^k(s)w^{k'}(s)\,ds.
$$

Therefore, if $\|w\|_T<\delta$ then $\langle J_{21},J_{21}\rangle\leqslant C\delta^2$. Hence

$$
\begin{aligned}
P\big(\|J_{21}\|_T>\varepsilon\delta^{-\frac{1}{2}},\|w\|_T<\delta\big)&\leqslant P\Big(\max_{0\leqslant t\leqslant C\delta^2}\big|B(t)\big|>\varepsilon\delta^{-\frac{1}{2}}\Big)\\
&=P\Big(\max_{0\leqslant t\leqslant 1}\big|B(t)\big|>C\varepsilon\delta^{-\frac{3}{2}}\Big)\\
&=2\int_{C\varepsilon\delta^{-\frac{3}{2}}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\Big\{-\frac{x^2}{2}\Big\}\,dx\\
&\leqslant C\exp\Big\{-C\frac{\varepsilon^2}{\delta^3}\Big\}.
\end{aligned}
$$

By Lemma 5.1, $P(\|w\|_T\leqslant\delta)\sim C\exp\{-C\frac{1}{\delta^2}\}$. Hence

$$
P\big(\|J_{21}\|_T>\varepsilon\delta^{-\frac{1}{2}}\;\big|\;\|w\|_T<\delta\big)\to 0.
$$

Since $P(\|J_{22}\|_T>\varepsilon\delta^{-\frac{1}{2}}\mid\|w\|_T<\delta)\to 0$ and $P(\|J_{23}\|_T>\varepsilon\delta^{-\frac{1}{2}}\mid\|w\|_T<\delta)\to 0$ as $\delta\downarrow 0$, we have

$$
P\big(\|J_2\|_T>\varepsilon\delta^{-\frac{1}{2}}\;\big|\;\|w\|_T<\delta\big)\to 0.
$$

Similarly, we have

$$
P\big(\|J_5\|_T>\varepsilon\delta^{-\frac{1}{2}}\;\big|\;\|w\|_T<\delta\big)\to 0.
$$

Since $|J_4|\leqslant C\|w\|_T|K|_T$ and $|J_7|\leqslant C\|w\|_T|K|_T$, we obtain by (41)

$$
P\big(\|J_4\|_T>\varepsilon\delta^{-\frac{1}{2}}\;\big|\;\|w\|_T<\delta\big)\to 0
$$

and

$$
P\big(\|J_7\|_T>\varepsilon\delta^{-\frac{1}{2}}\;\big|\;\|w\|_T<\delta\big)\to 0.
$$

Summing up gives

$$
P\big(\|M\|_T\geqslant\varepsilon\delta^{-\frac{1}{2}}\;\big|\;\|w\|_T<\delta\big)\to 0,\quad \delta\downarrow 0. \tag{43}
$$

Combining (43) with

$$|K|_T \leqslant CT + 2 \sup_{0\leqslant t\leqslant T} \left| \int_0^t \langle X(s) - a, \sigma(X(s)) \circ dw(s) \rangle \right| \tag{44}$$

yields

$$P\left(|K|_T \geqslant \varepsilon\delta^{-\frac{1}{2}} \;\middle|\; \|w\|_T < \delta\right) \to 0, \quad \delta \downarrow 0. \tag{45}$$

This completes the proof. □

Now we have

**Lemma 5.6.** *For every $\varepsilon > 0$,*

$$\lim_{\delta\downarrow 0} P\left(\left\|\zeta^{km}\right\|_T |K|_T > \varepsilon \;\middle|\; \|w\|_T < \delta\right) = 0.$$

**Proof.**

$$\begin{aligned}
&P\left(\left\|\zeta^{km}\right\|_T |K|_T > \varepsilon \;\middle|\; \|w\|_T < \delta\right) \\
&\quad \leqslant P\left(\left\|\zeta^{km}\right\|_T > \delta^{\frac{1}{2}} \;\middle|\; \|w\|_T < \delta\right) + P\left(|K|_T > \varepsilon\delta^{-\frac{1}{2}} \;\middle|\; \|w\|_T < \delta\right) \\
&\quad \to 0. \qquad \square
\end{aligned}$$

It follows immediately:

**Lemma 5.7.**

$$P\left(\left\| \int_0^t \zeta^{km}(s)\, dK(s) \right\|_T > \varepsilon \;\middle|\; \|w\|_T < \delta\right) \to 0. \tag{46}$$

We also need the following

**Lemma 5.8.** *Let $f(x) : \mathbb{R}^m \to \mathbb{R}$ be bounded and uniformly continuous. Then for all $\varepsilon > 0$ and $k, m = 1, 2, \ldots, d$*

$$P\left(\left\| \int_0^t f(X(s))\, d\zeta^{km}(s) \right\|_T > \varepsilon \;\middle|\; \|w\|_T < \delta\right) \to 0 \quad \text{as } \delta \downarrow 0. \tag{47}$$

**Proof.** First we assume that $f \in C_b^2(\mathbb{R}^m)$. Set $f_l(x) := \frac{\partial}{\partial x_l} f(x)$. By Itô's formula we have

$$\begin{aligned}
\int_0^t f(X(s))\, d\zeta^{km}(s) = {} & f(X(t))\zeta^{km}(t) - \int_0^t \zeta^{km}(s) f_l(X(s)) \sigma_i^l(X(s))\, dw^i(s) \\
& - \int_0^t (Lf)(X(s))\zeta^{km}(s)\, ds - \int_0^t f_l(X(s)) \sigma_m^l(X(s)) w^k(s)\, ds
\end{aligned}$$

$$+\int_0^t f_l\big(X(s)\big)\zeta^{km}(s)\,dK^l(s)$$
$$:= P_1 + P_2 + P_3 + P_4 + P_5.$$

We now prove that

$$\lim_{\delta\downarrow 0} P\big(\|P_i\|_T > \varepsilon \,\big|\, \|w\|_T < \delta\big) = 0 \tag{48}$$

for $i = 1, \ldots, 5$. This is deduced plainly from (35) for $i = 1$, $i = 3$ and is trivial for $i = 4$. Since

$$\left|\int_0^t f_l\big(X(s)\big)\zeta^{km}(s)\,dK^l(s)\right| \leqslant C\big\|\zeta^{km}\big\|_T \cdot |K|_T,$$

we obtain by Lemma 5.6

$$\lim_{\delta\downarrow 0} P\big(\|P_5\|_T > \varepsilon \,\big|\, \|w\|_T < \delta\big) = 0. \tag{49}$$

It remains to look at $P_2$. Set $\alpha_i(x) := -f_l(x)\sigma_i^l(x)$, $\alpha_{i,l} := \frac{\partial}{\partial x^l}\alpha_i(x)$. Then by Itô's formula,

$$\begin{aligned}
P_2 &= \int_0^t \alpha_i\big(X(s)\big)\zeta^{km}(s)\,dw^i(s)\\
&= \alpha_i\big(X(t)\big)\zeta^{km}(t)w^i(t) - \int_0^t \alpha_{i,l}\big(X(s)\big)\sigma_j^l\big(X(s)\big)\zeta^{km}(s)w^i(s)\,dw^j(s)\\
&\quad - \int_0^t (L\alpha_i)\big(X(s)\big)\zeta^{km}(s)w^i(s)\,ds - \int_0^t \alpha_i\big(X(s)\big)w^i(s)\,d\zeta^{km}(s)\\
&\quad - \int_0^t \alpha_m\big(X(s)\big)w^k(s)\,ds - \int_0^t \zeta^{km}(s)\alpha_{i,l}\big(X(s)\big)\sigma_j^l\big(X(s)\big)\delta^{ij}\,ds\\
&\quad - \int_0^t \alpha_{i,l}\big(X(s)\big)\sigma_m^l\big(X(s)\big)w^i(s)w^k(s)\,ds + \int_0^t \alpha_{i,l}\big(X(s)\big)\zeta^{km}(s)w^i(s)\,dK^l(s)\\
&:= \sum_{i=1}^8 L_i.
\end{aligned}$$

Again it is sufficient to show that $P(\|L_i\|_T > \varepsilon \mid \|w\|_T < \delta) \to 0$. The proof can be done in a similar way to the estimation of $I_2$ in [11, pp. 522–524] except an extra term $L_8$. But $L_8$ can be estimated easily by using Lemma 5.6.

The passage from $C_b^2$ functions to bounded and uniformly continuous functions is completely the same as on [11, p. 525] and so we omit it. □

**Lemma 5.9.** *Let $f$ be a $C_b^2$-function. Then for $k=1,\dots,d$,*

$$\lim_{\delta\downarrow 0} P\left(\left\|\int_0^t f\big(X(s)\big)\circ dw^k(s)\right\|_T > \varepsilon \,\Big|\, \|w\|_T < \delta\right) = 0.$$

**Proof.** Set $f_l := \partial_l f$. By Ito's formula we have

$$\begin{aligned}\int_0^t f\big(X(s)\big)\circ dw^k(s) &= f\big(X(t)\big)\circ dw^k(t) - \int_0^t \big[f_l\sigma_m^l\big]\big(X(s)\big)w^k(s)\circ dw^m(s)\\ &\quad - \int_0^t \big[f_l b^l\big]\big(X(s)\big)w^k(s)\,ds - \int_0^t f_l\big(X(s)\big)w^k(s)\,dK^l(s)\\ &= I_1(t)+I_2(t)+I_3(t)+I_4(t).\end{aligned}$$

It is sufficient to prove

$$\lim_{\delta\downarrow 0} P\big(\big\|I_i(t)\big\|_T > \varepsilon \,\big|\, \|w\|_T < \delta\big) = 0, \quad \forall i = 1,2,3,4.$$

$I_1$ and $I_3$ are trivial and $I_4$ will make no trouble by Lemma 5.5. It remains to look at $I_2$. We have

$$\begin{aligned}I_2(t) &= -\int_0^t \big[f_l\sigma_m^l\big]\big(X(s)\big)\circ d\zeta^{km}(s)\\ &= -\int_0^t \big[f_l\sigma_m^l\big]\big(X(s)\big)\,d\zeta^{km}(s) - \frac{1}{2}\int_0^t \frac{\partial}{\partial x_j}\big[f_l\sigma_m^l\big]\sigma_q^j\big(X(s)\big)w^k(s)\delta^{qm}\,ds\\ &= J_1(t)+J_2(t).\end{aligned}$$

We claim that

$$\lim_{\delta\downarrow 0} P\big(\big\|J_i(t)\big\|_T > \varepsilon \,\big|\, \|w\|_T < \delta\big) = 0, \quad \forall i = 1,2.$$

For $i=2$ this is obvious and for $i=1$ this follows from Lemma 5.8. The proof is now complete. □

Now we are in the position to state our main result of this section.

**Theorem 5.10.** $\forall h \in \mathscr{S}$ *and* $\varepsilon > 0$

$$P\big(\big\|X(t)-\xi(t)\big\|_T + \big|K(t)-\eta(t)\big|_T < \varepsilon \,\big|\, \|w-h\|_T < \delta\big) \to 1 \quad \text{as } \delta \downarrow 0. \tag{50}$$

**Proof.** By the standard argument (see [18] or [11, pp. 527–528]) it suffices to prove (50) for $h \equiv 0$. Then we have

$$X(t)-\xi(t) = \int_0^t \sigma\big(X(s)\big)\circ dw(s) + \int_0^t \big(b\big(X(s)\big)-b\big(\xi(s)\big)\big)\,ds - \int_0^t \big(dK(s)-\eta(s)\,ds\big).$$

Set $\psi(x) := 1 - e^{-|x|^2}$, $\psi_i(x) := \frac{\partial}{\partial x_i}\psi(x)$, $\psi_{i,l}(x) := \frac{\partial}{\partial x_l}\psi_i(x)$. Then $\psi \in C_b^2(\mathbb{R}^m)$ and there exists $C > 0$ such that $|\psi_i(x)| = |2x_i e^{-|x|^2}| \leqslant C\psi(x)$ and $|\psi_{i,l}(x)| \leqslant C\psi(x)$. Set again

$$G(t) := X(t) - \xi(t).$$

Since $\langle X(s) - \xi(s), dK(s) - \eta(s)ds\rangle \geqslant 0$, we can have

$$\begin{aligned}
\psi\big(G(t)\big) &= \int_0^t \psi_i\big(G(s)\big)\sigma_k^i\big(X(s)\big) \circ dw^k(s) \\
&\quad + \int_0^t \psi_i\big(G(s)\big)\big(b^i\big(X(s)\big) - b^i\big(\xi(s)\big)\big)\,ds \\
&\quad - 2\int_0^t e^{-|G(s)|^2}\big\langle X(s) - \xi(s), dK(s) - \eta(s)\,ds\big\rangle \\
&\leqslant \int_0^t \psi_i\big(G(s)\big)\sigma_k^i\big(X(s)\big) \circ dw^k(s) + C\int_0^t \psi\big(G(s)\big)\,ds \\
&:= I + C\int_0^t \psi\big(G(s)\big)\,ds.
\end{aligned}$$

Let $\sigma_{k,l}^i(x) := \frac{\partial}{\partial x_l}\sigma_k^i(x)$, then

$$\begin{aligned}
I &= \int_0^t \psi_i\big(G(s)\big)\sigma_k^i\big(X(s)\big) \circ dw^k(s) \\
&= \psi_i\big(G(t)\big)\sigma_k^i\big(X(t)\big)w^k(t) - \int_0^t w^k(s) \circ d\big[\psi_i\big(G(s)\big)\sigma_k^i\big(X(s)\big)\big] \\
&= \psi_i\big(G(t)\big)\sigma_k^i\big(X(t)\big)w^k(t) - \int_0^t \big[w^k(s)\psi_i\big(G(s)\big)\big] \circ d\sigma_k^i\big(X(s)\big) \\
&\quad - \int_0^t \big[w^k(s)\sigma_k^i\big(X(s)\big)\big] \circ d\big(\psi_i\big(G(s)\big)\big) \\
&= \psi_i\big(G(t)\big)\sigma_k^i\big(X(t)\big)w^k(t) \\
&\quad - \int_0^t \psi_i\big(G(s)\big)\sigma_{k,l}^i\big(X(s)\big)\sigma_m^l\big(X(s)\big)w^k(s) \circ dw^m(s) \\
&\quad - \int_0^t \psi_i\big(G(s)\big)\sigma_{k,l}^i\big(X(s)\big)b^l\big(X(s)\big)w^k(s)\,ds
\end{aligned}$$

$$
\begin{aligned}
&+\int_0^t \psi_i\big(G(s)\big)\sigma^i_{k,l}\big(X(s)\big)w^k(s)\,dK^l(s)\\
&-\int_0^t \psi_{i,l}\big(G(s)\big)\sigma^i_k\big(X(s)\big)\sigma^l_m\big(X(s)\big)w^k(s)\circ dw^m(s)\\
&-\int_0^t \psi_{i,l}\big(G(s)\big)\sigma^i_k\big(X(s)\big)w^k(s)\big[b^l\big(X(s)\big)-b^l\big(\xi(s)\big)\big]\,ds\\
&+\int_0^t \psi_{i,l}\big(G(s)\big)\sigma^i_k\big(X(s)\big)w^k(s)\big[dK^l(s)-\eta(s)\,ds\big]\\
&:= I_1+I_2+I_3+I_4+I_5+I_6+I_7.
\end{aligned}
$$

Obviously,

$$
\begin{aligned}
&I_1\leqslant C\|w\|_T,\qquad I_3\leqslant C\|w\|_T,\qquad I_4\leqslant C\|w\|_T|K|_T,\\
&I_6\leqslant C\int_0^t \psi\big(G(s)\big)\|w\|_T\,ds\leqslant C\|w\|_T
\end{aligned}
$$

and

$$
I_7\leqslant C\int_0^t \big|w(s)\big|\,d|K|^0_s+C\int_0^t \big|w(s)\big|\big|\eta(s)\big|\,ds\leqslant C\|w\|_T|K|_T+C\|w\|_T.
$$

We need to estimate $I_2$ and $I_6$. Clearly,

$$
\begin{aligned}
I_2&=-\int_0^t \psi_i\big(G(s)\big)\sigma^i_{k,l}\big(X(s)\big)\sigma^l_m\big(X(s)\big)\circ d\zeta^{km}(s)\\
&=-\int_0^t \psi_i\big(G(s)\big)\sigma^i_{k,l}\big(X(s)\big)\sigma^l_m\big(X(s)\big)\,d\zeta^{km}(s)\\
&\quad-\frac{1}{2}\int_0^t \frac{\partial}{\partial x_n}\big[\sigma^i_{k,l}\sigma^l_m\big]\big(X(s)\big)\psi_i\big(G(s)\big)\sigma^n_\alpha\big(X(s)\big)w^k(s)\delta^{\alpha m}\,ds\\
&\quad-\frac{1}{2}\int_0^t \big[\sigma^i_{k,l}\sigma^l_m\big]\big(X(s)\big)\psi_{i,j}\big(G(s)\big)\sigma^j_\alpha\big(X(s)\big)w^k(s)\delta^{\alpha m}\,ds\\
&\leqslant-\int_0^t \psi_i\big(G(s)\big)\sigma^i_{k,l}\big(X(s)\big)\sigma^l_m\big(X(s)\big)\,d\zeta^{km}(s)+C\|w\|_T
\end{aligned}
$$

and

$$\begin{aligned}
I_5 &= -\int_0^t \psi_{i,l}\big(G(s)\big)\sigma_k^i\big(X(s)\big)\sigma_m^l\big(X(s)\big) \circ d\zeta^{km}(s) \\
&= -\int_0^t \psi_{i,l}\big(G(s)\big)\sigma_k^i\big(X(s)\big)\sigma_m^l\big(X(s)\big)\, d\zeta^{km}(s) \\
&\quad - \frac{1}{2}\int_0^t \frac{\partial}{\partial x_n}\big[\sigma_k^i\sigma_m^l\big]\big(X(s)\big)\psi_{i,l}\big(G(s)\big)\sigma_\alpha^n\big(X(s)\big)w^k(s)\delta^{\alpha m}\, ds \\
&\quad - \frac{1}{2}\int_0^t \big[\sigma_k^i\sigma_m^l\big]\big(X(s)\big)\psi_{i,l,j}\big(G(s)\big)\sigma_\alpha^j\big(X(s)\big)w^k(s)\delta^{\alpha m}\, ds \\
&\leqslant -\int_0^t \psi_{i,l}\big(G(s)\big)\sigma_k^i\big(X(s)\big)\sigma_m^l\big(X(s)\big)\, d\zeta^{km}(s) + C\|w\|_T.
\end{aligned}$$

By all the above, we can get that

$$\begin{aligned}
\psi\big(G(t)\big) &\leqslant C\int_0^t \psi\big(G(s)\big)\, ds + C\|w\|_T|K|_T + C\|w\|_T \\
&\quad - \int_0^t \psi_i\big(G(s)\big)\sigma_{k,l}^i\big(X(s)\big)\sigma_m^l\big(X(s)\big)\, d\zeta^{km}(s) \\
&\quad - \int_0^t \psi_{i,l}\big(G(s)\big)\sigma_k^i\big(X(s)\big)\sigma_m^l\big(X(s)\big)\, d\zeta^{km}(s) \\
&:= C\int_0^t \psi\big(G(s)\big)\, ds + \sum_{i=1}^4 A_i.
\end{aligned}$$

Obviously, for every $\varepsilon > 0$, $P(\|A_i\|_T > \varepsilon | \|w\|_T < \delta) \to 0$ as $\delta \downarrow 0$ holds for $i = 1$ and $i = 2$. By Lemma 5.8, for $i = 3$ and $i = 4$ we can have

$$P\big(\|A_i\|_T > \varepsilon \bigm| \|w\|_T < \delta\big) \to 0 \quad \text{as } \delta \downarrow 0. \tag{51}$$

On the set $\{\omega;\ \|A_i\|_T < \varepsilon,\ i = 1, \dots, 4\}$, we have

$$\psi\big(G(t)\big) \leqslant \varepsilon \exp\{CT\} \leqslant C\varepsilon,$$

that is

$$\big|X(t) - \xi(t)\big| \leqslant \sqrt{-\ln(1 - C\varepsilon)}.$$

Combining this with (51), we get that

$$P\big(\|X - \xi\|_T > \varepsilon \bigm| \|w\|_T < \delta\big) \to 0 \quad \text{as } \delta \downarrow 0.$$

Finally, to see

$$P\big(|K-\eta|_T<\varepsilon \,\big|\, \|w-h\|_T<\delta\big)\to 1 \quad \text{as } \delta\downarrow 0,$$

it suffices to notice that

$$K(t)-\eta(t)=X(t)-\xi(t)+\int_0^t \sigma\big(X(s)\big)\,dw(s)+\int_0^t \big(b\big(X(s)\big)-b\big(\xi(s)\big)\big)\,ds$$

and use Lemma 5.9. □

## References


[1] J.-P. Aubin, A. Cellina, Differential Inclusions, Grundlehren Math. Wiss., vol. 264, Springer-Verlag, Berlin–Heidelberg, 1984.

[2] V. Barbu, Nonlinear Semigroups and Differential Equations in Banach Spaces, Noordhoff Internet Publishing, Leyden, Netherlands, 1976.

[3] V. Barbu, G. Da Prato, L. Tubaro, Kolmogorov equation associated to the stochastic reflection problem on a smooth convex set of a Hilbert space, Ann. Probab. 37 (4) (2009) 1427–1458.

[4] A. Bensoussan, A. Rascanu, Stochastic variational inequalities in infinite-dimensional spaces, Numer. Funct. Anal. Optim. 18 (1–2) (1997) 19–54.

[5] H. Brezis, Opérateurs Monotones et Semi-Groups de Contractions dans les Espaces de Hilbert, North-Holland, Amsterdam, 1973.

[6] N.L. Carothers, Real Analysis, Cambridge Univ. Press, 2000.

[7] E. Cépa, Équations différentielles stochasticcques multivoques, in: Lecture Notes in Math. Sém. Prob., vol. XXIX, Springer, Berlin, 1995, pp. 86–107.

[8] E. Cépa, Problème de Skorohod multivoque, Ann. Probab. 26 (2) (1998) 500–532.

[9] E. Cépa, S. Jacquot, Ergodicité d'inégalités variationnelles stochastiques, Stochastics and Stochastics Reports 63 (1997) 41–64.

[10] H. Doss, P. Priouret, Support d'un processus de réflexion, Z. Wahrsch. Verw. Gebiete 61 (3) (1982) 327–345.

[11] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland Publishing Company, 1989.

[12] J. Jacod, A.N. Shiryaev, Limit Theorems for Stochastic Processes, Springer-Verlag, Berlin, 1987.

[13] T. Kurtz, Random time change and convergence in distribution under the Meyer–Zheng conditions, Ann. Probab. 19 (3) (1991) 1010–1034.

[14] T. Nakayama, Support theorem for mild solutions of SDE's in Hilbert spaces, J. Math. Sci. Univ. Tokyo 11 (3) (2004) 245–311.

[15] E. Pardoux, A. Rascanu, Backward stochastic differential equations with subdifferential operator and related variational inequalities, Stochastic Process. Appl. 76 (2) (1998) 191–215.

[16] R. Pettersson, Wong–Zakai approximations for reflecting stochastic differential equations, Stoch. Anal. Appl. 17 (1999) 609–617.

[17] J. Ren, S. Xu, A transfer principle for multivalued stochastic differential equations, J. Funct. Anal. 256 (2009) 2780–2814.

[18] D.W. Stroock, S.R.S. Varadhan, On the support of diffusion processes with applications to the strong maximum principle, in: Proc. Sixth Berkeley Symp. Math. Statist. Prob. III, Univ. California Press, Berkeley, 1972, pp. 333–359.

[19] S. Xu, Denjoy's Approximate continuity for the solutions of multivalued stochastic differential equations, Stoch. Dyn. 1 (1) (2010) 1–23.