Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER Stochastic Processes and their Applications 114 (2004) 331–350
stochastic processes and their applications
www.elsevier.com/locate/spa

An approximation result for a nonlinear Neumann boundary value problem via BSDEs☆

B. Boufoussi[a], J. van Casteren[b,*]

[a] *Department of Mathematics, Faculty of Sciences Semlalia, Cadi Ayyad University 2390 Marrakesh, Morocco*
[b] *Department of Mathematics and Computer Science, University of Antwerp, Middelheimlaan 1, 2020 Antwerp, Belgium*

Received 16 November 2002; received in revised form 14 October 2003; accepted 14 June 2004

Available online 6 July 2004

Abstract

We prove a weak convergence result for a sequence of backward stochastic differential equations related to a semilinear parabolic partial differential equation; under the assumption that the diffusion corresponding to the PDEs is obtained by penalization method converging to a normal reflected diffusion on a smooth and bounded domain D. As a consequence we give an approximation result to the solution of semilinear parabolic partial differential equations with nonlinear Neumann boundary conditions. A similar result in the linear case was obtained by Lions et al. in 1981.

MSC: 60H99; 35K20

Keywords: Backward stochastic differential equation; Reflected diffusion

1. Introduction

Let $T > 0$ and let $(X_t^n, 0 \leqslant t \leqslant T)$ be a diffusion in $\mathbb{R}^d$, $d \geqslant 1$, starting at $x \in \bar{D}$ with drift coefficient $b(x) - n\delta(x)$ and diffusion term $\sigma(x)$. The functions σ and b are given, δ is a penalization factor and D is a regular convex and bounded domain in $\mathbb{R}^d$. Lions

☆ The first author is obliged to the University of Antwerp for its financial and material support. Both authors thank the Flemish Fund for Scientific Research (FWO Flanders) for its financial support. An anonymous referee is acknowledged for pointing out a serious error in an earlier version, and for a significant improvement of the overall presentation of the paper.
* Corresponding author. Tel.: +32-3-265-3889; fax: +32-3-265-3777.
E-mail addresses: boufoussi@ucam.ac.ma (B. Boufoussi), jan.vancasteren@ua.ac.be (J. van Casteren).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2004.06.003

et al. [6] (see also Menaldi [8]) proved that X^n converges to a $\bar{D}$-valued diffusion X which is normally reflected on the boundary of D. The convergence is established uniformly in $0 \leqslant t \leqslant T$, $x \in \bar{D}$ and in the L^p-norms of the probability space on which the process X^n is defined. In a more analytic point of view (cf. [6]), this means the following. If we consider the Cauchy problem

$$\frac{\partial u^n}{\partial t}(t,x) + Lu^n(t,x) - n\delta(x)\nabla u^n(t,x) = 0, \quad (t,x) \in (0,T] \times \mathbb{R}^d,$$

$$u^n(0,x) = u_0(x),$$

where L is the infinitesimal generator corresponding to the diffusion part of X and u_0 is a given regular initial condition, then for any $(t,x) \in [0,T] \times \bar{D}$, $u^n(t,x)$ converges to $u(t,x)$, which is the unique solution to the following Neumann problem:

$$\frac{\partial u}{\partial t}(t,x) + Lu(t,x) = 0, \quad (t,x) \in (0,T] \times D,$$

$$u(0,x) = u_0(x), \quad x \in D,$$

$$\frac{\partial u}{\partial n}(t,x) = 0, \quad (t,x) \in (0,T] \times \partial D.$$

This article is devoted to prove a similar result for systems of semi-linear partial differential equations—PDEs—with a nonlinear Neumann boundary condition h satisfying Lipschitz and linear growth conditions. For this purpose we use the theory of backward stochastic differential equations—in short BSDEs.

The concrete formulation of BSDEs was first introduced by Pardoux and Peng [11], who proved existence and uniqueness of adapted solutions for these equations, under suitable square-integrability assumptions on the coefficients and on the terminal condition. They provide probabilistic formulas for solution of systems of semilinear partial differential equations, both of parabolic and elliptic type; which can be considered as a generalization of the well known Feynman–Kac formula to semi-linear partial differential equations. The interest for this kind of stochastic equations has increased steadily, since it has been widely recognized that they provide a useful framework for formulating many problems in mathematical finance (see for example [2]), and they also appear to be useful in stochastic control and differential games (cf. [3]). Recently a new class of BSDEs, involving an integral with respect to a continuous increasing process, was studied by Pardoux and Zhang in [13]. They use this class of BSDEs to give a probabilistic formula for the solution of a system of parabolic or elliptic semi-linear partial differential equation with Neumann boundary condition.

To describe our result more precisely, we first recall some classical notions that will be used in this paper. We introduce the function $\rho \in \mathscr{C}_b^1(\mathbb{R}^d)$ such that $\rho = 0$ in $\bar{D}$, $\rho > 0$ in $\mathbb{R}^d \setminus \bar{D}$ and $\rho(x) = (d(x,\bar{D}))^2$ in a neighborhood of $\bar{D}$. On the other hand, since the domain D is smooth (say $\mathscr{C}^3$), it is possible to consider an extension $\phi \in \mathscr{C}_b^2(\mathbb{R}^d)$ of the function $d(\cdot,\partial D)$ defined on the restriction to D of a neighborhood of ∂D such that D and ∂D are characterized by

$$D = \{x \in \mathbb{R}^d : \phi(x) > 0\} \quad \text{and} \quad \partial D = \{x \in \mathbb{R}^d : \phi(x) = 0\}$$

and for all $x \in \partial D$, $\nabla\phi(x)$ coincides with the unit normal pointing toward the interior of D (see for example [7, Remark 3.1]). In particular we may and do choose ρ and ϕ such that

$$\langle \nabla\phi(x), \delta(x)\rangle \leqslant 0, \quad \text{for all } x \in \mathbb{R}^d, \tag{1.1}$$

where $\delta(x) := \nabla\rho(x)$ which we call the penalization term. We will consider the following sequence of semi-linear partial differential equations ($1 \leqslant i \leqslant k$, $0 \leqslant t \leqslant T$, $x \in \mathbb{R}^d$, $n \in \mathbb{N}$).

$$\frac{\partial u_i^n}{\partial t}(t,x) + Lu_i^n(t,x) + f_i(t,x,u^n(t,x))$$
$$-n\delta(x)\cdot(\nabla u_i^n(t,x) + \nabla\phi(x)h_i(t,x,u^n(t,x))) = 0,$$
$$u^n(T,x) = g(x).$$

Suppose that (t,x) belongs to $[0,T]\times\bar{D}$. Under suitable linear growth and Lipschitz conditions on the coefficients f, g and h, we will show—by using the connection of BSDEs with semi-linear PDEs—that the sequence $u^n(t,x)$ converges, as n goes to infinity, to a function $u(t,x)$, which is the solution, in viscosity sense, to the following PDE with Neumann boundary condition:

$$\frac{\partial u_i}{\partial t}(t,x) + Lu_i(t,x) + f_i(t,x,u(t,x)) = 0, \quad 1 \leqslant i \leqslant k,\ (t,x) \in [0,T)\times D,$$
$$u(T,x) = g(x),\ x \in D,$$
$$\frac{\partial u}{\partial n}(t,x) + h(t,x,u(t,x)) = 0, \quad \forall (t,x) \in [0,T)\times\partial D.$$

The case with Neumann boundary condition $h = 0$ was treated in [1]. This note is organized as follows. In the second section we recall, without proof, results due to Menaldi [8] for the penalization method to approach the reflected diffusion in a convex bounded domain D and some basic notions on BSDEs. Section 3 is devoted to prove our main result.

2. Assumptions and basic notions

2.1. Approximation of a reflected diffusion process

Throughout this paper $(B_t : t \geqslant 0)$ is an r-dimensional Brownian motion defined on a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and for $t \geqslant 0$, $\mathscr{F}_t$ is the σ-algebra $\sigma(B_s : s \leqslant t)$, augmented with the $\mathbb{P}$-null sets of $\mathscr{F}$. Let the domain D be an open subset of $\mathbb{R}^d$ ($d \geqslant 1$). Several authors have studied approximations of reflected diffusions in such domains. We refer for example to the paper by Menaldi [8] for the case of a convex and bounded domain D. The non-convex case was treated by Lions and Sznitman in [7]. For a bounded domain the construction by approximation of the stationary reflected Brownian motion can be found in the work by Williams and Zheng [14]. Afterwards this was generalized to reflected diffusions on not necessarily bounded domains by

Pardoux and Williams [12]. Here, we consider a class of reflected diffusion processes on a regular convex and bounded domain $D \subset \mathbb{R}^d$.

Let $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times r}$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ be uniformly bounded functions and satisfying the uniform Lipschitz condition with finite constant $C_0 > 0$

$$\forall x, y \in \mathbb{R}^d, \quad \|\sigma(x) - \sigma(y)\| + |b(x) - b(y)| \leqslant C_0|x - y|.$$

From Lions and Sznitman [7] we know that for every $x \in \bar{D}$ there exists a unique pair of progressively measurable processes (X_t, K_t) which is a solution to the following problem:

(i) $\mathbb{P}\{X \in \bar{D}\} = 1$, K has bounded variation on $[0, T]$, $0 < T < \infty$, $K_0 = 0$;

$$\text{(ii) } X_t = x + \int_0^t \sigma(X_s)\,\mathrm{d}B_s + \int_0^t b(X_s)\,\mathrm{d}s + K_t; \tag{2.1}$$

(iii) for all continuous and progressively measurable processes α_t attaining values in $\bar{D}$, the inequality $\int_0^T (X_t - \alpha_t)\,\mathrm{d}K_t \leqslant 0$ is valid.

In particular we have

$$K_t = \int_0^t \nabla\phi(X_s)\,\mathrm{d}k_s \quad \text{and} \quad \int_0^t I_{\{X_s \in D\}}\,\mathrm{d}K_s = 0, \tag{2.2}$$

where k_t stands for the total variation of K on the interval $[0, t]$; it is an increasing continuous process on the interval $[0, T]$. The meaning of the function ϕ has been made precise in Section 1.

We recall an approximation procedure of the reflected diffusion X by a penalization method. Fix $n \geqslant 1$, and consider the diffusion process $\{X_t^n : 0 \leqslant t \leqslant T\}$ in $\mathbb{R}^d$ given by the equation

$$\text{(i) } X_t^n = x + \int_0^t \sigma(X_s^n)\,\mathrm{d}B_s + \int_0^t b(X_s^n)\,\mathrm{d}s + K_t^n, \ t \geqslant 0;$$

$$\text{(ii) } x \in \bar{D}, \tag{2.3}$$

where $\delta(x)$ is defined as in the introduction and $K_t^n := -n\int_0^t \delta(X_s^n)\,\mathrm{d}s$. We refer to Menaldi [8] for the proof of the following convergence result. It is assumed that the stochastic processes X_s, X_s^n, $0 \leqslant s \leqslant T$, $n \in \mathbb{N}$, are defined on the same probability space, and that all of them are adapted with respect to a Brownian motion B_s, $0 \leqslant s \leqslant T$. Unless stated otherwise, this Brownian motion remains the same throughout the paper. In fact a similar remark applies to the processes Y_s, Y_s^n, Z_s, and Z_s^n, $0 \leqslant s \leqslant T$: see e.g. (2.7).

Theorem 2.1. *The problem in* (2.1) *has a unique solution. Moreover, for every* $1 \leqslant p < \infty$, $0 < T < \infty$, *the quantity* $\mathbb{E}[\sup_{0\leqslant s\leqslant T}|X_s^n - X_s|^p]$ *converges to* 0 *as* $n \to \infty$; *the limit is uniform in* x *for* $x \in \bar{D}$.

Remark 2.1. In particular, Menaldi shows that, for all $1 \leqslant p < \infty$,

$$\sup_{n\geqslant 1} \mathbb{E}\left[\sup_{0\leqslant s\leqslant T} |X_s^n|^p\right] < \infty \text{ and } \sup_{n\geqslant 1} \mathbb{E}\left[\left(n\int_0^T |\delta(X_s^n)|\,\mathrm{d}s\right)^p\right] < \infty. \tag{2.4}$$

In the proof of the identification of the limit, and in the proof of Lemma 2.2 the results in (2.4) will be used.

Assertion (i) of the following lemma is applied for $p = 1$ and 2:

Lemma 2.2. *Let K_t^n and K_t be as in* (2.1) *and* (2.3), *respectively. The following assertions are true*:

(i) $\lim_{n\to\infty} \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |K_t^n - K_t|^p\right] = 0,\ 1 \leqslant p < \infty;$

(ii) $\lim_{n\to\infty} \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} \left|\int_0^t \varphi(X_s^n)\,\mathrm{d}K_s^n - \int_0^t \varphi(X_s)\,\mathrm{d}K_s\right|^2\right] = 0,$ *for all* $\varphi \in \mathscr{C}_b^1(\mathbb{R}^d)$.

Proof. Assertion (i) is a simple consequence of Theorem 2.1, Lipschitz conditions on the coefficients b and σ together with an application of Burkholder–Davis–Gundy inequality. From assertion (i) it follows that the process $t \mapsto K_t$ is $\mathbb{P}$-almost surely continuous; so the same is true for the process $t \mapsto X_t$.

For a proof of assertion (ii) we proceed as follows. Let φ belong to $\mathscr{C}_b^1(\mathbb{R}^d)$, and fix a C^1-function $\eta : \mathbb{R} \to [0,\infty)$ with support in the open interval $(0,\delta)$, and with $\int_0^\delta \eta(\sigma)\,\mathrm{d}\sigma = 1$. Then we have

$$
\begin{aligned}
&\int_0^t \varphi(X_s^n)\,\mathrm{d}K_s^n - \int_0^t \varphi(X_s)\,\mathrm{d}K_s \\
&\quad = \int_0^t (\varphi(X_s^n) - \varphi(X_s))\,\mathrm{d}K_s^n \\
&\qquad + \int_0^{(t-\delta)\vee 0} \left(\int_0^\delta (\varphi(X_s) - \varphi(X_{s+\sigma}))\eta(\sigma)\,\mathrm{d}\sigma\right) \mathrm{d}(K_s^n - K_s) \\
&\qquad + \int_{(t-\delta)\vee 0}^t (\varphi(X_s) - \varphi(X_t))\,\mathrm{d}(K_s^n - K_s) \\
&\qquad + \varphi(X_t)(K_t^n - K_t - K_{(t-\delta)\vee 0}^n + K_{(t-\delta)\vee 0}) \\
&\qquad + \int_0^{(t-\delta)\vee 0} \int_s^{s+\delta} \varphi(X_\sigma)\eta(\sigma - s)\,\mathrm{d}\sigma\,\mathrm{d}(K_s^n - K_s) \\
&\qquad \text{(integration by parts : } K_0^n - K_0 = 0) \\
&\quad = \int_0^t (\varphi(X_s^n) - \varphi(X_s))\,\mathrm{d}K_s^n \\
&\qquad + \int_0^{(t-\delta)\vee 0} \left(\int_0^\delta (\varphi(X_s) - \varphi(X_{s+\sigma}))\eta(\sigma)\,\mathrm{d}\sigma\right) \mathrm{d}(K_s^n - K_s) \\
&\qquad + \int_{(t-\delta)\vee 0}^t (\varphi(X_s) - \varphi(X_t))\,\mathrm{d}(K_s^n - K_s)
\end{aligned}
$$

$$+\varphi(X_t)(K_t^n - K_t - K_{(t-\delta)\vee 0}^n + K_{(t-\delta)\vee 0})$$

$$+\int_{(t-\delta)\vee 0}^{t} \varphi(X_\sigma)\eta(\sigma - ((t-\delta)\vee 0))\,\mathrm{d}\sigma(K_{(t-\delta)\vee 0}^n - K_{(t-\delta)\vee 0})$$

$$+\int_0^{(t-\delta)\vee 0}\int_s^{s+\delta} \varphi(X_\sigma)\eta'(\sigma - s)\,\mathrm{d}\sigma(K_s^n - K_s)\,\mathrm{d}s.$$

The elementary inequality $|\sum_{j=1}^6 a_j|^2 \leqslant 6\sum_{j=1}^6 |a_j|^2$, $a_j \in \mathbb{R}$, $1 \leqslant j \leqslant 6$, yields

$$\begin{aligned}
&\frac{1}{6}\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\int_0^t \varphi(X_s^n)\,\mathrm{d}K_s^n - \int_0^t \varphi(X_s)\,\mathrm{d}K_s\right|^2 \\
&\leqslant \mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\int_0^t (\varphi(X_s^n) - \varphi(X_s))\,\mathrm{d}K_s^n\right|^2 \\
&\quad + \mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\int_0^{(t-\delta)\vee 0}\left(\int_0^\delta (\varphi(X_s) - \varphi(X_{s+\sigma}))\eta(\sigma)\,\mathrm{d}\sigma\right)\mathrm{d}(K_s^n - K_s)\right|^2 \\
&\quad + \mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\int_{(t-\delta)\vee 0}^t (\varphi(X_s) - \varphi(X_t))\mathrm{d}(K_s^n - K_s)\right|^2 \\
&\quad + \mathbb{E}\sup_{0\leqslant t\leqslant T}|\varphi(X_t)(K_t^n - K_t - K_{(t-\delta)\vee 0}^n + K_{(t-\delta)\vee 0})|^2 \\
&\quad + \mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\int_{(t-\delta)\vee 0}^t \varphi(X_\sigma)\eta(\sigma - ((t-\delta)\vee 0))\,\mathrm{d}\sigma(K_{(t-\delta)\vee 0}^n - K_{(t-\delta)\vee 0})\right|^2 \\
&\quad + \mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\int_0^{(t-\delta)\vee 0}\int_s^{s+\delta}\varphi(X_\sigma)\eta'(\sigma - s)\,\mathrm{d}\sigma(K_s^n - K_s)\,\mathrm{d}s\right|^2.
\end{aligned} \tag{2.5}$$

For the first term in the right-hand side of (2.5) we have, by Hölder's inequality and the fact that φ is a member of $\mathscr{C}_b^1(\mathbb{R}^d)$ with Lipschitz constant C

$$\begin{aligned}
&\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\int_0^t (\varphi(X_s^n) - \varphi(X_s))\,\mathrm{d}K_s^n\right|^2 \\
&\leqslant C^2\left(\mathbb{E}\sup_{0\leqslant s\leqslant T}|X_s^n - X_s|^4\right)^{1/2}\left(\sup_{n\in\mathbb{N}}\mathbb{E}\left(n\int_0^T |\delta(X_s^n)|\,\mathrm{d}s\right)^4\right)^{1/2},
\end{aligned}$$

where we used the representation $K_t^n = -n\int_0^t \delta(X_s^n)\,\mathrm{d}s$. Theorem 2.1 together with Remark 2.1 implies

$$\lim_{n\to\infty}\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\int_0^t (\varphi(X_s^n) - \varphi(X_s))\,\mathrm{d}K_s^n\right|^2 = 0.$$

The second term in the right-hand side of (2.5) is dominated by

$$\mathbb{E} \sup_{0\leqslant s\leqslant T-\delta} \sup_{0\leqslant \sigma<\delta} |\varphi(X_s) - \varphi(X_{s+\sigma})|^2 (K_T^n + K_T)^2$$

$$\leqslant \left(\mathbb{E} \sup_{0\leqslant s\leqslant T-\delta} \sup_{0\leqslant \sigma<\delta} |\varphi(X_s) - \varphi(X_{s+\sigma})|^4 \right)^{1/2} \left(\sup_{n\in\mathbb{N}} \mathbb{E}(K_T^n + K_T)^4 \right)^{1/2}.$$

Since the process $s \mapsto X_s$ is right-continuous, and the function φ is bounded and continuous, an appeal to Remark 2.1 shows that the second term converges to 0 (uniformly in n), whenever δ tends to 0. A similar argument, but now using the fact that the process $s \mapsto X_s$ possesses left limits, shows that the third term in the right-hand side of (2.5) tends, uniformly in n, to 0 as well whenever δ goes to 0. The sum of fourth and the fifth term is dominated by

$$5\|\varphi\|_\infty^2 \mathbb{E} \sup_{0\leqslant t\leqslant T} |K_t^n - K_t|^2$$

which tends to 0, if n tends to infinity. Finally, for η fixed, the sixth term is bounded by

$$\|\varphi\|_\infty^2 \left(\int_0^T \int_s^{s+\delta} |\eta'(\sigma - s)| \,\mathrm{d}\sigma \,\mathrm{d}s \right)^2 \mathbb{E} \sup_{0\leqslant t\leqslant T} |K_t^n - K_t|^2$$

and thus it converges to 0 by (i), provided n tends to ∞. These observations show the validity of assertion (ii) in Lemma 2.2. □

Remark 2.2. Recall that for any $x \in \partial D$, the boundary of the domain D, $\nabla\phi(x)$ coincides with the unit normal, so $\|\nabla\phi(x)\| = 1$ and $k_t = \int_0^t \nabla\phi(X_s)\,\mathrm{d}K_s$. Using Lemma 2.2 with $\varphi = \nabla\phi$, we obtain

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |A_t^n - k_t| \right] = 0,$$

where we wrote

$$A_t^n := \int_0^t \nabla\phi(X_s^n)\,\mathrm{d}K_s^n = -n \int_0^t \langle \nabla\phi(X_s^n), \delta(X_s^n)\rangle \,\mathrm{d}s.$$

We recall that the processes $t \mapsto A_t^n$ and $t \mapsto k_t$ are almost surely continuous and increasing: see (1.1) and (2.2). Moreover, there exists a subsequence $(A_t^{j_n})_{n\in\mathbb{N}}$ such that

$$\lim_{j\to\infty} \sup_{0\leqslant t\leqslant T} |A_t^{j_n} - k_t| = 0, \quad \mathbb{P}\text{-almost surely.} \tag{2.6}$$

2.2. Backward stochastic differential equations

Next fix $T > 0$ and let $f, h : [0,T] \times \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R}^k$ be continuous functions satisfying the following assumptions:

There exist positive constants C_i, $(i = 1,\dots,4)$, and $\beta < 0$ such that $\forall t,s \in [0,T]$, $\forall (x,x',y,y') \in (\mathbb{R}^d)^2 \times (\mathbb{R}^k)^2$ we have

(f-i) $|f(t,x,y) - f(t,x,y')| \leqslant C_1|y-y'|$,
(f-ii) $|f(t,x,y)| \leqslant C_2(1+|y|)$,
(h-i) $|h(t,x,y) - h(s,x',y')| \leqslant C_3(|t-s| + |y-y'| + |x-x'|)$,
(h-ii) $|h(t,x,y)| \leqslant C_4(1+|y|)$,
(h-iii) $\langle y-y', h(t,x,y) - h(t,x,y')\rangle \leqslant \beta|y-y'|^2$.

Moreover, let $g : \mathbb{R}^d \to \mathbb{R}^k$ be a continuous function for which there exist a constant $C_5 > 0$ and a real number $q \geqslant 1$ such that

(g) $|g(x)| \leqslant C_5(1+|x|^q)$ for any $x \in \mathbb{R}^d$.

Remark 2.3. To assure the existence and uniqueness of the solution to the BSDE (2.7) below only one of the conditions (h-i) or (h-iii) on h is needed. However, we will use (h-iii) to prove a tightness result in our convergence theorem (see Theorem 3.1), and (h-i) will be used for the identification of the limit.

For each $x \in \bar{D}$, the process $((X_s,K_s) : s \in [0,T])$ denotes the reflected diffusion solution to the SDE (2.1) starting at $x \in \bar{D}$ and adapted to the Brownian filtration $\mathscr{F}_t$. We will impose the following assumption:

for all $x \in \bar{D}$ the matrix $\sigma(x)$ is invertible.

Since f, h and g satisfy conditions (f-i), (f-ii), (h-i), (h-ii) and (g), it follows from Pardoux and Zhang [13] that there is a unique pair of $\mathscr{F}_t$-progressively measurable processes (Y,U) with values in $\mathbb{R}^k \times \mathbb{R}^{k\times d}$ such that

(i) $\mathbb{E}\left(\sup_{0\leqslant t\leqslant T}|Y_t|^2 + \operatorname{Tr}\int_0^T U_s \mathrm{d}\langle M^X\rangle_s U_s^*\right) < \infty$,

(ii) $Y_t = g(X_T) + \int_t^T f(s,X_s,Y_s)\,\mathrm{d}s - \int_t^T U_s\,\mathrm{d}M_s^X + \int_t^T h(s,X_s,Y_s)\,\mathrm{d}k_s$,
$0 \leqslant t \leqslant T$,

where M^X (resp. k) is the martingale part (resp. the continuous increasing process in the reflection part) of the reflected diffusion process X.

Furthermore, for any fixed $n \geqslant 1$, let $(X_t^n : t \in [0,T])$ be the solution to the SDE (2.3). Consider the following BSDE

$$Y_t^n = g(X_T^n) + \int_t^T f_n(s,X_s^n,Y_s^n)\,\mathrm{d}s - \int_t^T Z_s^n\,\mathrm{d}B_s, \quad 0 \leqslant t \leqslant T, \tag{2.7}$$

where we write $f_n(s,x,y) := f(s,x,y) - nh(s,x,y)\langle\nabla\phi(x),\delta(x)\rangle$. Recall that by construction δ is bounded on $\mathbb{R}^d$. By the assumptions on f and h we deduce that, for every $n \geqslant 1$, f_n is a continuous function satisfying: $\forall t \in [0,T], \forall (x,y,y') \in \mathbb{R}^d \times \mathbb{R}^k \times \mathbb{R}^k$,

(i) $|f_n(t,x,y) - f_n(t,x,y')| \leqslant K_1(n)|y-y'|$,
(ii) $|f_n(t,x,y)| \leqslant K_2(n)(1+|y|)$.

From Pardoux and Peng [11] we know that there exists a unique pair (Y^n, Z^n), which is $\mathscr{F}_t$-progressively measurable, satisfying (2.7) such that:

$$\mathbb{E}\left\{\sup_{0\leqslant s\leqslant T}|Y_s^n|^2+\int_0^T\|Z_s^n\|^2\,\mathrm{d}s\right\}<\infty.$$

3. Main result

In this section we will employ the following notation:

$$M_t^n=\int_0^t Z_s^n\,\mathrm{d}B_s,\qquad M_t=\int_0^t U_s\,\mathrm{d}M_s^X,$$
$$H_t^n=\int_0^t h(s,X_s^n,Y_s^n)\,\mathrm{d}A_s^n,\qquad H_t=\int_0^t h(s,X_s,Y_s)\,\mathrm{d}k_s,$$

where A^n is given in Remark 2.2. Our principal goal is here to show that (Y^n, M^n) converges to (Y, M), for n tending to ∞, in a weak sense in the Skorohod space $\mathbb{D} := \mathscr{D}([0,T],\mathbb{R}^k)$ (the space of "cadlag" functions), endowed with the so called S-topology which is weaker than the topology of convergence in dt measure introduced by Meyer–Zheng [9]. However, tightness criteria are easy to establish for this topology and are the same as for the topology of convergence in dt-measure. The reason of our choice is intimately related to our problem: we need to have the continuity of the mapping $y\mapsto\int_0^t h(s,y(s))\,\mathrm{d}k_s$, $y\in\mathbb{D}$, for any continuous function h and for any real continuous non-decreasing function k. This property is true for the S-topology, but for the Meyer–Zheng topology the continuity is only true for functions k_s of the form $k_s=\int_0^s\varphi(\tau)\,\mathrm{d}\tau$ (i.e. functions k_s which are absolutely continuous with respect to the Lebesgue measure). In terms of convergence for the S-topology our main result reads as follows:

Theorem 3.1. *Under the above conditions on the coefficients f, h and g, the following convergence takes place*:

$$(Y^n,M^n,H^n)\overset{\mathrm{S\times S\times S}}{\Longrightarrow}(Y,M,H),\quad as\quad n\to\infty. \tag{3.1}$$

Moreover, $\lim_{n\to\infty}Y_0^n=Y_0$.

The convergence in (3.1) means that the sequence Y^n converges to Y, that M^n converges to M, and that H^n converges to H for the S-topology.

Convergence for the S-topology will be explained in Section 3.1, Theorem 3.1, our main convergence result, will be proved in Section 3.2. In Section 3.3 we will give an application to a Neumann boundary value problem.

3.1. The S-topology

We recall here some relevant notions and results about the S-topology on the space $\mathbb{D}$; for more details on this subject we refer to Jakubowski [4]. Although we confine ourselves to $\mathbb{R}$-valued paths, the S-topology also extends easily to the finite dimensional Euclidean space $\mathbb{R}^k$. By $\mathscr{V}_+ \subset \mathbb{D}$ we denote the space of nonnegative and nondecreasing functions $V:[0,T]\to\mathbb{R}_+$ and $\mathscr{V}=\mathscr{V}_+-\mathscr{V}_+$. We know that any element $V\in\mathscr{V}_+$ determines a unique positive measure $\mathrm{d}V$ on $[0,T]$ and $\mathscr{V}$ can be equipped with the topology of weak convergence of measures. We have the following definition:

Definition 3.2. Let $(Y^n)_{1\leqslant n\leqslant\infty}\in\mathbb{D}$. We say that Y^n converges to Y^∞ with respect to the S-topology, and we write $Y^n\overset{\mathrm{S}}{\to}Y^\infty$, if for every $\varepsilon>0$ there exist elements $(V^{n,\varepsilon})_{1\leqslant n\leqslant\infty}\in\mathscr{V}$ such that

$$\|V^{n,\varepsilon}-Y^n\|_\infty\leqslant\varepsilon,\ n=1,\dots,\infty,\quad \text{and}$$

$$\mathrm{d}V^{n,\varepsilon}\overset{\mathrm{w}}{\Longrightarrow}\mathrm{d}V^{\infty,\varepsilon},\ \text{as } n \text{ tends to } \infty.$$

Here $\overset{\mathrm{w}}{\Longrightarrow}$ stands for the classical weak convergence of measures; in particular we have $\lim_{n\to\infty}\int_0^T\varphi(s)\,\mathrm{d}V^{n,\varepsilon}(s)=\int_0^T\varphi(s)\,\mathrm{d}V^{\infty,\varepsilon}(s)$, for all $\varphi\in\mathscr{C}([0,T],\mathbb{R})$.

We notice that the Riesz representation theorem together with the theorem of Banach –Steinhaus implies that, for a given $\varepsilon>0$, the total variation of the functions $V^{n,\varepsilon}(\cdot)$ in Definition 3.2 is uniformly bounded in n. Let

$$\|V\|_{\mathrm{BV}}=\sup\left\{\sum_{j=1}^{n}|V(s_{j+1})-V(s_j)|:0=s_1<s_2<\cdots<s_n<s_{n+1}=T\right\}$$

denote the total variation norm of the function $V\in\mathscr{V}$. We write $\mathscr{V}_+^c$ for the subspace of continuous functions $V\in\mathscr{V}_+$ vanishing at 0. We will need the following continuity lemma which should be compared to Corollary 2.11 in [4]:

Lemma 3.3. *Let $\Phi:[0,T]\times\mathbb{R}\to\mathbb{R}$ be a Lipschitz function and $(H^n)_{1\leqslant n\leqslant\infty}$ be a sequence of elements of $\mathscr{V}_+^c$ such that $\|H^n(\cdot)-H^\infty(\cdot)\|_\infty\overset{n\to\infty}{\to}0$. Let $(Y^n)_{1\leqslant n\leqslant\infty}$ be a sequence in $\mathbb{D}$ such that $Y^n\overset{\mathrm{S}}{\to}Y^\infty$ as $n\to\infty$. Then there exists a countable subset $Q\subset[0,T]$ such that for all $t\in[0,T]\setminus Q$:*

$$\lim_{n\to\infty}\int_{[0,t]}\Phi(s,Y^n(s))\,\mathrm{d}H^n(s)=\int_{[0,t]}\Phi(s,Y^\infty(s))\,\mathrm{d}H^\infty(s). \tag{3.2}$$

Proof. First notice that, $(H^n)_{1\leqslant n\leqslant\infty}$ is a sequence of elements of $\mathscr{V}_+^c$, which implies in particular that $\|H^\infty\|_{\mathrm{BV}}$ is finite. The uniform convergence of H^n to H^∞ entails that $\sup_{n\geqslant1}\|H^n\|_{\mathrm{BV}}$ is finite. Fix $\varepsilon>0$. By Definition 3.2 there exists a sequence $(V^{n,\varepsilon})_{0\leqslant n\leqslant\infty}$ in $\mathscr{V}$ such that for all $n=1,\dots,\infty$ we have $\|Y^n-V^{n,\varepsilon}\|_\infty\leqslant\varepsilon$ and $\mathrm{d}V^{n,\varepsilon}\overset{\mathrm{w}}{\Longrightarrow}\mathrm{d}V^{\infty,\varepsilon}$. For $t\in[0,T]$ we have

$$\left|\int_{[0,t]}\Phi(s,Y_s^n)\,\mathrm{d}H^n(s)-\int_{[0,t]}\Phi(s,Y_s^\infty)\,\mathrm{d}H^\infty(s)\right|$$

$$\leqslant\left|\int_{[0,t]}\Phi(s,Y_s^n)-\Phi(s,V_s^{n,\varepsilon})\,\mathrm{d}H^n(s)\right|$$

$$+\left|\int_{[0,t]} \Phi(s, Y_s^{\infty}) - \Phi(s, V_s^{\infty,\varepsilon})\,\mathrm{d}H^{\infty}(s)\right|$$

$$+\left|\int_{[0,t]} \Phi(s, V_s^{n,\varepsilon})\,\mathrm{d}H^{n}(s) - \int_{[0,t]} \Phi(s, V_s^{\infty\varepsilon})\,\mathrm{d}H^{\infty}(s)\right|. \tag{3.3}$$

As a consequence of the uniform boundedness of the total variation of H^n and the Lipschitz continuity of the function Φ with Lipschitz constant C, the first and the second term in (3.3) are bounded by $\varepsilon C(\sup_n \|H^n\|_{\mathrm{BV}} + \|H^{\infty}\|_{\mathrm{BV}})$. Let V be a member of $\mathscr{V}$. Since $\Phi : [0,T]\times\mathbb{R}\to\mathbb{R}$ is a Lipschitz function, the function $s \to \Phi(s, V(s))$ is of bounded variation on $[0,T]$ and we have

$$\|\Phi(\cdot, V(\cdot))\|_{\mathrm{BV}} \leqslant C(T + \|V\|_{\mathrm{BV}}).$$

First we will justify this statement. Let C be the Lipschitz constant of the function Φ, and let $0 = s_1 < s_2 < s_3 < \cdots < s_n < s_{n+1} = T$ be a sub-division of the interval $[0,T]$. Then

$$\sum_{j=1}^{n} |\Phi(s_{j+1}, V(s_{j+1})) - \Phi(s_j, V(s_j))|$$

$$\leqslant C\sum_{j=1}^{n}((s_{j+1} - s_j) + |V(s_{j+1}) - V(s_j)|) \leqslant C(T + \|V\|_{\mathrm{BV}}). \tag{3.4}$$

Since the sequence $\mathrm{d}V^{n,\varepsilon}(\cdot)$ converges weakly to $\mathrm{d}V^{\infty,\varepsilon}(\cdot)$, the theorem of Banach–Steinhaus implies that $\sup_{n\in\mathbb{N}} \|V^{n,\varepsilon}\|_{\mathrm{BV}} < \infty$. An appeal to (3.4) then shows that

$$\sup_{n\in\mathbb{N}} \|\Phi(\cdot, V^{n,\varepsilon}(\cdot))\|_{\mathrm{BV}} < \infty. \tag{3.5}$$

Hence, it makes sense to write $\mathrm{d}\Phi(s, V^{n,\varepsilon}(s))$: see (3.6) below. Next, let Q_ε denote the set of points of discontinuity of the function $V^{\infty,\varepsilon}$ and choose $t \in [0,T]\setminus Q_\varepsilon$. In the final term of (3.3) we just apply an integration by parts formula to obtain

$$\int_{[0,t]} \Phi(s, V_s^{n,\varepsilon})\,\mathrm{d}H^{n}(s) - \int_{[0,t]} \Phi(s, V_s^{\infty,\varepsilon})\,\mathrm{d}H^{\infty}(s)$$

$$= H^n(t)\Phi(t, V^{n,\varepsilon}(t)) - H^{\infty}(t)\Phi(t, V^{\infty,\varepsilon}(t))$$

$$-\left(\int_{[0,t]} H^n(s)\,\mathrm{d}\Phi(s, V_s^{n,\varepsilon}) - \int_{[0,t]} H^{\infty}(s)\,\mathrm{d}\Phi(s, V_s^{\infty,\varepsilon})\right). \tag{3.6}$$

Since t is a continuity point of $V^{\infty,\varepsilon}$, we have

$$\lim_{n\to\infty} H^n(t)\Phi(t, V^{n,\varepsilon}(t)) - H^{\infty}(t)\Phi(t, V^{\infty,\varepsilon}(t)) = 0.$$

However, since in addition to (3.5) we also have $\lim_{n\to\infty}\|H^n - H^{\infty}\|_{\infty} = 0$, it is a classical result to deduce that

$$\lim_n \int_{[0,t]} H^n(s)\,\mathrm{d}\Phi(s, V_s^{n,\varepsilon}) = \int_{[0,t]} H^{\infty}(s)\,\mathrm{d}\Phi(s, V_s^{\infty,\varepsilon}). \tag{3.7}$$

Then from (3.6) it is clear that

$$\lim_{n\to\infty}\int_{[0,t]}\Phi(s,V_s^{n,\varepsilon})\,\mathrm{d}H^n(s)=\int_{[0,t]}\Phi(s,V_s^{\infty,\varepsilon})\,\mathrm{d}H^\infty(s).$$

The corresponding countable set is then $Q=\cup_{\varepsilon\in\mathbb{Q}_+}Q_\varepsilon$. This finishes the proof. □

Remark 3.1. By Remark 2.2 we know that the sequence of non-decreasing processes $(A^n)_{n\in N}$ has the L^1-limit k:

$$\lim_{n\to\infty}\mathbb{E}\sup_{0\leqslant t\leqslant T}|A_t^n-k_t|=0.$$

Then, along a subsequence $(A_t^{j_n})_{n\in\mathbb{N}}$, this sequence of nondecreasing processes converges $\mathbb{P}$-almost surely uniformly to k_t. In order to identify the limit in Section 3.2, we will take this subsequence into consideration. To be precise we will employ Lemma 3.3 with $H^n(t)=A_t^{j_n}$, $H^\infty(t)=k_t$, and with $\Phi(s,Y^n(s))=h(s,X_s^n,Y_s^n)$; see the equalities (3.20) and (3.21) in Section 3.2. Here $A_t^{j_n}$ is as in Remark 2.2.

Next we give a tightness criterion with respect to the S-topology which we will employ further on in the present section. For $n\geqslant 1$, the symbol $N^{a,b}(Y^n)$ denotes the number of up-crossings for given levels $a<b$.

Definition 3.4. The family $\{Y_t^n, 0\leqslant t\leqslant T\}_{n\geqslant 1}$ is tight with respect to the S-topology if and only if $(N^{a,b}(Y^n))_{n\geqslant 1}$ and $(\sup_{0\leqslant t\leqslant T}|Y_t^n|)_{n\geqslant 1}$ are tight for all real pairs $a<b$.

On a probability space $(\Omega,\mathscr{F},\mathbb{P})$ with a filtration $\mathscr{F}_t$, let Y be an adapted process with paths a.s in $\mathbb{D}$. If Y_t is integrable for all $t\in[0,T]$, we define the conditional variation of Y by

$$\mathrm{CV}_T(Y)=\sup_\pi\sum_{i=1}^n\mathbb{E}[|\mathbb{E}[Y_{t_{i+1}}-Y_{t_i}|\mathscr{F}_{t_i}]|],\tag{3.8}$$

where the supremum is taken over all subdivisions π of the interval $[0,T]$. If $\mathrm{CV}_T(Y)<\infty$ then the process Y is called a quasi-martingale. Notice that for martingales Y the quantity $\mathrm{CV}_T(Y)=0$.

We will use the following criterion; for the proof we refer for example to LeJay [5] and the references therein.

Theorem 3.5. *Let $(Y^n)_{n\geqslant 1}$ be a family of stochastic process in $\mathbb{D}$. If*

$$\sup_{n\geqslant 1}\left(\mathrm{CV}_T(Y^n)+\mathbb{E}\left[\sup_{0\leqslant s\leqslant T}|Y_s^n|\right]\right)<\infty\tag{3.9}$$

then the sequence $(Y^n)_{n\geqslant 1}$ is S-tight and there exists a subsequence $(Y^{n_k})_{k\geqslant 1}$ of $(Y^n)_{n\geqslant 1}$, a process Y belonging to $\mathbb{D}$, and a countable subset $Q\subset[0,T)$ such that *for every* $j \geqslant 1$ *and for any finite subset* $\{t_1, \ldots, t_j\}$ *of* $[0, T] \setminus Q$ *the following convergence is true*:

$$(Y_{t_1}^{n_k}, \ldots, Y_{t_j}^{n_j}) \overset{\text{dist}}{\Longrightarrow} (Y_{t_1}, \ldots, Y_{t_j}) \text{ as } k \to \infty.$$

Here $\overset{\text{dist}}{\Longrightarrow}$ *means convergence in distributional (or weak) sense.*

Remark 3.2. Note that T is not in the countable subset Q. More precisely the projection $\pi_T : \mathscr{D}([0, T], \mathbb{R}) \to \mathbb{R}$, which assigns to x the value $x(T)$, is continuous with respect to the S-topology (c.f. [4, Remark 2.4, p. 8]).

We will use the following results: the first lemma will allow us to identify a weak limit of a sequence of martingales. For proofs we refer the reader to LeJay [5].

Lemma 3.6. *Let* (Y^n, M^n) *be a multi-dimensional process in* $\mathscr{D}([0, T], \mathbb{R}^k)$, $k \in \mathbb{N}^*$, *converging to* (Y, M) *with respect to the S-topology. Let* $(\mathscr{F}_t^{Y^n})_{t \geqslant 0}$ *(resp.* $(\mathscr{F}_t^{Y})_{t \geqslant 0}$*) be the minimal complete admissible filtration for* Y^n *(resp.* Y*). Assume that* M^n *is a* $\mathscr{F}_t^{Y^n}$*-martingale and* M *is* $\mathscr{F}_t^{Y}$*-adapted. Also suppose that*

$$\sup_{n \geqslant 1} \mathbb{E}\left[\sup_{0 \leqslant t \leqslant T} |M_t^n|^2\right] \leqslant C_T < \infty.$$

Then M *is an* $\mathscr{F}_t^{Y}$*-martingale.*

Lemma 3.7. *Let* $(Y^n)_{n \geqslant 1}$ *be a sequence of processes in* $\mathscr{D}([0, T], \mathbb{R}^k)$ *converging weakly to* Y *with respect to the S-topology. Assume that* $\sup_{n \geqslant 1} \mathbb{E}(\sup_{0 \leqslant s \leqslant T} |Y_s^n|^2) < \infty$. *Then* $\mathbb{E}(\sup_{0 \leqslant s \leqslant T} |Y_s|^2) < \infty$.

3.2. Proof of Theorem 3.1

We are now ready to prove Theorem 3.1, our main convergence result.

Proof. (1) *A priori estimates*: For $0 \leqslant t \leqslant T$ an application of Itô's formula yields

$$\begin{aligned}
&|Y_t^n|^2 + \int_t^T \|Z_s^n\|^2 \,\mathrm{d}s \\
&\quad = |g(X_T^n)|^2 + 2\int_t^T \langle Y_s^n, Z_s^n \,\mathrm{d}B_s\rangle + 2\int_t^T \langle Y_s^n, f(s, X_s^n, Y_s^n)\rangle \,\mathrm{d}s \\
&\qquad + 2\int_t^T \langle Y_s^n, h(s, X_s^n, Y_s^n) - h(s, X_s^n, 0)\rangle \mathrm{d}A_s^n + 2\int_t^T \langle Y_s^n, h(s, X_s^n, 0)\rangle \,\mathrm{d}A_s^n.
\end{aligned}$$

We recall that for every $n \geqslant 1$, $\mathrm{d}A^n$ is a positive measure, and that β in (h-iii) is a strictly negative real number. Taking into account (h-ii) and (h-iii) we obtain

$$|Y_t^n|^2 + \int_t^T \|Z_s^n\|^2\,\mathrm{d}s \leqslant |g(X_T^n)|^2 + 2\int_t^T \langle Y_s^n, Z_s^n\,\mathrm{d}B_s\rangle + 2\int_t^T \langle Y_s^n, f(s,X_s^n,Y_s^n)\rangle\,\mathrm{d}s$$
$$+ 2\beta\int_t^T |Y_s^n|^2\,\mathrm{d}A_s^n + 2C_4\int_t^T |Y_s^n|\,\mathrm{d}A_s^n.$$

Using the equality $\beta|Y_s^n|^2 + C_4|Y_s^n| = -|\beta|(|Y_s^n| - \frac{C_4}{2|\beta|})^2 + \frac{C_4^2}{4|\beta|}$ and taking expectations we get

$$\mathbb{E}|Y_t^n|^2 + \mathbb{E}\int_t^T \|Z_s^n\|^2\,\mathrm{d}s \leqslant \mathbb{E}|g(X_T^n)|^2 + 2\mathbb{E}\int_t^T \langle Y_s^n, Z_s^n\,\mathrm{d}B_s\rangle$$
$$+ 2\mathbb{E}\int_t^T \langle Y_s^n, f(s,X_s^n,Y_s^n)\rangle\,\mathrm{d}s + \frac{C_4^2}{2|\beta|}\mathbb{E}(A_T^n - A_t^n).$$

For every fixed $n \geqslant 1$, the process $(\int_0^t \langle Y_s^n, Z_s^n\,\mathrm{d}B_s\rangle : 0 \leqslant t \leqslant T)$ is a local martingale. From the Burkholder–Davis–Gundy inequality in the form

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |M_t|\right] \leqslant C\mathbb{E}[\langle M,M\rangle_T^{1/2}], \quad \text{with } C = 4\sqrt{2}, \text{ and } M_t = \int_0^t \langle Y_s^n, Z_s^n\,\mathrm{d}B_s\rangle$$

we infer

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\left|\int_0^t \langle Y_s^n, Z_s^n\,\mathrm{d}B_s\rangle\right|\right]$$
$$\leqslant C\mathbb{E}[(\langle M,M\rangle_T)^{1/2}] \leqslant C\mathbb{E}\left[\left(\sup_{0\leqslant s\leqslant T}|Y_s^n|^2\int_0^T \|Z_s^n\|^2\,\mathrm{d}s\right)^{1/2}\right]. \tag{3.10}$$

We use inequality (3.10) together with the elementary inequality $2ab \leqslant a^2 + b^2$ to obtain

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\left|\int_0^t \langle Y_s^n, Z_s^n\,\mathrm{d}B_s\rangle\right|\right] \leqslant \frac{C}{2}\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}|Y_s^n|^2 + \int_0^T |Z_s^n|^2\,\mathrm{d}s\right] < \infty.$$

It follows that the process $(\int_0^t \langle Y_s^n, Z_s^n \mathrm{d}B_s\rangle : 0 \leqslant t \leqslant T)$ is a uniformly integrable martingale; in particular it has zero expectation. So by conditions (f-i), (f-ii) and (g) we get

$$\mathbb{E}|Y_t^n|^2 + \mathbb{E}\int_t^T \|Z_s^n\|^2\,\mathrm{d}s$$
$$\leqslant \mathbb{E}|g(X_T^n)|^2 + 2C_1\mathbb{E}\int_t^T |Y_s^n|^2\,\mathrm{d}s + 2\mathbb{E}\int_t^T |Y_s^n|\|f(s,X_s^n,0)|\,\mathrm{d}s + \frac{C_4^2}{2|\beta|}\mathbb{E}|A_T^n|$$
$$\leqslant C_4^2\mathbb{E}[(1+|X_T^n|^q)^2] + C_2^2 T + (1+2C_1)\mathbb{E}\int_t^T |Y_s^n|^2\,\mathrm{d}s + \frac{C_4^2}{2|\beta|}\mathbb{E}|A_T^n|.$$

Remark 2.1 yields the uniform boundedness (in n) of the moments of X^n, and since $\phi \in \mathscr{C}_b^2(\mathbb{R}^d)$ we also have $\sup_{n\geqslant 1}\mathbb{E}|A_T^n| < \infty$. By Gronwall's lemma we then infer

$$\sup_{n\geqslant 1}\sup_{0\leqslant s\leqslant T}\mathbb{E}\left\{|Y_s^n|^2+\int_0^T\|Z_s^n\|^2\,\mathrm{d}s\right\}<\infty.$$

Consequently, using the Burkholder–Davis–Gundy inequality once again shows

$$\sup_{n\geqslant 1}\mathbb{E}\left\{\sup_{0\leqslant s\leqslant T}|Y_s^n|^2+\int_0^T\|Z_s^n\|^2\,\mathrm{d}s\right\}<\infty. \tag{3.11}$$

(2) *Tightness*: Since M^n is an $\mathscr{F}_t$-martingale we have, by using assumptions (f-ii) and (h-ii),

$$\begin{aligned}\mathrm{CV}_T(Y^n)&=\sup_\pi\sum_i\mathbb{E}(|\mathbb{E}(Y_{t_{i+1}}^n-Y_{t_i}^n|\mathscr{F}_{t_i})|)\\&\leqslant\mathbb{E}\int_0^T|f(s,X_s^n,Y_s^n)|\,\mathrm{d}s+\mathbb{E}\int_0^T|h(s,X_s^n,Y_s^n)|\,\mathrm{d}A_s^n\\&\leqslant C_2\mathbb{E}\int_0^T(1+|Y_s^n|)\,\mathrm{d}s+C_3\mathbb{E}\int_0^T(1+|Y_s^n|)\,\mathrm{d}A_s^n.\end{aligned}$$

The uniform boundedness of the moments of A^n together with (3.11) entails

$$\sup_{n\geqslant 1}\mathrm{CV}_T(Y^n)<C \tag{3.12}$$

for some constant C depending only on C_i, $i=1,\dots,5$, and q. With similar arguments one may show that

$$\sup_n\left\{\mathrm{CV}_T(H^n)+\mathbb{E}\sup_{0\leqslant s\leqslant T}|H_s^n|\right\}<\infty. \tag{3.13}$$

By using (3.11)–(3.13) we finally get

$$\sup_{n\geqslant 1}\left(\mathrm{CV}_T(Y^n)+\mathbb{E}\sup_{0\leqslant s\leqslant T}|Y_s^n|+\mathbb{E}\sup_{0\leqslant s\leqslant T}|M_s^n|+\mathrm{CV}_T(H^n)+\mathbb{E}\sup_{0\leqslant s\leqslant T}|H_s^n|\right)<\infty.$$

Hence (Y^n,M^n,H^n) satisfies condition (3.9) which gives the tightness of the sequence (Y^n,M^n,H^n) with respect to the S-topology.

(3) *Identification of the limit*: By Theorem 3.5 and after extracting a subsequence, the notation of which being suppressed, there exists a process, denoted as an ordered triple $(\bar{Y},\bar{M},\bar{H})$ in $(\mathscr{D}([0,T],\mathbb{R}^k))^3$, such that

$$(X^n,Y^n,M^n,H^n)\overset{\mathrm{U}\times\mathrm{S}\times\mathrm{S}\times\mathrm{S}}{\Longrightarrow}(X,\bar{Y},\bar{M},\bar{H}), \tag{3.14}$$

where $\overset{\mathrm{U}\times\mathrm{S}\times\mathrm{S}\times\mathrm{S}}{\Longrightarrow}$ designates weak convergence in the space $\mathscr{C}([0,T],\mathbb{R}^d)\times\mathscr{D}([0,T],\mathbb{R}^k)^3$ endowed with the product topology of the uniform norm topology for the space $\mathscr{C}([0,T],\mathbb{R}^d)$ and the S- topology for $\mathscr{D}([0,T],\mathbb{R}^k)$. Since f is continuous, the mapping $(x,y)\to\int_0^T f(s,x(s),y(s))\,\mathrm{d}s$ is continuous from $\mathscr{C}([0,T],\mathbb{R}^d)\times\mathscr{D}([0,T],\mathbb{R}^k)$, equipped with the topology described above, to $\mathbb{R}^k$. Furthermore, by Remark 2.2 we have $\mathbb{E}[\sup_{0\leqslant s\leqslant T}|A_s^n-k_s|]\overset{n\to\infty}{\to}0$ and, since h is Lipschitz, an appeal to Theorem 3.5 and Lemma 3.3 shows the existence of a countable subset $Q \subset [0,T)$ such that for all $t \in [0,T] \setminus Q$ the equalities $\bar{H}_t = \int_0^t h(s, X_s, \bar{Y}_s)\,\mathrm{d}k_s$ and

$$\bar{Y}_t = g(X_T) + \int_t^T f(s, X_s, \bar{Y}_s)\,\mathrm{d}s - (\bar{M}_T - \bar{M}_t) + \int_t^T h(s, X_s, \bar{Y}_s)\,\mathrm{d}k_s$$

are valid. The $\mathbb{P}$-almost sure right continuity of $\bar{Y}$, $\bar{M}$ and $\bar{H}$ yields, for all $0 \leqslant t \leqslant T$,

$$\bar{Y}_t = g(X_T) + \int_t^T f(s, X_s, \bar{Y}_s)\,\mathrm{d}s - (\bar{M}_T - \bar{M}_t) + \int_t^T h(s, X_s, \bar{Y}_s)\,\mathrm{d}k_s. \tag{3.15}$$

We still have to show that M^X and $\bar{M}$ are martingales with respect to the same filtration. From (3.15) it is also clear that

$$\bar{M}_t = \bar{Y}_t - \bar{Y}_0 + \int_0^t f(s, X_s, \bar{Y}_s)\,\mathrm{d}s + \int_0^t h(s, X_s, \bar{Y}_s)\,\mathrm{d}k_s + \bar{M}_0,$$

which implies in particular that $\bar{M}_t$ is $\mathscr{F}_t^{X,\bar{Y},\bar{M}}$-adapted. At this moment this is the only available information about the measurability of $\bar{M}$; that is why we are obliged to use the latter filtration instead of the one generated by Brownian motion. In order to show that $\bar{M}$ is in fact a martingale with respect to Brownian motion, for each $n \in \mathbb{N}$, we introduce the complete filtration $\mathscr{F}_t^{X^n,Y^n,M^n}$ generated by the process (X^n, Y^n, M^n). Since Y^n is $\mathscr{F}_t$-adapted and M^n is an $\mathscr{F}_t$-martingale, it clearly follows that M^n is an $\mathscr{F}_t^{X^n,Y^n,M^n}$-martingale. Since $\sup_{n \geqslant 1} \mathbb{E}(\sup_{0 \leqslant t \leqslant T} |M_t^n|^2) < \infty$, by Lemma 3.6 we deduce that $\bar{M}$ is an $\mathscr{F}^{X,\bar{Y},\bar{M}}$-martingale.

On the other hand, for any $0 \leqslant t \leqslant s \leqslant T$, let ψ_t be a bounded continuous mapping in $\mathscr{C}([0,t], \mathbb{R}^d) \times \mathscr{D}([0,t], \mathbb{R}^k)^2$ and let φ belong to $\mathscr{C}_c^\infty(\mathbb{R}^d)$. Denote by L^n the infinitesimal generator of the diffusion process X^n

$$L^n = L - n\delta(x) \cdot \nabla \quad \text{and put } M_s^{n,\varphi} = \varphi(X_s^n) - \varphi(X_0^n) - \int_0^s L^n \varphi(X_\tau^n)\,\mathrm{d}\tau,$$

where $L = \frac{1}{2} \sum_{i,j} a_{i,j}(\cdot) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b_i(\cdot) \frac{\partial}{\partial x_i}$, and $(a_{i,j}(\cdot))_{ij} = (\sigma(\cdot)\sigma^*(\cdot))_{ij}$. Since the process $M_s^{n,\varphi}$ is a martingale we see

$$\begin{aligned}
&\mathbb{E}\left(\psi_t(X^n, Y^n, M^n)\left(\varphi(X_s^n) - \varphi(X_t^n) - \int_t^s L\varphi(X_t^n)\,\mathrm{d}r\right)\right)\\
&\quad = \mathbb{E}\left(\psi_t(X^n, Y^n, M^n)\left(M_s^{n,\varphi} - M_t^{n,\varphi} - n\int_t^s \delta(X_r^n).\nabla\varphi(X_t^n)\,\mathrm{d}r\right)\right)\\
&\quad = \mathbb{E}\left(\psi_t(X^n, Y^n, M^n)\left(-n\int_t^s \delta(X_r^n).\nabla\varphi(X_r^n)\,\mathrm{d}r\right)\right)\\
&\quad = \mathbb{E}(\psi_t(X^n, Y^n, M^n)\int_t^s \nabla\varphi(X_r^n) \cdot \mathrm{d}K_r^n).
\end{aligned} \tag{3.16}$$

By Lemma 2.2 for any $\varphi\in\mathscr{C}_c^\infty(\mathbb{R}^d)$ we have

$$\lim_{n\to\infty}\mathbb{E}\psi_t(X^n,Y^n,M^n)\left(\varphi(X_s^n)-\varphi(X_t^n)-\int_t^s L\varphi(X_r^n)\,\mathrm{d}r\right)$$

$$=\mathbb{E}(\psi_t(X,\bar{Y},\bar{M})\int_t^s\nabla\varphi(X_r)\,\mathrm{d}K_r). \tag{3.17}$$

On the other hand, by weak convergence and Itô's formula we also have for all $\varphi\in C_c^\infty(\mathbb{R}^d)$

$$\mathbb{E}\left(\psi_t(X,\bar{Y},\bar{M})\left(\varphi(X_s)-\varphi(X_t)-\int_t^s L\varphi(X_r)\,\mathrm{d}r\right)\right)$$

$$=\mathbb{E}\left(\psi_t(X,\bar{Y},\bar{M})\left(\int_t^s\nabla\varphi(X_r)\,\mathrm{d}M_r^X+\int_t^s\nabla\varphi(X_r)\,\mathrm{d}K_r\right)\right).$$

Thus, in view of (3.17), for all $\varphi\in\mathscr{C}_c^\infty(\mathbb{R}^d)$ we get

$$\mathbb{E}\left(\psi_t(X,\bar{Y},\bar{M})\int_t^s\nabla\varphi(X_r)\,\mathrm{d}M_r^X\right)=0. \tag{3.18}$$

From (3.18) one may deduce that M^X is a $\mathscr{F}^{X,\bar{Y},\bar{M}}$-martingale.

Recall that

(i) $Y_t=g(X_T)+\int_t^T f(s,X_s,Y_s)\,\mathrm{d}s-\int_t^T U_s\,\mathrm{d}M_s^X+\int_t^T h(s,X_s,Y_s)\,\mathrm{d}k_s,$

(ii) $\mathbb{E}\left(\sup_s|Y_s|^2+\mathrm{Tr}\int_0^T U_s d\langle M^X\rangle_s U_s^*\right)<\infty,$

(iii) $\bar{Y}_t=g(X_T)+\int_t^T f(s,X_s,\bar{Y}_s)-(\bar{M}_T-\bar{M}_t)-\int_t^T h(s,X_s,\bar{Y}_s)\,\mathrm{d}k_s.$

Since Y and U are $\mathscr{F}_t$-adapted and M^X is an $\mathscr{F}^{X,\bar{Y},\bar{M}}$-martingale, so is $M_t=\int_0^t U_s\,\mathrm{d}M_s^X$. On the other hand, by Lemma 3.7, $\bar{Y}$ and $\bar{M}$ are square-integrable. An application of the Burkholder–Davis–Gundy inequality shows that $\int_0^t\bar{Y}_s\,\mathrm{d}\bar{M}_s$ is an $\mathscr{F}^{X,\bar{Y},\bar{M}}$ martingale. Now combining all these facts and using Itô's formula for possibly discontinuous semi-martingales we get

$$|Y_t-\bar{Y}_t|^2+([M-\bar{M}]_T-[M-\bar{M}]_t)$$

$$=2\int_t^T\langle Y_s-\bar{Y}_s,f(s,X_s,Y_s)-f(s,X_s,\bar{Y}_s\rangle\,\mathrm{d}s$$

$$+2\int_t^T\langle Y_s-\bar{Y}_s,h(s,X_s,Y_s)-h(s,X_s,\bar{Y}_s)\rangle\,\mathrm{d}k_s$$

$$+2\int_t^T\langle Y_s-\bar{Y}_s,\,\mathrm{d}(M-\bar{M})_s\rangle. \tag{3.19}$$

Taking expectation in (3.19) and using condition (h-iii) and (f-i) we obtain

$$\mathbb{E}|Y_t-\bar{Y}_t|^2+\mathbb{E}([M-\bar{M}]_T-[M-\bar{M}]_t)\leqslant 2C\int_t^T|Y_s-\bar{Y}_s|^2\,\mathrm{d}s.$$

Hence, from Gronwall's lemma we have $Y=\bar{Y}$ and $M=\bar{M}$ which ends the first part of the proof of Theorem 3.1.

We have

$$Y_0^n = g(X_T^n) + \int_0^T f(s, X_s^n, Y_s^n)\,\mathrm{d}s - M_T^n + \int_0^T h(s, X_s^n, Y_s^n)\,\mathrm{d}A_s^n. \tag{3.20}$$

Since under the S-topology, the projection $\pi_T : \mathscr{D}([0,T], \mathbb{R}^k) \mapsto \mathbb{R}$, given by $\pi_T(y) = y(T)$, is continuous (see Remark 3.2), the sequence M_T^n, $n \in \mathbb{N}$, converges in distribution to M_T. By Lemma 3.3 and since Y_0^n and Y_0 are deterministic we conclude that

$$Y_0^n \to Y_0 = g(X_T) + \int_0^T f(s, X_s, Y_s)\,\mathrm{d}s - M_T + \int_0^T h(s, X_s, Y_s)\,\mathrm{d}k_s. \tag{3.21}$$

This proves the second claim in Theorem 3.1 as well. $\square$

3.3. Application to nonlinear Neumann boundary value problems

For any $x \in \mathbb{R}^d$, $0 \leqslant t \leqslant T$ and $n \geqslant 1$ we consider the process $\{X_s^{n,t,x},\ t \leqslant s \leqslant T\}$ (resp. $\{X_s^{t,x},\ t \leqslant s \leqslant T\}$), the solution to (2.3) (resp. (2.1)) starting in x at time t. Denote by $\{(Y_s^{n,t,x}, Z_s^{n,t,x}) : t \leqslant s \leqslant T\}$ the unique solution to the BSDE: $t \leqslant s \leqslant T$,

$$Y_s^{n,t,x} = g(X_T^{n,t,x}) + \int_s^T f(r, X_r^{n,t,x}, Y_r^{n,t,x})\,\mathrm{d}r - \int_s^T Z_r^{n,t,x}\,\mathrm{d}B_r$$
$$+ \int_s^T h(r, X_r^{n,t,x}, Y_r^{n,t,x})\,\mathrm{d}A_r^n.$$

Using the notation $f_n(t,x,y) = f(t,x,y) - nh(t,x,y)\langle \nabla\phi(x), \delta(x)\rangle$, we may also write

$$Y_s^{n,t,x} = g(X_T^{n,t,x}) + \int_s^T f_n(r, X_r^{n,t,x}, Y_r^{n,t,x})\,\mathrm{d}r - \int_s^T Z_r^{n,t,x}\,\mathrm{d}B_r.$$

Existence and uniqueness are assured by Pardoux [10] (see also [11]). We define

$$u^n(t,x) := Y_t^{n,t,x},$$

which is obviously a deterministic term, because $Y_s^{n,t,x}$ is measurable with respect to the σ-algebra $\sigma\{B_r - B_t, t \leqslant r \leqslant s\}$. As is well known (see for example [10]) the function $u^n(t,x)$ is continuous in (t,x) and it is a viscosity solution to the problem: $(1 \leqslant i \leqslant k,\ n \in \mathbb{N},\ 0 < t < T,\ x \in \mathbb{R}^d)$

$$\frac{\partial u_i^n}{\partial t}(t,x) + Lu_i^n(t,x) + f_i(t,x,u^n(t,x))$$
$$-n\delta(x)(\nabla u_i^n(t,x) + \nabla\phi(x)h_i(t,x,u^n(t,x))) = 0,$$
$$u^n(T,x) = g(x),$$

where $L = \frac{1}{2}\sum_{i,j}(\sigma(\cdot)\sigma^*(\cdot))_{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b_i(\cdot)\frac{\partial}{\partial x_i}$ is the infinitesimal generator of the process X.

Under the present assumptions on the coefficients f, g and h and from Theorem 1.7 in [13] we know that for any $0 \leqslant t \leqslant T$ and any $x \in \bar{D}$ there exists a process $\{(Y_s^{t,x}, Z_s^{t,x}),\ t \leqslant s \leqslant T\}$, which is the unique solution to the BSDE: $t \leqslant s \leqslant T$

$$Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x})\,\mathrm{d}r - \int_s^T Z_r^{t,x}\sigma(X_r^{t,x})\,\mathrm{d}B_r + \int_s^T h(r, X_r^{t,x}, Y_r^{t,x})\,\mathrm{d}k_r.$$

By Theorem 4.3 in [13], $u(t,x) := Y_t^{t,x}$ is deterministic and, in viscosity sense, $u(t,x)$ is a solution to the following Neumann problem $(1 \leqslant i \leqslant k)$

$$\begin{aligned} &\frac{\partial u_i}{\partial t}(t,x) + Lu_i(t,x) + f_i(t,x,u(t,x)) = 0, \quad 0 < t < T;\ x \in D,\\ &u(T,x) = g(x), x \in \bar{D},\\ &\frac{\partial u_i}{\partial n}(t,x) + h_i(t,u(t,x)) = 0, \quad 0 < t < T,\ x \in \partial D. \end{aligned} \tag{3.22}$$

As a consequence of Theorem 3.1 we have the following corollary:

Corollary 3.8. *For all* $t \in [0,T]$ *and* $x \in \bar{D}$ *the equality* $\lim_{n\to\infty} u^n(t,x) = u(t,x)$ *is true.*

Remark 3.3. Suppose that f and h do not depend on t and consider

$$v(t,x) := u(T-t,x), \quad (t,x) \in [0,T] \times \bar{D}.$$

From Corollary 3.8 we may deduce a similar result for the following system of forward parabolic PDEs $(1 \leqslant i \leqslant k)$:

$$\begin{aligned} &\frac{\partial v_i}{\partial t}(t,x) + Lv_i(t,x) + f_i(x,u(t,x)) = 0, \quad 0 < t < T;\ x \in D,\\ &u(0,x) = g(x), \quad x \in \bar{D},\\ &\frac{\partial v_i}{\partial n}(t,x) + h_i(x,u(t,x)) = 0, \quad 0 < t < T,\ x \in \partial D. \end{aligned}$$

Our result is then a generalization of the linear case proved by Lions–Menaldi–Sznitman [6]. Originally we intended to treat the problem in the present paper without condition (h.iii), but since some classical techniques, like Gronwall's lemma, could not be used because of the presence of the random measure $\mathrm{d}k_s$, we adopted it.

To remove the condition that the diffusion matrix σ is invertible, we need a representation theorem with respect to the martingale part M^X of a given reflected diffusion process X.

References

[1] B. Boufoussi, J.A. van Casteren, An approximation result for solutions to semilinear PDE's with Neumann boundary conditions via BSDE's, Preprint of UIA 2002.

[2] N. El Karoui, S. Peng, M.C. Quenez, Backward stochastic differential equation in finance, Math. Finance 7 (1997) 1–71.

[3] S. Hamadène, J.P. Lepeltier, Zero-sum stochastic differential games and backward equations, Systems Control Lett. 24 (1995) 259–263.

[4] A. Jakubowski, A non-Skorohod topology on the Skorohod space, Electron. J. Probab. 2 (4) (1997) 1–21.
[5] A. LeJay, BSDE driven by Dirichlet process and semi-linear parabolic PDE. Application to homogenization, Stochastic Process. Appl. 97 (1) (2002) 1–39.
[6] P.L. Lions, J.L. Menaldi, A.S. Sznitman, Construction de processus de diffusion réfléchis par pénalisation du domaine, C. R. Acad. Sci. Paris, t. 229, Série I (1981) 459–462.
[7] P.L. Lions, A.S. Sznitman, Stochastic differential equations with reflecting boundary conditions, Comm. Pure Appl. Math. XXXVII (1984) 511–537.
[8] J.L. Menaldi, Stochastic variational inequality for reflected diffusion, Indiana Univ. Math. J. 32 (5) (1983).
[9] P.A. Meyer, W.A. Zheng, Tightness criteria for laws of semimartingales, Ann. Inst. H. Poincaré 20 (1981) 353–372.
[10] É. Pardoux, BSDEs, weak convergence and homogenization of semi-linear PDEs, in: F.H. Clarke, R.J. Stern (Eds.), Non Linear Analysis, Differential Equations and Control, Proc. Séminaires de Mathématiques Supérieures Montréal 1998, Kluwer Academic Publishers, Dordrecht, 1999, pp. 503–549.
[11] É. Pardoux, S. Peng, Adapted solution of a backward stochastic differential equation, Systems Control Lett. 14 (1990) 55–61.
[12] É. Pardoux, R.J. Williams, Symmetric reflected diffusions, Ann. Inst. H. Poincaré 30 (1) (1994) 13–62.
[13] É. Pardoux, S. Zhang, Generalized BSDEs and nonlinear Neumann boundary value problems, Probab. Theory Related Fields 110 (1998) 535–558.
[14] R.J. Williams, W. Zheng, On reflecting Brownian motion—a weak convergence approach, Ann. Inst. H. Poincaré 26 (1990) 461–468.