Stochastics and Stochastics Reports, Vol. 60, pp. 255–270
Reprints available directly from the publisher
Photocopying permitted by license only

AVERAGING OF BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS, WITH APPLICATION TO SEMI-LINEAR PDE'S

E. PARDOUX[a,*] and A. YU. VERETENNIKOV[b]

[a] *L A T P, UMR-CNRS, Centre de Mathématiques et d'Informatique, Université de Provence, 39, rue F. Joliot Curie, F 13453 Marseille cedex 13, France;*
[b] *Institute of Information Transmission Problems, 19, Bolshoy Karetnii, 101447 Moscow, Russia*

(Received 15 June 1996; In final form 17 December 1996)

We establish an averaging result for a backward SDE, which implies an averaging result for a system of parabolic semilinear PDEs. The method of proof uses the Meyer-Zheng convergence, and an extended uniqueness result for BSDEs for identifying the limit.

Keywords: Backward stochastic differential equations; Feynman-Kac formula; diffusion approximation; averaging for parabolic PDEs

AMS Classification: 60H10, 35K60

INTRODUCTION

A new type of stochastic differential equations, called backward stochastic differential equations, has been introduced by Pardoux, Peng [8]. Soon after, it was noticed that these equations provide probabilistic formulas for solutions of certain semilinear partial differential equations, see [9].

It was then tempting to exploit this new tool for proving results concerning semilinear PDEs, similarly as the Feynman-Kac formula has been used for proving results on linear PDEs of second order. This program has already been successfully followed by Pradeilles [10], [11], who proves new asymptotic results concerning reaction-diffusion equations.

*Corresponding author.

In this paper, we combine BSDEs with the theory of diffusion approximation, as exposed e.g. in Papanicolaou, Stroock, Varadhan [7] and in Ethier, Kurtz [3], in order to prove averaging type of results for systems of semilinear parabolic partial differential equations, similar to some of the results in Bensoussan [1], Bensoussan, Lions, Papanicolaou [2].

The weak convergence of the solution of the BSDE is proved in the sense of the topology of Meyer, Zheng [6], which appears to be ideally suited for our purpose. Note that we have been able to treat only BSDEs where the coefficient in front of the Brownian motion does not enter the nonlinear term. In terms of the associated semilinear PDE, this means that the nonlinear term is a function of the solution, not of its gradient. We believe that BSDEs is a powerful tool for proving various types of asymptotic results for semilinear PDES, along the lines e.g. of the results presented in Freidlin [4]. For this programm, one needs to be able to take weak limits in BSDEs. One of the aims of the present paper is to develop a methodology for taking such limits.

The paper is organized as follows. The problem is stated in section 1. The main results are stated in section 2. Section 3 is devoted to the proof of an auxiliary theorem, while the main result, the convergence of the solution of the BSDE, is proved in section 4. Finally the corollaries, in particular the convergence of the PDE, are proved in section 5.

1 STATEMENT OF THE PROBLEM

We now consider the SDE system, defined on some probability space $(\Omega, \mathcal{F}, P)$

$$\begin{aligned} dX_t^{1,\varepsilon} &= \varepsilon^{-1}F(X_t^{1,\varepsilon}, X_t^{2,\varepsilon})dt + G(X_t^{1,\varepsilon}, X_t^{2,\varepsilon})dt, \quad X_0^{1,\varepsilon} = x_0^1, \\ dX_t^{2,\varepsilon} &= \varepsilon^{-2}H(X_t^{2,\varepsilon})dt + \varepsilon^{-1}K(X_t^{2,\varepsilon})dW_t, \quad X_0^{2,\varepsilon} = x_0^2, \end{aligned} \tag{1}$$

where $X^{1,\varepsilon} \in \mathbb{R}^d, X^{2,\varepsilon} \in \mathbb{R}^\ell, F, G, H, K$ are measurable functions with values in $\mathbb{R}^d, \mathbb{R}^\ell$ and $\mathbb{R}^\ell \times \mathbb{R}^\ell$ correspondently, $W_t, t \geq 0$ is an ℓ-dimensional Wiener process.

We assume that the coefficients F, G, H and K are periodic (of period one in each direction) functions of the variable x_2, so that the process $\{X_t^{2,\varepsilon}\}$ can be considered as taking values in the ℓ-dimensional torus T^ℓ.

We assume moreover that H, K are bounded and $KK^* \geq \alpha I > 0$, so that the process $\{X_t^{2,\varepsilon}\}$ possesses a unique invariant probability measure μ on T^ℓ (uniqueness will follow from assumption $(A.1.5)$).

The basic assumption concerning $\{X_t^{1,\varepsilon}\}$ is that

$$\int F(x_1, x_2)\mu(dx_2) = 0, \quad \forall x_1 \in \mathbb{R}^d. \tag{A.1.1}$$

We can then solve the following Poisson equation for each $1 \le i \le d$, $x_1 \in \mathbb{R}^d$:

$$(L_2 J_i(x_1, \cdot))(x_2) = -F_i(x_1, x_2), \quad x_2 \in T^\ell,$$

where L_2 denotes the infinitesimal generator of the diffusion process $\{X_t^{2,\varepsilon}\}$ in case $\varepsilon = 1$ (see its explicit expression below).

We assume in addition that

$$F \in C^{2,0}(\mathbb{R}^d \times \mathbb{R}^\ell; \mathbb{R}^d), \quad F \text{ and } \nabla_{x_1} F \text{ being bounded;} \tag{A.1.2}$$

$$|G(x_1, x_2) - G(x'_1, x_2)| \le K_n |x_1 - x'_1|, \quad \forall |x_1|, |x'_1| \le n, x_2 \in \mathbb{R}^\ell; \tag{A.1.3}$$

$$\sup_{(x_1,x_2)\in\mathbb{R}^{d+\ell}} (1 + |x_1|)^{-1} |G(x_1, x_2)| < \infty. \tag{A.1.4}$$

$$K \in C(T^\ell) \tag{A.1.5}$$

It will be shown in section 3 that under the above assumptions $X^{1,\varepsilon}$ converges in distribution to X^1 which is a d-dimensional diffusion process with the generator

$$\overline{\mathcal{L}} = \frac{1}{2}\sum_{i,j=1}^{d} \overline{a}_{ij}(x_1)\frac{\partial^2}{\partial x_{1i}\partial x_{1j}} + \sum_{i=1}^{d} \overline{b}_i(x_1)\frac{\partial}{\partial x_{1i}},$$

where

$$\overline{a}_{ij}(x_1) = \int_{T^\ell} [F_i(x_1, x_2)J_j(x_1, x_2) + F_j(x_1, x_2)J_i(x_1, x_2)]\mu(dx_2),$$

$$\overline{b}_i(x_1) = \int_{T^\ell} G_i(x_1, x_2)\mu(dx_2) + \int_{T^\ell} \langle F(x_1, x_2), \nabla_{x_1} J_i(x_1, x_2)\rangle \mu(dx_2).$$

In other words, there exists a d-dimensional Brownian motion $\{B_t, t \ge 0\}$ such that

$$X_t^1 = x_0^1 + \int_0^t \overline{b}(X_s^1)ds + \int_0^t \overline{\sigma}(X_s^1)dB_s, \tag{2}$$

where $\overline{\sigma}(x_1) = [\overline{a}(x_1)]^{1/2}$. Moreover, it will follow from the results of section 3 that equation (2) has a unique non-exploding strong solution.

We consider the k-dimensional BSDE

$$Y_t^\varepsilon = g(X_T^{1,\varepsilon}) + \int_t^T f(X_s^{1,\varepsilon}, X_s^{2,\varepsilon}, Y_s^\varepsilon)ds - \int_t^T Z_s^\varepsilon dW_s, \tag{3}$$

which has a unique $\mathcal{F}_t^W$-adapted $\mathbb{R}^k \times \mathbb{R}^{k\times d}$-valued solution, provided we add the restriction that

$$E\int_0^T |Z_t^\varepsilon|^2 dt < \infty.$$

Denote $M_t^\varepsilon = \int_0^t Z_s^\varepsilon dW_s$.

The problem under consideration is the averaging of the process Y^ε. One immediate application is the averaging for the following system of parabolic semi-linear PDEs.

Let $\mathcal{L}_\varepsilon$ denote the infinitesimal generator of the diffusion process $\{X^{1,\varepsilon}, X^{2,\varepsilon}\}$, i.e.

$$\mathcal{L}_\varepsilon = \varepsilon^{-2}L_2 + \sum_{i=1}^d [\varepsilon^{-1}F_i(x) + G_i(x)]\frac{\partial}{\partial x_{1i}},$$

where

$$L_2 = \frac{1}{2}\sum_{i,j=1}^{\ell}(KK^*(x_2))_{ij}\frac{\partial^2}{\partial x_{2i}\partial x_{2j}} + \sum_{i=1}^{\ell} H_i(x_2)\frac{\partial}{\partial x_{2i}},$$
$$x = (x_1, x_2)', \quad x_1 = (x_{11}, \ldots, x_{1d})' \quad x_2 = (x_{21}, \ldots, x_{2l})'$$

The system of PDEs to be averaged reads

$$\frac{\partial u^\varepsilon}{\partial t}(t,x) = \mathcal{L}_\varepsilon u^\varepsilon(t,x) + f(x, u(t,x)), \quad 0 \le t \le T, x \in \mathbb{R}^{d+\ell}$$
$$u^\varepsilon(0,x) = g(x_1), x \in \mathbb{R}^{d+\ell}, \tag{4}$$

where the solution u^ε maps $[0,T]\times\mathbb{R}^{d+\ell}$ into $\mathbb{R}^k$, and $\mathcal{L}_\varepsilon u^\varepsilon$ denotes the k-dimensional vector whose i-th component equals $\mathcal{L}_\varepsilon u_i^\varepsilon, 0 \le i \le k$.

We now formulate our assumptions on f and g. For some real number $K, p \in \mathbb{N}$, and α a continuous increasing function from $\mathbb{R}_+$ into itself, such that $\alpha(0) = 0$,

(A.1.6) (a) The functions f and g are continuous; f is a periodic function (of period one in each direction) of the variable x_2;

(b) $|f(x_1,x_2,y) - f(x_1,x_2,y')| \le K|y-y'|, \forall x_1 \in \mathbb{R}^d, x_2 \in \mathbb{R}^\ell, y, y' \in \mathbb{R}^k$;

(c) $|f(x_1, x_2, y) - f(x'_1, x_2, y)| \leq \alpha(|x_1 - x'_1|)(1 + |y|), \forall x_1, x'_1 \in \mathbb{R}^d$, $x_2 \in \mathbb{R}^\ell, y \in \mathbb{R}^k$;

(d) $|g(x_1)| + |f(x_1, x_2, 0)| \leq K(1 + |x_1|^p), \forall x_1 \in \mathbb{R}^d, x_2 \in \mathbb{R}^\ell$.

2 STATEMENT OF THE RESULTS

Denote

$$\bar{f}(x_1, y) = \int_{\mathbb{R}^\ell} f(x_1, x_2, y)\mu(dx_2).$$

Let $\{(Y_t, Z_t), 0 \leq t \leq T\}$ be the unique $\mathcal{F}_t^B$-adapted $\mathbb{R}^k \times \mathbb{R}^{k \times d}$-valued process which solves the BSDE

$$Y_t = g(X_T^1) + \int_t^T \bar{f}(X_s^1, Y_s)ds - \int_t^T Z_s dB_s, \tag{5}$$

together with the condition

$$E\int_0^T |Z_t|^2 dt < \infty,$$

and $u : [0, T] \times \mathbb{R}^d \to \mathbb{R}^k$ be the viscosity solution of the system of parabolic semi-linear PDEs

$$\begin{aligned} \frac{\partial u}{\partial t}(t, x) &= \bar{\mathcal{L}}u(t, x) + \bar{f}(x, u(t, x)), 0 \leq t \leq T, x \in \mathbb{R}^d \\ u(0, x) &= g(x), x \in \mathbb{R}^d \end{aligned} \tag{6}$$

Before stating our results, let us precise a notion of convergence which will be used repeatedly below.

DEFINITION 1 Let $\{(U_t^n, n \in \mathbb{N}, U_t), 0 \leq t \leq T\}$ denote k-dimensional measurable processes. We say that "$U^n \Longrightarrow U$ in the sense of Meyer-Zheng" whenever U^n converges to U in distribution in $M([0, T], \mathbb{R}^k)$, the space of (equivalence classes of) Borel measurable functions topologized by convergence in measure.

The notation $U^n \Longrightarrow U$ without any additional mention will be used to mean that the processes are continuous, and the sequence converges in distribution in $C([0, T], \mathbb{R}^k)$ equipped with the usual sup-norm topology.

THEOREM 1 *Assume that the conditions (A.1.1)–(A.1.6) are in force. Then* $Y^\varepsilon \Longrightarrow Y$ *in the sense of Meyer-Zheng, as* $\varepsilon \to 0$.

COROLLARY 1 *Under the assumptions of theorem 1,* $Y_0^\varepsilon \longrightarrow Y_0$, *as* $\varepsilon \to 0$.

COROLLARY 2 *Under the same assumptions,* $u^\varepsilon(t,x_1,x_2) \to u(t,x_1)$, $\forall (t,x_1,x_2) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^\ell$ *as* $\varepsilon \to 0$.

3 CONVERGENCE OF $X^{1,\varepsilon}$ AS $\varepsilon \to 0$

The aim of this section is to establish the following result:

THEOREM 2 *Under the assumptions (A.1.1), (A.1.2), (A.1.3), (A.1.4), assuming moreover the boundedness of H and K and the nondegeneracy of K (i.e.,* $KK^*(x_2) \geq \alpha I > 0 \ \forall x_2 \in T^\ell$*),*

$$X^{1,\varepsilon} \Longrightarrow X^1, \quad \varepsilon \to 0,$$

where X^1 *is the unique strong solution of the SDE (2).*

Proof Convergence and identification of the limit follow from theorem 12.2.4 in Ethier, Kurtz [3]. We now prove that the SDE (2) has a unique non-exploding strong solution. As already noted, the T^ℓ-valued homogeneous Markov process $\{X_t^{2,\varepsilon}\}$ possesses a unique invariant measure μ, and from the fact that L_2 is anticompact (i.e. the inverse of a compact operator) and (A.1.1), the Poisson equation

$$(L_2 J_i(x_1,\cdot))(x_2) = -F_i(x_1,x_2), \quad x_2 \in T^\ell$$

has a unique solution in $L^2(T^\ell,\mu)$, for each $1 \leq i \leq d$, $x_1 \in \mathbb{R}^d$. Moreover, it follows from Corollary 6.5.3 in Revuz [12] that J_i is bounded since F_i is bounded, from (A.1.2). Indeed, to use Corollary 6.5.3 from [12] it suffices to verify the inequality

$$\sup_{z \in T^l} E_z \int_0^\infty \exp\left(-\int_0^t h(Z_s)ds \right) < \infty \tag{7}$$

for any $h : T^\ell \to \mathbb{R}_+$ which is bounded, continuous and such that $\mu(z : h(z) \neq 0) > 0$, where $Z_s \equiv X_s^{2,\varepsilon}$ with $\varepsilon = 1$. Since h is continuous, there exists $0 < \nu \leq 1$ such that $\text{mes}(z \in T^\ell : h(z) \geq \nu) > 0$ (here "mes" stands for Lebesgue's measure). One has for $t > 1$,

$$E_z \exp\Big(- \int_0^t h(Z_s)ds\Big) \leq E_z \exp\Big(- \int_0^{t-1} h(Z_s)ds\Big) \cdot \sup_{z' \in T^\ell} E_{z'} \exp\Big(- \nu \int_0^1 I(h(Z_s) \geq \nu)ds\Big).$$

Notice that

$$\Big(\int_0^1 I(h(Z_s) \geq \nu)ds\Big)^2 \leq \int_0^1 I(h(Z_s) \geq \nu)ds.$$

So,

$$\begin{aligned} &\exp\Big(- \nu \int_0^1 I(h(Z_s) \geq \nu)ds\Big) \\ &\quad \leq 1 - \nu \int_0^1 I(h(Z_s) \geq \nu)ds + (\nu^2/2) \int_0^1 I(h(Z_s) \geq \nu)ds) \\ &\quad \leq 1 - (\nu - \nu^2/2) \int_0^1 I(h(Z_s) \geq \nu)ds. \end{aligned}$$

It follows from the Harnack inequality for nondegenerate diffusions (see the results in Krylov and Safonov [15]) that

$$\inf_{z \in T^\ell} E_s \int_0^1 I(h(Z_s) \geq \nu)ds > \beta > 0$$

Hence

$$E_{z'} \exp\Big(- \nu \int_0^1 I(h(Z_s) \geq \nu)ds\Big) \leq 1 - \beta(\nu - \nu^2/2) \leq 1 - \beta\nu/2$$

By induction one gets

$$\sup_z E_s\Big[\exp\Big(- \int_0^t h(Z_s)ds\Big)\Big] \leq (1 - \beta\nu/2)^{[t]},$$

$[t]$ denoting the integer part of t, from which (7) follows.

Now, because the mapping $F_i(x_1, \cdot) \to J_i(x_1, \cdot)$ is linear and continuous in $L^2(T^\ell, \mu)$ and $x_1 \to F_i(x_1, \cdot)$ is twice differentiable, one gets that $x_1 \to J_i(x_1, \cdot)$ is twice differentiable, and in particular $\nabla_{x_1} J(x_1, \cdot)$, which is the solution of

$$(L_2 \nabla_{x_1} J(x_1, \cdot))(x_2) = -\nabla_{x_1} F_1(x_1, x_2), x_2 \in T^\ell,$$

is bounded.

From this and (A.1.3), (A.1.4), $\overline{b}_i$ is locally Lipschitz and grows at most linearly at infinity, for each $1 \leq i \leq d$. Moreover, $\overline{a}_{ij}$ is bounded and of class C^2. Hence from theorem 5.2.3 in [13], $\sigma = [\overline{a}]^{1/2}$ is Lipschitz continuous. The proof is complete.

4 PROOF OF THEOREM 1

a A Priory Estimates

LEMMA 1 *There exists C such that for every $0 < \varepsilon \leq 1$,*

$$E\int_0^T |f(X_t^{1,\varepsilon}, X_t^{2,\varepsilon}, Y_t^\varepsilon)|dt \leq C.$$

Proof One has

$$E|Y_t^\varepsilon|^2 + E\int_t^T |Z_s^\varepsilon|^2 ds = E|g(X_T^{1,\varepsilon})|^2 + 2E\int_t^T \langle Y_s^\varepsilon, f(X_s^{1,\varepsilon}, X_s^{2,\varepsilon}, Y_s^\varepsilon)\rangle ds.$$

We claim that there exists a constant c such that $\forall \varepsilon > 0$,

$$E|g(X_T^{1,\varepsilon})|^2 \leq c, \quad E\int_0^T |f(X_t^{1,\varepsilon}, X_t^{2,\varepsilon}, 0)|^2 dt \leq c.$$

Indeed, from the assumption (A.1.6.d), it suffices to prove the

LEMMA 2 *For any $p \in \mathbb{N}\backslash\{0\}$,*

$$\sup_{0<\varepsilon\leq 1}\sup_{0\leq t\leq T} E\left(|X_t^{1,\varepsilon}|^{2p}\right) < \infty.$$

Proof Define $\Phi : \mathbb{R}^d \to \mathbb{R}_+$ by $\Phi(x_1) = |x_1|^{2p}$, and $\Psi : \mathbb{R}^d \times \mathbb{R}^\ell \to \mathbb{R}$ by $\Psi(x_1, x_2) = 2p|x_1|^{2(p-1)}\langle x_1, J(x_1, x_2)\rangle$. We note that

$$L_2\Psi(x_1, x_2) + \langle \nabla\Phi(x_1), F(x_1, x_2)\rangle = 0. \tag{8}$$

Since J is bounded, there exists $\alpha > 0$ such that

$$\alpha\Phi(x_1) \leq \Phi(x_1) + \varepsilon\Psi(x_1, x_2), \ \forall (x_1, x_2) \in \mathbb{R}^d \times \mathbb{R}^\ell, 0 \leq \varepsilon \leq 1.$$

Hence, if we define

$$U_t^\varepsilon = \Phi(X_t^{1,\varepsilon}) + \varepsilon\Psi(X_t^{1,\varepsilon}, X_t^{2,\varepsilon}),$$

the result will follow from

$$\sup_{0<\varepsilon\leq 1}\sup_{0\leq t\leq T} EU_t^\varepsilon < \infty.$$

But it follows from equation (1), Itô's formula and (8) that

$$\begin{aligned}\frac{d}{dt}EU_t^\varepsilon &= E\langle \nabla\Phi(X_t^{1,\varepsilon}) + \varepsilon\nabla_{x_1}\Psi(X_t^{1,\varepsilon}, X_t^{2,\varepsilon}), G(X_t^{1,\varepsilon}, X_t^{2,\varepsilon})\rangle \\ &\quad + \langle \nabla_{x_1}\Psi(X_t^{1,\varepsilon}, X_t^{2,\varepsilon}), F(X_t^{1,\varepsilon}, X_t^{2,\varepsilon})\rangle \\ &\leq c\big(1 + E\Phi(X_t^{1,\varepsilon})\big) \\ &\leq c'(1 + EU_t^\varepsilon),\end{aligned}$$

where we have used (A.1.2), (A.1.4), and the fact that for each $1 \leq r < q$, there exists a constant k such that $|x|^r \leq k(1 + |x|^q)$ for the first inequality, and the last inequality above for the second one. The result now follows from Gronwall's lemma.

We now return to the proof of lemma 1.We have that

$$E|Y_t^\varepsilon|^2 \leq K\left(1 + E\int_t^T |Y_s^\varepsilon|^2 ds\right),$$

and so

$$E|Y_t^\varepsilon|^2 \leq K\exp(KT)$$

The lemma now follows from the above estimates and the assumption (A.1.6.b).

Note that the argument in this proof together with Burkholder's inequality yields

$$\sup_{0<\varepsilon\leq 1} E\left(\sup_{0\leq t\leq T}|Y_t^\varepsilon|^2 + \int_0^T |Z_s^\varepsilon|^2 ds\right) < \infty$$

b Meyer-Zheng Method

Let us estimate that "conditional variation" of $\{Y_t^\varepsilon\}$. Denote $f_s^\varepsilon = f(X_s^{1,\varepsilon}X_s^{2,\varepsilon}, Y_s^\varepsilon)$ and take any partition $0 = t_0 < t_1 < \cdots < t_n = T$. One has

$$\begin{aligned}\sum_i E|E(Y_{t_{i+1}}^\varepsilon - Y_{t_i}^\varepsilon|F_{t_i})| &= \sum_i E\left|E\left(\int_{t_i}^{t_{i+1}} f_s^\varepsilon ds|F_{t_i}\right)\right| \\ &\leq E\int_0^T |f_s^\varepsilon| ds.\end{aligned}$$

Hence the conditional variation of Y^ε

$$V(Y^\varepsilon) = \sup E\left[\sum_i |E(Y^\varepsilon_{t_{i+1}} - Y^\varepsilon_{t_i}|F_{t_i})|\right]$$

(where the sup is taken over all partitions of the interval $[0, T]$) is bounded uniformly w.r.t ε. Consequently, from theorem 4 in Meyer-Zheng [6] and lemma 5 below, we can pass to the limit in the BSDE as $\varepsilon \to 0$, in the sense of Meyer-Zheng (along some subsequence):

$$(Y^\varepsilon, M^\varepsilon) \xrightarrow[M-Z]{\mathcal{L}} (\overline{Y}, \overline{M}),$$

where

$$\overline{Y}_t = g(X^1_T) + \int_t^T \bar{f}(X^1_s, \overline{Y}_s)ds + \overline{M}_T - \overline{M}_t,$$

and $\{\overline{M}_t, 0 \leq t \leq T\}$ is a martingale. Moreover, we have that (again along a subsequence)

$$(X^{1,\varepsilon}, Y^\varepsilon) \xrightarrow[M-Z]{\mathcal{L}} (X^1, \overline{Y}).$$

c Identification of the Limit

We need to show that $\overline{Y}_t = Y_t, \overline{M}_t = -\int_0^t Z_s dB_s$. The convergence of $X^{1,\varepsilon}$ tells us that $\forall \varphi \in C_b^\infty(\mathbb{R}^d), 0 < s < t < T$, and any function Φ_s of $X^{1,\varepsilon}_r, Y^\varepsilon_r$, $0 \leq r < s$ which is bounded and continuous in the Meyer-Zheng topology,

$$E\left\{\Phi_s(X^{1,\varepsilon} Y^\varepsilon)\left[\varphi(X^{1,\varepsilon}_t) - \varphi(X^{1,\varepsilon}_s) - \int_s^t \overline{\mathcal{L}}\varphi(X^{1,\varepsilon}_r)dr\right]\right\} \to 0, \quad \varepsilon \to 0,$$

due to [3], page 477 and lemma 5 below. Hence

$$E\left\{\Phi_s(X^1, \overline{Y})\left[\varphi(X^1_t) - \varphi(X^1_s) - \int_s^t \overline{\mathcal{L}}\varphi(X^1_r)ds\right]\right\} = 0.$$

and $M^\varphi_t \triangleq \varphi(X^1_t) - \varphi(x^1_0) - \int_0^t \overline{\mathcal{L}}\varphi(X^1_r)dr$ is a $\mathcal{F}^{X^1,\overline{Y}}_t$-martingale for any $\varphi \in C_b^\infty(\mathbb{R}^d)$, and there exists a $\mathcal{F}^{X^1\overline{Y}}_t$-Brownian motion $\{B_t\}$ s.t.

$$dX^1_t = \bar{b}(X^1_t)dt + \bar{\sigma}(X^1_t)dB_t.$$

For a fixed filtration $\mathcal{F}_t = \mathcal{F}^{X^1,\overline{Y}_t}$, we have:
(Y, Z) is the $\mathcal{F}_t^B$-adapted solution (hence it is also $\mathcal{F}_t$-adapted) of the BSDE

$$Y_t = g(X_T^1) + \int_t^T \bar{f}(X_s^1, Y_s)ds - \int_t^T Z_s dB_s,$$
$$E\int_0^T |Z_t|^2 dt < \infty,$$

$(\overline{Y}, \overline{M})$ is $\mathcal{F}_t$-adapted, $\overline{M}$ is a square-integrable $\mathcal{F}_t$-martingale, and

$$\overline{Y}_t = g(X_T^1) + \int_t^T \bar{f}(X_s^1, \overline{Y}_s)ds + \overline{M}_T - \overline{M}_t, E(\sup_t |\overline{Y}_t|^2) < \infty.$$

It remains to prove the uniqueness result: $Y_t = \overline{Y}_t$. If we define $M_t \triangleq -\int_0^t Z_s dB_s$, we deduce from Itô's formula that

$$|Y_t - \overline{Y}_t|^2 + [M - \overline{M}]_T - [M - \overline{M}]_t = 2\int_t^T \langle Y_s - \overline{Y}_s, \bar{f}(X_s^1, Y_s)$$
$$-\bar{f}(X_s^1, \overline{Y}_s)\rangle ds + 2\int_t^T \langle Y_s - \overline{Y}_s, dM_s - d\overline{M}_s\rangle.$$

Taking expectation and using Gronwall's lemma one obtains the uniqueness result.

d Three Lemmas

Now it remains to formulate and prove several lemmas from which we have exploited lemma 5 essentially.

LEMMA 3 *For any $\delta > 0$ there exist N_1 and $x^1, \ldots, x^{N_1} \in C([0,T];\mathbb{R}^d)$ s.t.*

$$P\left(\bigcap_{k=1}^{N_1}\{\sup_{0\le t\le T} |X_t^{1,\varepsilon} - x_t^k| > \delta\}\right) < \delta, \quad \forall\varepsilon > 0.$$

Proof The result follows from the tightness of the sequences $\{X^{1,\varepsilon}\}$ and the separability of $C([0,T];\mathbb{R}^d)$

LEMMA 4 *For any $\delta > 0$ there exist N_2 and $y^1, \ldots, y^{N_2} \in D([0,T];\mathbb{R}^\ell)$ s.t.*

$$P\left(\bigcap_{k=1}^{N_2}\{\lambda(0 \le t \le T, |Y_t^\varepsilon - y_t^k| > \delta) > \delta\}\right) < \delta, \quad \forall\varepsilon > 0,$$

where λ denotes Lebesgue's measure.

Proof For each $\varepsilon > 0$, let

$$R_\varepsilon \triangleq \sup_{0\le t\le T} |Y_t^\varepsilon|, \quad N_\varepsilon^{h,K,r} \triangleq \sum_{i=-K}^{K} \left(N^{\varepsilon,r}_{i/h,(i+1)/h} + N^{\varepsilon,r}_{i/h,(i-1)/h} \right),$$

where $N^{\varepsilon,r}_{i/h,(i+1)/h}$ is the number of upcrossings of $[i/h,(i+1)/h]$ and $N^{\varepsilon,r}_{i/h,(i-1)/h}$ the number of downcrossings of $[(i-1)/h, i/h]$ by $Y^{\varepsilon,r}$ – the rth component of Y^ε – on $[0,T]$.

We choose $h \in \mathbb{N}, h \ge 3/\delta$, and $K \in \mathbb{N}$ s.t. $P(R_\varepsilon > K/h) < \delta/2\ \forall \varepsilon > 0$.

For each $n \in \mathbb{N}$ we let $y^1, \ldots, y^{N(n)}$ denote the sequence of all step functions from $[0,T]$ into $\mathbb{R}^\ell$ which are constant on each $[j/n,(j+1)/n[, j = 0,1,\ldots,[T_n]$, and whose components take all possible values $\ell/n, -Kn/h \le l \le Kn/h, l \in \mathbb{N}$.

Note that, provided $n > h$,

$$\bigcap_{k=1}^{N(n)} \{\lambda(0\le t\le T; |Y_t^{\varepsilon,r} - y_t^{k,r}| > \delta) > \delta\} \subset \{R_\varepsilon > K/h\} \cup \{N_\varepsilon^{n,K,r} > 3n/h\}.$$

Indeed, to each $\{Y_t^{\varepsilon,r}(\omega)\}$ such that $R_\varepsilon(\omega) \le K/h$, and each $1 \le r \le l$, we can associate one number $k, 1 \le k \le N(n)$ such that $|Y^{\varepsilon,r}_{j/n}(\omega) - y^{k,r}_{j/n}| < 1/h, 0 \le j \le [T_n]$. Let $t \in [j/n,(j+1)/n[$ be such that $|Y_t^{\varepsilon,r}(\omega) - y_t^{k,r}| > \delta \ge 3/h$. Then $|Y_t^{\varepsilon,r}(\omega) - Y^{\varepsilon,r}_{j/n}(\omega)| > 2/h$, which gives at least one contribution to the number $N_\varepsilon^{h,K,r}$. Thus, if $R_\varepsilon(\omega) \le K/h$ and $\lambda(0 \le t \le T; |Y_t^{\varepsilon,r}(\omega) - y_t^{k,r}| > \delta) > \delta$ for each $k, 1 \le k \le N(n)$, one gets $N_\varepsilon^{h,K,r} \ge n\delta \ge 3n/h$.

Finally, it follows from the tightness of the sequence $\{Y^\varepsilon\}_{\varepsilon>0}$ that we can choose n large enough such that $P(N_\varepsilon^{h,K,r} \ge 3n/h) < (2l)^{-1}\delta, \forall \varepsilon > 0, 1 \le r \le l$ see [6]. The lemma is proved.

LEMMA 5 *For any $t < T$*

$$F_\varepsilon = \int_t^T [f(X_s^{1,\varepsilon}, X_s^{2,\varepsilon}, Y_s^\varepsilon) - \bar{f}(X_s^{1,\varepsilon}, Y_s^\varepsilon)]ds \to 0$$

in probability, as $\varepsilon \to 0$ (in fact the convergence holds in $L^1(\Omega)$).

Proof Let $h(x_1,x_2,y) \triangleq f(x_1,x_2,y) - \bar{f}(x_1,y)$. From lemmas 3 and 4, to each $\delta > 0$ we associate $N_1, x^1, \ldots, x^{N_1}, N_2, y^1 \ldots, y^{N_2}$ such that $P(A_1) < \delta$, where $A_1 = \cap_{k=1}^{N_1}\{\sup_t |X_t^{1,\varepsilon} - x_t^k| > \delta\}$ and $P(A_2) < \delta$, where $A_2 = \cap_{k=1}^{N_2} \{\lambda(t; |Y_t^\varepsilon - y_t^k| > \delta) > \delta\}$.

We have

$$A_1^c = \bigcup_{k=1}^{N_1}\{\sup_t |X_t^{1,\varepsilon} - x_t^k| \le \delta\} = \bigcup_{k=1}^{N_1} B_1^k,$$

and we can rewrite

$$A_1^c = \bigcup_{k=1}^{N_1} \overline{B}_1^k,$$

where $\overline{B}_1^k \subset B_1^k \forall k$, and the $\overline{B}_1^k$'s are disjoint.

Following the proof of lemma 4, if $N_2 = N(n)$, to each ω belonging to $\{R_\varepsilon \le K/h\}$, we associate one (measurable) $k(\omega), 1 \le k(\omega) \le N_2$, such that $\sup_j |Y_{j/n}^\varepsilon(\omega) - y_{j/n}^{k(\omega)}| < \sqrt{l}/n$.

Let

$$B_2^\ell = \{\lambda(t; |Y_t^\varepsilon - y_t^\ell| > \delta) \le \delta\} \cap \{k(\omega) = \ell\} \cap \{R_\varepsilon \le K/h\}.$$

Then the B_2^ℓ's are disjoint, and $A_2^c \cap \{R_\varepsilon \le K/h\} = \bigcup_{\ell=1}^{N_2} B_2^\ell$. We have that

$$F_\varepsilon \le 1_{A_1} F_\varepsilon + 1_{A_2} F_\varepsilon + 1_{\{R_\varepsilon > K/h\}} F_\varepsilon + \sum_{k=1}^{N_1}\sum_{\ell=1}^{N_2} 1_{\overline{B}_1^k \cap B_2^\ell} F_\varepsilon$$

$$E|1_{A_1} F_\varepsilon| \le ||F_\varepsilon||_{L^2(\Omega)} [P(A_1)]^{1/2} \le c\sqrt{\delta}.$$

$$E|1_{A_2} F_\varepsilon| \le ||F_\varepsilon||_{L^2(\Omega)} [P(A_2)]^{1/2} \le c\sqrt{\delta}.$$

$$E|1_{\{R_\varepsilon > K/h\}} F_\varepsilon| \le ||F_\varepsilon||_{L^2(\Omega)} [P(\{R_\varepsilon > K/h\})]^{1/2} \le c\sqrt{\delta}.$$

On $\overline{B}_1^k \cap B_2^\ell$ we write

$$\begin{aligned} F_\varepsilon &= \int_t^T [h(X_s^{1,\varepsilon}, X_s^{2,\varepsilon}, Y_s^\varepsilon) - h(x_2^k, X_s^{2,\varepsilon}, Y_s^\varepsilon)]ds \\ &\quad + \int_t^T [h(x_s^k, X_s^{2,\varepsilon}, Y_s^\varepsilon) - h(x_s^k, X_s^{2,\varepsilon}, y_s^\ell)]ds + \int_t^T h(x_s^k, X_s^{2,\varepsilon}, y_s^\ell)ds \\ &= F_\varepsilon^{1,k} + F_\varepsilon^{2,k,\ell} + F_\varepsilon^{3,k,\ell}. \end{aligned}$$

We get

$$E|\sum_{k,\ell} 1_{\overline{B}_1^k \cap B_2^\ell} F_\varepsilon^{1,k}| = E|\sum_k 1_{\overline{B}_1^k} F_\varepsilon^{1,k}| \le 2\alpha(\delta) E\int_t^T (1 + |Y_s^\varepsilon|)ds \le c\alpha(\delta),$$

$$E|\sum_{k,\ell} 1_{\overline{B}_1^k \cap B_2^\ell} F_\varepsilon^{2,k,\ell}| \le 2KE\sum_{k,\ell} 1_{\overline{B}_1^k \cap B_2^\ell} \int_t^T |Y_s^\varepsilon - y_s^\ell| ds$$

$$\leq 2K\sum_\ell E\left(1_{B_2^\ell}\int_t^T |Y_s^\varepsilon - y_s^\ell|ds\right)$$

$$\leq 2K\sum_\ell \left[\delta TP(B_2^\ell) + E\left(1_{B_2^\ell}\int_{\{t\leq s\leq T;|Y_s^\varepsilon - y_s^\ell|>\delta\}} |Y_s^\varepsilon - y_s^\ell|ds\right)\right]$$

$$\leq 2K\left[\delta T + \delta\left\{E\|Y^\varepsilon\|_\infty + E\sum_\ell(\|y^\ell_\cdot\|_\infty 1_{B_2^\ell})\right\}\right]$$

$$\leq 2K\delta\left(T + E\|Y^\varepsilon\|_\infty + E\sum_\ell(\|Y^\varepsilon\|_\infty + 1/n)1_{B_2^\ell}\right)$$

$$\leq 4K\delta(T + E\|Y^\varepsilon\|_\infty) \leq c\delta.$$

Finally, ergodicity implies that for each k, l $F_\varepsilon^{3,k,l} \to 0$ in $L^1(\Omega)$ as $\varepsilon \to 0$. Hence,

$$\limsup_{\varepsilon\to 0} E|F_\varepsilon| \leq c(3\sqrt{\delta} + \alpha(\delta) + \delta), \forall\delta > 0,$$

and the result follows.

5 PROOF OF THE COROLLARIES

Proof of corollary 1 Note that since Y_0^ε is deterministic,

$$Y_0^\varepsilon = E\left[g(X_T^{1,\varepsilon}) + \int_0^T f(X_t^{1,\varepsilon}, X_t^{2,\varepsilon}, Y_t^\varepsilon)dt\right].$$

It follows from the results proved above that

$$\xi_\varepsilon \triangleq g(X_T^{1,\varepsilon}) + \int_0^T f(X_t^{1,\varepsilon}, X_t^{2,\varepsilon}, Y_t^\varepsilon)dt$$

converges in law, as $\varepsilon \to 0$, towards

$$g(X_T^1) + \int_0^T \bar{f}(X_t^1, Y_t)dt,$$

and moreover from some arguments in lemma 1 we have that $E(|\xi_\varepsilon|^2) \leq c$. Hence by uniform integrability $\lim_{\varepsilon\to 0} E(\xi_\varepsilon) = E(\lim_{\varepsilon\to 0} \xi_\varepsilon)$, i.e.

$$Y_0^\varepsilon \to Y_0 = E\left[g(X_T^1) + \int_0^T \bar{f}(X_t^1, Y_t)dt\right].$$

Proof of corollary 2 For each $x = (x_1, x_2)' \in \mathbb{R}^{d+\ell}$, let $\{X_s^{x,\varepsilon} = (X_s^{1,x,\varepsilon}, X_s^{2,x,\varepsilon})', s \geq 0\}$ denote the solution of the SDE (1) starting from the initial condition $(X_0^{1,x,\varepsilon}, X_0^{2,x,\varepsilon}) = (x_1, x_2)$. Let t be an arbitrary positive real number. We denote by $\{Y_s^{t,x,\varepsilon}; 0 \le s \le t\}$ the solution of the BSDE (3) where $(X^{1,\varepsilon}, X^{2,\varepsilon})$ is replaced by $(X^{1,x,\varepsilon}, X^{2,x,\varepsilon})$. and T by t. In other words, $\{Y_s^{t,x,\varepsilon}; 0 \le s \le t\}$ solves the BSDE

$$Y_s^{t,x,\varepsilon} = g(X_t^{1,x,\varepsilon}) + \int_s^t f(X_r^{1,x,\varepsilon}, X_r^{2,x,\varepsilon}, Y_r^{t,x,\varepsilon})dr - \int_s^t Z_s^\varepsilon dW_r, 0 \le s \le t.$$

We have that the function $u^\varepsilon : \mathbb{R}_+ \times \mathbb{R}^{d+\ell} \to \mathbb{R}^k$ defined by

$$u^\varepsilon(t,x) \triangleq Y_0^{t,x,\varepsilon}, (t,x) \in \mathbb{R}_+ \times \mathbb{R}^{d+\ell}$$

is the unique viscosity solution of the systems of PDEs (4), see [9], [10].

Similarly, for each $x \in \mathbb{R}^d$, let $\{X_s^{1,x}, s \ge 0\}$ denote the solution of the SDE associated to the limiting process X^1, with initial condition $X_0^{1,x} = x$, and for each t positive, let $\{Y_s^{t,x}; 0 \le s \le t\}$ be the solution of the BSDE (5) with X^1 replaced by $X^{1,x}$, and T by t. In other words, $\{Y_s^{t,x}; 0 \le s \le t\}$ solves the BSDE

$$Y_s^{t,x} = g(X_t^{1,x}) + \int_s^t \bar{f}(X_r^{1,x}, Y_r^{t,x})dr - \int_s^t Z_r dB_r, \quad 0 \le s \le t.$$

We have that the function $u : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^k$ defined by

$$u(t,x) \triangleq Y_0^{t,x}, \quad (t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$$

is the unique viscosity solution of the systems of PDEs (6).

Hence corollary 2 follows easily from corollary 1.

References

[1] Bensoussan, A. (1988). *Perturbation methods in optimal control*, Dunod Gauthier-Villars.

[2] Bensoussan, A., Lions, L. and Papanicolau, G. (1978). *Asymptotic analysis for periodic structures*, North Holland.

[3] Ethier, S. N. and Kurtz, T. G. (1986). *Markov processes*, John Wiley & Sons, N. Y. *et al.*

[4] Freidlin M. (1996). *Markov processes and differential equations: asymptotic problems*, Birkhäuser.

[5] Krylov, N. V. and Safonov, M. V. (1981). A certain property of solutions of parabolic equations with measurable coefficients, *Math. USSR Izvestija*, **16**(1), 151–164.

[6] Meyer, P. A. and Zheng, W.A. (1984). Tightness criteria for laws of semimartingales, *Ann. Inst. Henri Poincaré*, **20**, 353–372.

[7] Papanicolau, G. C., Stroock, D. W. and Varadhan, S. R. S. (1977). Martingale approach to some limit theorems, in *Conf. on Statistical Mechanics, Dynamical systems and Turbulence, Duke Univ.*, Reed, M., ed., *Duke Univ. Math. series* **3**.

[8] Pardoux, E. and Peng, S. (1990). Adapted solution of a backward stochastic differential equations, *Systems & Control Letters*, **14**, 55–61.

[9] Pardoux, E. and Peng, S. (1992). Backward SDEs and quasilinear PDEs, in *Stochastic partial differential equations and their applications*, Rozovskii, B. L., Sowers, R. B., eds., *LNCIS*, **176**, Springer.

[10] Pradeilles, F., Propagation d'ondes dans les équations de réaction–diffusion, submitted.
[11] Pradeilles, F., Wavefront propagation for reaction-diffusion systems and backward SDE's, *Annals of Probab.* to appear.
[12] Revuz, D. (1975). *Markov Chains*, North-Holland and American Elsevier.
[13] Stroock, D.W. and Varadhan, S. R. S. (1979). *Multidimensional diffusion processes*, Springer.