Optimal control of two scale stochastic systems in infinite dimensions: the BSDE approach

Giuseppina Guatteri,
Dipartimento di Matematica,
Politecnico di Milano,
Piazza Leonardo da Vinci 32,
20133 Milano, Italia.
e-mail: giuseppina.guatteri@polimi.it

Gianmario Tessitore,
Dipartimento di Matematica e Applicazioni
Universitá Milano-Bicocca,
via R. Cozzi 53 - Edificio U5,
20125 Milano, Italia.
e-mail: gianmario.tessitore@unimib.it

Abstract

In this paper we study, by probabilistic techniques, the convergence of the value function for a two-scale, infinite-dimensional, stochastic controlled system as the ratio between the two evolution speeds diverges. The value function is represented as the solution of a *backward stochastic differential equation* (BSDE) that it is shown to converge towards a *reduced* BSDE. The noise is assumed to be additive both in the slow and the fast equations for the state. Some non degeneracy condition on the slow equation are required. The limit BSDE involves the solution of an *ergodic* BSDE and is itself interpreted as the value function of an auxiliary stochastic control problem on a reduced state space.

1 Introduction

In this paper we study the convergence of the value function of an optimal control problem for a singularly perturbed infinite dimensional state equation as

$$\begin{cases} dX_t = AX_t^{\varepsilon,\alpha} + b(X_t^{\varepsilon,\alpha}, Q_t^{\varepsilon,\alpha}, \alpha_t)dt\ + Rd\mathcal{W}_t^1, & X_0^{\varepsilon,\alpha} = x_0, \\ \varepsilon dQ_t^{\varepsilon,\alpha} = (BQ_t^{\varepsilon,\alpha} + F(X_t^{\varepsilon,\alpha}, Q_t^{\varepsilon,\alpha})\,dt + G\rho(\alpha_t)dt + \varepsilon^{1/2} G\, d\mathcal{W}_t^2, & Q_0^{\varepsilon} = q_0, \end{cases}$$

where the state processes X and Q are Hilbert valued, A and B are unbounded linear operators, α represents the control, $(\mathcal{W}_t^1)_{t\geq 0}$, $(\mathcal{W}_t^2)_{t\geq 0}$ are infinite dimensional cylindrical Wiener processes, b, F, ρ are functions satisfying suitable assumptions. We notice that the presence of the constant ε in the second equation corresponds to the fact that Q evolves with a speed which is larger by a factor $1/\varepsilon$ then the speed of evolution of the component X. In other words the above equation is a good model for a so called *two scale system*. The optimal control problem is then completed by a standard cost functional of the form :

$$J^{\varepsilon}(x_0, q_0, \alpha) := \mathbb{E}\left(\int_0^1 l(X_t^{\varepsilon,\alpha}, Q_t^{\varepsilon,\alpha}, \alpha_t)dt + h(X_1^{\varepsilon,\alpha}) \right)$$

Several authors have studied the convergence of singular stochastic control problems in finite dimensions, see for instance [1], [2], [13], [14], [16]. In particular [1] has been an inspiration for the present work. In that paper authors represent the value function of a singular stochastic control problem, in finite dimensions, by the solution, in viscosity sense, of an Hamilton Jacobi Bellman equation. Then they show, by PDE methods their convergence towards the solution, again in viscosity sense, of a *reduced* parabolic PDE with smaller state space and a new nonlinearity usually called *effective Hamiltonian*. Such analysis is performed in the case of periodic boundary conditions. Although PDE techniques perfectly fit the finite dimensional case allowing to cover general situations, including state equations with control dependent diffusions that require introduction of fully non-linear H.J.B. equations, they seem not to be adaptable to the infinite dimensional case, and consequently to the case of two scale stochastic control problems for stochastic PDEs. The reason essentially is the difficulty of handling, by analytic tools and viscosity solutions, parabolic equations in infinite variables, see the discussion in the Introduction of [11].

The purpose of the present paper is twofold. On one side we wish to show that Backward Stochastic Differential Equations (BSDEs) are, in general, an efficient way to represent the limit of the value functions of two scale systems when the ratio between the two evolutions' speed diverge. On the other we wish to show that, in such a way, we can cover the case of infinite dimensional state equations (that is the case of two scale systems described by stochastic PDEs) that, at our best knowledge, was not considered in the existing letterature. As a counterpart we notice that we consider state equation in which the control only affects the drift and in which the noise of the slow component is assumed to be non-degenerate. Such restrictions seem not to be intrinsic in the BSDEs approach but allows essential technical simplifications.
To be more specific our main result will be to prove that if

$$v^{\varepsilon}(x_0, q_0) := \inf_{\alpha} J^{\varepsilon}(x_0, q_0, \alpha)$$

then

$$v^{\varepsilon}(x_0, q_0) \to Y_0$$

where $(X, \bar{Y}, \bar{Z})$ is the unique solution of the following decoupled forward backward system of stochastic differential equations

$$\begin{cases} dX_t = AX_t\,dt + R\,dW^1_t, \\ -d\bar{Y}_t = \lambda(X_t, \bar{Z}_t)\,dt - \bar{Z}\,dW^1_t, \\ X_0 = x_0, \quad \bar{Y}_1 = h(X_1). \end{cases}$$

It is important to notice that the 'reduced nonlinearity' λ is itself a component of the unique solution $(\check{Y}, \check{Z}, \lambda)$ of a parametrized *ergodic* BSDE (see (4.1) and Theorem 4.2) as they were introduced in [9] (see [8] and [15] as well). Function λ can itself be interpreted as the optimal cost of an ergodic optimal control problem. Moreover, as it happens in the finite dimensional case, the space in which the above reduced BSDE lives is a subspace of the original one (corresponding to the slow evolution). As a by-product of our main result, using the Bismut Elworthy formula in in [12] we immediately get that the solution of the reduced BSDE, and therefore the limit value function, depends on x_0 in a differentiable way and is linked to the unique *mild* solution of a semilinear parabolic PDE in infinite dimensional spaces:

$$\begin{cases} \dfrac{\partial v(t,x)}{\partial t} + \dfrac{1}{2}\mathrm{Tr}[\nabla^2_x v(t,x)] = \lambda(x, \nabla v(t,x)), \qquad t \in [0,1],\, x \in H, \\ v(1,x) = h(x). \end{cases}$$

Finally, exploiting the concavity of λ we give a representation of $\bar{Y}_t$ as the value function of an auxiliar stochastic control problem on a reduced state space.

The paper is organized in the following way. In Section 2 we report some notation and assumptions while Section 3 contains some estimates on the two scale state equation that will be useful in the paper. In Section 4 we introduce parametrized ergodic BSDEs and study their regularity with respect to parameters. In Section 5 we state the form of the limit equations and prove a convergence result for BSDEs that represents the main technical issue of this paper. In Section 6, we apply our results to a stochastic singular control problem. Finally, in section 7 we interpret the solution of the reduced BSDE in terms of a stochastic optimal control problem.

2 Notation and preliminary results

Given a Banach space E, the norm of its elements x will be denoted by $|x|_E$, or even by $|x|$ when no confusion is possible. If F is another Banach space, $L(E,F)$ denotes the space of bounded linear operators from E to F, endowed with the usual operator norm. When $F=\mathbb{R}$ the dual space $L(E,\mathbb{R})$ will be denoted by E^*. The letters Ξ, H and K will always be used to denote Hilbert spaces. The scalar product is denoted $\langle\cdot,\cdot\rangle$, equipped with a subscript to specify the space, if necessary. All Hilbert spaces are assumed to be real and separable and the dual of a Hilbert space will never be identified with the space itself. By $L_2(\Xi,H)$ and $L_2(\Xi,K)$ we denote the spaces of Hilbert-Schmidt operators from Ξ to H and to K, respectively. Finally $\mathcal{G}(K,H)$ is the space of all Gateaux differentiable mappings ϕ from K to H such that the map $(k,v)\to\nabla\phi(k)v$ is continuous from $K\times K$ to H; see [11] for details.

Let $W^1=(W^1_t)_{t\geq 0}$ and $W^2=(W^2_t)_{t\geq 0}$ be two independent cylindrical Wiener processes with values in Ξ, defined on a complete probability space $(\Omega,\mathcal{F},\mathbb{P})$. By $\{\mathcal{F}_t,\ t\in[0,T]\}$ we will denote the natural filtration of (W^1,W^2), augmented with the family $\mathcal{N}$ of $\mathbb{P}$- null sets of $\mathcal{F}$. Obviously, the filtration $(\mathcal{F}_t)$ satisfies the usual conditions of right-continuity and completeness. All the concepts of measurability for stochastic processes will refer to this filtration. By $\mathcal{P}$ we denote the predictable σ-algebra on $\Omega\times[0,T]$ and by $\mathcal{B}(\Lambda)$ the Borel σ-algebra of any topological space Λ.
Next we define the following two classes of stochastic processes with values in a Hilbert space V. Given an arbitrary time horizon T and constant $p\geq 1$:

- $L^p_{\mathcal{P}}(\Omega\times[0,T];V)$ denotes the space of equivalence classes of processes $Y\in L^p(\Omega\times[0,T];V)$ admitting a predictable version. It is endowed with the norm
$$|Y|=\left(\mathbb{E}\int_0^T|Y_s|^p\,ds\right)^{1/p}.$$
- $L^{p,loc}_{\mathcal{P}}(\Omega\times[0,+\infty[;V)$ denotes the set of processes defined on $\mathbb{R}^+$ such that their restriction to an arbitrary $[0,T]$ belongs to $L^p_{\mathcal{P}}(\Omega\times[0,T];V)$.
- $L^p_{\mathcal{P}}(\Omega;C([0,T];V))$ denotes the space of predictable processes Y with continuous paths in V, such that the norm
$$\|Y\|_p=(\mathbb{E}\sup_{s\in[0,T]}|Y_s|^p)^{1/p}$$
is finite. The elements of $L^p_{\mathcal{P}}(\Omega;C([0,T];V))$ are identified up to indistinguishability.
- $L^{p,loc}_{\mathcal{P}}(\Omega:\mathcal{C}[0,+\infty[;V)$ denotes the set of processes defined on $\mathbb{R}^+$ such that their restriction to an arbitrary $[0,T]$ belongs to $L^p_{\mathcal{P}}(\Omega;C([0,T];V))$.

Given Φ in $L^2_{\mathcal{P}}(\Omega\times[0,T];L_2(\Xi,V))$, the Itô stochastic integrals $\int_0^t\Phi_s\,dW^1_s$ and $\int_0^t\Phi_s\,dW^2_s$, $t\in[0,T]$, are V-valued martingales belonging to $L^2_{\mathcal{P}}(\Omega;C([0,T];V))$. The previous definitions have obvious extensions to processes defined on the entire positive real line $\mathbb{R}^+$.

3 The forward system

For arbitrarily fixed $x_0\in H$ and $q_0\in K$ we consider the following system of controlled stochastic differential equations in $H\times K$:

$$\begin{cases} dX_t=AX_tdt+R\,dW^1_t, & X_0=x_0,\ t\geq 0,\\ \varepsilon dQ^\varepsilon_t=(BQ^\varepsilon_t+F(X^\varepsilon_t,Q^\varepsilon_t))\,dt+\varepsilon^{1/2}G\,dW^2_t, & Q^\varepsilon_0=q_0,\ t\geq 0\end{cases}\tag{3.1}$$

where the "slow" variable X takes its values in H and the "fast" variable Q^ε takes its values in K, $\varepsilon \in]0,1]$ is a small parameter.
Finally $A : D(A) \subset H \to H$ and $B : D(B) \subset K \to K$ are unbounded linear operators generating C_0-semigroups $\{e^{tA}\}_{t\geq 0}$ and $\{e^{tB}\}_{t\geq 0}$ over H and K, respectively, while R and G are linear bounded operators from Ξ to H (respectively to K).
Moreover, we make the following, standard assumptions:

Hypothesis 3.1 *$A : D(A) \subset H \to H$ is the generator of a semigroup, $\{e^{tA}\}_{t\geq 0}$, such that $|e^{tA}|_{L(H,H)} \leq Me^{\omega_A t}, t \geq 0$ for some positive constants M_A and ω_A. $B : D(B) \subset K \to K$ is a linear, unbounded operator that generates a C_0- semigroup $\{e^{tB}\}_{t\geq 0}$ such that $|e^{tB}|_{L(K,K)} \leq M_B e^{\omega_B t}, t \geq 0$ for some $M_B > 0$. Moreover there exist constants $L > 0$ and $\gamma \in [0, \frac{1}{2}[$, s.t.:*

$$|e^{sA}|_{L_2(\Xi,H)} + |e^{sB}|_{L_2(\Xi,K)} \leq \frac{L}{(1 \wedge s)^\gamma},$$

We also assume that $R \in L(\Xi; H)$ and admits a bounded right inverse $R^{-1} \in L(H; \Xi)$.

Hypothesis 3.2 *$F : H \times K \to K$ is bounded and there exists a constant L_F for which:*

$$|F(x, y) - F(u, v)|_K \leq L_F(|x - u|_H + |y - v|_K)$$

for every $x, u \in H$, $y, v \in K$. Moreover we assume that for every $x \in H$, $F(x, \cdot)$ is Gateaux differentiable, more precisely, $F(x, \cdot) \in \mathcal{G}^1(K, K)$.

Hypothesis 3.3 *We ask $B + F$ to be dissipative i.e. there exists some $\mu > 0$ such that:*

$$\langle Bq + F(x, q) - (Bq' + F(x, q')), q - q'\rangle \leq -\mu|q - q'|^2$$

for all $x \in H, q, q' \in D(B)$.

Hypothesis 3.4 *$G \in \mathcal{L}(\Xi, K)$.*

Given any cylindrical Wiener process $(\beta_t)_{t\geq 0}$ with values in Ξ we denote by $(\beta_t^B)_{t\geq 0}$ the stochastic convolution

$$\beta_s^B = \int_0^s e^{(s-\ell)B} d\beta_\ell.$$

In the following we shall assume as in [9] that:

Hypothesis 3.5 *:* $\sup_{s>0} \mathbb{E}|\beta_s^B|^2 < \infty$.

Remark 3.6 *Notice that since (β_t) is a centered gaussian process this implies that, $\forall p \geq 1$ it holds* $\sup_{s>0} \mathbb{E}|\beta_s^B|^p < \infty$*. Moreover Hypothesis* 3.5 *is verified whenever B is a strongly dissipative operator.*

Remark 3.7 *Under Hypothesis 3.1 equation (3.1) admits a unique mild solution $(X_t^{x_0})$ that has continuous trajectories and satisfies $\mathbb{E}(\sup_{t\in[0,1]} |X_t^{x_0}|^p) \leq c_p(1 + |x_0|^p)$. See [11] for the proof.*

Lemma 3.8 *Let $(\Gamma_s)_{s\geq 0}$ be a given, K-valued, predictable process with $\Gamma \in L^{p,loc}_{\mathcal{P}}(\Omega \times [0,\infty[, H)$ and let $(g)_{s\geq 0}$ be a given, Ξ-valued, process with $g \in L^{p,loc}_{\mathcal{P}}(\Omega \times [0, +\infty[, K)$. We introduce the following equation:*

$$dQ_s = (BQ_s + F(\Gamma_s, Q_s))\, ds + g_s ds + G d\beta_s,\ s \geq 0, \qquad Q_0 = q_0, \tag{3.2}$$

that admits a unique mild solution $Q \in L^{p,loc}_{\mathcal{P}}(\Omega; C[0, +\infty[, K)$.

Under hypotheses (3.1)–(3.4), for all $p \geq 1$, there exists a constant k_p (independent on T) such that for all $T > 0$:

$$\sup_{s\in[0,T]} \mathbb{E}|Q_s|^p \leq k_p(1 + |q_0|^p + \sup_{s\in[0,T]} \mathbb{E}|\Gamma_s|^p + \sup_{s\in[0,T]} \mathbb{E}|\beta^B_s|^p + \sup_{s\in[0,T]} \mathbb{E}|g_s|^p) \tag{3.3}$$

Moreover if $(\Gamma'_s)_{s\geq 0}$ is another K-valued, predictable processes in $L^2_{\mathcal{P}}(\Omega\times[0,T], H)$ and Q' is the mild solution of equation:

$$dQ'_s = (BQ'_s + F(\Gamma'_s, Q'_s))\, ds + g_s ds + G d\beta_s,\ s \geq 0, \qquad Q_0 = q_0,$$

then, for all $T > 0$,

$$|Q_T - Q'_T| \leq K \int_0^T e^{-\mu(T-\ell)} |\Gamma_\ell - \Gamma'_\ell| d\ell \quad \mathbb{P}\text{-a.s.}$$

where again K does not depend on T.

Proof. For the reader's convenience we briefly report the argument which is a slight modification of the one in [7], Section 6.3.2.
Let $Z_s = e^{\mu s}(Q_s - \beta^B_s)$. By Ito rule (going through Yosida approximations) we deduce that (Z) is the mild solution of the following equation

$$dZ_s = \mu Z_s + BZ_s + e^{\mu s} F(\Gamma_s, e^{-\mu s} Z_s + \beta^B_s))ds + e^{\mu s} g_s ds.$$

Differentiating $\sqrt{|Z_s|^2 + \varepsilon}$ (going, once more, through Yosida approximations), using dissipativity of $B + F$ as in Hypothesis 3.3 we obtain

$$|Z_s| \leq \sqrt{|Z_s|^2 + \varepsilon} \leq \sqrt{q_0^2 + \varepsilon} + \int_0^s e^{\mu\ell} \left| F(\Gamma_\ell, \beta^B_\ell) + g_\ell \right| d\ell + \mu \int_0^s \left[\sqrt{|Z_\ell|^2 + \varepsilon} - |Z_\ell| \right] d\ell$$

Letting $\varepsilon \to 0$, by dominated convergence we obtain:

$$|Z_s| \leq |q_0| + \int_0^s e^{\mu\ell} \left| F(\Gamma_\ell, \beta^B_\ell) + g_\ell \right| d\ell.$$

Recalling the definition of Z we conclude

$$|Q_s| \leq |\beta^B_s| + e^{-\mu t}|q_0| + \int_0^s e^{-\mu(s-\ell)} \left| F(\Gamma_\ell, \beta^B_\ell) + g_\ell \right| d\ell$$

and by Holder inequality (for the last term)

$$|Q_s|^p \leq 3^p|\beta^B_s|^p + 3^p e^{-p\mu t}|q_0|^p + \left(\int_0^s e^{-p^* \frac{\mu}{2}(s-\ell)} d\ell \right)^{p/p^*} 3^p \int_0^s e^{-p\frac{\mu}{2}(s-\ell)} |F(\Gamma_\ell, \beta^B_\ell) + g_\ell|^p d\ell$$

The claim then follows by Hypothesis 3.2
The proof of the last statement is similar (and easier) noticing that

$$d_s(Q_s - Q'_s) = B(Q_s - Q'_s)ds + [F(\Gamma_s, Q_s) - F(\Gamma'_s, Q'_s)]ds$$

and arguing as before □

As a special case of equation (3.2), for all $x \in H$ and $q_0 \in K$ we denote by $\hat{Q}^{x,q_0}$ the unique mild solution of equation:

$$d\hat{Q}^{x,q_0}_s = (B\hat{Q}^{x,q_0}_s + F(x, \hat{Q}^{x,q_0}_s))\, ds + d\hat{W}^2_s,\ s \geq 0, \qquad \hat{Q}^{x,q_0}_0 = q_0. \tag{3.4}$$

where $\hat{W}^2_s = \varepsilon^{-1/2} W^2_{\varepsilon s}$ is a cylindrical Wiener process.

4 The ergodic BDSE parametrized

We introduce a function $\psi : H \times K \times \Xi^* \times \Xi^* \to \mathbb{R}$ and assume the following

Hypothesis 4.1 *Function ψ is measurable and there exist $L_q, L_\xi, L_x, L_z > 0$ such that $\forall\, q, q' \in K$, $\xi, \xi' \in \Xi^*$, $x, x' \in H$, $z, z' \in \Xi^*$:*

$$|\psi(x,q,z,\xi) - \psi(x',q',z',\xi')| \leq L_x(1+|z|)|x-x'| + L_z|z-z'| + L_q(1+|z|)|q-q'| + L_\xi|\xi-\xi'|.$$

Moreover we assume that $\sup_{x\in H, q\in K} |\psi(x,q,0,0)| < +\infty$

The next result states existence of a solution to the so called *ergodic backward stochastic differential equation.*

$$-\mathrm{d}\check{Y}_t = [\psi(x, \hat{Q}^{x,q_0}, z, \check{\Xi}_t) - \lambda(x,z)]\,\mathrm{d}t - \check{\Xi}_t \mathrm{d}W_t^2, \quad \forall\, t \geq s \tag{4.1}$$

Its proof is in large part contained in [9] Theorem 4.4 and Corollary 5.9.

Theorem 4.2 *Under Hypotheses 3.1, 3.2, 3.3, 3.4, 3.5 and 4.1 there exist measurable functions $\check{v} : H \times K \times \Xi^* \to \mathbb{R}$, $\check{\zeta} : H \times K \times \Xi^* \to \mathbb{R}$, $\lambda : H \times \Xi^* \to \mathbb{R}$ with*

$$|\check{v}(x,q,z)| \leq c(1+|z|)|q| \tag{4.2}$$

(where $c > 0$ depends only on the constants introduced in the above mentioned Hypotheses) such that the following holds. If we set:

$$\check{Y}_t^{x,q_0,z} = \check{v}(x, \hat{Q}_t^{x,q_0}, z), \quad \check{\Xi}_t^{x,q_0,z} = \check{\zeta}(x, \hat{Q}_t^{x,q_0}, z) \tag{4.3}$$

then $\check{\Xi}^{x,q_0,z}$ is in $L^{2,loc}([0,+\infty[, \Xi^*)$ and $(\check{Y}_t^{x,q,z}, \check{\Xi}^{x,q,z}, \lambda(x,z))$ is a solution to equation (4.1)*.*
Moreover we have:

$$|\lambda(x,z) - \lambda(x',z')| \leq L_x^1(1+|z|)|x-x'| + L_z^1|z-z'|. \tag{4.4}$$

for some posive constants L_x^1 and L_z^1.

Proof. Fix $x \in H$ and $z \in \Xi^*$; in [9] Theorem 4.4, authors prove existence of the function $\check{v}(x,\cdot,z)$, $\check{\zeta}(x,\cdot,z)$ and $\lambda(x,z)$ such that (4.2) holds and, if $\check{Y}^{x,q_0,z}$, $\check{\Xi}^{x,q_0,z}$ are defined as in (4.3), then $\check{\Xi}^{x,q_0,z}$ is in $L^{2,loc}([0,+\infty[, \Xi^*)$ and $(\check{Y}_t^{x,q_0,z}, \check{\Xi}^{x,q_0,z}, \lambda(x,z))$ is a solution to equation (4.1).
Measurably of $\check{v}$, $\check{\zeta}$ and λ with respect to all parameters follows by their construction (see again [9] Theorem 4.4).

We only need to prove (4.4). Fixed $x, x' \in H$ and $z, z' \in \Xi^*$ we set $\widetilde{\lambda} = \lambda(x,z) - \lambda(x',z')$, $\widetilde{Y} = Y^{x,0,z} - Y^{x',0,z'}$, $\widetilde{\Xi} = \check{\Xi}^{x,0,z} - \check{\Xi}^{x',0,z'}$,

$$\theta_t = \begin{cases} \dfrac{\psi(x, \hat{Q}_r^{x,0}, z, \check{\Xi}_r^{x,0,z}) - \psi(x, \hat{Q}_r^{x,0}, z, \check{\Xi}_r^{x',0,z'})}{|\check{\Xi}_r^{x,0,z} - \check{\Xi}_r^{x',0,z'}|_{K^*}^2} (\check{\Xi}_r^{x,0,z} - \check{\Xi}_r^{x',0,z'}), & \text{if } \check{\Xi}_r^{x,0,z} \neq \check{\Xi}_r^{x',0,z'} \\ 0 & \text{elsewhere} \end{cases}$$

and

$$f_t = \psi(x, \hat{Q}_r^{x,0}, z, \check{\Xi}_r^{x',0,z'}) - \psi(x', \hat{Q}_r^{x',0}, z', \check{\Xi}_r^{x',0,z'})$$

Then we have

$$\widetilde{Y}_0 + \lambda T = \widetilde{Y}_T + \int_t^T f_r \,\mathrm{d}r - \int_t^T \widetilde{\Xi}_r(\theta_t dt + \mathrm{d}\hat{W}_r^2), \quad \forall\, T \geq t \geq 0$$

So, by Girsanov theorem, there exists a probability $\widetilde{\mathbb{P}}$ (mean value denoted by $\widetilde{\mathbb{E}}$) such that $\widetilde{W}_t = \int_0^t \theta_\ell d\ell + \hat{W}_t^2$, $t \geq 0$, is a cylindrical Wiener process. We notice that

$$\lambda T = \widetilde{Y}_T - \widetilde{Y}_0 + \int_0^T f_r \,\mathrm{d}r - \int_0^T \widetilde{\Xi}_r \mathrm{d}\widetilde{W}_r, \quad \forall\, T \geq t \geq 0$$

and consequently:

$$|\lambda| \le T^{-1}|\widetilde{Y}_0| + T^{-1}\widetilde{\mathbb{E}}|\widetilde{Y}_T| + T^{-1}\int_0^T \widetilde{\mathbb{E}}|f_s|\,\mathrm{d}s \tag{4.5}$$

Thanks to Hypothesis 4.1 we get that for all $t \ge 0$:

$$|f_t| \le L_x(1+|z|)|x-x'| + L_z|z-z'| + L_q(1+|z|)|\hat{Q}_t^{x,0} - \hat{Q}_t^{x',0}|, \qquad \mathbb{P}-a.s.$$

We notice that with respect to $(\widetilde{W}_t)$ processes $\hat{Q}^{x,0}$ and $\hat{Q}^{x',0}$ satisfy respectively

$$d\hat{Q}_s^{x,0} = (B\hat{Q_s^{x,0}} + F(x, \hat{Q}_s^{x,0}))\,ds + \theta_s ds + d\widetilde{W}_s,\ s \ge 0,$$
$$d\hat{Q}_s^{x',0} = (B\hat{Q^{x',0}}_s + F(x', \hat{Q}_s^{x',0}))\,ds + \theta_s ds + d\widetilde{W}_s,\ s \ge 0,$$

and Lemma 3.8 yields $|\hat{Q}_s^{x,0} - \hat{Q}_s^{x',0}| \le (K/\mu)|x-x'|$ thus

$$|f_t| \le (L_x + L_q K/\mu)(1+|z|)|x-x'| + L_z|z-z'| \qquad \mathbb{P}-a.s. \text{ for all } t \ge 0. \tag{4.6}$$

Moreover under Hypothesis 3.5 Lemma 3.8 also yields $sup_{t\in[0,\infty[}\widetilde{E}(|Q_t^{x,0}| + |Q_t^{x',0}|) < \infty$. Thus by (4.2) we get $sup_{t\in[0,\infty[}\widetilde{E}(|\widetilde{Y}_t|) < \infty$. Consequently $T^{-1}\widetilde{E}(|\widetilde{Y}_T|) \to 0$ as $T \to \infty$ and the claim follows by (4.5) and (4.6) letting $T \to \infty$.

□

5 Limit equation and convergence of singular BSDEs

We are interested in the following *forward-backward system* for $t \in [0,1]$

$$\begin{cases} dX_t = AX_t + R\,dW_t^1, \\ \varepsilon dQ_t^\varepsilon = (BQ_t^\varepsilon + F(X_t^\varepsilon, Q_t^\varepsilon)\,dt + \varepsilon^{1/2}\,G\,dW_t^2, \\ -dY_t^\varepsilon = \psi(X_t^\varepsilon, Q_t^\varepsilon, Z_t^\varepsilon, \Xi_t^\varepsilon/\sqrt{\varepsilon})\,dt - Z_t^\varepsilon d\,W_t^1 - \Xi_t^\varepsilon dW_t^2, \\ X_0^\varepsilon = x_0 \quad Q_0^\varepsilon = q_0, \quad Y_1^\varepsilon = h(X_1), \end{cases} \tag{5.1}$$

that, as we will see in the sequel, is associated to a controlled multiscale dynamics. Function $h : H \to \mathbb{R}$ satisfies:

Hypothesis 5.1 *h is Lipschitz continuous with constant $L > 0$.*

We have that:

Theorem 5.2 *Assume 3.1–3.4, 4.1 and 5.1.*
Then for every $\varepsilon > 0$ (5.1) has a unique solution $(X, Q^\varepsilon, Y^\varepsilon, Z^\varepsilon, \Xi^\varepsilon)$, with $X \in L^2_{\mathcal{P}}(\Omega; C([0,1];H))$, $Q^\varepsilon \in L^2_{\mathcal{P}}(\Omega; C([0,1];K))$, $Y^\varepsilon \in L^2_{\mathcal{P}}(\Omega; C([0,1];\mathbb{R}))$, $Z^\varepsilon \in L^2_{\mathcal{P}}(\Omega \times [0,1]; \Xi^*)$, $\Xi^\varepsilon \in L^2_{\mathcal{P}}(\Omega \times [0,1]; \Xi^*)$.

Proof. The proof is contained in [11], Propositions 3.2 and 5.2, we just notice that the system is decoupled, so once the forward equation is solved then it becomes a parameter in the backward equation. □

We remark that if we slow down time, that is, for $s \in [0, 1/\varepsilon[$ we set: $\hat{Q}_s^\varepsilon = Q_{\varepsilon s}^\varepsilon$, $\hat{Y}_s^\varepsilon = Y_{\varepsilon s}^\varepsilon$, $\hat{\Xi}_s^\varepsilon = \varepsilon^{-1/2}\Xi_{\varepsilon s}^\varepsilon$ then the following holds:

$$\begin{cases} d\hat{Q}_s^\varepsilon = (B\hat{Q}_s^\varepsilon + F(X_{\varepsilon s}, \hat{Q}_s^\varepsilon)\,dt + G\,d\hat{W}_s^2, \\ -d\hat{Y}_s^\varepsilon = \psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, \hat{Z}_s^\varepsilon, \hat{\Xi}_s^\varepsilon)\,ds - \sqrt{\varepsilon}\hat{Z}_{\varepsilon s}^\varepsilon d\,\hat{W}_s^1 - \hat{\Xi}_s^\varepsilon d\hat{W}_s^2, \\ X_0 = x_0 \quad \hat{Q}_0^\varepsilon = q_0, \quad \hat{Y}_{1/\varepsilon}^\varepsilon = h(X_1). \end{cases} \tag{5.2}$$

where $\hat{W}^\ell_s = \varepsilon^{-1/2} W^\ell_{\varepsilon s}$, $\ell = 1, 2$.

The purpose of our work is to study the limit behaviour of Y^ε as ε tends to 0.
We introduce the candidate limit equation, that turns out to be a forward-backward system on the *finite horizon* $[0,1]$ and on the reduced state space H.

$$\begin{cases} dX_t = AX_t\,dt + R\,dW^1_t, \\ -d\bar{Y}_t = \lambda(X_t, \bar{Z}_t)\,dt - \bar{Z}_t\,dW^1_t, \\ X_0 = x_0, \quad \bar{Y}_1 = h(X_1). \end{cases} \tag{5.3}$$

where λ is defined in Theorem 4.2. Thanks to (4.4) one has that

Theorem 5.3 *System* (5.3) *has only one solution* (X, Y, Z) *with* $X \in L^p_{\mathcal{P}}(\Omega; C([0,1]; H))$, $Y \in L^p_{\mathcal{P}}(\Omega; C([0,1]; \mathbb{R}))$, $Z \in L^p_{\mathcal{P}}(\Omega \times [0,1]; \Xi^*)$.

Proof. Thank to (4.4) the proof of existence and uniqueness of the solution to equation (5.3) is standard (see, for instance [11]). □

We can now state our main result:

Theorem 5.4 *Under Hypothesis 3.1—3.5, 4.1 and 5.1, if* $\bar{Y}$ *is the solution to equation* (5.3) *and* Y^ε_t *is the solution to equation* (5.1), *we have that the following holds*

$$\lim_{\varepsilon \to 0} Y^\varepsilon_0 = \bar{Y}_0 \tag{5.4}$$

Proof. We must compare

$$Y^\varepsilon_0 - \bar{Y}_0 = h(X_1) - h(X_1) + \int_0^1 (\psi(X_t, Q^\varepsilon_t, Z^\varepsilon_t, \Xi^\varepsilon_t/\sqrt{\varepsilon}) - \lambda(X_t, \bar{Z}_t))\,dt - \int_0^1 (Z^\varepsilon_t - \bar{Z}_t)\,dW^1_t$$
$$- \int_0^1 \Xi^\varepsilon_t\,dW^2_t.$$

By adding and subtracting we get:

$$\int_0^1 (\psi(X_t, Q^\varepsilon_t, Z^\varepsilon_t, \Xi^\varepsilon_t/\sqrt{\varepsilon}) - \lambda(X_t, \bar{Z}_t))\,dt = \int_0^1 [(\psi(X_t, Q^\varepsilon_t, Z^\varepsilon_t, \Xi^\varepsilon_t/\sqrt{\varepsilon}) - \psi(X_t, Q^\varepsilon_t, \bar{Z}_t, \Xi^\varepsilon_t/\sqrt{\varepsilon})]\,dt$$
$$+ \int_0^1 (\psi(X_t, Q^\varepsilon_t, \bar{Z}_t, \Xi^\varepsilon_t/\sqrt{\varepsilon}) - \lambda(X_t, \bar{Z}_t)\,dt = I_1 + I_2. \tag{5.5}$$

We leave I_1 for the moment
As far as I_2 is concerned we have to use a discretization argument.
Let us now introduce for every N positive integer, a partition of the interval $[0,1]$ of the form $t_k = k2^{-N}$, $k = 0, 1, \dots, 2^N$ and define a couple of step processes X^N and $\widetilde{Z}^N$ defined as follows:

$$X^N(t) = X(t_k), \quad t \in [t_k, t_{k+1}[, \quad k = 0, \dots 2^N - 1, \tag{5.6}$$

$$\widetilde{Z}^N(t) = 2^N \int_{t_{k-1}}^{t_k} \bar{Z}_\ell\,d\ell, \text{ for } t \in [t_k, t_{k+1}[,\ k = 1, \dots 2^N - 1,\ \widetilde{Z}_0 = 0 \text{ for } t \in [0, t_1[, \tag{5.7}$$

where $X, \bar{Z}$ are part of the solution of (5.3). By construction

$$\lim_{N \to \infty} \mathbb{E} \int_0^1 |\widetilde{Z}^N_t - \bar{Z}_t|^2\,dt \tag{5.8}$$

We fix N, then for $k = 0, 1, \dots, 2^N - 1$ we consider the following, iteratively defined, class of forward SDE:

$$d\hat{\mathcal{Q}}^{N,k}_s = (B\hat{\mathcal{Q}}^{N,k}_s + F(X_{t_k}, \hat{\mathcal{Q}}^{N,k}_s))\,dt + Gd\hat{W}^2_s, \quad s \geq t_k/\epsilon, \qquad \hat{\mathcal{Q}}^{N,k}_{t_k/\epsilon} = \hat{\mathcal{Q}}^{N,k-1}_{t_k/\epsilon}, \tag{5.9}$$

Moreover we define

$$\check{Y}_s^{N,k} = \check{v}(X_{t_k}, \hat{\mathcal{Q}}_s^{N,k}, \widetilde{Z}_{t_k}^N), \quad \check{\Xi}_s^{N,k} = \check{\zeta}(X_{t_k}, \hat{\mathcal{Q}}_s^{N,k}, \widetilde{Z}_{t_k}^N), \quad \text{for } s \geq t_k/\epsilon$$

so that the triplet $((\check{Y}_s^{N,k})_{s\geq t_{k/\varepsilon}}, \lambda(X_{t_k}, \widetilde{Z}_{t_k}^N), (\check{\Xi}_s^{N,k})_{s\geq t_{k/\varepsilon}})$ verifies:

$$-d\check{Y}_s^{N,k} = [\psi(X_{t_k}, \hat{\mathcal{Q}}_s^{N,k}, \widetilde{Z}_{t_k}^N, \check{\Xi}_s^{N,k}) - \lambda(X_{t_k}, \widetilde{Z}_{t_k}^N)]\,\mathrm{d}s - \check{\Xi}_s^{N,k}\mathrm{d}\hat{W}_s^2, \quad \forall\, s \geq t_k/\varepsilon \tag{5.10}$$

and

$$|\check{Y}_t^{N,k}| \leq c(1 + |\widetilde{Z}_{t_k}^N|)|\hat{\mathcal{Q}}_s^{N,k}| \qquad \text{for all } s \geq t_k/\epsilon \quad \mathbb{P} - a.s. \tag{5.11}$$

for some positive constant $c > 0$ independent of k and N.
We also set $X_t^N = \sum_{k=0}^{2^N-1} X_{t_k} I_{[t_k,t_{k+1}[}(t)$, for $t \in [0,1[$ and, for $s \in [0,1/\epsilon[$:

$$\hat{\mathcal{Q}}_s^N = \sum_{k=0}^{2^N-1} \hat{\mathcal{Q}}_s^{N,k} I_{[t_k/\epsilon, t_{k+1}/\epsilon[}(s), \qquad \check{\Xi}_s^N = \sum_{k=0}^{2^N-1} \check{\Xi}_s^{N,k} I_{[t_k/\epsilon, t_{k+1}/\epsilon[}(s).$$

so that, for all $N \in \mathbb{N}$ and $k = 0, ..., 2^N - 1$ have

$$\check{Y}_{t_k/\varepsilon}^{N,k} - \check{Y}_{t_{k+1}/\varepsilon}^{N,k} - \int_{t_k/\varepsilon}^{t_{k+1}/\varepsilon} [\psi(X_{\varepsilon s}^N, \hat{\mathcal{Q}}_s^N, \widetilde{Z}_{\varepsilon s}^N, \check{\Xi}_s^N) - \lambda(X_{\varepsilon s}^N, \widetilde{Z}_{\varepsilon s}^N)]\, ds + \int_{t_k/\varepsilon}^{t_{k+1}/\varepsilon} \check{\Xi}_s^N\, d\hat{W}_s^2 = 0. \tag{5.12}$$

So, coming back to the term I_2, we have, after rescaling of time:

$$I_2 = \varepsilon \sum_{k=0}^{2^N-1} \int_{t_k/\varepsilon}^{t_{k+1}/\varepsilon} [\psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, \bar{Z}_{\varepsilon s}, \hat{\Xi}_s^\varepsilon) - \lambda(X_{\varepsilon s}, \bar{Z}_{\varepsilon s})]\, ds$$

and, adding the null terms in (5.12) for $k = 1, ..., 2^N$:

$$\begin{aligned} I_2 = \varepsilon \sum_{k=0}^{2^N-1} \int_{t_k/\varepsilon}^{t_{k+1}/\varepsilon} [\psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, \bar{Z}_{\varepsilon s}, \hat{\Xi}_s^\varepsilon) - \psi(X_{\varepsilon s}^N, \hat{\mathcal{Q}}_s^N, \widetilde{Z}_{\varepsilon s}^N, \check{\Xi}_s^N)]\, ds + \varepsilon \sum_{k=0}^{2^N-1} \int_{t_k/\varepsilon}^{t_{k+1}/\varepsilon} \check{\Xi}_s^N\, d\hat{W}_s^2 \\ - \varepsilon \sum_{k=0}^{2^N-1} \int_{t_k/\varepsilon}^{t_{k+1}/\varepsilon} [\lambda(X_{\varepsilon s}, \bar{Z}_{\varepsilon s}) - \lambda(X_{\varepsilon s}^N, \widetilde{Z}_{\varepsilon s}^N)]\, ds + \varepsilon \sum_{k=1}^{2^N-1} (\check{Y}_{t_k/\varepsilon}^{N,k} - \check{Y}_{t_{k+1}/\varepsilon}^{N,k}). \end{aligned} \tag{5.13}$$

Therefore coming back to our original term $Y_0^\varepsilon - \bar{Y}_0$ we have

$$\begin{aligned} Y_0^\varepsilon - \bar{Y}_0 = & \varepsilon \sum_{k=1}^{N} (\check{Y}_{t_k/\varepsilon}^{N,k} - \check{Y}_{t_{k+1}/\varepsilon}^{N,k}) + \varepsilon \int_0^{1/\varepsilon} \Big[\psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, Z_{\varepsilon s}^\varepsilon, \hat{\Xi}_s^\varepsilon) - \psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, \bar{Z}_{\varepsilon s}, \hat{\Xi}_s^\varepsilon)\Big]\, ds \\ & + \varepsilon \int_0^{1/\varepsilon} [\psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, \bar{Z}_{\varepsilon s}, \hat{\Xi}_s^\varepsilon) - \psi(X_{\varepsilon s}^N, \hat{\mathcal{Q}}_s^N, \widetilde{Z}_{\varepsilon s}^N, \check{\Xi}_s^N)]\, ds \\ & - \varepsilon \int_0^{1/\varepsilon} [\lambda(X_{\varepsilon s}, \bar{Z}_{\varepsilon s}) - \lambda(X_{\varepsilon s}^N, \widetilde{Z}_{\varepsilon s}^N)]\, ds \\ & - \sqrt{\varepsilon} \int_0^{1/\varepsilon} (Z_{\varepsilon s}^\varepsilon - \bar{Z}_{\varepsilon s}) d\hat{W}_s^1 - \varepsilon \int_0^{1/\varepsilon} (\hat{\Xi}_s^\varepsilon - \check{\Xi}_s^N) d\hat{W}_s^2. \end{aligned}$$

By lipscitzianity of ψ and λ (see Theorem 4.2)

$$\begin{aligned}Y_0^\varepsilon - \bar{Y}_0 =&\varepsilon \int_0^{1/\varepsilon} \mathcal{R}_s^{\varepsilon,N}\, ds + \varepsilon \sum_{k=1}^N (\check{Y}^{N,k}_{t_k/\varepsilon} - \check{Y}^{N,k}_{t_{k+1}/\varepsilon})\\
&+ \varepsilon \int_0^{1/\varepsilon} [\psi(X_{\varepsilon t}, \hat{Q}_s^\varepsilon, Z^\varepsilon_{\varepsilon s}, \hat{\Xi}_s^\varepsilon) - \psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, \bar{Z}_{\varepsilon s}, \hat{\Xi}_s^\varepsilon)]\, ds\\
&+ \varepsilon \int_0^{1/\varepsilon} [\psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}_s^N, \widetilde{Z}^N_{\varepsilon s}, \hat{\Xi}_s^\varepsilon) - \psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}_s^N, \widetilde{Z}^N_{\varepsilon s}, \check{\Xi}_t^N)]\, ds\\
&+ \varepsilon \int_0^{1/\varepsilon} (\check{\Xi}_s^N - \hat{\Xi}_s^\varepsilon)\, d\hat{W}_s^2 + \sqrt{\varepsilon} \int_0^{1/\varepsilon} (Z^\varepsilon_{\varepsilon s} - \bar{Z}_{\varepsilon s})\, d\hat{W}_s^1\end{aligned}$$

where, see Hypothesis 4.1, for a suitable constant c it holds:

$$|\mathcal{R}_s^{\varepsilon,N}| \le c(1 + |\bar{Z}_{\varepsilon s}|)|X_{\varepsilon s} - X^N_{\varepsilon s}| + c(1 + |\bar{Z}_{\varepsilon s}|)|\hat{Q}_s^\varepsilon - \hat{\mathcal{Q}}_s^N| + c|\bar{Z}_{\varepsilon s} - \widetilde{Z}^N_{\varepsilon s}| \tag{5.14}$$

We need to get rid of some terms using a Girsanov argument, thus we introduce:

$$\delta^{1,\varepsilon}(s) = \begin{cases} \dfrac{[\psi(X_{\varepsilon t}, \hat{Q}_s^\varepsilon, Z^\varepsilon_{\varepsilon s}, \hat{\Xi}_s^\varepsilon) - \psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, \bar{Z}_{\varepsilon s}, \hat{\Xi}_s^\varepsilon)]}{|Z^\varepsilon_{\varepsilon s} - \bar{Z}_{\varepsilon s}|^2}(Z^\varepsilon_{\varepsilon s} - \bar{Z}_{\varepsilon s})^* & \text{if } |Z^\varepsilon_{\varepsilon s} - \bar{Z}_{\varepsilon s}| \neq 0\\ 0 & \text{if } |Z^\varepsilon_{\varepsilon s} - \bar{Z}_{\varepsilon s}| = 0\end{cases} \tag{5.15}$$

and

$$\delta^{2,\varepsilon,N}(s) = \begin{cases} \dfrac{\psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}_s^N, \widetilde{Z}^N_{\varepsilon s}, \hat{\Xi}_s^\varepsilon) - \psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}_s^N, \widetilde{Z}^N_{\varepsilon s}, \check{\Xi}_s^N)}{|\hat{\Xi}_s^\varepsilon - \check{\Xi}_s^N|^2}(\hat{\Xi}_s^\varepsilon - \check{\Xi}_s^N)^* & \text{if } |\hat{\Xi}_s^\varepsilon - \check{\Xi}_s^N| \neq 0\\ 0 & \text{if } |\hat{\Xi}_s^\varepsilon - \check{\Xi}_s^N| = 0\end{cases} \tag{5.16}$$

We notice that processes $(\delta_s^{1,\varepsilon})_{s\in[0,1/\varepsilon]}$ and $(\delta_s^{2,N,\varepsilon})_{s\in[0,1/\varepsilon]}$ are bounded uniformly with respect to N and ϵ. We have

$$\begin{aligned}Y_0^\varepsilon - \bar{Y}_0 =& \varepsilon \int_0^{1/\varepsilon} \delta^{1,\varepsilon}(s)[Z^\varepsilon_{\varepsilon s} - \bar{Z}_{\varepsilon s}]\, ds + \varepsilon \int_0^{1/\varepsilon} \delta^{2,\varepsilon,N}(s)[\check{\Xi}_s^N - \hat{\Xi}_s^\varepsilon]\, ds\\
&+ \varepsilon \int_0^{1/\varepsilon} (\check{\Xi}_s^N - \check{\Xi}_s^\varepsilon)\, d\hat{W}_s^2 + \sqrt{\varepsilon} \int_0^{1/\varepsilon} (Z^\varepsilon_{\varepsilon s} - \bar{Z}_{\varepsilon s})\, d\hat{W}_s^1\\
&+ \varepsilon \int_0^{1/\varepsilon} \mathcal{R}_s^{\varepsilon,N}\, ds + \varepsilon \sum_{k=1}^N (\check{Y}^{N,k}_{t_k/\varepsilon} - \check{Y}^{N,k}_{t_{k+1}/\varepsilon})\end{aligned}$$

and rescaling time (speeding it up this time)

$$\begin{aligned}Y_0^\varepsilon - \bar{Y}_0 =& \int_0^1 \delta^{1,\varepsilon}(t/\varepsilon)[Z_t^\varepsilon - \bar{Z}_t]\, dt + \int_0^1 \delta^{2,\varepsilon,N}(t/\varepsilon)[\check{\Xi}^N_{\varepsilon^{-1}t} - \hat{\Xi}^\varepsilon_{\varepsilon^{-1}t}]\, dt\\
&+ \sqrt{\varepsilon} \int_0^1 (\check{\Xi}^N_{\varepsilon^{-1}t} - \check{\Xi}^\varepsilon_{\varepsilon^{-1}t})\, dW_t^2 + \int_0^1 (Z_t^\varepsilon - \bar{Z}_t)\, dW_t^1\\
&+ \int_0^1 \mathcal{R}^{\varepsilon,N}_{\varepsilon^{-1}t}\, dt + \varepsilon \sum_{k=1}^N (\check{Y}^{N,k}_{t_k/\varepsilon} - \check{Y}^{N,k}_{t_{k+1}/\varepsilon})\end{aligned}$$

We set for $t \in [0,1]$:

$$\widetilde{W}_t^1 =: \int_0^t \delta^{1,\varepsilon}(r/\varepsilon)\, dr + W_t^1 \tag{5.17}$$

$$\widetilde{W}_t^2 =: \varepsilon^{-1/2} \int_0^t \delta^{2,\varepsilon,N}(r/\varepsilon)\, dr + W_t^2 \tag{5.18}$$

We denote by $\widetilde{\mathbb{E}}^\varepsilon$ the expectation under the new probability $\mathbb{P}^\varepsilon$ with respect to which $(\widetilde{W}^1_t, \widetilde{W}^2_t)_{t\in[0,1]}$ is a $H \times K$ valued cylindrical Wiener process (recall that $(W^1_t, W^2_t)_{t\in[0,1]}$ is a $H \times K$ valued cylindrical Wiener process). Since the left hand side is deterministic, we have:

$$Y^\varepsilon_0 - \bar{Y}_0 = \widetilde{\mathbb{E}}^\varepsilon \int_0^1 \mathcal{R}^{\varepsilon,N}_{t/\varepsilon}\, dt + \varepsilon \widetilde{\mathbb{E}}^\varepsilon \sum_{k=1}^N [\check{Y}^{N,k}_{t_k/\varepsilon} - \check{Y}^{N,k}_{t_{k+1}/\varepsilon}] \tag{5.19}$$

Thus by (5.14) we get

$$\widetilde{\mathbb{E}}^\varepsilon \int_0^1 |\mathcal{R}^{\varepsilon,N}_{t/\varepsilon}|\, dt \le c\widetilde{\mathbb{E}}^\varepsilon \int_0^1 \Big((1+|\bar{Z}_t|)|X_t - X^N_t| + (1+|\bar{Z}_t|)|Q^\varepsilon_t - \hat{\mathcal{Q}}^N_{t/\varepsilon}| + |\bar{Z}_t - \widetilde{Z}^N_t| \Big)\, dt$$

Let us start from

$$\widetilde{\mathbb{E}}^\varepsilon \int_0^1 (1+|\bar{Z}_t|)|X_t - X^N_t| dt$$

We notice that, with respect to $\widetilde{W}^1$ we have

$$\begin{cases} dX_t = AX_t\, dt - R\delta^{1,\varepsilon}(t/\varepsilon)dt + R\, d\widetilde{W}^1_t, \\ -d\bar{Y}_t = \lambda(X_t, \bar{Z}_t)\, dt - \bar{Z}_t[-\delta^{1,\varepsilon}(t/\varepsilon)dt + d\widetilde{W}^1_t], \\ \bar{Y}_1 = h(X_1), \quad X_0 = x_0. \end{cases}$$

Define

$$\rho := \exp\left(\int_0^1 \delta^{1,\varepsilon}(s/\varepsilon)\, d\widetilde{W}^1_s - \frac{1}{2}\int_0^1 |\delta^{1,\varepsilon}(s/\varepsilon)|^2\, ds \right)$$

then, by Holder inequality, setting $\Delta_{X,N} := \sup_{t\in[0,1]} |X_t - X^N_t|$

$$\widetilde{\mathbb{E}}^\varepsilon \int_0^1 (1+|\bar{Z}_t|)|X_t - X^N_t| dt \le \widetilde{\mathbb{E}}^\varepsilon \left[\Delta_{X,N} \int_0^1 (1+|\bar{Z}_t|)dt \right] \le$$

$$\widetilde{\mathbb{E}}^\varepsilon \left[\rho^{-3/4}(\rho^{1/4}\Delta_{X,N})\rho^{1/2}\int_0^1 (1+|\bar{Z}_t|)dt \right] \le \left[\widetilde{\mathbb{E}}^\varepsilon \rho^{-3} \right]^{1/4} \left[\widetilde{\mathbb{E}}^\varepsilon(\rho\Delta^4_{X,N}) \right]^{1/4} \left[\widetilde{\mathbb{E}}^\varepsilon \left(\rho \int_0^1 (1+|\bar{Z}_t|^2)dt \right) \right]^{1/2}$$

Again by Girsanov the process $\left(-\int_0^t \delta^1 \psi^\epsilon_{t/\varepsilon} dt + \widetilde{W}^1_t \right)_{t\in[0,1]}$ is a cylindrical Wiener process with respect to $\rho\, d\mathbb{P}^\varepsilon$. By uniqueness of the solution to the forward backward system (5.3) the law of the processes $(X_t)_{t\ge 0}$ and $(\bar{Z}_t)_{t\ge 0}$ under $\rho d\mathbb{P}^\varepsilon$ coincides with its law with respect to $\mathbb{P}$. Recalling that $\delta^{1,\varepsilon}$ is uniformly bounded and consequently (with respect to ε as well) we have $\widetilde{\mathbb{E}}^\varepsilon \rho^{-3} \le c$ (where c does not depend on ϵ). Moreover

$$\widetilde{\mathbb{E}}^\varepsilon \left(\rho \int_0^1 |\bar{Z}_t|^2 \right) dt = \mathbb{E} \left(\int_0^1 |\bar{Z}_t|^2 \right) dt < +\infty.$$

Thus we can conclude

$$\widetilde{\mathbb{E}}^\varepsilon \int_0^1 (1+|\bar{Z}_t|)|X_t - X^N_t| dt \le C[\mathbb{E}\Delta^4_{X,N}]^{1/4}$$

where C is independent of N and ε.
By the continuity of trajectories of $(X_t)_{t\ge 0}$ since $\mathbb{E}\sup_{t\in[0,1]} |X_t|^4 < \infty$ we get

$$\mathbb{E}\Delta^4_{X,N} \to 0, \qquad \text{as} \qquad N \to \infty. \tag{5.20}$$

Concerning the term

$$\widetilde{\mathbb{E}}^\varepsilon \int_0^1 |\bar{Z}_t - \widetilde{Z}^N_t| dt$$

we notice that being $\bar{Z}_t = \zeta(X_t)$ where ζ is a deterministic Borel function $H \to \Xi^*$ then the law of $(\bar{Z})$ and $(\tilde{Z}^N)$ depend only on the law of (X) in a non anticipating way. So, the above argument still works and yields:

$$\widetilde{\mathbb{E}}^\varepsilon \int_0^1 |\bar{Z}_t - \widetilde{Z}_t^N| dt \le C \left[\mathbb{E} \int_0^1 |\bar{Z}_t - \widetilde{Z}_t^N|^2 dt\right]^{1/2} = C(\mathbb{E}\Delta_{Z,N})^{1/2}$$

where $\Delta_{Z,N} = \int_0^1 |\bar{Z}_t - \widetilde{Z}_t^N|^2 dt$ and by (5.8)

$$\mathbb{E}\Delta_{Z,N} \to 0, \qquad \text{as} \qquad N \to \infty. \tag{5.21}$$

Coming to the term:

$$\widetilde{\mathbb{E}}^\varepsilon \int_0^1 |Q_t^\varepsilon - \mathcal{Q}^N_{\varepsilon^{-1}t}|\, dt = \widetilde{\mathbb{E}}^\varepsilon \int_0^1 |\hat{Q}^\varepsilon_{\varepsilon^{-1}t} - \hat{\mathcal{Q}}^N_{\varepsilon^{-1}t}|\, dt$$

With respect to $\widetilde{W}^2$ the process $(Q_t^\varepsilon)_{t\in[0,1]}$ solves

$$\varepsilon\, dQ^\varepsilon{}_t = (BQ^\varepsilon{}_t + F(X_t, Q^\varepsilon{}_t))\, dt - \delta^{2,\varepsilon,N}(t/\varepsilon)\, dt + \sqrt{\varepsilon}\, d\widetilde{W}_t^2, \quad t \ge 0, \qquad Q^\varepsilon{}_0 = q_0,$$

thus introducing the $\mathbb{P}^\varepsilon$ Wiener process $\hat{\widetilde{W}}_s := (\varepsilon)^{-1/2}\widetilde{W}_{\varepsilon s}$ the process $(\hat{Q}^\varepsilon_s)_{s\in[0,1/\varepsilon]}$ solves:

$$d\hat{Q}^\varepsilon_s = (B\hat{Q}^\varepsilon_s + F(X_{\varepsilon s}, \hat{Q}^\varepsilon_s))\, ds - \delta^{2,\varepsilon,N}(s)\, ds - \, d\hat{\widetilde{W}}{}^2_s, \quad s \ge 0, \qquad \hat{Q}^\varepsilon_0 = q_0, \tag{5.22}$$

moreover $(\hat{\mathcal{Q}}^N_s)$ solves

$$d\hat{\mathcal{Q}}^N_s = (B\hat{\mathcal{Q}}^N_s + F(X^N_{\varepsilon s}, \hat{\mathcal{Q}}^N_s))\, dt - \delta^{2,\varepsilon,N}(s)\, ds + d\hat{\widetilde{W}}{}^2_s, \quad s \ge 0, \qquad \hat{\mathcal{Q}}_0 = q_0, \tag{5.23}$$

By Lemma 3.8 and Hypothesis 3.5 we have, for all $p \ge 1$:

$$\sup_{s\in[0,1/\varepsilon]} \widetilde{\mathbb{E}}^\varepsilon[|\hat{\mathcal{Q}}^N_t|^p] \le c_p(1 + |q_0|^p + |x_0|^p) \tag{5.24}$$

for a constant c independent of ϵ and N.
Moreover again by Lemma 3.8 for all $s > 0$,

$$|\hat{Q}^\varepsilon_s - \hat{\mathcal{Q}}^N_s| \le c \int_0^s e^{-\eta(s-\ell)} |X_{\varepsilon\ell} - X^N_{\varepsilon\ell}| d\ell \le c\Delta_{X,N}$$

thus

$$\widetilde{\mathbb{E}}^\varepsilon \int_0^1 (1 + |\bar{Z}_t|) |\hat{Q}^\varepsilon_{\varepsilon^{-1}t} - \hat{\mathcal{Q}}^N_{\varepsilon^{-1}t}| dt \le c\widetilde{\mathbb{E}}^\varepsilon \left[\Delta_{X,N} \int_0^1 (1 + |\bar{Z}_t|) dt\right] \le C[\mathbb{E}\Delta^4_{X,N}]^{1/4}$$

as above.
Now we come to the last term. By (5.11)

$$|\varepsilon\widetilde{\mathbb{E}}^\varepsilon \sum_{k=1}^{2^N} (\hat{Y}^{N,k}_{t_k/\varepsilon} - \hat{Y}^{N,k}_{t_{k+1}/\varepsilon})| \le c\varepsilon \sum_{k=1}^{2^N} \widetilde{\mathbb{E}}^\varepsilon \left[(1 + |\widetilde{Z}^N_{t_k}|)(1 + |\hat{\mathcal{Q}}^N_{t_k/\varepsilon}| + |\hat{\mathcal{Q}}^N_{t_{k+1}/\varepsilon}|)\right]$$

$$\le c\varepsilon \left[\widetilde{\mathbb{E}}^\varepsilon (1 + |\widetilde{Z}^N_{t_k}|)^{4/3}\right]^{3/4} \left[\widetilde{\mathbb{E}}^\varepsilon (1 + |\hat{\mathcal{Q}}^N_{t_k/\varepsilon}| + |\hat{\mathcal{Q}}^N_{t_{k+1}/\varepsilon}|)^4\right]^{1/4} \le c\varepsilon \left[\widetilde{\mathbb{E}}^\varepsilon (1 + |\widetilde{Z}^N_{t_k}|)^{4/3}\right]^{3/4}.$$

Proceeding as above, recalling that the law of $\widetilde{Z}^N_{t_k}$ depends only on the law of the process (X_t) we have:

$$\widetilde{\mathbb{E}}^\varepsilon(|\widetilde{Z}^N_{t_k}|^{4/3}) \le \left[\widetilde{\mathbb{E}}^\varepsilon \rho^{-2}\right]^{1/3} \left[\mathbb{E}(|\widetilde{Z}^N_{t_k}|^2))\right]^{2/3} \le c2^{N/3}\mathbb{E} \int_{t_k}^{t_{k+1}} |\bar{Z}_t|^2 dt \le c2^{N/3}\mathbb{E} \int_0^1 |\bar{Z}_t|^2$$

where in the above formulae the value of the constant c can change from line to line but does not depend neither on k nor on N or on ε.
At last we sum up all results to get

$$|Y_0^\varepsilon - \bar{Y}_0| \leq \widetilde{\mathbb{E}}^\varepsilon \int_0^1 |\mathcal{R}_{t/\varepsilon}^{\varepsilon,N}|\, dt + \varepsilon \widetilde{\mathbb{E}}^\varepsilon \sum_{k=1}^{2^N} |\hat{Y}_{t_k/\varepsilon}^{N,k} - \hat{Y}_{t_{k+1}/\varepsilon}^{N,k}|$$
$$\leq C[\mathbb{E}\Delta_{X,N}^4]^{1/4} + C(\mathbb{E}\Delta_{Z,N})^{1/2} + \varepsilon c 2^{N/4} \left[\mathbb{E}\int_0^1 |\bar{Z}_t|^2\right]^{3/4}$$

So letting first ε tend to 0 and then N to ∞ the claim follows, by (5.20) and (5.21). □

Remark 5.5 Consider the following class of forward backward systems with initial time $\tau \in [0,1]$

$$\begin{cases} dX_t^{\tau,x} = AX_t^{\tau,x}\, dt + \, RdW_t^1, \quad t \leq \tau \leq 1 \\ -d\bar{Y}_t^{\tau,x} = \lambda(X_t^{\tau,x}, \bar{Z}_t^{\tau,x})\, dt - \, \bar{Z}^{\tau,x}\, dW_t^1, \\ X_\tau^{\tau,x} = x, \quad \bar{Y}_1^{\tau,x} = h(X_1^{\tau,x}). \end{cases} \tag{5.25}$$

If we set $v(\tau,x) = \bar{Y}_\tau^{\tau,x}$ then it is shown in [12] that v is a deterministic continuous function $[0,1]\times H \to \mathbb{R}$ being Gateaux differentiable with respect to the second variable. Moreover it is the unique mild solution of the nonlinear Kolmogorov equation

$$\begin{cases} \dfrac{\partial v(t,x)}{\partial t} + \mathcal{L}v(t,x) = \lambda(x, \nabla v(t,x)R), \qquad t \in [0,T],\, x \in H, \\ v(1,x) = h(x), \end{cases}$$

where $\mathcal{L}$ is the second order operator

$$\mathcal{L}g(x) = \frac{1}{2}\mathrm{Tr}[R^* \nabla^2 g(x) R], \quad g \in \mathcal{C}^2(H)$$

$\nabla^2 g(x) \in \mathcal{L}(H)$ being the second derivative of g in x.
In particular the limit $\lim_{\epsilon\to 0} Y_0^\epsilon$ can also be represented by the solution of the above HJB equation as:

$$\lim_{\epsilon\to 0} Y_0^\epsilon = \bar{Y}_0^{0,x_0} = v(0,x_0)$$

6 Application to control

Given the solution (X, Q^ε) of system (3.1) and an adapted process $(\alpha_t)_{t\in[0,1]}$ taking its values in a complete metric space U we denote by $\Theta^{\varepsilon,\alpha}$ the density

$$\Theta^{\varepsilon,\alpha} = \exp\left(\int_0^1 \left[R^{-1}b(X_t, Q_t^\varepsilon, \alpha_t) + \frac{1}{\sqrt{\varepsilon}}\rho(\alpha_t)\right] dW_t^1 - \frac{1}{2}\int_0^1 \left[|R^{-1}b(X_t, Q_t^\varepsilon, \alpha_t)|^2 + \frac{1}{\varepsilon}|\rho(\alpha_t)|^2\right] dt\right)$$

where $b : H \times K \times U \to H$ and $\rho : U \to K$ are measurable functions satisfying suitable assumptions listed below.
We also consider the following cost functional:

$$J^\varepsilon(x_0, q_0, \alpha) = \mathbb{E}\left[\Theta^{\varepsilon,\alpha}\left(\int_0^1 l(X_t, Q_t^\varepsilon, \alpha_t) dt + h(X_1)\right)\right] \tag{6.1}$$

where $l : H \times K \times U \to \mathbb{R}$ and $\phi : H \to \mathbb{R}$ are measurable and satisfy the assumptions below:

Hypothesis 6.1 *There are positive constants* L *and* M *such that*

$$|b(x,q,u)-b(x',q',\alpha)|\le L(|x-x'|+|q-q'|) \qquad \forall\, q,q'\in K, x,x'\in H,\ \alpha\in U$$

$$|l(x,q,\alpha)-l(x',q',\alpha)|\le L(|x-x'|+|q-q'|) \qquad \forall\, q,q'\in K, x,x'\in H,\ \alpha\in U$$

$$|h(x)-h(x')|\le L|x-x'| \qquad \forall x,x'\in H,$$

$$|b(x,q,\alpha)|,|l(x,q,\alpha)|,|\rho(\alpha)|,|h(x)|\le M \qquad \forall q\in K, x\in H,\ \alpha\in U.$$

We recall that if $d\mathbb{P}^{\varepsilon,\alpha}:=\Theta^{\varepsilon,\alpha}d\mathbb{P}$ then under probability $\mathbb{P}^{\varepsilon,\alpha}$ the process:

$$(\mathcal{W}^1_t,\mathcal{W}^2_t)=(-\int_0^t R^{-1}b(X_r,Q^\varepsilon_r,\alpha_r)dr+W^1_t,-\int_0^t\rho(\alpha_r)dr+W^2_t)$$

is a cylindrical Wiener process in $\Xi\times\Xi$ and that with respect to $(\mathcal{W}^1_t,\mathcal{W}^2_t)$ the couple of processes (X_t,Q^ε_t) satisfies the controlled system:

$$\begin{cases} dX(t)=AX(t)+b(X_t,Q^\varepsilon_t,\alpha_t)dt\ +Rd\mathcal{W}^1(t), & X_0=x_0,\\ \\ \varepsilon dQ^\varepsilon(t)=(BQ^\varepsilon(t)+F(X^\varepsilon(t),Q^\varepsilon(t))\,dt+G\rho(\alpha_t)dt+\varepsilon^{1/2}G\,d\mathcal{W}^2(t), & Q^\varepsilon_0=q_0,\end{cases} \tag{6.2}$$

and

$$J^\varepsilon(x_0,q_0,\alpha)=\mathbb{E}^{\mathbb{P}^{\varepsilon,\alpha}}\left(\int_0^1 l(X_t,Q^\varepsilon_t,\alpha_t)dt+h(X_1)\right)$$

Finally we define, for $x\in H$, $q\in K$ and $z,\xi\in\Xi^*$

$$\psi(x,q,z,\xi)=\inf_{\alpha\in U}\{l(x,q,\alpha)+z[R^{-1}b(x,q,\alpha)]+\xi\rho(v)\} \tag{6.3}$$

and notice that, under Hypothesis (6.1), the Hamiltonian ψ verifies Hypothesis (4.1).
The following is an immediate consequence of our general results

Theorem 6.2 *Denote by* V^ε *the value function of our control problem, that is:*

$$V^\varepsilon(x_0,q_0):=\inf_\alpha J^\varepsilon(x_0,q_0,\alpha)$$

where the infimum is taken over all adapted processes α *with value in* U.
The sequence $V^\varepsilon(x_0,q_0)$ *converges to the solution* $\bar{Y}_0$ *of equation* (5.3) *evaluated at zero.*

Proof. In [10] it is shown that $V^\varepsilon(x_0,q_0)=Y^\varepsilon_0$ (see (5.1)). The claim then follows by Theorem 5.4

Remark 6.3 The nonlinearity λ in the limit equation (5.3) has itself a control theoretic interpretation. Namely, fixed $x\in H$ and $z\in\Xi^*$, let us consider the following ergodic control problem with *state equation*

$$d\hat{Q}^\beta_s=B\hat{Q}^\beta_s ds+F(x,\hat{Q}^\beta_s)\,ds+G\rho(\beta_s)ds+GdW^2_s \tag{6.4}$$

and *ergodic cost functional*

$$\check{J}(x,z,\beta)=\liminf_{\delta\to 0}\mathbb{E}\,\delta\int_0^\infty e^{-\delta s}[zR^{-1}b(x,\hat{Q}^\beta_s,\beta_s)+l(x,\hat{Q}^\beta_s,\beta_s)]ds. \tag{6.5}$$

Then $\lambda(x,z)$ is the value function of the ergodic control problem the we have just described, that is:

$$\lambda(x,z)=\inf_\beta\check{J}(x,z,\beta)$$

where the infimum is taken over all adapted processes $\beta:[0,\infty[\to U$.

Notice that, in particular, beeing the infimum of linear functionals, the map $z\to\lambda(x,z)$ is concave.

Moreover notice that the result was proven in [9] with lim inf replaced by lim sup in the definition (6.5) of the ergodic cost nevertheless, as it can be easily verified, this substitution is inessential in the argument reported in [9]

7 Control interpretation of the limit forward-backward system

Most of our analysis in this section is based on the fact that λ is concave with respect to z. In particular, by Fenchel-Moreau theorem (translated in the obvious way for concave functions instead than for convex ones), we can write $\lambda = \lambda_{**}$ where for all $x \in H$:

$$\lambda_*(x,p) = \inf_{z\in\Xi^*}\big(-zp - \lambda(x,z)\big), \quad p \in \Xi$$

and the map $\lambda_*(x,\cdot)$ is an upper semicontinuous concave function with non empty domain in Ξ. Thus for all $x \in H$, $z \in \Xi^*$:

$$\lambda(x,z) = \inf_{p\in\Xi}\big(-zp - \lambda_*(x,p)\big)$$

Recalling that λ is Lipschitz continuous with respect to z uniformly in x and denoting by L the Lipschitz constant we have

$$\lambda_*(x,p) = -\infty, \text{ whenever } |p| > L.$$

Moreover, $\lambda_*(x,p) \le -\lambda(x,0) \le c(1+|x|)$, in particular, for any process $(\mathfrak{p}_t)_{0\le t\le 1}$ with values in Ξ, the process

$$\left(\int_0^t \lambda_*(X_s,\mathfrak{p}_s)ds\right)_{0\le t\le 1}$$

is well-defined and takes values in $[-\infty,\infty)$.
Given any Ξ valued progressively measurable process $(\mathfrak{p}_t)_{t\ge 0}$ with $|\mathfrak{p}_t| \le L$:

$$\begin{aligned}\bar{Y}_t &= h(X_1) + \int_t^1 \lambda(X_s,\bar{Z}_s)ds - \int_t^1 \bar{Z}_s dW_s \\ &\le h(X_1) - \int_t^1 \big(\bar{Z}_s\mathfrak{p}_s + \lambda_*(X_s,\mathfrak{p}_s)\big)\, ds - \int_t^1 \bar{Z}_s dW_s\end{aligned}$$

Introducing $W_t^{\mathfrak{p}} = \int_0^t \mathfrak{p}_s ds + W_t$ and the probability $\mathbb{P}^{\mathfrak{p}}$ under which it is a Wiener process we get:

$$dX_t = AX_t dt - R\mathfrak{p}_t dt + RdW_t^{\mathfrak{p}}$$

$$\bar{Y}_t \le h(X_1) - \int_t^1 \lambda_*(X_s,\mathfrak{p}_s)ds - \int_0^1 \bar{Z}_s dW_s^{\mathfrak{p}}$$

which shows that:

$$\bar{Y}_t \le \mathbb{E}^{\mathfrak{p}}\left(h(X_1) - \int_t^1 \lambda_*(X_s,\mathfrak{p}_s)ds\middle|\mathcal{F}_t\right).$$

Conversely, we may call, for any $n \ge 1$, $(\mathfrak{p}_t^n)_{0\le t\le 1}$ such that $-\bar{Z}_t\mathfrak{p}_t^n - \lambda_*(X_t,\mathfrak{p}_t^n) - 1/n \le \lambda(X_t,\bar{Z}_t)$. Clearly we have $|\mathfrak{p}_t^n| \le L$. Using a measurable selection Theorem, see for instance Theorem 6.9.13 in [4], one can chose the process $\mathfrak{p}^n$ to be progressive measurable.
Then, we have

$$\bar{Y}_t \ge h(X_1) - \int_t^1 \left(\bar{Z}_s\mathfrak{p}_s^n + \lambda_*(X_s,\mathfrak{p}_s^n) + \frac{1}{n}\right) ds - \int_t^1 \bar{Z}_s dW_s$$

and rewriting the above in terms of $W^{\mathfrak{p}^n}$

$$\bar{Y}_t + \frac{1-t}{n} \ge h(X_1) - \int_t^1 \lambda_*(X_t,\mathfrak{p}_t^n)dt - \int_t^1 \bar{Z}_t dW_t^{\mathfrak{p}^n}.$$

Therefore we can conclude that $\bar{Y}_t$ is the value function of a stochastic optimal control problem in the sense that:

$$\bar{Y}_t = \inf_{\mathfrak{p}} \mathbb{E}^{\mathfrak{p}}\left(h(X_1) - \int_0^1 \lambda_*(X_t,\mathfrak{p}_t)dt\middle|\mathcal{F}_t\right).$$

where (X) is the solution of the following controlled stochastic differential equation:

$$dX_t = AX_t dt - R\mathfrak{p}_t dt + RdW_t^{\mathfrak{v}},\ X_0 = x_0;$$

the supremum is extended to all Ξ-valued, predictable processes $(\mathfrak{p}_s)_{0\leq s\leq 1}$ that are bounded by L and finally $W^{\mathfrak{v}}$ is a Ξ-valued Wiener process with respect to $\mathbb{P}^{\mathfrak{v}}$.

Acknowledgements: The authors wish to thank Francois Delarue for several enlightening discussions on the content of the present paper.

References

[1] O. Alvarez and M. Bardi. Viscosity solutions methods for singular perturbations in deterministic and stochastic control. *SIAM J. Control Optim.* 40 No.4 (2001), pp. 11591188

[2] A. Bensoussan. *Perturbation Methods in Optimal Control.* Wyley/Gauthiers-Villars, U.K., 1988.

[3] P. Briand and Y. Hu Stability of BSDEs with random terminal time and homogenization of semilinear elliptic PDEs. *J.Func. Anal.,* 155 (1998), pp. 455-494.

[4] V. I. Bogachev. *Measure Theory.* Springer-Verlag, Berlin, Heidelberg, 2007

[5] G. Da Prato, S. Kwapien and J. Zabczyk. Regularity of solutions of linear stochatsic equations in Hilbert spaces. *Stochastics*, 23, 1-23.

[6] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions.* Cambridge University Press, Cambridge, 1992.

[7] G. Da Prato and J. Zabczyk. *Ergodicity for infinite-dimensional systems.* London Mathematical Society Lecture Note Series, 229. Cambridge University Press, Cambridge, 1996.

[8] A. Debussche, Y. Hu and G. Tessitore. Ergodic BSDEs under weak dissipative assumptions. *Stoc. Pro. Appl.* 121 (2011), 407-426.

[9] M. Fuhrman, H. Ying, G. Tessitore, Ergodic BSDES and optimal ergodic control in Banach spaces. Siam Journal On Control and Optimization, 48(3) (2009), 1542-1566.

[10] M. Fuhrman, G. Tessitore, Generalized directional gradients, backward stochastic differential equations and mild solutions of semilinear parabolic equations. Applied Matematics and Optimization, 51(3) (2005), 279-332.

[11] M. Fuhrman and G. Tessitore. Nonlinear Kolmogorov equations in infinite dimensional spaces: The backward stochastic differential equations approach and applications to optimal control. *Ann. Probab.* 30 (2002), 1397–1465.

[12] M. Fuhrman and G. Tessitore. The Bismut-Elworthy formula for backward SDE's and applications to nonlinear Kolmogorov equations and control in infinite dimensional spaces. *Stochastics and Stochastic Reports* 74 (2002), 429–464.

[13] Y. Kabanov and S. Pergamenshchikov. *Two-scale Stochastic Systems.* Applications of Mathematics, 49. Springer-Verlag, Berlin, Heidelberg, New York, 2003.

[14] Y. Kabanov, W. Runggaldier, On Control of Two-scale Stochastic Systems with Linear Dynamics in the Fast Variables. *Math. Control Signals Systems* 9 (1996), 107–122.

[15] Y. Hu, P.-Y. Madec and A. Richou. A probabilistic approach to large time behavior of mild solutions of HJB equations in infinite dimension. *SIAM J. Control Optim.* 53 (2015), 378398.

[16] H. Kushner. *Weak convergence Methods and Singularly Perturbed Stochatsic Control And Filtering Problem.* Birkauser, Boston, 1990.