Electronic Journal of Probability

Vol. 13 (2008), Paper no. 54, pages 1529–1561.

Journal URL
http://www.math.washington.edu/~ejpecp/

Quadratic BSDEs with random terminal time and elliptic PDEs in infinite dimension.

Philippe Briand

IRMAR, Université Rennes 1, 35042 Rennes Cedex, FRANCE

philippe.briand@univ-rennes1.fr

Fulvia Confortola

Dipartimento di Matematica, Politecnico di Milano

Piazza Leonardo da Vinci 32, 20133 Milano, ITALY

fulvia.confortola@polimi.it

Abstract

In this paper we study one dimensional backward stochastic differential equations (BSDEs) with random terminal time not necessarily bounded or finite when the generator $F(t,Y,Z)$ has a quadratic growth in Z. We provide existence and uniqueness of a bounded solution of such BSDEs and, in the case of infinite horizon, regular dependence on parameters. The obtained results are then applied to prove existence and uniqueness of a mild solution to elliptic partial differential equations in Hilbert spaces. Finally we show an application to a control problem.

Key words: Backward stochastic differential equations, quadratically growing driver, elliptic partial differential equation, stochastic optimal control.

AMS 2000 Subject Classification: Primary 60H10; 60H3.

Submitted to EJP on June 25, 2007, final version accepted June 10, 2008.

1 Introduction

In this paper, we are mainly interested in finding a probabilistic representation for the solution to the following elliptic PDE

$$\mathscr{L}u(x) + F(x, u(x), \nabla u(x)\sigma) = 0, \quad x \in H, \tag{1}$$

where H is a Hilbert space and $\mathscr{L}$ is a second order differential operator of type

$$\mathscr{L}\phi(x) = \frac{1}{2}\mathrm{Trace}(\sigma\sigma^*\nabla^2\phi(x)) + \langle Ax, \nabla\phi(x)\rangle + \langle b(x), \nabla\phi(x)\rangle.$$

with A being the generator of a strongly continuous semigroup of bounded linear operators $(e^{tA})_{t\geq 0}$ in H and F being a nonlinear function.

It is by now well known that this kind of Feynman-Kac formula involves Markovian backward stochastic differential equations (BSDEs for short in the remaining of the paper) with infinite horizon, which, roughly speaking, are equations of the following type

$$Y_t^x = \int_t^\infty F\left(X_s^x, Y_s^x, Z_s^x\right) ds - \int_t^\infty Z_s^x \, dW_s \tag{2}$$

where $\{X_t^x\}_{t\geq 0}$ stands for the mild solution to the SDE

$$dX_t^x = AX_t^x \, dt + b\left(X_t^x\right) dt + \sigma \, dW_t, \quad t \geq 0, \quad X_0^x = x, \tag{3}$$

W being a cylindrical Wiener process with values in some Hilbert space Ξ (see Section 2 for details). With these notations, the solution u to the PDE (1) is given by

$$\forall x \in H, \quad u(x) = Y_0^x, \tag{4}$$

where (Y^x, Z^x) is the solution to the previous BSDE. For this infinite dimensional setting, we refer to the article [14] in which the authors deal with functions F being Lipschitz with respect to z.

One of the main objective of this study is to obtain this nonlinear Feynman-Kac formula in the case where the function F is not Lipschitz continuous with respect to z but has a quadratic growth with respect to this variable meaning that the PDE is quadratic with respect to the gradient of the solution. In particular, in order to derive this formula in this setting, we will have to solve quadratic BSDEs with infinite horizon.

BSDEs with infinite horizon are a particular class of BSDEs with random terminal time which have been already studied in several paper. Let us recall some basic facts about these equations. Let τ be a stopping time which is not assumed to be bounded or $\mathbb{P}$–a.s. finite. We are looking for a pair of processes $(Y_t, Z_t)_{t\geq 0}$, progressively measurable, which satisfy, $\forall t \geq 0, \forall T \geq t$,

$$\begin{cases} Y_{t\wedge\tau} = Y_{T\wedge\tau} + \displaystyle\int_{t\wedge\tau}^{T\wedge\tau} F(s, Y_s, Z_s)ds - \int_{t\wedge\tau}^{T\wedge\tau} Z_s dW_s \\ Y_\tau = \xi \text{ on } \{\tau < \infty\} \end{cases} \tag{5}$$

where the terminal condition ξ is $\mathscr{F}_\tau$-measurable, $\{\mathscr{F}_t\}_{t\geq 0}$ being the filtration generated by W. As mentioned before, there exists a wide literature about the problem, mainly when the generator F has a sublinear growth with respect to z. There are two classical assumptions on the generator F in order to solve such BSDEs, we refer to [7], [23] and [3]:

- F is Lipschitz with respect to z: $|F(t,y,z)-F(t,y,z')| \le K|z-z'|$;
- F is monotone in y: $(y-y')\left(F(t,y,z)-F(t,y',z)\right) \le -\lambda|y-y'|^2$.

Of course, one also needs some integrability conditions on the data namely

$$\mathbb{E}\left[e^{\rho\tau}|\xi|^2 + \int_0^\tau e^{\rho s}|F(s,0,0)|^2 ds\right] < +\infty$$

for some $\rho > K^2 - 2\lambda$. Under these assumptions, the BSDE (5) has a unique solution (Y,Z) which satisfies

$$\mathbb{E}\left[\int_0^\tau e^{\rho s}\left(|Y_s|^2 + |Z_s|^2\right) ds\right] < \infty.$$

Thus, solving BSDEs with random terminal time requires a "structural" condition on the coefficient F which links the constant of monotonicity and the Lipschitz constant of F in z, that is $\rho > K^2 - 2\lambda$. In particular, if $\tau = +\infty$ and $F(s,0,0)$ bounded (there is no terminal condition in this case), one needs $\lambda > K^2/2$, in order to construct a solution. Let us point out that, under this structural condition, BSDE (5) can be solved when the process Y takes its values in $\mathbb{R}^k$ with $k \ge 1$ and also in an infinite dimensional framework (see e.g. [14]).

For real-valued BSDEs, in other words when the process Y takes its values in $\mathbb{R}$, Briand and Hu in [4] and, afterward Royer in [25], improve these results by removing the structural condition on the generator F. In the real case, they require that $F(t,0,0)$ is bounded and use the Girsanov transform to prove that the equation (5) has unique solution (Y,Z) such that Y is a bounded process as soon as $\lambda > 0$. The same arguments are handled by Hu and Tessitore in [17] in the case of a cylindrical Wiener process. The main idea which allows to avoid this structural condition is to get rid of the dependence of the generator F with respect to z with a Girsanov transformation. To be more precise, the main point is to write the equation (5) in the following way

$$\begin{aligned} Y_{t\wedge\tau} &= Y_{T\wedge\tau} + \int_{t\wedge\tau}^{T\wedge\tau} \left(F(s,Y_s,0) + \langle b_s, Z_s\rangle\right) ds - \int_{t\wedge\tau}^{T\wedge\tau} Z_s dW_s, \\ &= Y_{T\wedge\tau} + \int_{t\wedge\tau}^{T\wedge\tau} F(s,Y_s,0)\, ds - \int_{t\wedge\tau}^{T\wedge\tau} Z_s\, d\widehat{W_s}, \end{aligned} \tag{6}$$

where $\widehat{W_t} = W_t - \int_0^t b_s\, ds$ and the process b is given by

$$b_s = \frac{F(s,Y_s,Z_s) - F(s,Y_s,0)}{|Z_s|^2} Z_s \mathbf{1}_{|Z_s|>0}.$$

When F is Lipschitz with respect to z, the process b is bounded and, for each $T > 0$,

$$\mathscr{E}_t = \exp\left(\int_0^t b_s\, dW_s - \frac{1}{2}\int_0^t |b_s|^2 ds\right), \quad 0 \le t \le T,$$

is a uniformly integrable martingale. If $\mathbb{P}$ stands for the probability under which W is a Wiener process, the probability measure $\mathbb{Q}_T$, whose density with respect to the restriction, $\mathbb{P}_T$, of $\mathbb{P}$ to $\mathscr{F}_T^W$ is $\mathscr{E}_T$, is equivalent to $\mathbb{P}_T$ and, under $\mathbb{Q}_T$, $\left\{\widehat{W_t}\right\}_{0\le t\le T}$ is a Wiener process. Coming back to (6) and working under $\mathbb{Q}_T$, we see that the dependence of the generator with respect to z has been removed allowing finally to get rid of the structure condition.

As mentioned before, we are interested in the case where F has a quadratic growth with respect to z and F is strictly monotone in y without any structure condition. We will assume more precisely that

$$\left|F(t,y,z)-F(t,y,z')\right| \leq C(1+|z|+|z'|)|z-z'|,$$

and we will apply more or less the same approach we have just presented when F is Lipschitz with respect to z. In this quadratic setting, the process b will not be bounded in general. However, we will still be able to prove that $\{\mathscr{E}_t\}_{0\leq t\leq T}$ is a uniformly integrable martingale for each $T>0$. This will result from the fact that $\left\{\int_0^t b_s\,dW_s\right\}_{0\leq t\leq T}$ is a BMO-martingale. We refer to [18] for the theory of BMO-martingales.

Let us also mention that M. Kobylanski in [19] considers also quadratic BSDEs with random terminal time. However, she requires that the stopping time is bounded or $\mathbb{P}$-a.s finite. Her method, based on a Hopf-Cole transformation together with some sharp approximations of the generator F, do not allow to treat the case we have in mind, precisely the case where the stopping time τ is almost surely equal to $+\infty$.

The results on quadratic BSDEs on infinite horizon that we will obtain in Section 3 will be exploited to study existence and uniqueness of a mild solution (see Section 5 for the definition) to the PDE (1) where F is a function strictly monotone with respect the second variable and with quadratic growth in the gradient of the solution. Existence and uniqueness of a mild solution of equation (1) in infinite dimensional spaces have been recently studied by several authors employing different techniques (see [6], [15], [11] and [20]). Following several papers (see, for instance [5], [7], [22] for finite dimensional situations, or [14], [17] for infinite dimensional case), we will use a probabilistic approach based on the nonlinear Feynman-Kac formula (4).

The main technical point here will be proving the differentiability of the bounded solution of the backward equation (2) with respect to the initial datum x of the forward equation (3). The proof is based on an a-priori bound for suitable approximations of the equations for the gradient of Y with respect to x and to this end we need to require that the coefficient σ in the forward equation is constant and $A+\nabla b$ is dissipative. We use arguments based on Girsanov transform that we have previously described. We stress again that doing this way we need only the monotonicity constant of F to be positive. The same strategy is applied by Hu and Tessitore [17] to solve the equation (1) when the generator has sublinear growth with respect to the gradient.

The mild solutions to (1), together with their probabilistic representation formula, are particularly suitable for applications to optimal control of infinite dimensional nonlinear stochastic systems. In Section 6 we consider a controlled process X solution of

$$\begin{cases} dX_s = AX_s d\tau + b(X_s)ds + \sigma r(X_s,u_s)ds + \sigma dW_s, \\ X_0 = x \in H, \end{cases} \tag{7}$$

where u denotes the control process, taking values in a given closed subset $\mathscr{U}$ of a Banach space U. The control problem consists of minimizing an infinite horizon cost functional of the form

$$J(x,u) = \mathbb{E}\int_0^\infty e^{-\lambda s} g(X_s^u,u_s)\,ds.$$

Due to the special structure of the control term, the Hamilton-Jacobi-Bellman equation for the value function is of the form (1), provided we set, for $x \in H$ and $z \in \Xi^*$,

$$F(x,y,z) = \inf\{g(x,u) + zr(x,u) : u \in \mathcal{U}\} - \lambda y \tag{8}$$

We suppose that r is a function with values in Ξ^* with linear growth in u and g is a given real function with quadratic growth in u. λ is any positive number. We assume that neither $\mathcal{U}$ nor r is bounded. In this way the Hamiltonian F has quadratic growth in the gradient of the solution and consequently the associated BSDE has quadratic growth in the variable Z. Hence the results obtained on equation (1) allow to prove that the value function of the above problem is the unique mild solution of the corresponding Hamilton-Jacobi-Bellman equation. We adapt the same procedure used in [12] in finite dimension to our infinite dimensional framework. We stress that the usual application of the Girsanov technique is not allowed (since the Novikov condition is not guaranteed) and we have to use specific arguments both to prove the fundamental relation and to solve the closed loop equation. The substantial differences, in comparison with the cited paper, consist in the fact that we work on infinite horizon and we are able to characterize the optimal control in terms of a feedback that involves the gradient of that same solution to the Hamilton-Jacobi-Bellman equation. At the end of the paper we provide a meaningful example for this control problem. We wish to mention that application to stochastic control problem is presented here mainly to illustrate the effectiveness of our results on nonlinear Kolmogorov equation.

Such type of control problems are studied by several authors (see [13],[12]). We underline that the particular structure of the control problem permits that no nondegeneracy assumptions are imposed on σ. In [13] the reader can find a model of great interest in mathematical finance, where absence of nondegeneracy assumptions reveals to be essential.

The paper is organized as follows: the next Section is devoted to notations; in Section 3 we deal with quadratic BSDEs with random terminal time; in Section 4 we study the forward backward system on infinite horizon; in Section 5 we show the result about the solution to PDE. The last Section is devoted to the application to the control problem.

Ackwoledgments. The authors would like to thank the anonymous referee for his careful reading of this manuscript. His remarks and comments allowed to improve this paper.

2 Notations

The norm of an element x of a Banach space E will be denoted $|x|_E$ or simply $|x|$, if 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.

The letters Ξ, H, U will always denote Hilbert spaces. Scalar product is denoted $\langle \cdot, \cdot \rangle$, with a subscript to specify the space, if necessary. All Hilbert spaces are assumed to be real and separable. $L_2(\Xi,U)$ is the space of Hilbert-Schmidt operators from Ξ to U, i.e.

$$L^2(\Xi,U) = \left\{T \in L(\Xi,U) : |T|^2 < +\infty\right\}, \quad \text{with} \quad |T|^2 = \sum\nolimits_{n\geq 1} |Te_n|_U^2,$$

where $\{e_n\}_{n\geq 1}$ is a orthonormal basis of U. $L^2(\Xi,U)$ is a Hilbert space, and the norm $|T|$ defined above makes it a separable Hilbert space. We observe that if $U = \mathbb{R}$ the space $L_2(\Xi,\mathbb{R})$ is the space

$L(\Xi,\mathbb{R})$ of bounded linear operators from Ξ to $\mathbb{R}$. By the Riesz isometry the dual space $\Xi^* = L(\Xi,\mathbb{R})$ can be identified with Ξ.

By a cylindrical Wiener process with values in a Hilbert space Ξ, defined on a probability space $(\Omega,\mathscr{F},\mathbb{P})$, we mean a family $\{W_t,\ t\geq 0\}$ of linear mappings from Ξ to $L^2(\Omega)$, denoted $\xi \mapsto \langle \xi, W_t\rangle$, such that

(i) for every $\xi\in\Xi$, $\{\langle\xi,W_t\rangle,\ t\geq 0\}$ is a real (continuous) Wiener process;

(ii) for every $\xi_1,\xi_2\in\Xi$ and $t\geq 0$, $\mathbb{E}\left(\langle\xi_1,W_t\rangle\cdot\langle\xi_2,W_t\rangle\right)=\langle\xi_1,\xi_2\rangle_{\Xi}\, t$.

$\{\mathscr{F}_t\}_{t\geq 0}$ will denote, the natural filtration of W, augmented with the family of $\mathbb{P}$-null sets. The filtration $\{\mathscr{F}_t\}_{t\geq 0}$ satisfies the usual conditions. All the concepts of measurability for stochastic processes refer to this filtration. By $\mathscr{B}(\Lambda)$ we mean the Borel σ-algebra of any topological space Λ.

We introduce now some classes of stochastic processes with values in a Hilbert space K which we use in the sequel.

- $L^p\left(\Omega;L^2(0,s;K)\right)$ defined for $s\in]0,+\infty]$ and $p\in[1,\infty)$, denotes the space of equivalence classes of progressively measurable processes $\psi:\Omega\times[0,s[\to K$, such that
$$|\psi|^p_{L^p(\Omega;L^2(0,s;K))}=\mathbb{E}\left[\int_0^s|\psi_r|^2_K\,dr\right]^{p/2}<\infty.$$
Elements of $L^p(\Omega;L^2(0,s;K))$ are identified up to modification.
- $L^p(\Omega;C(0,s;K))$, defined for $s\in]0,+\infty[$ and $p\in[1,\infty[$, denotes the space of progressively measurable processes $\{\psi_r,\ r\in[0,s]\}$ with continuous paths in K, such that the norm
$$|\psi|^p_{L^p(\Omega;C([0,s];K))}=\mathbb{E}\left[\sup_{r\in[0,s]}|\psi_r|^p_K\right]$$
is finite. Elements of $L^p(\Omega;C(0,s;K))$ are identified up to indistinguishability.
- $\mathrm{L}^2_{\mathrm{loc}}(K)$ denotes the space of equivalence classes of progressively measurable processes $\psi:\Omega\times[0,\infty)\to K$ such that
$$\forall t>0,\quad \mathbb{E}\left[\int_0^t|\psi_r|^2\,dr\right]<\infty.$$
- If ε is a real number, $\mathrm{M}^{2,\varepsilon}(K)$ denotes the set of $\{\mathscr{F}_t\}_{t\geq 0}$-progressively measurable processes $\{\psi_t\}_{t\geq 0}$ with values in K such that
$$\mathbb{E}\left[\int_0^{+\infty}e^{-2\varepsilon s}|\psi_s|^2\,ds\right]<\infty.$$

We also recall notations and basic facts on a class of differentiable maps acting among Banach spaces, particularly suitable for our purposes (we refer the reader to [13] for details and properties).

Let now X, Z, V denote Banach spaces. We say that a mapping $F : X \to V$ belongs to the class $\mathscr{G}^1(X,V)$ if it is continuous, Gâteaux differentiable on X, and its Gâteaux derivative $\nabla F : X \to L(X,V)$ is strongly continuous.

The last requirement is equivalent to the fact that for every $h \in X$ the map $\nabla F(\cdot)h : X \to V$ is continuous. Note that $\nabla F : X \to L(X,V)$ is not continuous in general if $L(X,V)$ is endowed with the norm operator topology; clearly, if this happens then F is Fréchet differentiable on X. It can be proved that if $F \in \mathscr{G}^1(X,V)$ then $(x,h) \mapsto \nabla F(x)h$ is continuous from $X \times X$ to V; if, in addition, G is in $\mathscr{G}^1(V,Z)$ then $G(F)$ belongs to $\mathscr{G}^1(X,Z)$ and the chain rule holds: $\nabla(G(F))(x) = \nabla G(F(x))\nabla F(x)$.

When F depends on additional arguments, the previous definitions and properties have obvious generalizations.

3 Quadratic BSDEs with random terminal time

In all this section, let τ be an $\{\mathscr{F}_t\}_{t\geq 0}$ stopping time where $\{\mathscr{F}_t\}_{t\geq 0}$ is the filtration generated by the Wiener process defined in the previous section. We use also the following notation.

Definition 3.1. A couple (ξ, F) is said to be a standard quadratic parameter if:

1. the *terminal condition* ξ is a bounded, $\mathscr{F}_\tau$–measurable, real valued random variable;
2. the *generator* F is a function defined on $\Omega \times [0,\infty) \times \mathbb{R} \times \Xi^*$ with values in $\mathbb{R}$, measurable with respect to $\mathscr{P} \otimes \mathscr{B}(\mathbb{R}) \otimes \mathscr{B}(\Xi^*)$ and $\mathscr{B}(\mathbb{R})$ where $\mathscr{P}$ stands for the σ-algebra of progressive sets and such that, for some constant $C \geq 0$, $\mathbb{P}$–a.s. and for all $t \geq 0$,
   (a) $(y,z) \longmapsto F(t,y,z)$ is continuous;
   (b) $\forall y \in \mathbb{R}$, $\forall z \in \Xi^*$, $|F(t,y,z)| \leq C\left(1 + |y| + |z|^2\right)$.

Let us mention that these conditions are the usual ones for studying quadratic BSDEs.

Let (ξ, F) be a standard quadratic parameter. We want to construct an adapted solution $(Y_t, Z_t)_{t\geq 0}$ to the BSDE

$$-\,dY_t = \mathbf{1}_{t\leq\tau}\, F(t, Y_t, Z_t)\, dt - Z_t\, dW_t, \quad Y_\tau = \xi \text{ on } \{\tau < \infty\}. \tag{9}$$

Let us first recall that by a solution to the equation (9) we mean a pair of progressively measurable processes $(Y_t, Z_t)_{t\geq 0}$ with values in $\mathbb{R} \times \Xi^*$ such that:

1. Y is a continuous process, $\mathbb{P}$–a.s., for each $T > 0$, $t \longmapsto Z_t$ belongs to $\mathrm{L}^2((0,T);\Xi^*)$ and $t \longmapsto F(t, Y_t, Z_t) \in \mathrm{L}^1((0,T);\mathbb{R})$;
2. on the set $\{\tau < \infty\}$, we have, for $t \geq \tau$, $Y_t = \xi$ and $Z_t = 0$;
3. for each nonnegative real T, $\forall t \in [0,T]$,

$$Y_t = Y_T + \int_t^T \mathbf{1}_{s\leq\tau} F(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dW_s.$$

Remark 3.2. In the case of a deterministic and finite stopping time, this definition is the usual one except that we define the process (Y, Z) on the whole time axis.

Since the stopping time τ is not assumed to be bounded or $\mathbb{P}$–a.s. finite we will need a further assumption on the generator.

Assumption A1. There exist two constants, $C \geq 0$ and $\lambda > 0$, such that, $\mathbb{P}$–a.s., for all $t \geq 0$,

(i) for all real y,

$$\forall z \in \Xi^*,\ \forall z' \in \Xi^*, \quad \left|F(t,y,z) - F(t,y,z')\right| \leq C\left(1 + |z| + \left|z'\right|\right)\left|z - z'\right|;$$

(ii) F is strictly monotone with respect to y: for all $z \in \Xi^*$,

$$\forall y \in \mathbb{R},\ \forall y' \in \mathbb{R}, \quad (y - y')\ (F(t,y,z) - F(t,y',z)) \leq -\lambda\left|y - y'\right|^2.$$

Theorem 3.3. *Let (ξ, F) be a standard quadratic parameter such that F satisfies A1.*

Then, the BSDE (9) has a unique solution (Y,Z) such that Y is a bounded process and Z belongs to $\mathrm{L}^2_{\text{loc}}(\Xi^*)$. Moreover, $Z \in \mathrm{M}^{2,\varepsilon}(\Xi^*)$ for all $\varepsilon > 0$.

Before proving this result, let us state a useful lemma.

Lemma 3.4. *Let $0 \leq S < T$ and $\left(\xi^1, F^1\right)$, $\left(\xi^2, F^2\right)$ be two standard quadratic parameters. Let, for $i = 1,2$, $\left(Y^i, Z^i\right)$ be a solution to the BSDE*

$$Y^i_t = \xi^i\,\mathbf{1}_{\tau \leq T} + \int_t^T \mathbf{1}_{s \leq \tau} F^i(s, Y^i_s, Z^i_s)\,ds - \int_t^T Z^i_s\,dW_s, \tag{10}$$

such that Y^i is a bounded process and $Z^i \in \mathrm{L}^2((0,T)\times\Omega)$.
If A1 holds for F^1, $\xi^1 - \xi^2 = 0$ on the set $\{S < \tau\}$ and $\left|F^1 - F^2\right|\left(s, Y^2_s, Z^2_s\right) \leq \rho(s)$ where ρ is a deterministic Borelian function then

$$\forall t \in [0,T], \quad \left|Y^1_t - Y^2_t\right| \leq \left\|\xi^1 - \xi^2\right\|_\infty e^{-\lambda_1(S-t)_+} + \int_t^T e^{-\lambda_1(s-t)}\rho(s)\,ds,$$

where $\lambda_1 > 0$ is the constant of monotonicity of F^1.

Proof. Let us start with a simple remark. Let $i \in \{1,2\}$. Since (ξ^i, F^i) is a standard quadratic parameter, and (Y^i, Z^i) a solution to (10) with Y^i bounded and Z^i square integrable, it is by know well known (see e.g. [2]) that the martingale $\left\{N^i_t = \int_0^t Z^i_s\,dW_s\right\}_{0\leq t\leq T}$ has the following property: there exists a constant γ_i such that, for each stopping time $\sigma \leq T$,

$$\mathbb{E}\left(\left|N^i_T - N^i_\sigma\right|^2 \,\middle|\, \mathscr{F}_\sigma\right) = \mathbb{E}\left(\int_\sigma^T |Z^i_s|^2\,ds \,\middle|\, \mathscr{F}_\sigma\right) \leq \gamma_i.$$

In other words (we refer to N. Kamazaki [18] for the notion of BMO–martingales), $\{N^i_t\}_{0\leq t\leq T}$ is a BMO–martingale.

With this observation in hands, let us prove our lemma. Since F^1 satisfies A1, F^1 is strictly monotone with respect to y. Let us denote $\lambda_1 > 0$ the constant of monotonicity of F^1 end let us fix $t \in [0,T]$. We set, for $s \in [0,T]$,

$$E_s = \exp\left(-\lambda_1 \int_0^s \mathbf{1}_{u\leq\tau}\mathbf{1}_{u>t}\,du\right) = \exp\left(-\lambda_1(\tau\wedge s - \tau\wedge t)_+\right).$$

We have, from Itô–Tanaka formula applied to $E_s\left|Y_s^1 - Y_s^2\right|$,

$$\begin{aligned}\left|Y_t^1 - Y_t^2\right| = E_T\left|\xi^1 - \xi^2\right|\mathbf{1}_{\tau\leq T} - \int_t^T E_s\,\mathrm{sgn}\left(Y_s^1 - Y_s^2\right)\left(Z_s^1 - Z_s^2\right)dW_s - \int_t^T dL_s \\ + \int_t^T E_s\mathbf{1}_{s\leq\tau}\left[\lambda_1\left|Y_s^1 - Y_s^2\right| + \mathrm{sgn}\left(Y_s^1 - Y_s^2\right)\left(F^1\left(s,Y_s^1,Z_s^1\right) - F^2\left(s,Y_s^2,Z_s^2\right)\right)\right]ds,\end{aligned}$$

where L is the local time at 0 of the semimartingale $Y^1 - Y^2$ and $\mathrm{sgn}(x) = -\mathbf{1}_{x\leq 0} + \mathbf{1}_{x>0}$. Now, we use the usual decomposition

$$\begin{aligned}F^1\left(s,Y_s^1,Z_s^1\right) - F^2\left(s,Y_s^2,Z_s^2\right) = F^1\left(s,Y_s^1,Z_s^1\right) - F^1\left(s,Y_s^2,Z_s^1\right) \\ + F^1\left(s,Y_s^2,Z_s^1\right) - F^1\left(s,Y_s^2,Z_s^2\right) + \left(F^1 - F^2\right)\left(s,Y_s^2,Z_s^2\right).\end{aligned}$$

By assumption, we have $\left|F^1 - F^2\right|\left(s,Y_s^2,Z_s^2\right) \leq \rho(s)$. Moreover, since F^1 is λ_1–monotone,

$$\mathrm{sgn}\left(Y_s^1 - Y_s^2\right)\left(F^1\left(s,Y_s^1,Z_s^1\right) - F^1\left(s,Y_s^2,Z_s^1\right)\right) \leq -\lambda_1\left|Y_s^1 - Y_s^2\right|.$$

Thus we get, L being nondecreasing,

$$\begin{aligned}\left|Y_t^1 - Y_t^2\right| \leq E_T\left|\xi^1 - \xi^2\right|\mathbf{1}_{\tau\leq T} - \int_t^T E_s\,\mathrm{sgn}\left(Y_s^1 - Y_s^2\right)\left(Z_s^1 - Z_s^2\right)dW_s + \int_t^T E_s\mathbf{1}_{s\leq\tau}\rho(s)\,ds \\ + \int_t^T E_s\mathbf{1}_{s\leq\tau}\,\mathrm{sgn}\left(Y_s^1 - Y_s^2\right)\left(F^1\left(s,Y_s^2,Z_s^1\right) - F^1\left(s,Y_s^2,Z_s^2\right)\right)ds.\end{aligned}$$

To go further, let us remark, for $s \in [t,T]$, $E_s\mathbf{1}_{s\leq\tau} = e^{-\lambda_1(\tau\wedge s - \tau\wedge t)}\mathbf{1}_{s\leq\tau} \leq e^{-\lambda_1(s-t)}$. Moreover, since $\xi^1 - \xi^2 = 0$ on the set $\{S < \tau \leq T\}$, we have

$$E_T\left|\xi^1 - \xi^2\right|\mathbf{1}_{\tau\leq T} = e^{-\lambda_1(T\wedge\tau - t\wedge\tau)_+}\left|\xi^1 - \xi^2\right|\mathbf{1}_{S<\tau\leq T} \leq e^{-\lambda_1(S-t)_+}\|\xi^1 - \xi^2\|_\infty,$$

from which we deduce the following inequality

$$\begin{aligned}\left|Y_t^1 - Y_t^2\right| \leq e^{-\lambda_1(S-t)_+}\|\xi^1 - \xi^2\|_\infty + \int_t^T e^{-\lambda_1(s-t)}\rho(s)\,ds - \int_t^T E_s\,\mathrm{sgn}\left(Y_s^1 - Y_s^2\right)\left(Z_s^1 - Z_s^2\right)dW_s \\ + \int_t^T E_s\mathbf{1}_{s\leq\tau}\,\mathrm{sgn}\left(Y_s^1 - Y_s^2\right)\left(F^1\left(s,Y_s^2,Z_s^1\right) - F^1\left(s,Y_s^2,Z_s^2\right)\right)ds.\end{aligned}$$

To conclude, the proof of this lemma, le us define the process $\{b_s\}_{0\leq s\leq T}$ with values in Ξ^*, by setting

$$b_s = \mathbf{1}_{s\leq\tau}\frac{F^1\left(s,Y_s^2,Z_s^1\right) - F^1\left(s,Y_s^2,Z_s^2\right)}{\left|Z_s^1 - Z_s^2\right|^2}\left(Z_s^1 - Z_s^2\right)\mathbf{1}_{\left|Z_s^1 - Z_s^2\right|>0}.$$

We can rewrite the previous inequality in the following way

$$\left|Y_t^1 - Y_t^2\right| \leq e^{-\lambda_1(S-t)_+}\|\xi^1 - \xi^2\|_\infty + \int_t^T e^{-\lambda_1(s-t)}\rho(s)\,ds \\ - \int_t^T E_s \operatorname{sgn}\left(Y_s^1 - Y_s^2\right)\left(Z_s^1 - Z_s^2\right) dW_s + \int_t^T E_s \operatorname{sgn}\left(Y_s^1 - Y_s^2\right)\left\langle b_s, Z_s^1 - Z_s^2\right\rangle ds.$$

Let us observe that, since F^1 satisfies A1.1, we have $|b_s| \leq C\left(1 + \left|Z_s^1\right| + \left|Z_s^2\right|\right)$. Since we know that the stochastic integral (as process on $[0,T]$) of Z^1 and Z^2 are BMO–martingales we deduce that $\left\{\int_0^t b_s\,dW_s\right\}_{0\leq t\leq T}$ is also a BMO–martingale. As a byproduct, see [18, Theorem 2.3], the exponential martingale,

$$\mathscr{E}_t = \exp\left(\int_0^t b_s dW_s - \frac{1}{2}\int_0^t |b_s|^2 ds\right), 0 \leq t \leq T,$$

is a uniformly integrable martingale. Let us consider the probability measure $\mathbb{Q}_T$ on $(\Omega, \mathscr{F}_T)$ whose density with respect to $\mathbb{P}_{|\mathscr{F}_T}$ is given by $\mathscr{E}_T$. Then $\mathbb{Q}_T$ and $\mathbb{P}_{|\mathscr{F}_T}$ are equivalent on $(\Omega, \mathscr{F}_T)$, and under $\mathbb{Q}_T$, by Girsanov theorem, the process $\left\{\widehat{W_t} = W_t - \int_0^t b_s ds\right\}_{0\leq t\leq T}$ is a Wiener process.

To conclude, let us write the last inequality in the following way

$$\left|Y_t^1 - Y_t^2\right| \leq e^{-\lambda_1(S-t)_+}\|\xi^1 - \xi^2\|_\infty + \int_t^T e^{-\lambda_1(s-t)}\rho(s)\,ds - \int_t^T E_s \operatorname{sgn}\left(Y_s^1 - Y_s^2\right)\left(Z_s^1 - Z_s^2\right) d\widehat{W_s};$$

taking the conditional expectation under $\mathbb{Q}_T$ with respect to $\mathscr{F}_t$, we obtain the result of the lemma. □

Now we can prove the main result of this section, concerning the existence and uniqueness of solutions of BSDE (9).

Proof of Theorem 3.3.

Existence. We adopt the same strategy as in [4] and [25], with some significant modifications.

Let us denote by γ a positive constant such that

$$\|\xi\|_\infty \leq \gamma, \quad |F(t,y,z)| \leq \gamma(1 + |y| + |z|), \quad |F(t,y,z) - F(t,y,z')| \leq \gamma(1 + |z| + |z'|)|z - z'|, \quad (11)$$

and by $\lambda > 0$ the monotonicity constant of F.

Fore each integer n, let us denote (Y^n, Z^n) the unique solution to the BSDE

$$Y_t^n = \xi\mathbf{1}_{\tau\leq n} + \int_t^n \mathbf{1}_{s\leq\tau}F(s, Y_s^n, Z_s^n)ds - \int_t^n Z_s^n dW_s, \qquad 0 \leq t \leq n. \quad (12)$$

We know from results of [19] (these results can be easily generalized to the case of cylindrical Wiener process) that, (ξ, F) being a standard quadratic parameter, the BSDE (12) has a unique bounded solution under A1. Moreover we have $Y_t^n = Y_{t\wedge\tau}^n$, $Z_t^n\mathbf{1}_{t>\tau} = 0$, see e.g. [25].

We define, (Y^n, Z^n) on the whole time axis by setting,

$$\forall t > n, \quad Y_t^n = Y_n^n = \xi \mathbf{1}_{\tau \leq n}, \quad Z_t^n = 0.$$

First of all we prove, thanks to the assumption of monotonicity A1.2, that Y^n is bounded by a constant independent of n. Let us apply Lemma 3.4, with $S = 0$, $T = n$, $F^1 = F$, $F^2 = 0$, $\xi^1 = \xi$ and $\xi^2 = 0$. We get, for all $t \in [0, n]$,

$$|Y_t^n| \leq \|\xi\|_\infty + \gamma \int_t^n e^{-\lambda(s-t)} \, ds \leq \gamma \left(1 + \frac{1}{\lambda}\right). \tag{13}$$

In all the remaining of the proof, we will denote $C(\gamma, \lambda)$ a constant depending on γ and λ which may change from line to line.

Moreover we can show that, for each $\epsilon > 0$,

$$\sup_{n \geq 1} \mathbb{E}\left[\int_0^\infty e^{-2\epsilon s} \left|Z_s^n\right|^2 ds\right] < \infty. \tag{14}$$

To obtain this estimate we consider the function $\varphi(x) = \left(e^{2\gamma x} - 2\gamma x - 1\right)/(2\gamma)^2$, where $\gamma > 0$ is the constant defined in (11) which has the following properties: for $x \geq 0$,

$$\varphi'(x) \geq 0, \qquad \varphi''(x) - 2\gamma\varphi'(x) = 1.$$

The function $\varphi(|x|)$ is $\mathscr{C}^2$ and the estimate follows directly from the computation of the Itô differential of $e^{-2\epsilon t}\varphi(|Y_t^n|)$.

Now we study the convergence of the sequence $(Y^n)_{n \geq 0}$. By construction we have, for $n < m$,

$$Y_t^m = \xi \mathbf{1}_{\tau \leq m} + \int_t^m \mathbf{1}_{s \leq \tau} F\left(s, Y_s^m, Z_s^m\right) ds - \int_t^m Z_s^m \, dW_s, \quad 0 \leq t \leq m,$$
$$Y_t^n = \xi \mathbf{1}_{\tau \leq n} + \int_t^m \mathbf{1}_{s \leq \tau} \widehat{F}\left(s, Y_s^n, Z_s^n\right) ds - \int_t^m Z_s^n \, dW_s, \quad 0 \leq t \leq m,$$

where $\widehat{F}(s, y, z) = \mathbf{1}_{s<n} F(s, y, z)$. Let us apply Lemma 3.4 with $T = m$, $(\xi^1, F^1) = (\xi, F)$, $(\xi^2, F^2) = (\xi \mathbf{1}_{\tau \leq n}, \widehat{F})$. We have $\xi - \xi \mathbf{1}_{\tau \leq n} = \xi \mathbf{1}_{\tau > n}$, and

$$\left|F - \widehat{F}\right|(s, Y_s^n, Z_s^n) = \mathbf{1}_{s>n} \left|F(s, Y_s^n, Z_s^n)\right| = \mathbf{1}_{s>n} \left|F(s, \xi \mathbf{1}_{\tau \leq n}, 0)\right| \leq C(\gamma, \lambda) \mathbf{1}_{s>n}.$$

Choosing $S = n$, we get, for $t \in [0, m]$,

$$\left|Y_t^m - Y_t^n\right| \leq C(\gamma, \lambda) \left(e^{-\lambda(n-t)_+} + \int_t^m e^{-\lambda(s-t)} \mathbf{1}_{s>n} ds\right) \leq C(\gamma, \lambda) e^{-\lambda(n-t)_+}.$$

Since both processes Y^n and Y^m are bounded by a constant depending only on γ and λ, the previous inequality holds for all nonnegative real t, namely

$$\forall t \geq 0, \qquad \left|Y_t^m - Y_t^n\right| \leq C(\gamma, \lambda) e^{-\lambda(n-t)_+}. \tag{15}$$

We deduce immediatly from the previous estimate that the sequence $(Y^n)_{n\geq 0}$ converges uniformly on compacts in probability (ucp for short) since, for any $a \geq 0$, we have

$$\sup_{0\leq t\leq a} \left|Y_t^m - Y_t^n\right| \leq C(\gamma,\lambda)e^{-\lambda(n-a)},$$

as soon as $a \leq n \leq m$. Let Y be the limit of $(Y^n)_{n\geq 0}$. Since, for each n, Y^n is continuous and bounded by $\gamma(1+1/\lambda)$ the same is true for Y, and sending m to infinity in (15), we get

$$\forall t \geq 0, \qquad \left|Y_t - Y_t^n\right| \leq C(\gamma,\lambda)e^{-\lambda(n-t)_+}.$$

It follows that the convergence of $(Y^n)_{n\geq 0}$ to Y holds also in $\mathrm{M}^{2,\varepsilon}(\mathbb{R})$ for all $\varepsilon > 0$. Indeed, it is enough to prove this convergence for $0 < \varepsilon < \lambda$ and in this case we have

$$\begin{aligned}\mathbb{E}\left[\int_0^\infty e^{-2\varepsilon s}\left|Y_t - Y_t^n\right|^2 ds\right] &\leq C(\gamma,\lambda)\int_0^\infty e^{-2\varepsilon s}e^{-2\lambda(n-s)_+}\,ds \\ &= C(\gamma,\lambda)\left(\frac{1}{2(\lambda-\varepsilon)}\left(e^{-2\varepsilon n} - e^{-2\lambda n}\right) + \frac{1}{2\varepsilon}e^{-2\varepsilon n}\right).\end{aligned}$$

Let us show that the sequence $(Z_n)_{n\geq 0}$ is a Cauchy sequence in the space $\mathrm{M}^{2,\varepsilon}(\Xi^*)$, for all $\varepsilon > 0$. Let $\varepsilon > 0$, and $m > n$ be two integers. Applying Ito's formula to the process $e^{-2\varepsilon t}\left|Y_t^m - Y_t^n\right|^2$ we get

$$\begin{aligned}&\left|Y_0^m - Y_0^n\right|^2 + \int_0^m e^{-2\varepsilon s}\left|Z_s^m - Z_s^n\right|^2 ds \qquad (16)\\ &= e^{-2\varepsilon m}|\xi|^2\,\mathbf{1}_{n<\tau\leq m} - \int_0^m 2e^{-2\varepsilon s}\left(Y_s^m - Y_s^n\right)\left(Z_s^m - Z_s^n\right)dW_s \\ &\quad +2\int_0^m e^{-2\varepsilon s}\left[\varepsilon\left|Y_s^m - Y_s^n\right|^2 + \left(Y_s^m - Y_s^n\right)\mathbf{1}_{s\leq\tau}\left(F(s,Y_s^m,Z_s^m) - \widehat{F}(s,Y_s^n,Z_s^n)\right)\right]ds.\end{aligned}$$

Since Y^n and Y^m are bounded by $C(\gamma,\lambda)$, we have in view of the growth assumption on F, for a constant D depending on γ, λ and ε (and changing from line to line if necessary),

$$\begin{aligned}&\varepsilon\left|Y_s^m - Y_s^n\right|^2 + \left(Y_s^m - Y_s^n\right)\mathbf{1}_{s\leq\tau}\left(F(s,Y_s^m,Z_s^m) - \widehat{F}(s,Y_s^n,Z_s^n)\right) \\ &\qquad\qquad \leq D\left|Y_s^m - Y_s^n\right|\left(1 + \left|Z_s^m - Z_s^n\right|^2\right).\end{aligned}$$

Coming back to (16) and taking the expectation, we obtain the inequality, since $Z_s^m = Z_s^n = 0$ for $s > m$ and ξ is bounded by γ,

$$\begin{aligned}\mathbb{E}\left[\int_0^\infty e^{-2\varepsilon s}\left|Z_s^m - Z_s^n\right|^2 ds\right] &\leq \gamma e^{-2\varepsilon m} + D\,\mathbb{E}\left[\int_0^\infty e^{-2\varepsilon s}\left|Y_s^m - Y_s^n\right|\left(1+\left|Z_s^m - Z_s^n\right|^2\right)ds\right] \\ &\leq \gamma e^{-2\varepsilon m} + D\,\mathbb{E}\left[\int_0^\infty e^{-2\varepsilon s}e^{-\lambda(n-s)_+}\left(1+\left|Z_s^m - Z_s^n\right|^2\right)ds\right],\end{aligned}$$

where we have used (15) to get the last upper bound. We have, finally

$$\mathbb{E}\left[\int_0^{n/2} e^{-2\varepsilon s}e^{-\lambda(n-s)_+}\left(1+\left|Z_s^m - Z_s^n\right|^2\right)ds\right] \leq e^{-\lambda n/2}\,\mathbb{E}\left[\int_0^{n/2} e^{-2\varepsilon s}\left(1+\left|Z_s^m - Z_s^n\right|^2\right)ds\right],$$

$$\mathbb{E}\left[\int_{n/2}^\infty e^{-2\varepsilon s}e^{-\lambda(n-s)_+}\left(1+\left|Z_s^m - Z_s^n\right|^2\right)ds\right] \leq e^{-\varepsilon n/2}\,\mathbb{E}\left[\int_{n/2}^\infty e^{-\varepsilon s}\left(1+\left|Z_s^m - Z_s^n\right|^2\right)ds\right],$$

from which we get the result since we have already shown that the sequence $(Z^n)_{n\geq 0}$ is bounded in $\mathrm{M}^{2,\varepsilon}(\Xi^*)$ for each $\varepsilon > 0$. We call Z the limit of $(Z^n)_{n\geq 0}$.

It remains to show that the process (Y,Z) satisfies the BSDE (9). We already know that Y is continuous and bounded and Z belongs to $\mathrm{M}^{2,\varepsilon}(\Xi^*)$. Let us fix $0 \leq t \leq T$. By definition, for each $n \geq T$, we have

$$Y_t^n = Y_T^n + \int_t^T \mathbf{1}_{s\leq\tau} F\left(s, Y_s^n, Z_s^n\right) ds - \int_t^T Z_s^n \, dW_s. \tag{17}$$

The sequence $(Y^n)_{n\geq 0}$ converges to Y ucp and is bounded by $\gamma(1+1/\lambda)$ uniformly in n. Thus $\sup_{0\leq t\leq T} \left|Y_t^n - Y_t\right|$ converges to 0 in L^1. Moreover, from Doob's inequality, we get

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|\int_0^t \left(Z_s^n - Z_s\right) dW_s\right|^2\right] \leq 4e^{2\lambda T}\mathbb{E}\left[\int_0^\infty e^{-2\lambda s}\left|Z_s^n - Z_s\right|^2 ds\right] \to 0.$$

Finally, let us notice that $\int_t^T \mathbf{1}_{s\leq\tau} F(s, Y_s^n, Z_s^n)\, ds$ converges to $\int_t^T \mathbf{1}_{s\leq\tau} F(s, Y_s, Z_s)\, ds$ in L^1. Indeed,

$$\mathbb{E}\left[\left|\int_{t\wedge\tau}^{T\wedge\tau} F(s, Y_s^n, Z_s^n) ds - \int_{t\wedge\tau}^{T\wedge\tau} F(s, Y_s, Z_s) ds\right|\right] \leq \mathbb{E}\left[\int_0^T \left|F(s, Y_s^n, Z_s^n) - F(s, Y_s, Z_s)\right| ds\right],$$

and, by the growth assumption on F, the map $(Y,Z) \to F(\cdot, Y, Z)$ is continuous from the space $L^1(\Omega; L^1([0,T];\mathbb{R})) \times L^2(\Omega; L^2([0,T];\Xi^*)$ to $L^1(\Omega; L^1([0,T];\mathbb{R}))$, by classical result on continuity of evaluation operators, see e.g. [1]. Hence, passing to the limit in the equation (17), we obtain forall t and all T such that $0 \leq t \leq T$

$$Y_t = Y_T + \int_t^T \mathbf{1}_{s\leq\tau} F(s, Y_s, Z_s)\, ds - \int_t^T Z_s \, dW_s.$$

So to conclude the proof, it only remains to check that, on the set $\{\tau < +\infty\}$, we have $Y_\tau = \xi$. Let us fix $a > 0$. For each $n \geq a$, we have in view of (15),

$$\left|Y_a - \xi \mathbf{1}_{\tau\leq n}\right| = \left|Y_a - Y_n^n\right| \leq \left|Y_a - Y_a^n\right| + \left|Y_a^n - Y_n^n\right| \leq C(\gamma,\lambda)e^{-\lambda(n-a)} + \left|Y_a^n - Y_n^n\right|.$$

Let us recall that, for each t, $Y_t^n = Y_{t\wedge\tau}^n$ and $Y_t = Y_{t\wedge\tau}$. Hence, on the event $\{\tau \leq a\}$, we have, since $n \geq a$, $Y_a^n = Y_n^n = Y_\tau^n$ and $Y_a = Y_\tau$. Thus, we deduce from the previous inequality, that $|Y_\tau - \xi| \leq C(\gamma,\lambda)e^{-\lambda(n-a)}$ on the set $\{\tau \leq a\}$. It follows that $Y_\tau = \xi$ $\mathbb{P}$-a.s. on the set $\{\tau < \infty\}$, and the process (Y,Z) is solution for BSDE (9).

Uniqueness. Suppose that (Y^1, Z^1) and (Y^2, Z^2) are both solutions of the BSDE (9) such that Y^1 and Y^2 are continuous and bounded and Z^1 and Z^2 belong to $\mathrm{L}^2_{\mathrm{loc}}(\Xi^*)$. It follows directly from Lemma 3.4 that, $\mathbb{P}$–a.s.,

$$\forall t \geq 0, \qquad Y_t^1 = Y_t^2.$$

Applying Ito's formula to $\left|Y_t^1 - Y_t^2\right|^2$, we have that $d\mathbb{P}\otimes dt$–a.e. $Z_t^1 = Z_t^2$. □

Let us finish this section by the following remark.

Remark 3.5. Let $\{(Y_t, Z_t)\}_{0\le t\le T}$ be a solution, with Y bounded and Z square integrable, to the linear BSDE

$$Y_t = \xi + \int_t^T \left(\psi_s + a_s Y_s + \langle b_s, Z_s\rangle\right) ds - \int_t^T Z_s\, dW_s, \quad 0 \le t \le T,$$

where the processes ψ, a and b are progressively measurable with values in $\mathbb{R}$, $\mathbb{R}$ and Ξ^* respectively. Let us assume that ξ is bounded, for some $\gamma \ge 0$, $|\psi_s| \le \gamma$ and $a_s \le -\lambda$ for $\lambda > 0$. Then, arguing as in the proof of Lemma 3.4, when

$$\mathscr{E}_t = \exp\left(\int_0^t b_s dW_s - \frac{1}{2}|b_s|^2\, ds\right), \quad 0 \le t \le T,$$

is a uniformly integrable martingale, we have, $\mathbb{P}$–a.s.,

$$\forall t \in [0, T], \qquad |Y_t| \le \left(\|\xi\|_\infty + \frac{\gamma}{\lambda}\right) e^{-\lambda(T-t)}.$$

4 The forward-backward system on infinite horizon

In this Section we use the previous result to study a forward-backward system on infinite horizon, when the backward equation has quadratic generator.

We consider the Itô stochastic equation for an unknown process $\{X_s, s \ge 0\}$ with values in a Hilbert space H:

$$X_s = e^{sA}x + \int_0^s e^{(s-r)A} b(X_r) dr + \int_0^s e^{(s-r)A}\sigma dW_r,\ s \ge 0. \tag{18}$$

Our assumptions will be the following:

Assumption A2. (i) The operator A is the generator of a strongly continuous semigroup e^{tA}, $t \ge 0$, in a Hilbert space H. We denote by m and a two constants such that $|e^{tA}| \le m e^{at}$ for $t \ge 0$.

(ii) $b : H \to H$ satisfies, for some constant $L > 0$,

$$|b(x) - b(y)| \le L|x - y|,\ x, y \in H.$$

(iii) σ belongs to $L(\Xi, H)$ such that $e^{tA}\sigma \in L_2(\Xi, H)$ for every $t > 0$, and

$$|e^{tA}\sigma|_{L_2(\Xi,H)} \le L t^{-\gamma} e^{at},$$

for some constants $L > 0$ and $\gamma \in [0, 1/2)$.

(iv) $b(\cdot) \in \mathscr{G}^1(H, H)$.

(v) Operators $A + b_x(x)$ are dissipative: $\langle Ay, y\rangle + \langle b_x(x)y, y\rangle \le 0$ for all $x \in H$ and $y \in D(A)$.

Remark 4.1. This kind of requests is usual if we wish to study the problem of the regular dependence on the data in a forward-backward system in the degenerate case on infinite horizon (compare with [17]).

We note that the assumptions (i)-(iii) are the classical assumptions to prove existence and uniqueness of the solution of equation (18) (see [9], Theorem 5.3.1, for the theory and §11.2, or [13],[14] for some typical examples). In general the coefficient σ depends on the process X. We need that σ is constant and moreover we have to assume assumption (v) to obtain the following estimate

$$\mathbb{P}\text{-a.s. } |\nabla_x X_t^x h| \leq K\,|h|, \quad \forall t > 0.$$

We stress that the previous inequality is crucial in order to show the regular dependence with respect to x of the process Y in the forward-backward system (Theorem 4.6 below). Assumption (iv) is clearly natural to have differentiable dependence on x.

We start by recalling a well known result on solvability of equation (18) on a bounded interval, see e.g. [13].

Proposition 4.2. *Under the assumption A2, for every $p \in [2,\infty)$ and $T > 0$ there exists a unique process $X^x \in L^p(\Omega; C(0,T;H))$ solution of (18). Moreover, for all fixed $T > 0$, the map $x \to X^x$ is continuous from H to $L^p(\Omega; C(0,T;H))$ and*

$$\mathbb{E}\left[\sup_{r\in[0,T]} |X_r^x|^p\right] \leq C(1+|x|)^p,$$

for some constant C depending only on p, γ, T, L, a and m.

We need to state a regularity result on the process X. The proof of the following lemma can be found in [17]. In the sequel X^x denotes the unique mild solution to (18) starting from $X_0 = x$.

Lemma 4.3. *Under the assumption A2, the map $x \to X^x$ is Gâteaux differentiable and belongs to $\mathcal{G}(H, L^p(\Omega, C(0,T;H))$. Moreover denoting by $\nabla_x X^x$ the partial Gâteaux derivative, then for every direction $h \in H$, the directional derivative process $\nabla_x X^x h, t \in \mathbb{R}$, solves, $\mathbb{P} - a.s.$, the equation*

$$\nabla_x X_t^x h = e^{tA} h + \int_0^t e^{sA} \nabla_x b(X_s^x) \nabla_x X_s^x h\, ds, \quad t \in \mathbb{R}^+.$$

Finally, $\mathbb{P}$-a.s., $|\nabla_x X_t^x h| \leq K|h|$, for all $t > 0$.

The associated BSDE is:

$$Y_t^x = Y_T^x + \int_t^T F(X_s^x, Y_s^x, Z_s^x) ds - \int_t^T Z_s^x dW_s, \quad 0 \leq t \leq T < \infty. \tag{19}$$

Here Y is real valued and Z takes values in Ξ^*, $F : H \times \mathbb{R} \times \Xi^* \to \mathbb{R}$ is a given measurable function. The notation Y^x and Z^x stress the dependence of the processes Y and Z solution to the backward equation by the starting point x in the forward equation.

We assume the following on F:

Assumption A3. There exist $C \geq 0$ and $\alpha \in (0,1)$ such that

1. $|F(x,y,z)| \leq C\left(1+|y|+|z|^2\right)$;
2. $F(\cdot,\cdot,\cdot)$ is $\mathcal{G}^{1,1,1}(H \times \mathbb{R} \times \Xi^*; \mathbb{R})$;

3. $\left|\nabla_x F(x,y,z)\right| \le C$;

4. $\left|\nabla_z F(x,y,z)\right| \le C(1+|z|)$;

5. $\left|\nabla_y F(x,y,z)\right| \le C(1+|z|)^{2\alpha}$;

6. F is monotone in y with constant of monotonicity $\lambda > 0$ in the following sense:

$$\forall x \in H, y, y' \in \mathbb{R}, z \in \Xi^*, \quad (y-y')(F(x,y,z)-F(x,y',z)) \le -\lambda|y-y'|^2.$$

Remark 4.4. In comparison with the assumptions of the previous section, we add mainly the differentiability of F A3.2. We use again an approximation procedure in order to prove the regular dependence on the parameter x of the solution to the BSDE (19). Hence to use known results on differentiability for BSDEs with quadratic generator on finite time interval (see [2]) we need A3.5. Finally we use A3.3 to obtain a uniform estimate on $\nabla_x Y_0^x$ (see Theorem 4.6 below).

Applying Theorem 3.3, we obtain:

Proposition 4.5. *Let us suppose that Assumptions A2 and A3 hold. Then we have:*

(i) *For any $x \in H$, there exists a solution (Y^x, Z^x) to the BSDE (19) such that Y^x is a continuous process bounded by C/λ, and $Z \in \mathrm{M}^{2,\varepsilon}(\Xi^*)$ for each $\varepsilon > 0$. The solution is unique in the class of processes (Y,Z) such that Y is continuous and bounded, and Z belongs to $\mathrm{L}^2_{\mathrm{loc}}(\Xi^*)$.*

(ii) *For all $T > 0$ and $p \ge 1$, the map $x \to (Y^x\big|_{[0,T]}, Z^x\big|_{[0,T]})$ is continuous from H to the space $L^p(\Omega; C(0,T;\mathbb{R})) \times L^p(\Omega; L^2(0,T;\Xi^*))$.*

Proof. Statement (i) is an immediate consequences of Theorem 3.3. Let us prove (ii). Denoting by $(Y^{n,x}, Z^{n,x})$ the unique solution of the following BSDE (with finite horizon):

$$Y_t^{n,x} = \int_t^n F(X_s^x, Y_s^{n,x}, Z_s^{n,x})ds - \int_t^n Z_s^{n,x} dW_s, \tag{20}$$

then, from Theorem 3.3 again, $|Y_t^{n,x}| \le \frac{C}{\lambda}$ and the following convergence rate holds:

$$|Y_t^{n,x} - Y_t^x| \le \frac{C}{\lambda}\exp\{-\lambda(n-t)\}.$$

Now, if $x_i' \to x$ as $i \to +\infty$ then

$$\begin{aligned} |Y_T^{x_i'} - Y_T^x| &\le |Y_T^{x_i'} - Y_T^{n,x_i'}| + |Y_T^{n,x} - Y_T^x| + |Y_T^{n,x_i'} - Y_T^{n,x}| \\ &\le 2\frac{C}{\lambda}\exp\{-\lambda(n-T)\} + |Y_T^{n,x_i'} - Y_T^{n,x}|. \end{aligned}$$

Moreover for fixed n, as $i \to \infty$, $Y_T^{n,x_i'} \to Y_T^{n,x}$ in $L^p(\Omega, \mathscr{F}_T, \mathbb{P}; \mathbb{R})$ for all $p > 1$, by Proposition 4.2 in [2] Thus $Y_T^{x_i'} \to Y_T^x$ in $L^p(\Omega, \mathscr{F}_T, \mathbb{P}; \mathbb{R})$.

Now we can notice that $(Y^x\big|_{[0,T]}, Z^x\big|_{[0,T]})$ is the unique solution of the following BSDE (with finite horizon):

$$Y_t^x = Y_T^x + \int_t^T F(X_s^x, Y_s^x, Z_s^x)ds - \int_t^T Z_s^x dW_s,$$

and the same holds for $(Y^{x_i'}\big|_{[0,T]}, Z^{x_i'}\big|_{[0,T]})$. By similar argument as in [2] we have

$$\mathbb{E}\left[\sup_{t\in[0,T]}|Y_t^x - Y_t^{x_i'}|^p\right]^{1\wedge 1/p} + \mathbb{E}\left[\left(\int_0^T |Z_s^x - Z_s^{x_i'}|ds\right)^{p/2}\right]^{1\wedge 1/p}$$
$$\leq C\mathbb{E}\left[\left|Y_T^x - Y_T^{x_i'}\right|^{p+1}\right]^{\frac{1}{p+1}} + \mathbb{E}\left[\left(\int_0^T \left|F(s,X_s^x,Y_s^x,Z_s^x) - F(s,X_s^{x_i'},Y_s^{x_i'},Z_s^{x_i'})\right| ds\right)^{p+1}\right]^{\frac{1}{p+1}}$$

and we can conclude that $(Y^{x_i'}\big|_{[0,T]}, Z^{x_i'}\big|_{[0,T]}) \longrightarrow (Y^x\big|_{[0,T]}, Z^x\big|_{[0,T]})$ in $L^p(\Omega; C(0,T;\mathbb{R})) \times L^p(\Omega; L^2(0,T;\Xi^*))$. □

We need to study the regularity of Y^x. More precisely, we would like to show that Y_0^x belongs to $\mathcal{G}^1(H,\mathbb{R})$.

We are now in position to prove the main result of this section.

Theorem 4.6. *Let A2 and A3 hold. The map $x \to Y_0^x$ belongs to $\mathcal{G}^1(H,\mathbb{R})$. Moreover $|Y_0^x| + |\nabla_x Y_0^x| \leq c$, for a suitable constant c.*

Proof. Fix $n \geq 1$, let us consider the solution $(Y^{n,x}, Z^{n,x})$ of (20). Then, see [2], Proposition 4.2, the map $x \to (Y^{n,x}(\cdot), Z^{n,x}(\cdot))$ is Gâteaux differentiable from H to $L^p(\Omega, C(0,T;\mathbb{R})) \times L^p(\Omega; L^2(0,T;\Xi^*))$, $\forall p \in (1,\infty)$. Denoting by $(\nabla_x Y^{n,x}h, \nabla_x Z^{n,x}h)$ the partial Gâteaux derivatives with respect to x in the direction $h \in H$, the processes $\{\nabla_x Y_t^{n,x}h, \nabla_x Z_t^{n,x}h, t \in [0,n]\}$ solves the equation, $\mathbb{P} - a.s.$,

$$\begin{aligned}\nabla_x Y_t^{n,x}h &= \int_t^n \nabla_x F(X_s^x, Y_s^{n,x}, Z_s^{n,x})\nabla_x X_s^{n,x}h\,ds \\ &\quad + \int_t^n \nabla_y F(X_s^x, Y_s^{n,x}, Z_s^{n,x})\nabla_x Y_s^{n,x}h\,ds \\ &\quad + \int_t^n \nabla_z F(X_s^x, Y_s^{n,x}, Z_s^{n,x})\nabla_x Z_s^{n,x}h\,ds - \int_t^n \nabla_x Z_s^{n,x}h\,dW_s.\end{aligned} \tag{21}$$

We note that we can write the generator of the previous equation as

$$\phi_s^n(u,v) = \psi_s^n + a_s^n u + < b_s^n, v >$$

where ψ and a are real processes defined respectively by

$$\psi_s^n = \nabla_x F(X_s^x, Y_s^{n,x}, Z_s^{n,x})\nabla_x X_s^{n,x}h, \quad a_s^n = \nabla_y F(X_s^x, Y_s^{n,x}, Z_s^{n,x})$$

and b^n is given by

$$b_s^n = \nabla_z F(X_s^x, Y_s^{n,x}, Z_s^{n,x}).$$

b^n belongs to the space $L(\Xi^*,\mathbb{R})$ and by Riesz isometry can be identified with an element of Ξ^*. By Assumption A3 and Lemma 4.3, we have that for all $x,h\in H$ the following holds $\mathbb{P}$-a.s. for all $n\in\mathbb{N}$ and all $s\in[0,n]$:

$$|\psi_s^n| = \left|\nabla_x F(X_s^x, Y_s^{n,x}, Z_s^{n,x})\nabla_x X_s^x h\right| \le C|h|,$$
$$a_s^n = \nabla_y F(X_s^x, Y_s^{n,x}, Z_s^{n,x}) \le -\lambda \le 0, \qquad |b_s^n| = \left|\nabla_z F(X_s^x, Y_s^{n,x}, Z_s^{n,x})\right| \le C(1+|Z_s^{n,x}|).$$

As mentionned before, $\int_0^\cdot Z_s^{n,x}dW_s$ is a BMO–martingale. Hence,

$$\mathscr{E}_t = \exp\left(\int_0^t b_s dW_s - \frac{1}{2}\int_0^t |b_s|^2 ds\right), \quad 0\le t\le n,$$

$\int_0^\cdot b_s^n dW_s$ is also a is a uniformly integrable martingale. By Remark 3.5, we obtain

$$\sup_{t\in[0,n]} |\nabla_x Y_t^{n,x}| \le \overline{C}|h|, \quad \mathbb{P}-\text{a.s.}$$

We recall that (see (14))

$$\sup_{n\ge1}\mathbb{E}(\int_0^\infty e^{-\epsilon s}|Z_s^{n,x}|^2 ds) < \infty. \tag{22}$$

Hence, applying Itô's formula to $e^{-2\lambda t}|\nabla_x Y_t^{n,x}h|^2$ and arguing as in the proof of Theorem 3.3, thanks to the (22), we get:

$$\mathbb{E}\int_0^\infty e^{-2\lambda t}(|\nabla_x Y_t^{n,x}h|^2 + |\nabla_x Z_t^{n,x}h|^2)dt \le C_1|h|^2.$$

Fix $x,h\in H$, there exists a subsequence of $\{(\nabla_x Y^{n,x}h, \nabla_x Z^{n,x}h, \nabla_x Y_0^{n,x}h) : n\in\mathbb{N}\}$ which we still denote by itself, such that $(\nabla_x Y^{n,x}h, \nabla_x Z^{n,x}h)$ converges weakly to $(U^1(x,h), V^1(x,h))$ in $\mathrm{M}^{2,\lambda}(\Xi^*)$ and $\nabla_x Y_0^{n,x}h$ converges to $\xi(x,h)\in\mathbb{R}$.

Now we write the equation (21) as follows: $t\in[0,n]$

$$\begin{aligned}\nabla_x Y_t^{n,x}h &= \nabla_x Y_0^{n,x}h - \int_0^t \nabla_x F(X_s^x, Y_s^{n,x}, Z_s^{n,x})\nabla_x X_s^x h ds \\ &\quad - \int_0^t \nabla_y F(X_s^x, Y_s^{n,x}, Z_s^{n,x})\nabla_x Y_s^{n,x}h ds \\ &\quad - \int_0^t \nabla_z F(X_s^x, Y_s^{n,x}, Z_s^{n,x})\nabla_x Z_s^{n,x}h ds + \int_0^t \nabla_x Z_s^{n,x}h dW_s\end{aligned} \tag{23}$$

and we define an other process $U_t^2(x,h)$ by

$$\begin{aligned}U_t^2(x,h) &= \xi(x,h) - \int_0^t \nabla_x F(X_s^x, Y_s^x, Z_s^x)\nabla_x X_s^x h ds \\ &\quad - \int_0^t \nabla_y F(X_s^x, Y_s^x, Z_s^x)U_s^1(x,h)ds \\ &\quad - \int_0^t \nabla_z F(X_s^x, Y_s^x, Z_s^x)V_s^1(x,h)ds + \int_0^t V_s^1(x,h)dW_s,\end{aligned} \tag{24}$$

where (Y^x, Z^x) is the unique bounded solution to the backward equation (19), see Proposition 4.5. Passing to the limit in the equation (23) it is easy to show that $\nabla_x Y_t^{n,x} h$ converges to $U_t^2(x,h)$ weakly in $L^1(\Omega)$ for all $t > 0$.

Thus $U_t^2(x,h) = U_t^1(x,h)$, $\mathbb{P}$-a.s. for a.e. $t \in \mathbb{R}^+$ and $|U_t^2(x,h)| \leq \overline{C}|h|$.

Now consider the following equation on infinite horizon

$$\begin{aligned} U_t(x,h) &= U_0(x,h) - \int_0^t \nabla_x F(X_s^x, Y_s^x, Z_s^x) \nabla_x X_s^x h ds \\ &\quad - \int_0^t \nabla_y F(X_s^x, Y_s^x, Z_s^x) U_s(x,h) ds \\ &\quad - \int_0^t \nabla_z F(X_s^x, Y_s^x, Z_s^x) V_s(x,h) ds + \int_0^t V_s(x,h) dW_s. \end{aligned} \tag{25}$$

We claim that this equation has a solution.

For each $n \in \mathbb{N}$ consider the finite horizon BSDE (with final condition equal to zero):

$$\begin{aligned} U_t^n(x,h) &= \int_t^n \nabla_x F(X_s^x, Y_s^x, Z_s^x) \nabla_x X_s^x h ds \\ &\quad + \int_t^n \nabla_y F(X_s^x, Y_s^x, Z_s^x) U_s^n(x,h) ds \\ &\quad + \int_t^n \nabla_z F(X_s^x, Y_s^x, Z_s^x) V_s^n(x,h) ds - \int_t^n V_s^n(x,h) dW_s, \end{aligned}$$

By the result in [2] we know that this equation has a unique solution $(U_n(\cdot,x,h), V_n(\cdot,x,h)) \in L^p(\Omega; C(0,n;\mathbb{R})) \times L^p(\Omega; L^2(0,n;\Xi^*))$. The generator of this equation can be rewritten as

$$\phi_t(u,v) = \psi_t + a_t u + b_t v$$

where $\psi_t = \nabla_x F(X_t^x, Y_t^x, Z_t^x) \nabla_x X_t^x$ and $|\psi_t| \leq C|h|$, $a_t = \nabla_y F(X^s, Y^s, Z^s) \leq -\lambda$, $b_t = \nabla_z F(X^s, Y^s, Z^s)$ and $|b_t| \leq C(1 + |Z_t^x|)$. On the interval $[0,n]$ the process $\int_0^{\cdot} Z_s^x dWs$ is a BMO-martingale. Hence, from Remark 3.5 it follows that $\mathbb{P}$-a.s. $\forall n \in \mathbb{N}$, $\forall t \in [0,n]$ $|U_t^n| \leq \frac{\overline{C}}{\lambda}|h|$ and as in the proof of existence in the Theorem 3.3, we can conclude that

1. for each $t \geq 0$ $U_t^n(x,h)$ is a Cauchy sequence in $L^\infty(\Omega)$ which converges to a process U and $\mathbb{P}$-a.s., $\forall t \in [0,n]$
$$|U_t^n(x,h) - U_t(x,h)| \leq \frac{C}{\lambda}|h| e^{-\lambda(n-t)}; \tag{26}$$
2. $V_\cdot^n(x,h)$ is a Cauchy sequence in $\mathrm{L}^2_{\mathrm{loc}}(\Xi^*)$;
3. The limit processes $(U_\cdot(x,h), V_\cdot(x,h))$ satisfy the BSDE (25).

Moreover still from Remark 3.5 we get that the solution is unique.

Coming back to equation (24), we have that $(U^2(x,h), V^1(x,h))$ is solution in $\mathbb{R}^+$ of the equation (25).

In particular we notice that $U_0(x,h) = \xi(x,h)$ is the limit of $\nabla_x Y_0^{n,x} h$ (along the chosen subsequence). The uniqueness of the solution to (25) implies that in reality $U_0(x,h) = \lim_{n\to\infty} \nabla_x Y_0^{n,x} h$ along the original sequence.

Now let $x_i \to x$.

$$\begin{aligned}|U_0(x,h) - U_0(x_i,h)| &\leq |U_0(x,h) - U_0^n(x,h)| + |U_0^n(x,h) - U_0^n(x_i,h)| + \\ +|U_0^n(x_i,h) - U_0(x_i,h)| &\leq \frac{2C}{\lambda} e^{-\lambda n}|h| + |U_0^n(x,h) - U_0^n(x_i,h)|,\end{aligned} \tag{27}$$

where we have used the inequality (26). We now notice that $\nabla_x F$, $\nabla_y F$, $\nabla_z F$ are, by assumptions, continuous and $|\nabla_x F| \leq C$, $|\nabla_y F| \leq C(1+|Z|)^{2\alpha}$, $|\nabla_z F| \leq C(1+|Z|)$. Moreover the following statements on continuous dependence on x hold:

maps $x \to X^x$, $x \to \nabla_x X^x h$ are continuous from $H \to L^p_{\mathscr{P}}(\Omega; C(0,T;H))$ (see [13] Proposition 3.3);

the map $x \to Y^x\big|_{[0,T]}$ is continuous from H to $L^p_{\mathscr{P}}(\Omega; C(0,T;\mathbb{R}))$ (see Proposition 4.5 here);

the map $x \to Z^x\big|_{[0,T]}$ is continuous from H to $L^p_{\mathscr{P}}(\Omega; L^2(0,T;\Xi))$ (see Proposition 4.5 here).

We can therefore apply to (26) the continuity result of [13] Proposition 4.3 to obtain in particular that $U_0^n(x_i,h) \to U_0^n(x,h)$ for all fixed n as $i \to \infty$. And by (27) we can conclude that $U_0(x_i,h) \to U_0(x,h)$ as $i \to \infty$.

Summarizing $U_0(x,h) = \lim_{n\to\infty} \nabla_x Y_0^{n,x} h$ exists, moreover it is clearly linear in h and verifies $|U_0(x,h)| \leq C|h|$, finally it is continuous in x for every h fixed.

Finally, for $t > 0$,

$$\begin{aligned}\lim_{t\searrow 0} \frac{1}{t}[Y_0^{x+th} - Y_0^x] &= \lim_{t\searrow 0} \frac{1}{t} \lim_{n\to+\infty} [Y_0^{n,x+th} - Y_0^{n,x}] = \lim_{t\searrow 0} \lim_{n\to+\infty} \int_0^1 \nabla_x Y_0^{n,x+\theta th} h d\theta \\ &= \lim_{t\searrow 0} \int_0^1 U_0(x+\theta th, h) d\theta = U_0(x,h)\end{aligned}$$

and the claim is proved. □

5 Mild Solution of the elliptic PDE

Now we can proceed as in [14]. Let us consider the forward equation

$$X_s^x = e^{sA}x + \int_0^s e^{(s-r)A} b(X_r^x) dr + \int_0^s e^{(s-r)A} \sigma dW_r, \; s \geq 0. \tag{28}$$

Assuming that Assumption A2 holds, we define in the usual way the transition semigroup $(P_t)_{t\geq 0}$, associated to the process X:

$$P_t[\phi](x) = \mathbb{E}\, \phi(X_t^x), \qquad x \in H,$$

for every bounded measurable function $\phi : H \to \mathbb{R}$. Formally, the generator $\mathscr{L}$ of (P_t) is the operator

$$\mathscr{L}\phi(x) = \frac{1}{2}\text{Trace}\left(\sigma\sigma^* \nabla^2 \phi(x)\right) + \langle Ax + b(x), \nabla \phi(x)\rangle.$$

In this section we address solvability of the non linear stationary Kolmogorov equation:

$$\mathscr{L} v(x) + F(x, v(x), \nabla v(x)\sigma) = 0, \qquad x \in H, \tag{29}$$

when the coefficient F verifies Assumption A3. Note that, for $x \in H$, $\nabla v(x)$ belongs to H^*, so that $\nabla v(x)\sigma$ is in Ξ^*.

Definition 5.1. We say that a function $v : H \to \mathbb{R}$ is a mild solution of the non linear stationary Kolmogorov equation (29) if the following conditions hold:

(i) $v \in \mathscr{G}^1(H, \mathbb{R})$ and $\exists\, C > 0$ such that $|v(x)| \le C$, $|\nabla_x v(x)h| \le C\,|h|$, for all $x, h \in H$;

(ii) the following equality holds, for every $x \in H$ and $T \ge 0$:

$$v(x) = e^{-\lambda T}\, P_T[v](x) + \int_0^T e^{-\lambda t}\, P_t\Big[F\Big(\cdot, v(\cdot), \nabla v(\cdot)\sigma\Big) + \lambda v(\cdot)\Big](x)\, dt. \tag{30}$$

where λ is the monotonicity constant in Assumption A3.

Together with equation (28) we also consider the backward equation

$$Y_t^x = Y_T^x + \int_t^T F(X_s^x, Y_s^x, Z_s^x) ds - \int_t^T Z_s^x dW_s, \qquad 0 \le t \le T < \infty, \tag{31}$$

where $F : H \times \mathbb{R} \times \Xi^* \to \mathbb{R}$ is the same occurring in the nonlinear stationary Kolmogorov equation. Under the Assumptions A2, A3, Propositions 4.2-4.5 give a unique solution $\{X_t^x, Y_t^x, Z_t^x\}$, for $t \ge 0$, of the forward-backward system (28)-(31). We can now state the following

Theorem 5.2. *Let Assumption A2 and hold.*
Then equation (29) has a unique mild solution given by the formula

$$v(x) = Y_0^x.$$

where $\{X_t^x, Y_t^x, Z_t^x, t \ge 0\}$ is the solution of the forward-backward system (28)-(31). Moreover the following holds:

$$Y_t^x = v(X_t^x), \quad Z_t^x = \nabla v(X_t^x)\sigma.$$

Proof. Let us recall that for $s \ge 0$, Y_s^x is measurable with respect to $\mathscr{F}_{[0,s]}$ and $\mathscr{F}_s$; it follows that Y_0^x is deterministic (see also [8]). Moreover, as a byproduct of Proposition 4.6, the function v defined by the formula $v(x) = Y_0^x$ has the regularity properties stated in Definition 5.1. The proof of the equality (30) and of the uniqueness of the solution is identical to the proof of Theorem 6.1 in [14].

6 Application to optimal control

We wish to apply the above results to perform the synthesis of the optimal control for a general nonlinear control system on an infinite time horizon. To be able to use non-smooth feedbacks we settle the problem in the framework of weak control problems. Again we follow [14] with slight modifications. We report the argument for reader's convenience.

As above by H, Ξ we denote separable real Hilbert spaces and by U we denote a Banach space.

For fixed $x_0 \in H$ an *admissible control system* (a.c.s) is given by $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \ge 0}, \mathbb{P}, \{W_t, t \ge 0\}, u)$ where

- $(\Omega, \mathscr{F}, \mathbb{P})$ is a complete probability space and $(\mathscr{F}_t)_{t\geq 0}$ is a filtration on it satisfying the usual conditions.

- $\{W_t : t \geq 0\}$ is a Ξ-valued cylindrical Wiener process relatively to the filtration $(\mathscr{F}_t)_{t\geq 0}$ and the probability $\mathbb{P}$.

- $u : \Omega \times [0, \infty[\to U$ is a predictable process (relatively to $(\mathscr{F}_t)_{t\geq 0}$) that satisfies the constraint: $u_t \in \mathscr{U}$, $\mathbb{P}$-a.s. for a.e. $t \geq 0$, where $\mathscr{U}$ is a fixed closed subset of U.

To each a.c.s. we associate the mild solution $X \in L^p_{\mathscr{P}}(\Omega; C(0, T; H))$ (for arbitrary $T > 0$ and arbitrary $p \geq 1$) of the state equation:

$$\begin{cases} dX^x_s = \left(AX^x_s + b(X^x_s) + \sigma r(X^x_s, u_s)\right)\, ds + \sigma\, dW_s, \qquad s \geq 0, \\ X_0 = x \in H, \end{cases} \tag{32}$$

and the cost:

$$J(x, u) = \mathbb{E} \int_0^{+\infty} e^{-\lambda t} g(X^x_t, u_t)\, dt, \tag{33}$$

where $g : H \times U \to \mathbb{R}$. Our purpose is to minimize the functional J over all a.c.s. Notice the occurrence of the operator σ in the control term: this special structure of the state equation is imposed by our techniques.

We work under the following assumptions.

Assumption A4. 1. The process W is a Wiener process in Ξ, defined on a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with respect to a filtration $(\mathscr{F}_t)$ satisfying the usual conditions.

2. A, b and σ verify Assumption A2.

3. The set $\mathscr{U}$ is a nonempty closed subset of U .

4. The functions $r : H \times U \to \Xi$, $g : H \times U \to \mathbb{R}$ are Borel measurable and for all $x \in H$, $r(x, \cdot)$ and $g(x, \cdot)$ are continuous functions from U to Ξ and from U to $\mathbb{R}$, respectively.

5. There exists a constant $C \geq 0$ such that for every $x, x' \in H$, $u \in K$ it holds that

$$|r(x, u) - r(x', u)| \leq C(1 + |u|)|x - x'|,$$

$$|r(x, u)| \leq C(1 + |u|), \tag{34}$$

$$0 \leq g(x, u) \leq C(1 + |u|^2), \tag{35}$$

6. There exist $R > 0$ and $c > 0$ such that for every $x \in H$ $u \in U$ satisfying $|u| \geq R$,

$$g(x, u) \geq c|u|^2. \tag{36}$$

We will say that an $(\mathscr{F}_t)$-adapted stochastic process $\{u_t, t \geq 0\}$ with values in U is an admissible control if it satisfies

$$\mathbb{E} \int_0^{\infty} e^{-\lambda t} |u_t|^2 dt < \infty. \tag{37}$$

This square summability requirement is justified by (36): a control process which is not square summable would have infinite cost. Moreover it follows by (35) that the cost functional is well defined and $J(x, u) < \infty$ for all $x \in H$ and all a.c.s.

Remark 6.1. We set the optimal control problem for a general nonlinear control system on an infinite time horizon in such way to use the previous results on forward-backward system and elliptic partial differential equations. In particular the assumptions A4.4 and A4.5 are needed to have the Hamiltonian corresponding to the control problem with quadratic growth in the gradient and consequently the associated BSDEs with quadratic growth in the z variable.

Now we state that for every admissible control the solution to (32) exists.

Proposition 6.2. *Let u be an admissible control. Then there exists a unique, continuous, $(\mathscr{F}_t)$-adapted process X satisfying $\mathbb{E}\sup_{t\in[0,T]}|X_t|^2 < \infty$, and $\mathbb{P}$-a.s., $t \in [0,T]$*

$$X_t^x = e^{tA}x + \int_0^t e^{(t-s)A} b(X_s^x)ds + \int_0^t e^{(t-s)A}\sigma dW_s + \int_0^t e^{(t-s)A}\sigma r(X_s^x, u_s)ds.$$

Proof. The proof is an immediate extension to the infinite dimensional case of the Proposition 2.3 in [12]. □

By the previous Proposition and the arbitrariness of T in its statement, the solution is defined for every $t \geq 0$. We define in a classical way the Hamiltonian function relative to the above problem: for all $x \in H$, $y \in \mathbb{R}$, $z \in \Xi^*$,

$$F(x,y,z) = \inf\{g(x,u) + zr(x,u) : u \in \mathscr{U}\} - \lambda y \tag{38}$$

$$\Gamma(x,y,z) = \{u \in U : g(x,u) + zr(x,u) - \lambda y = F(x,y,z)\}. \tag{39}$$

$\Gamma(x,y,z)$ is the set of minimizers in (38).

The proof of the following Lemma can be found in [12] Lemma 3.1.

Lemma 6.3. *The map F is a Borel measurable function from $H \times \Xi^*$ to $\mathbb{R}$. There exists a constant $C > 0$ such that*

$$-C(1+|z|^2) - \lambda y \leq F(x,y,z) \leq g(x,u) + C|z|(1+|u|) - \lambda y \quad \forall u \in \mathscr{U}. \tag{40}$$

We require moreover that

Assumption A5. F satisfies assumption A3 2-3-4.

Example 6.4. Let consider the following situation: $H = \Xi = U = \mathscr{U} = L^2(0,1)$, and $r(x,u) = u$, $g(x,u) = q(x) + |u|^2/2$, where $q : H \to \mathbb{R}$ such that $0 \leq q(x) \leq C$, $q \in \mathscr{G}^1(H,\mathbb{R})$ with $|\nabla_x q(x)h| \leq C|h|$. Then we have

$$F(x,y,z) = q(x) - \frac{|z|^2}{2} - \lambda y.$$

We note that F is Fréchet differentiable with respect to z and $\Gamma(x,y,z) = -z$ turns out to be a continuous function of z only.

By Theorem 5.2, the stationary Hamilton-Jacobi-Bellman equation relative to the above stated problem, namely:

$$\mathscr{L}v(x) + F(x, v(x), \nabla v(x)\sigma) = 0, \qquad x \in H, \tag{41}$$

admits a unique mild solution, in the sense of Definition 5.1.

The fundamental relation

Proposition 6.5. *Let v be the solution of (41). For every admissible control u and for the corresponding trajectory X starting at x we have*

$$J(x,u) = v(x) + \mathbb{E}\int_0^\infty e^{-\lambda t}\Big(- F(X_t^x, \nabla v(X_t^x)\sigma) - \lambda v(X_t^x) + \nabla_x v(X_t^x)\sigma r(X_t^x, u_t) + g(X_t^x, u_t)\Big)\, dt.$$

Proof. We introduce the sequence of stopping times

$$\tau_n = \inf\{t \in [0,T] : \int_0^t |u_s|^2 ds \geq n\},$$

with the convention that $\tau_n = T$ if the indicated set is empty. By (37), for $\mathbb{P}$-almost every $\omega \in \Omega$, there exists an integer $N(\omega)$ depending on ω such that

$$n \geq N(\omega) \Longrightarrow \tau_n(\omega) = T. \tag{42}$$

Let us fix $u_0 \in K$, and for every n, let us define

$$u_t^n = u_t 1_{t \leq \tau_n} + u_0 1_{t > \tau_n}$$

and consider the equation

$$\begin{cases} dX_t^{n,x} = b(X_t^{n,x})dt + \sigma[dW_t + r(X_t^{n,x}, u_t^n)dt], \quad 0 \leq t \leq T \\ X_0^{n,x} = x. \end{cases} \tag{43}$$

Let us define

$$W_t^n = W_t + \int_0^t r(X_s^{n,x}, u_s^n) ds, \quad 0 \leq t \leq T.$$

From the definition of τ_n and from (34), it follows that

$$\int_0^T |r(X_s^{n,x}, u_s^n)|^2 ds \leq C \int_0^T (1 + |u_s^n|)^2 ds \leq C \int_0^{\tau_n} (1 + |u_s|)^2 ds + C \leq C + Cn. \tag{44}$$

Therefore defining

$$\rho_n = \exp\left(\int_0^T -r(X_s^{n,x}, u_s^n) dW_s - \frac{1}{2}\int_0^T |r(X_s^{n,x}, u_s^n)|^2 ds\right)$$

the Novikov condition implies that $\mathbb{E}\rho_n = 1$. Setting $d\mathbb{P}_T^n = \rho_n d\mathbb{P}_{|\mathscr{F}_T}$, by the Girsanov theorem W^n is a Wiener process under $\mathbb{P}_T^n$. Relatively to W^n the equation (43) can be written:

$$\begin{cases} dX_t^{n,x} = b(X_t^{n,x})dt + \sigma dW_t^n, \quad 0 \leq t \leq T \\ X_0^{n,x} = x. \end{cases} \tag{45}$$

Consider now the following finite horizon Markovian forward-backward system (with respect to probability $\mathbb{P}^n_T$ and to the filtration generated by $\{W^n_r : r \in [0,T]\}$):

$$\begin{cases} X^{n,x}_r = e^{rA}x + \int_0^r e^{(r-s)A} b(X^{n,x}_s)\, ds + \int_0^r e^{(r-s)A}\sigma\, dW^n_s, \quad r \geq 0, \\ Y^{n,x}_r - v(X^{n,x}_T) + \int_r^T Z^{n,x}_s dW^n_s = \int_r^T F(X^{n,x}_s, Y^{n,x}_s, Z^{n,x}_s) ds,\ 0 \leq r \leq T, \end{cases} \tag{46}$$

and let $(X^{n,x}, Y^{n,x}, Z^{n,x})$ be its unique solution. $(X^{n,x}, Y^{n,x}, Z^{n,x})$ is predictable with respect to the filtration generated by $\{W^n_r : r \in [0,T]\}$, $X^{n,x}$ is continuous and $\mathbb{E}^n_T \sup_{t\in[0,T]} |X^{n,x}_t|^2 < +\infty$, $Y^{n,x}$ is bounded and continuous, and $\mathbb{E}^n_T \int_0^T |Z^{n,x}_t|^2 dt < +\infty$. Moreover, Theorem 5.2 and uniqueness of the solution of system (46), yields that

$$Y^{n,x}_t = v(X^{n,x}_t), \qquad Z^{n,x}_t = \nabla v(X^{n,x}_t)\sigma. \tag{47}$$

Applying the Itô formula to the process $e^{-\lambda t} Y^{n,x}_t$, and restoring the original noise W we get

$$\begin{aligned} e^{-\lambda\tau_n} Y^{n,x}_{\tau_n} = Y^{n,x}_0 + \int_0^{\tau_n} -e^{-\lambda t} F(X^{n,x}_t, Y^{n,x}_t, Z^{n,x}_t) dt + \\ + \int_0^{\tau_n} -\lambda e^{-\lambda t} Y^{n,x}_t dt + \int_0^{\tau_n} e^{-\lambda t} Z^{n,x}_t dW_t + \int_0^{\tau_n} e^{-\lambda t} Z^{n,x}_t r(X^{n,x}_t, u^n_t) dt. \end{aligned} \tag{48}$$

We note that for every $p \in [1,\infty)$ we have

$$\begin{aligned} \rho_n^{-p} = \exp\left(p \int_0^T r(X^{n,x}_s, u^n_s) dW^n_s - \frac{p^2}{2} \int_0^T |r(X^{n,x}_s, u^n_s)|^2 ds \right) \\ \cdot \exp\left(\frac{p^2 - p}{2} \int_0^T |r(X^{n,x}_s, u^n_s)|^2 ds \right). \end{aligned} \tag{49}$$

By (44) the second exponential is bounded by a constant depending on n and p, while the first one has $\mathbb{P}^n$-expectation, equal to 1. So we conclude that $\mathbb{E}^n \rho_n^{-p} < \infty$. It follows that

$$\begin{aligned} \mathbb{E}\left(\int_0^T e^{-2\lambda t} |Z^{n,x}_t)|^2 dt \right)^{1/2} \leq \mathbb{E}^n \left(\int_0^T \rho_n^{-2} |Z^{n,x}_t|^2 dt \right)^{1/2} \leq \\ \leq (\mathbb{E}^n \rho_n^{-2})^{1/2} \mathbb{E}^n \left(\int_0^T |Z^{n,x}_t|^2 dt \right)^{1/2} < \infty, \end{aligned}$$

and the stochastic integral in (48) has zero expectation. By identification in (47) we have $Y^n_0 = v(x)$ and, for $t \leq \tau_n$, we also have $u^n_t = u_t$, $X^{n,x}_t = X^x_t$, $Y^{n,x}_t = v(X^{n,x}_t) = v(X^x_t)$ and $Z^{n,x}_t = \nabla_x v(X^x_t)$. Thus, taking the expectation in (48), we obtain

$$\begin{aligned} \mathbb{E}[e^{-\lambda\tau_n} Y^{n,x}_{\tau_n}] = v(x) + \\ + \mathbb{E} \int_0^{\tau_n} e^{-\lambda t} \Big(-F(X^x_t, v(X^x_t), \nabla_x v(X^x_t)\sigma) - \lambda v(X^x_t) + \nabla_x v(X^x_t) \sigma r(X^x_t, u_t) \Big) dt \end{aligned} \tag{50}$$

and, adding to the both sides $\mathbb{E}\int_0^{\tau_n} e^{-\lambda t} g(X_t^x, u_t)dt$,

$$\mathbb{E}\int_0^{\tau_n} e^{-\lambda t} g(X_t^x, u_t)dt + \mathbb{E}[e^{-\lambda\tau_n} Y_{\tau_n}^{n,x}] = v(x)+$$
$$+\mathbb{E}\int_0^{\tau_n} e^{-\lambda t}\Big(-F(X_t^x, v(X_t^x), \nabla_x v(X_t^x)\sigma) - \lambda v(X_t^x) + \nabla_x v(X_t^x)\sigma r(X_t^x, u_t) + g(X_t^x, u_t)\Big)dt. \quad (51)$$

Now we let $n \to \infty$. For $n \geq N(\omega)$ we have $\tau_n(\omega) = T$ and $e^{-\lambda\tau_n} Y_{\tau_n}^{n,x} = e^{-\lambda T} Y_T^{n,x} = e^{-\lambda T} v(X_T^{n,x}) = e^{-\lambda T} v(X_T^x)$. Since $Y^{n,x}$ is bounded, by the dominated convergence theorem it follows that

$$\lim_{n\to\infty} \mathbb{E} e^{-\lambda\tau_n} Y_{\tau_n}^{n,x} = e^{-\lambda T}\mathbb{E} v(X_T^x). \quad (52)$$

Moreover, by definition of F, $-F(x,y,z) - \lambda y + zr(x,u) + g(x,u) \geq 0$ and by (35) $g(x,u) \geq 0$. Hence, thanks to (52) and the monotone convergence theorem, we obtain for $n \to \infty$,

$$\mathbb{E}\int_0^{T} e^{-\lambda t} g(X_t^x, u_t)dt + e^{-\lambda T}\mathbb{E} v(X_T^x) = v(x)+$$
$$+\mathbb{E}\int_0^{T} e^{-\lambda t}\Big(-F(X_t^x, \nabla_x v(X_t^x)\sigma) - \lambda v(X_t^x) + \nabla_x v(X_t^x)\sigma r(X_t^x, u_t) + g(X_t^x, u_t)\Big)dt. \quad (53)$$

Recalling that v is bounded, letting $T \to \infty$, we have that $e^{-\lambda T}\mathbb{E} v(X_T^x) \to 0$ and thanks to the monotone convergence theorem we conclude that

$$J(x,u) = v(x)+$$
$$\mathbb{E}\int_0^{\infty} e^{-\lambda t}\left[-F(X_t^x, v(X_t^x), \nabla v(X_t^x)\sigma) - \lambda v(X_t^x) + \nabla_x v(X_t^x)\sigma r(X_t^x, u_t) + g(X_t^x, u_t)\right]dt.$$

□

The above equality is known as the *fundamental relation* and immediately implies the following

Corollary 6.6. *For every admissible control u and any initial datum x, we have $J(x,u) \geq v(x)$ and that the equality holds if and only if the following feedback law holds $\mathbb{P}$-a.s. for almost every $t \geq 0$:*

$$F(X_t^x, v(X_t^x), \nabla_x v(X_t^x)\sigma) = \nabla_x v(X_t^x)\sigma + g(X_t^x, u_t) - \lambda v(X_t^x)$$

where X is the trajectory starting at x and corresponding to control u.

Existence of optimal controls: the closed loop equation.

Next we address the problem of finding a weak solution to the so-called closed loop equation (see equation (59) below). We recall the definitions of the Hamiltonian function and of the set Γ: for all $x \in H$, $y \in \mathbb{R}$, $z \in \Xi^*$

$$F(x,y,z) = \inf\{g(x,u) + zr(x,u) : u \in \mathscr{U}\} - \lambda y, \quad (54)$$

$$\Gamma(x,y,z) = \{u \in U : g(x,u) + zr(x,u) - \lambda y = F(x,y,z)\}. \quad (55)$$

We have to require the following

Assumption A6. $\Gamma(x, y, z)$ is non empty for all $x \in H$ and $z \in \Xi^*$.

By simple calculation (see [12] Lemma 3.1), we can prove that, if Assumption A6 holds, then

$$F(x, y, z) = \min_{u \in \mathscr{U}, |u| \leq C(1+|z|)} [g(x, u) + zr(x, u)] - \lambda y, \quad x \in H,\ y \in \mathbb{R},\ z \in \Xi^*,$$

that is the infimum in (54) is attained in a ball of radius $C(1 + |z|)$, and

$$F(x, y, z) < g(x, u) + zr(x, u) - \lambda y \quad \text{ if } |u| > C(1 + |z|). \tag{56}$$

Moreover, by the Filippov Theorem (see, e.g., [1, Thm. 8.2.10, p. 316]) there exists a measurable selection of Γ, a Borel measurable function $\gamma : H \times \Xi^* \to \mathscr{U}$ such that

$$F(x, y, z) = g(x, \gamma(x, z)) + zr(x, \gamma(x, z)) - \lambda y, \quad x \in H,\ y \in \mathbb{R},\ z \in \Xi^*. \tag{57}$$

By (56), we have

$$|\gamma(x, z)| \leq C(1 + |z|). \tag{58}$$

The closed loop equation is

$$\begin{cases} dX_t^x = AX_t^x dt + b(X_t^x)dt + \sigma[r(X_t^x, \underline{u}(X_t^x))dt + dW_t] \quad t \geq 0 \\ X_0 = x \end{cases} \tag{59}$$

where $\underline{u}$ is defined by

$$\underline{u}(x) = \gamma(x, \nabla_x v(X_t^x)\sigma) \quad \mathbb{P}\text{-}a.s.\ for\ a.e\ t \geq 0.$$

By a weak solution we mean a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with a filtration $(\mathscr{F}_t)$ satisfying the usual conditions, a Wiener process W in Ξ with respect to $\mathbb{P}$ and $(\mathscr{F}_t)$, and a continuous $(\mathscr{F}_t)$-adapted process X with values in H satisfying, $\mathbb{P}$-a.s.,

$$\mathbb{E}\int_0^\infty e^{-\lambda t}|\underline{u}(X_t^x)|^2 dt < \infty$$

and such that (59) holds. In other word we mean an admissible control system for which the closed-loop equation has solution.

Proposition 6.7. *Assume that b, σ, g satisfy Assumption A4, F verifies Assumption A5 and Assumption A6 holds. Then there exists a weak solution of the closed loop equation, satisfying in addition*

$$\mathbb{E}\int_0^\infty e^{-\lambda t}|\underline{u}(X_t^x)|^2 dt < \infty. \tag{60}$$

Proof. We start by constructing a canonical version of a cylindrical Wiener process in Ξ. An explicit construction is needed to clarify the application of an infinite-dimensional version of the Girsanov theorem that we use below. We choose a larger Hilbert space $\Xi' \supset \Xi$ in such a way that Ξ is continuously and densely embedded in Ξ' with Hilbert-Schmidt inclusion operator $\mathscr{J}$. By Ω we denote the space $C([0, \infty[, \Xi')$ of continuous functions $\omega : [0, \infty[\to \Xi'$ endowed with the usual locally convex topology that makes Ω a Polish space, and by $\mathscr{B}$ its Borel σ-field. Since $\mathscr{J}\mathscr{J}^*$ has finite trace on Ξ', it is well known that there exists a probability $\mathbb{P}^\circ$ on $\mathscr{B}$ such that the canonical processes $W_t'(\omega) := \omega(t)$, $t \geq 0$, is a Wiener process with continuous paths in Ξ' satisfying

$\mathbb{E}[\langle W'_t, \xi' \rangle_{\Xi'} \langle W'_s, \eta' \rangle_{\Xi'}] = \langle \mathscr{J}\mathscr{J}^* \xi', \eta' \rangle_{\Xi'} (t \wedge s)$ for all $\xi', \eta' \in \Xi'$, $t, s \geq 0$. This is called a $\mathscr{J}\mathscr{J}^*$-Wiener processes in Ξ' in [10], to which we refer the reader for preliminary material on Wiener processes on Hilbert spaces. Let us denote by $\mathscr{G}^\circ$ the $\mathbb{P}^\circ$-completion of $\mathscr{B}$ and by $\mathscr{N}$ the family of sets $A \in \mathscr{G}^\circ$ with $\mathbb{P}^\circ(A) = 0$. Let $\mathscr{B}_t = \sigma\{W'_s : s \in [0, t]\}$ and $\mathscr{F}^\circ_t = \sigma(\mathscr{B}_t, \mathscr{N})$, $t \geq 0$, where as usual $\sigma(\cdot)$ denotes the σ-algebra in Ω generated by the indicated collection of sets or random variables. Thus $(\mathscr{F}^\circ_t)_{t \geq 0}$ is the Brownian filtration of W'.

The Ξ-valued cylindrical Wiener process $\{W^\xi_t : t \geq 0, \xi \in \Xi\}$ can now be defined as follows. For ξ in the image of $\mathscr{J}^*\mathscr{J}$ we take η such that $\xi = \mathscr{J}^*\mathscr{J}\eta$ and define $W^\xi_s = \langle W'_s, \mathscr{J}\eta \rangle_{\Xi'}$. Then we notice that $\mathbb{E}|W^\xi_t|^2 = t|\mathscr{J}\eta|^2_{\Xi'} = t|\xi|^2_\Xi$, which shows that the mapping $\xi \to W^\xi_s$, defined for $\xi \in \mathscr{J}^*\mathscr{J}(\Xi) \subset \Xi$ with values in $L^2(\Omega, \mathscr{F}, \mathbb{P}^\circ)$, is an isometry for the norms of Ξ and $L^2(\Omega, \mathscr{F}, \mathbb{P}^\circ)$. Consequently, noting that $\mathscr{J}^*\mathscr{J}(\Xi)$ is dense in Ξ, it extends to an isometry $\xi \to L^2(\omega, \mathscr{F}, \mathbb{P}^\circ)$, still denoted $\xi \to W^\xi_s$. An appropriate modification of $\{W^\xi_t : t \geq 0, \xi \in \Xi\}$ gives the required cylindrical Wiener process, which we denote by W°. We note that the Brownian filtration of W° coincides with $(\mathscr{F}^\circ_t)_{t \geq 0}$.

Now let $X \in L^p_{\mathrm{loc}}(\Omega, C(0, +\infty; H))$ be the mild solution of

$$\begin{cases} dX^x_s = AX^x_s \, d\tau + b(X^x_s) \, ds + \sigma \, dW^\circ_s \quad s \geq 0 \\ X_0 = x \end{cases} \tag{61}$$

If together with the previous forward equation we consider the backward equation

$$Y^x_t - Y^x_T + \int_t^T Z^x_s dW^\circ_s = \int_t^T F(X^x_s, Y^x_s, Z^x_s) ds, \qquad 0 \leq t \leq T < \infty, \tag{62}$$

we know that there exists a unique solution $\{X^x_t, Y^x_t, Z^x_t, t \geq 0\}$ to the forward-backward system (61)-(62) and by Proposition 5.2, the function

$$v(x) = Y^x_0.$$

is the solution to the nonlinear Kolmogorov equation:

$$\mathscr{L} v(x) + F(x, v(x), \nabla v(x) \sigma) = 0, \qquad x \in H. \tag{63}$$

Moreover the following holds:

$$Y^x_t = v(X^x_t), \qquad Z^x_t = \nabla v(X^x_t) \sigma. \tag{64}$$

We have

$$\mathbb{E}^\circ \int_0^\infty e^{-(\lambda + \epsilon)t} |Z^x_t|^2 dt < \infty. \tag{65}$$

and hence, for each $T > 0$,

$$\mathbb{E}^\circ \int_0^T |Z^x_t|^2 dt < \infty. \tag{66}$$

By (34) we have

$$|r(X^x_t, \underline{u}(X^x_t))| \leq C(1 + |\underline{u}(X^x_t)|), \tag{67}$$

and by (58),

$$|\underline{u}(X_t^x)| = |\gamma(X_t^x, \nabla v(X_t^x)\sigma)| \leq C(1 + |\nabla v(X_t^x)\sigma|) = C(1 + |Z_t^x|). \tag{68}$$

Let us define, for each $T > 0$,

$$M_T = \exp\left(\int_0^T \langle r(X_s^x, \underline{u}(X_s^x)), dW_s^\circ\rangle_\Xi - \frac{1}{2}\int_0^T |r(X_s^x, \underline{u}(X_s^x)|_\Xi^2\, ds\right). \tag{69}$$

Now, arguing exactly as in the proof of Proposition 5.2 in [12], we can prove that $\mathbb{E}^\circ M_T = 1$, and M is a $\mathbb{P}^\circ$-martingale. Hence there exists a probability $\mathbb{P}_T$ on $\mathscr{F}_T^\circ$ admitting M_T as a density with respect to $\mathbb{P}^\circ$, and by the Girsanov Theorem we can conclude that the process $\{W_t, t \in [0,T]\}$ given by $W_t = W_t^\circ - \int_0^t r(X_s^x, \underline{u}(X_s^x))ds$ is a Wiener process with respect to $\mathbb{P}_T$ and $(\mathscr{F}_t^\circ)_{t\geq 0}$. Since $\Xi^{'}$ is a Polish space and $\mathbb{P}_{T+h}$ coincide with $\mathbb{P}_T$ on $\mathscr{B}_T$, $T, h \geq 0$, by known results (see [24], Chapter VIII, §1, Proposition (1.13)) there exists a probability $\mathbb{P}$ on $\mathscr{B}$ such that the restrictions on $\mathscr{B}_T$ of $\mathbb{P}_T$ and that of $\mathbb{P}$ coincide, $T \geq 0$. Let $\mathscr{G}$ be the $\mathbb{P}$-completion of $\mathscr{B}$ and $\mathscr{F}_T$ be the $\mathbb{P}$-completion of $\mathscr{B}_T$. Moreover, since for all $T > 0$, $\{W_t : t \in [0,T]\}$ is a Ξ-valued cylindrical Wiener process under $\mathbb{P}_T$ and the restriction of $\mathbb{P}_T$ and of $\mathbb{P}$ coincide on $\mathscr{B}_T$ modifying $\{W_t : t \geq 0\}$ in a suitable way on a $\mathbb{P}$-null probability set we can conclude that $(\Omega, \mathscr{G}, \{\mathscr{F}_t : t \geq 0\}, \mathbb{P}, \{W_t : t \geq 0\}, \gamma(X^x, \nabla v(X^x)\sigma))$ is an admissible control system. The above construction immediately ensures that, if we choose such an admissible control system, then (59) is satisfied. Indeed if we rewrite (61) in terms of $\{W_t : t \geq 0\}$ we get

$$\begin{cases} dX_s^x = AX_s^x + b(X_s^x)\, d\tau + \sigma\, [r(X_s^x, \underline{u}(X_s^x))d\tau + dW_s] \\ X_0 = x. \end{cases}$$

It remains to prove (60). Let us introduce, for each integer n, the following stopping time

$$\sigma_n = \inf\left\{t \geq 0 : \int_0^t e^{-2\lambda s}|Z_s^x|^2 ds \geq n\right\},$$

with the convention that $\sigma_n = \infty$ if the indicated set is empty. Of course $\sigma_n \leq \sigma_{n+1}$ and by (65), $\sup_{n\geq 1}\sigma_n = \infty$ $\mathbb{P}^\circ$–a.s. Let us prove that $\sup_{n\geq 1}\sigma_n = +\infty$ $\mathbb{P}$–a.s. For each $T > 0$, since M_T is $\mathbb{P}^\circ$–integrable, the bounded dominated convergence theorem gives

$$\mathbb{P}\left(\sup_{n\geq 1}\sigma_n \leq T\right) = \lim_{n\to\infty}\mathbb{P}(\sigma_n \leq T) = \lim_{n\to+\infty}\mathbb{E}^\circ\left[\mathbf{1}_{\sigma_n \leq T}\, M_T\right] = 0.$$

Hence, $\sup_{n\geq 1}\sigma_n = \infty$ $\mathbb{P}$–a.s.

Let us fix $T > 0$ and $n \geq 1$; we set $\tau = \sigma_n \wedge T$. Applying Itô's formula to $e^{-\lambda t}Y_t^x$, we get

$$Y_0^x = e^{-\lambda\tau}Y_\tau^x + \int_0^\tau e^{-\lambda s}\left[F(X_s^x, Y_s^x, Z_s^x) + \lambda Y_s^x\right] ds - \int_0^\tau e^{-\lambda s} Z_s^x\, dW_s^\circ,$$

and coming back to the definition of W,

$$Y_0^x = e^{-\lambda\tau}Y_\tau^x + \int_0^\tau e^{-\lambda s}\left[F(X_s^x, Y_s^x, Z_s^x) + \lambda Y_s^x - Z_s r\left(X_s^x, \underline{u}(X_s^x)\right)\right] ds - \int_0^\tau e^{-\lambda s} Z_s^x\, dW_s.$$

By definition of $\underline{u}$, we have

$$F(X_s^x, Y_s^x, Z_s^x) + \lambda Y_s^x - Z_s r\left(X_s^x, \underline{u}(X_s^x)\right) = g\left(X_s^x, \underline{u}(X_s^x)\right),$$

and hence

$$Y_0^x = e^{-\lambda\tau}Y_\tau^x + \int_0^\tau e^{-\lambda s} g\left(X_s^x, \underline{u}(X_s^x)\right) - \int_0^\tau e^{-\lambda s} Z_s^x \, dW_s.$$

Taking the expectation with respect to $\mathbb{P}$ (actually with respect to $\mathbb{P}_T$), we get, since Y^x is a bounded process,

$$\mathbb{E}\left[\int_0^\tau e^{-\lambda s} g\left(X_s^x, \underline{u}(X_s^x)\right)\right] = Y_0^x - \mathbb{E}\left[e^{-\lambda\tau}Y_\tau^x\right] \le C,$$

where C is independent of n and T. Taking into account (36), we finally prove (60) by sending n and then T to infinity. □

Corollary 6.8. *By Corollary 6.6 it immediately follows that if X is the solution to (59) and we set $\widetilde{u}_s = \underline{u}(X_s^x)$, then $J(x, \widetilde{u}) = v(x)$, and consequently X^x is an optimal state, $\widetilde{u}_s$ is an optimal control, and $\underline{u}$ is an optimal feedback.*

Example 6.9. Finally we briefly show that our results can be applied to perform the synthesis of optimal controls for infinite horizon costs when the state equation is a general semilinear heat equation with additive noise. Namely, for $t \ge 0$, $\xi \in [0,1]$

$$\begin{cases} \frac{\partial}{\partial t}X(t,\xi) = \frac{\partial^2}{\partial \xi^2}X(t,\xi) + \widetilde{b}(\xi, X(t,\xi)) + \widetilde{\sigma}(\xi)\widetilde{r}(X(t,\xi), u(t,\xi)) + \widetilde{\sigma}(\xi)\frac{\partial}{\partial t}\mathscr{W}(t,\xi) \\ X(t,0) = X(t,1) = 0, \\ X(0,\xi) = x_0(\xi) \end{cases} \tag{70}$$

where $\mathscr{W}$ is a space-time white-noise on $\mathbb{R}_+ \times [0,1]$. Moreover we introduce the cost functional:

$$J(x_0, u) = \mathbb{E}\int_0^\infty \int_0^1 e^{-\lambda t}[l(\xi, X(t,\xi)) + u^2(t,\xi)]\, d\xi \, dt,$$

that we minimize over all adapted controls u such that $\mathbb{E}\int_0^\infty \int_0^1 e^{-\lambda t}|u(t,\xi)|^2 d\xi \, dt < \infty$. To fit the assumptions of our abstract results we will suppose that the functions $\widetilde{b}$, $\widetilde{\sigma}$ and $\widetilde{r}$ are all measurable and real-valued and moreover:

- $\widetilde{b}$ is defined on $[0,1] \times \mathbb{R}$ and

$$|\widetilde{b}(\xi, \eta_1) - \widetilde{b}(\xi, \eta_1)| \le L|\eta_1 - \eta_2|, \qquad \int_0^1 |\widetilde{b}(\xi, 0)|^2 d\xi < \infty$$

  for a suitable constant L, almost all $\xi \in [0,1]$, and all $\eta_1, \eta_2 \in \mathbb{R}$. Moreover for a.a. $\xi \in [0,1]$, $\widetilde{b}(\xi, \cdot) \in C^1(\mathbb{R})$ with $\nabla_\eta \widetilde{b}(\xi, \eta) \le 0$ for a.a. $\xi \in [0,1]$ and all $\eta \in \mathbb{R}$.

- $\widetilde{\sigma}$ is defined on $[0,1]$ and there exists a constant K such that $|\widetilde{\sigma}(\xi)| \le K$ for a.a. $\xi \in [0,1]$.

- $\widetilde{r}$ is defined on $\mathbb{R} \times \mathbb{R}$ and

$$|\widetilde{r}(\theta, \eta)| \le C(1 + |\eta|) \quad |\widetilde{r}(\theta_1, \eta) - \widetilde{r}(\theta_2, \eta)| \le C(1 + |\eta|)|\theta_1 - \theta_2|,$$

  for a suitable constant C, for all $\theta, \theta_1, \theta_2 \in \mathbb{R}$ and for all $\eta \in \mathbb{R}$.

- l is defined on $[0,1]\times\mathbb{R}$ and $0\le l(\xi,\eta)\le c_1(\xi)$ for a.a. $\xi\in[0,1]$ and all $\eta\in\mathbb{R}$ with $c_1\in L^1(0,1)$. Moreover for a.a. $\xi\in[0,1]$ the map $l(\xi,\cdot)\in C^1(\mathbb{R},\mathbb{R})$ and

$$\left|\frac{\partial}{\partial\eta}l(\xi,\eta)\right|\le c_2(\xi)$$

  wit $c_2\in L^2(0,1)$.

Finally we assume hat $x_0\in L^2(0,1)$.

To rewrite the above problem in abstract way we set $H=\Xi=U=L^2[0,1]$. By $\{W_t : t\ge 0\}$ we denote a cylindrical Wiener process in $L^2[0,1]$. Moreover we define the operator A with domain $D(A)$ by

$$D(A)=H^2[0,1]\cap H^1_0[0,1],\qquad (Ay)(\xi)=\frac{\partial^2}{\partial\xi^2}y(\xi),\quad \forall y\in D(A)$$

where $H^2[0,1]$ and $H^1_0[0,1]$ are the usual Sobolev spaces, and we set

$$b(x)(\xi)=\widetilde{b}(\xi,x(\xi)),\quad (\sigma z)(\xi)=\widetilde{\sigma}(\xi)z(\xi),\quad r(x,u)=\widetilde{r}(x(\xi),u(\xi))$$

$$g(x,u)=|u|^2_U+q(x)=\int_0^1[|u(\xi)|^2+l(\xi,x(\xi))]d\xi$$

for all $x,z,u\in L^2[0,1]$ and a.a. $\xi\in[0,1]$. Under previous assumptions we know, (see [10], §11.2.1) that A,b,σ verify Assumptions A2. Moreover noticing that

$$\nabla_x q(x)h=\int_0^1\frac{\partial}{\partial\eta}l(\xi,x(\xi))h(\xi)d\xi$$

and recalling the result in Example 6.4 it can be easily verified that Assumptions A4, A5 and A6 are satisfied.

References

[1] A. Ambrosetti, G. Prodi. *A primer of nonlinear analysis*, Cambridge Studies in Advanced Mathematics, 34, Cambridge University Press, 1995. MR1336591

[2] Ph. Briand, F. Confortola. BSDEs with stochastic Lipschitz condition and quadratic PDEs in Hilbert spaces. *Stochastic Process. Appl.* **118** (2008), no. 5, 818–838. MR2411522

[3] Ph. Briand, B. Delyon, Y. Hu, É. Pardoux, and L. Stoica. L^p solutions of backward stochastic differential equations. *Stochastic Process. Appl.* **108** (2003), no. 1, 109–129. MR2008603

[4] P. Briand, Y. Hu. Stability of BSDEs with random terminal time and homogenization of semilinear elliptic PDEs. *J. Funct. Anal.* **155** (1998), 455-494. MR1624569

[5] R. Buckdahn, S. Peng. Stationary backward stochastic differential equations and associated partial differential equations. *Probab. Theory Related Fields* **115** (1999), 383-399. MR1725404

[6] S. Cerrai, *Second order PDE's in finite and infinite dimensions. A probabilistic approach*. Lecture Notes in Mathematics **1762**, Springer, Berlin, 2001. MR1840644

[7] R. W. R. Darling, E. Pardoux. Backwards SDE with random terminal time and applications to semilinear elliptic PDE, *Ann. Probab.* **25** (1997), 1135-1159. MR1457614

[8] N. El Karoui. Backward stochastic differential equations: a general introduction. In *Backward stochastic differential equations (Paris, 1995–1996)*, volume 364 of *Pitman Res. Notes Math. Ser.*, pages 7–26. Longman, Harlow, 1997. MR1752672

[9] G. Da Prato and J. Zabczyk. *Ergodicity for infinite-dimensional systems*, London Mathematical Society Lecture Notes Series, 229, Cambridge University Press, 1996. MR1417491

[10] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions.* Cambridge University Press, Cambridge, 1992 MR1207136

[11] G. Da Prato and J. Zabczyk. *Second Order Partial Differential Equations in Hilbert Spaces*. Cambridge University Press, Cambridge, 2002. MR1985790

[12] M. Fuhrman,Y. Hu and G. Tessitore. On a class of stochastic optimal control problems related to BSDEs with quadratic growth. *SIAM J. Control Optim.* **45** (2006), no. 4, 1279–1296. MR2257222

[13] M. Fuhrman and G. Tessitore. Nonlinear Kolmororov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control. *Ann. Probab.* **30** (2002), 1397-1465. MR1920272

[14] M. Fuhrman and G. Tessitore, Infinite horizon backward stochastic differential equations and elliptic equations in Hilbert spaces. *Ann. Probab.* **30** (2004), 607-660. MR2039938

[15] F. Gozzi, E. Rouy. Regular solutions of second-order stationary Hamilton-Jacobi equations. *J. Differential Equations* **130**, 201-234, 1996. MR1409030

[16] Y. Hu, P. Imkeller, and M. Müller. Utility maximization in incomplete markets. *Ann. Appl. Probab.*, 15(3):1691–1712, 2005. MR2152241

[17] Y. Hu and G. Tessitore. BSDE on an infinite horizon and elliptic PDEs in infinite dimension. *NoDEA Nonlinear Differential Equations Appl.* To appear. MR2374211

[18] N. Kazamaki. *Continuous exponential martingales and BMO*, volume 1579 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1994. MR1299529

[19] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.*, 28(2):558–602, 2000. MR1782267

[20] F. Masiero. Infinite horizon stochastic optimal control problems with degenerate noise and elliptic equations in Hilbert spaces. *Appl. Math. Optim.*, 55, no. 3: 285–32, 2007. MR2313330

[21] J.-P. Lepeltier and J. San Martin. Existence for BSDE with superlinear-quadratic coefficient. *Stochastic Stochastics Rep.*, 63(3-4):227–240, 1998. MR1658083

[22] E. Pardoux. Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order, in: *Stochastic Analysis and related topics*, the Geilo workshop 1996, eds. L. Decreusefond, J. Gjerde, B Øksendal, A.S. Üstünel, 79-127, Progress in Probability **42**, Birkhäuser, Boston, 1998. MR1652339

[23] É. Pardoux. BSDEs, weak convergence and homogenization of semilinear PDEs. *Nonlinear analysis, differential equations and control (Montreal, QC, 1998), 503–549, NATO Sci. Ser. C Math. Phys. Sci., 528,* Kluwer Acad. Publ., Dordrecht, 1999. MR1695013

[24] D. Revuz and M. Yor. *Continuous martingales and Brownian motion.* Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 293. Springer-Verlag, Berlin, (1999). MR1725357

[25] M. Royer. BSDEs with a random terminal time driven by a monotone generator and their links with PDEs. *Stochastics Stochastics Rep.* **76** (2004) 281-307. MR2075474