ON THE BACKWARD STOCHASTIC RICCATI EQUATION IN INFINITE DIMENSIONS

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

Gianmario Tessitore
Dipartimento di Matematica, Università di Parma,
via D'Azeglio 85, 43100 Parma, Italia
e-mail: gianmario.tessitore@unipr.it

Abstract. We study backward stochastic Riccati equations (BSREs) arising in quadratic optimal control problems with infinite dimensional stochastic differential state equations. We allow the coefficients, both in the state equation and in the cost, to be random. In such a context BSREs are backward stochastic differential equations living in a non-Hilbert space and involving quadratic non-linearities. We propose two different notions of solutions to BSREs and prove, for both of them, existence and uniqueness results. We also show that such solutions allow to perform the synthesis of the optimal control. Finally we apply our results to the optimal control of a delay equation and of a wave equation with random damping.

1. Introduction

Backward stochastic Riccati differential equations (BSREs) naturally arise in the study of stochastic optimal linear quadratic control problems with stochastic coefficients.

The interest of proving existence and uniqueness results for such a class of equations was firstly addressed by Bismut in [2]. It was clear from the beginning that to study those highly non-linear backward stochastic differential equations was a challenging task, already in the finite dimensional case (see [3], [21] or the historical review in [12]). The difficulty comes essentially from the fact that, in its general formulation, the BSRE involves quadratic terms in both the unknowns (in particular in the, so called, 'martingale' term). Moreover the non linearity can be well defined only in a subset of the space of non-negative matrices (where the equation naturally lives).

Several works followed the pioneering paper [2] (see [20] [12], [13], [14] [15]). In particular only very recently, in [22], the proof of the existence and uniqueness of a solution of the BSRE was given in the general case corresponding to a finite dimensional, linear quadratic problem with random coefficients and state and control-dependent noise. This last result, somehow, completes the theory of finite dimensional BSREs. We remark that in all the above literature it is clear that the treatment of the equation can not be solely based on general backward stochastic differential equation techniques but needs to exploit the interplay between the Riccati equation and its control theoretic interpretation (for results on general backward stochastic differential equation with quadratic nonlinearities see [11] and [17]).

On the other side several works, motivated by control of stochastic partial differential equations, have been devoted to linear quadratic optimal control problems for infinite dimensional stochastic differential equations with deterministic coefficients (see for instance [23] and references within). The corresponding Riccati equation is a deterministic nonlinear ODE in a suitable space of symmetric, non-negative, Hilbert valued operators.

The present paper is, as far as we know, the first attempt to consider infinite dimensional BSREs. Such equations naturally arise in several models; namely they appear in all the situations in which one has to perform the synthesis of the optimal control for a linear quadratic problem having, as state equation, an infinite dimensional stochastic evolution equation with random coefficients (see examples in Sections 9 and 10). We also underline that the study of infinite dimensional BSREs introduces specific new difficulties in the theory of backward stochastic differential equations. Namely these are non-linear backward stochastic differential equations that involve unbounded linear terms and quadratic nonlinearities. Moreover, and this is the main difficulty, they naturally live in a non-Hilbertian infinite dimensional space.

In order to separate difficulties we consider here only the case in which the non-linearity does not depend on the 'martingale term' of the backward equation. In other words we consider the infinite dimensional analogue of the equation considered, in the finite dimensional case in [20]. We believe that, as we explain in the following, this case already presents serious new difficulties.

To be more precise: in this paper we consider a quadratic optimal control problem for a system governed by the following *state equation*:

$$\begin{cases} dy(t) = (Ay(t) + A_\sharp(t)y(t) + B(t)u(t))\,dt + C(t)y(t)\,dW(t) \quad t \in [0,T] \\ y(0) = x \end{cases} \tag{1.1}$$

In the above equation y is the *state* of the system and u is the *control*; y has values in an Hilbert space H and u has values in another Hilbert space U; W is a cylindrical Ξ-valued Wiener process defined on a probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$, where Ξ is a third Hilbert space. Expanding notation with respect to an orthonormal basis $\{f_i : i \in \mathbb{N}\}$ in Ξ we have $C(t)y(t)\,dW(t) = \sum_{i=1}^\infty C_i(t)y(t)d\beta_i(t)$ where $\{\beta_i : i \in \mathbb{N}\} := \{(f_i, W)_\Xi : i \in \mathbb{N}\}$ is a family of standard independent brownian motions.

We assume that the unbounded operator $A : \mathcal{D}(A) \subset H \to H$ is independent of $\omega \in \Omega$ and $t \in [0,T]$ and is the infinitesimal generator of a C_0-semigroup. On the contrary $A_\sharp$, B and C are allowed to be random; namely they are bounded, operator valued, stochastic processes that we assume to be predictable relatively to the filtration $\mathcal{F} = \{\mathcal{F}_t : t \geq 0\}$ generated by W (this last condition is not restrictive see Remark 2.4).

Our purpose is to minimize, over all predictable controls u, the quadratic cost functional:

$$\mathbb{E} \int_0^T \left(|\sqrt{S}(s)y(s)|_H^2 + |u(s)|_U^2 \right) ds + \mathbb{E}(P_T y(T), y(T))_H \tag{1.2}$$

where S is a predictable stochastic processes and P_T is a random variable both taking values in the set of linear, symmetric, non negative and bounded operators from H into H.

If we define the *stochastic value function* by:

$$(P(t)x, x)_H \doteq \inf_u \mathbb{E}^{\mathcal{F}_t,\, y(t)=x} \left[\int_t^T \left(|\sqrt{S}(s)y(s)|_H^2 + |u(s)|_U^2 \right) ds + (P_T y(T), y(T))_H \right] \tag{1.3}$$

then P solves, at least in a formal way, the following backward stochastic differential equation:

$$\begin{cases} -dP(t) = \Big(A^*P(t) + P(t)A + A_\sharp^*(t)P(t) + P(t)A_\sharp(t) - P(t)B(t)B^*(t)P(t) + S(t) \Big)\, dt \\ \qquad\qquad + \mathrm{Tr}[C^*(t)P(t)C(t) + C^*(t)Q(t) + Q(t)C(t)]\,dt + Q(t)\,dW(t) \qquad\qquad t \in [0,T] \\ P(T) = P_T \end{cases} \tag{1.4}$$

We notice that the unknowns in (1.4) are the two processes P and Q (the second one is sometimes referred at as *martingale* term). Process P has values in the cone $\Sigma^+(H)$ of bounded, non-negative, linear symmetric operators in H and process Q in the space $L_2(\Xi, \Sigma(H))$ of Hilbert-Schmidt operators from Ξ to the space $\Sigma(H)$ of bounded, linear symmetric operators in H. Moreover making again notation explicit we have

$$\mathrm{Tr}\Big[C^*(t)PC(t) + C^*(t)Q + QC(t)\Big] = \sum_{i=1}^\infty \Big[C_i^*(t)PC_i(t) + C_i^*(t)(Qf_i) + (Qf_i)C_i(t)\Big].$$

The specificity of our situation resides in the fact that the above equation involves both the unbounded term $A^*P + PA$ and quadratic term PBB^*P. Moreover $\Sigma(H)$ is not an Hilbert space, thus some essential tools in stochastic calculus, commonly used in the theory of backward stochastic differential equations, such as the Kunita-Watanabe martingale representation theorem, fail to hold. To overcome this difficulty one could try to compute the operator valued random variables on the vectors of a basis and then apply classical representation results to each component; but this procedure does not seem to allow the reconstruction of a suitable operator valued process Q. The point is that, due to the presence of an unbounded term, we can not consider equation (1.4) in its classical sense. Normally this leads to a mild formulation of the equations. Here, due to the difficulty of handling the martingale representation term Q, this approach causes problems. As a matter of fact mild formulation requires to give sense to the process like $s \to e^{(s-t)A^*}Q(s)e^{(s-t)A}h$, $h \in H$, while only the processes $Q(\cdot)h$ with h independent on t are well defined.

For the same reason, in the generality considered here, it seems difficult to show uniqueness of weak solutions of BSREs.

To cope with such a roadblock we propose the following strategy inspired by the notion of 'strong solution' for partial differential equations, see [1] or [16] and references therein. Roughly speaking the method consists in considering first equations with more regular data an then defining the solution in the general case by a limiting procedure.

To carry on with this programm we devote a first part of the paper (up to section 5) to the case in which the process S and the random variable P_T (corresponding respectively to the running and final cost) take values in the Hilbert space $L_2(H)$ of Hilbert-Schmidt operators $H \to H$ (see assumption A5). To start with we prove existence, uniqueness and stability with respect to approximations of the solution to a class of infinite dimensional backward stochastic differential equations with unbounded linear term and lipschitz nonlinearity, see Theorem 4.4. This result is essentially included in [10] as far as existence and uniqueness is concerned (except that we find a slightly more regular solution) while the part dealing with stability seems to be new and of independent interest.

The above general result is then applied to the affine Lyapunov equation:

$$\begin{cases} -dP(t) = (A^*P(t) + P(t)A + \mathrm{Tr}[C^*(t)P(t)C(t) + C^*(t)Q(t) + Q(t)C(t)])\,dt \\ \qquad\qquad +(A_\sharp^*(t)P(t) + P(t)A_\sharp(t) + L(t))\,dt + Q(t)\,dW(t) \qquad\qquad t \in [0,T] \\ P(T) = P_T \end{cases} \tag{1.5}$$

when L is a given Hilbert-Schmidt valued predictable process.

Then by fixed point technique and a priory estimates (see also [20]) we are able to show that if S and P_T take values in the Hilbert space $L_2(H)$ then equation (1.4) has, in $L_2(H)$, a unique *mild* solution (P, Q). By that we mean a pair of processes verifying $\mathbb{P}$-a.s. for all $t \in [0,T]$:

$$\begin{aligned} P(t) = &\int_t^T e^{(s-t)A^*}\,\mathrm{Tr}\Big[C^*(s)P(s)C(s) + C^*(s)Q(s) + Q(s)C(s)\Big]e^{(s-t)A}\,ds \\ &+ e^{(T-t)A^*}P_T e^{(T-t)A} + \int_t^T e^{(s-t)A^*}Q(s)e^{(s-t)A}\,dW(s) + \int_t^T e^{(s-t)A^*}S(s)e^{(s-t)A}\,ds \\ &+ \int_t^T e^{(s-t)A^*}\left(A_\sharp^*(s)P(s) + P(s)A_\sharp(s) - P(s)B(s)B^*(s)P(s)\right)e^{(s-t)A}\,ds \end{aligned} \tag{1.6}$$

Moreover, we prove that, such a solution can be approximated by the classical solutions of the equations obtained replacing A by its Yosida approximations. Once we have a solution to the Riccati equation it is easy to perform in this Hilbertian framework, the standard synthesis of the optimal control: that is to verify that $(P(0)x, x)_H$ is the optimal cost and that the unique optimal control $\overline{u}$ verifies the *feedback law* $\overline{u}(t) = -B^*(t)P(t)\overline{y}(t)$(see Theorem 5.14).

Hilbert-Schmidt assumption A5 is too restrictive in many of the concrete applications (see the Example in Section 10 and Remark 10.1) so it is necessary to complete the above mentioned programm in order to include in the theory general running costs S and final conditions P_T. In Section 6 we introduce the concept of *generalized solutions* of (1.4). By this we mean limits (in a suitable sense) of solutions corresponding to Hilbert-Schmidt data S and P_T. We are able to prove, under fairly general assumptions, that a generalized solution, in the above sense, exists and is unique (see Theorem 6.6). Notice that if existence of a generalized solution is somehow expected uniqueness seems a more interesting result; its proof is largely based on the control-theoretic interpretation of equation (1.4). Moreover we show that such a solution still allows to perform the synthesis of the optimal control as in the Hilbert-Schmidt case (see again Theorem 6.6). We also notice that their control theoretic interpretation imply that generalized solutions enjoy 'strong continuity' property (see Lemma 6.5).

In section 7 we prove that generalized solutions verify the following *variation of constants* formula

$$(P(t)x, x)_H = (L_{t,T}P_Tx, x)_H + \int_t^T (L_{t,s}S(s)x, x)_H\,ds - \int_t^T (L_{t,s}P(s)B(s)B^*(s)P(s)x, x)_H\,ds \quad \mathbb{P}-\text{a.s.} \tag{1.7}$$

where $L_{t,s}$ is the evolution operator corresponding to the Lyapunov equation (1.5) with $L = 0$. We are also able to show that there exists a unique process P verifying (1.7). Thus (1.7) can be regarded as an alternative definition of solution to the BSRE (1.4). We notice that in both the definitions of solution we propose only the P term in the BSRE is characterized. This is natural by the point of view of control theory and, in any case, is enough to complete the synthesis of the optimal control, see also Remark 6.3.

In section 9 and 10 we show that our general results can be applied to a variety of concrete examples. The first example is a minimization of variance problem for a delay equation with a stochastic coefficient. The interest of such example is that on one side it is extremely simple (and consequently applicable to a wide range of concrete situations) on the other it is connected with financial applications. Namely it is a firs step towards a mean variance hedging problem for a market with stochastic variance and memory effects. The second example is an optimal control problem for a wave equation in random media. In this case a stochastic coefficient is introduced, in a realistic way, assuming that the equation is subject to a stochastic damping due to the media. We notice that for the example in section 9 the Hilbert-Schmidt assumptions A5 is verified and we obtain mild solutions of the corresponding Riccati equation. On the contrary for the example in section 9 the Hilbert-Schmidt assumptions A5 is never verified and we have to use the concept of generalized solution of the Riccati equation.

2. Main Notation and Assumptions

By H, U and Ξ we will always indicate real separable Hilbert spaces.

If K is an Hilbert space its inner scalar product and norm will be denoted by $(\cdot,\cdot)_K$ and $|\cdot|_K$ omitting the K when no confusion is possible.

For any Banach space E by $\mathcal{B}(E)$ we denote its Borel σ-field.

For any pair K_1 and K_2 of separable real Hilbert spaces we denote by $L(K_1,K_2)$ the Banach space of linear and bounded operators from K_1 to K_2 endowed by the norm $|T|_{L(K_1,K_2)} = \sup_{\{x\in K_1, |x|_{K_1}=1\}} |Tx|_{K_2}$ (as usual $L(H) = L(H,H)$).

By $\Sigma(H)$ we denote the subspace of all symmetric and bounded operators and by $\Sigma^+(H)$ the cone of $\Sigma(H)$ that contains all positive semidefinite operators.

$L_2(K,H)$ denotes the Hilbert space of Hilbert-Schmidt operators from K to H, endowed with the Hilbert-Schmidt norm $|T|^2_{L_2(K,H)} = \sum_{i=1}^\infty |Te_i|^2_H$ ($\{e_i : i\in\mathbb{N}\}$ being an orthonormal basis in K) and we set $L_2(H,H) = L_2(H)$. $\Sigma_2(H)$ is the subset of $L_2(H)$ that consists in all linear and symmetric operators and $\Sigma_2^+(H)$ is the cone of $\Sigma_2(H)$ that consists in all non negative operators.

The cylindrical Wiener Process

We fix a probability basis $(\Omega,\mathcal{F},\mathbb{P})$. A cylindrical Wiener process with value in Ξ is a family $W(t)$, $t\geq 0$, of linear mappings $\Xi\to L^2(\Omega)$ such that:

i) for every $h\in\Xi$, $\{W(t)h,\ t\geq 0\}$ is a real (continuous) Wiener process;
ii) for every $h,k\in\Xi$ and $t,s\geq 0$, $\mathbb{E}(W(t)h\cdot W(s)k) = (t\wedge s)(h,k)_\Xi$.

We denote by $\mathcal{F}_t$ its natural filtration augmented with the set $\mathcal{N}$ of $\mathbb{P}$-null sets of $\mathcal{F}$. As it is well known the filtration $\mathcal{F}_t$ satisfies the usual conditions. By $\mathbb{E}^{\mathcal{F}_t}$ we denote the conditional expectation with respect to $\mathcal{F}_t$.

Finally by $\mathcal{P}$ we denote the predictable σ-field on $\Omega\times[0,T]$.

Some classes of stochastic process

Let K be any separable Hilbert space and let $\mathcal{B}(K)$ be its Borel σ-field on K. The following classes of processes will be used in this work:

- $L^p_{\mathcal{P}}(\Omega\times[0,T];K)$, $p\in[1,+\infty]$ denotes the subset of $L^p(\Omega\times[0,T];K)$, given by all equivalence classes admitting a predictable version. This space is endowed with the natural norm
$$|Y|^p_{L^p_{\mathcal{P}}(\Omega\times[0,T];K)} = \mathbb{E}\int_0^T |Y_s|^p_K\,ds$$
Elements of this space are defined up to modification.
- $L^p_{\mathcal{P}}(\Omega;L^2([0,T];K))$ denotes the space of equivalence classes of processes Y, admitting a predictable version such that the norm:
$$|Y|^p_{L^p_{\mathcal{P}}(\Omega;L^2([0,T];K))} = \mathbb{E}\Big(\int_0^T |Y_s|^2_K\,ds\Big)^{p/2}$$
is finite. Elements of this space are defined up to modification.
- $C_{\mathcal{P}}([0,T];L^p(\Omega;K))$ denotes the space of K-valued processes Y such that $Y:[0,T]\to L^p(\Omega,K)$ is continuous and Y has a predictable modification, endowed with the norm:
$$|Y|^p_{C_{\mathcal{P}}([0,T];L^p(\Omega;K))} = \sup_{t\in[0,T]} \mathbb{E}|Y_t|^p_K$$
Elements of $C_{\mathcal{P}}([0,T];L^p(\Omega;K))$ are identified up to modification.
- $L^p_{\mathcal{P}}(\Omega;C([0,T];K))$ denotes the space of predictable processes Y with continuous paths in K, such that the norm
$$|Y|^p_{L^p_{\mathcal{P}}(\Omega;C([0,T];K))} = \mathbb{E}\sup_{t\in[0,T]} |Y_t|^p_K$$
is finite. Elements of this space are defined up to indistinguishability.

Now let us consider the space $L(H)$ of linear and bounded operators from a separable Hilbert space H to H. Moreover it turns out that the σ-field generated by the operator norm in $L(H)$ is too large. For instance if A

generates a C_0 semigroup the map $t \to e^{tA}$ is not even measurable with respect to such σ- field, see [5, pag. 23-24]. We are therefore led to introduce the σ-field:

$$\mathcal{L}_S = \sigma\Big\{\{T \in L(H) : Tu \in A\}, \text{ where } u \in H \text{ and } A \in \mathcal{B}(H)\Big\}$$

Following again [5] the elements of $\mathcal{L}_S$ are called *strongly measurable*.

We notice that the maps $P \to |P|_{L(H)}$ and $(P, u) \to Pu$ are measurable from $(L(H), \mathcal{L}_S)$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and from $(L(H) \times H, \mathcal{L}_S \otimes \mathcal{B}(H))$ to $(H, \mathcal{B}(H))$ respectively.

Moreover $\mathcal{L}_S$ is identical to the weak σ-field:

$$\mathcal{L}_S = \sigma\Big\{\{T \in L(H) : (Tu, x)_H \in A\}, \text{ where } u, x \in H \text{ and } A \in \mathcal{B}(\mathbb{R})\Big\}$$

We define the following spaces:

- $L^\infty_{\mathcal{P},S}(\Omega \times [0,T]; L(H))$ the space of essentially bounded, strongly measurable predictable processes $Y : \Omega \times [0,T] \to L(H)$. That is Y is measurable from $(\Omega \times [0,T], \mathcal{P})$ to $(L(H), \mathcal{L}_S)$ and the real valued random valued $|Y|_{L(H)}$ is in $L^\infty(\Omega \times [0,T]; \mathbb{R})$. By $|Y|_{L^\infty_{\mathcal{P},S}(\Omega\times[0,T];L(H))}$ we indicate the norm of $|Y|_{L(H)}$ in $L^\infty(\Omega \times [0,T]; \mathbb{R})$. Elements of this space are identified up to modification.
- $L^\infty_S(\Omega, \mathcal{F}_t; L(H))$ is the space of measurable maps $Y : (\Omega, \mathcal{F}_t) \to (L(H), \mathcal{L}_S)$ such that $|Y|_{L(H)}$ is in $L^\infty(\Omega; \mathbb{R})$. By $|Y|_{L^\infty_S(\Omega;L(H))}$ we indicate the norm of $|Y|_{L(H)}$ in $L^\infty(\Omega; \mathbb{R})$.
- $L^1_{\mathcal{P},S}([0,T]; L^\infty(\Omega, L(H)))$ is the space of predictable, strongly measurable processes such that $|Y|_{L(H)}$ is in $L^1([0,T]; L^\infty(\Omega; \mathbb{R}))$. By $|Y|_{L^1_{\mathcal{P},S}([0,T];L^\infty(\Omega,L(H)))}$ we indicate the norm of $|Y|_{L(H)}$ in $L^1([0,T]; L^\infty(\Omega; \mathbb{R}))$. Elements of this space are identified up to modification.

We identically define, with trivial changes the spaces: $L^\infty_{\mathcal{P},S}(\Omega \times [0,T]; \Sigma^+(H))$, $L^\infty_{\mathcal{P},S}(\Omega \times [0,T]; L(U,H)))$, $L^1_{\mathcal{P},S}([0,T]; L^\infty(\Omega, \Sigma^+(H)))$ and $L^\infty_S(\Omega, \mathcal{F}_t; \Sigma^+(H))$. Elements of these spaces are identified up to modification.

Statement of the problem and general assumptions on the coefficients

We consider the following infinite dimensional stochastic differential equation:

$$\begin{cases} dy(s) = (Ay(s) + A_\sharp(s)y(s) + B(s)u(s))\,ds + C(s)y(s)\,dW(s) \quad s \in [t,T] \\ y(t) = x \end{cases} \tag{2.1}$$

where y is an H valued process that represents the *state* of the system and is our unknown, u is the *control* and the initial data x is in H. To stress its dependence on u, t and x we will denote the (mild, see Definition 3.1) solution of equation (2.1) by $y^{t,x,u}$ when needed.

Our purpose is to minimize with respect to u the cost functional,

$$J(0,x,u) = \mathbb{E}\left[\int_0^T \left((S(s)y^{0,x,u}(s), y^{0,x,u}(s))_H + |u(s)|^2_U\right)\,ds + (P_T y^{0,x,u}(T), y^{0,x,u}(T))_H\right] \tag{2.2}$$

We also introduce the following random variables, for $t \in [0,T]$:

$$J(t,x,u) = \mathbb{E}^{\mathcal{F}_t}\left[\int_t^T \left((S(s)y^{t,x,u}(s), y^{t,x,u}(s))_H + |u(s)|^2_U\right)ds + (P_T y^{t,x,u}(T), y^{t,x,u}(T))_H\right]$$

We will work under the following general assumptions on A, B and C that will hold throughout the paper.

Hypothesis 2.1.

A1) *$A : D(A) \subset H \to H$ is the infinitesimal generator of a C_0 semigroup $e^{tA} : H \to H$.*

A2) *We assume that $A_\sharp \in L^\infty_{\mathcal{P},S}(\Omega \times [0,T]; L(H))$. We denote by $M_{A_\sharp}$ a positive constant such that:*

$$|A_\sharp(t,\omega)|_{L(U,H)} \le M_{A_\sharp}, \quad \mathbb{P}-a.s. \text{ and for a.e. } t \in (0,T).$$

Moreover $B \in L^\infty_{\mathcal{P},S}(\Omega \times [0,T]; L(U,H))$. We denote by M_B a positive constant such that:

$$|B(t,\omega)|_{L(U,H)} \le M_B, \quad \mathbb{P}-a.s. \text{ and for a.e. } t \in (0,T).$$

A3) *We assume that C is of the form: $C = \sum_{i=1}^\infty C_i(\cdot, f_i)_\Xi$, where $\{f_i : i \in \mathbb{N}\}$ is an orthonormal basis in Ξ. Moreover we suppose that*

$$C_i \in L^\infty_{\mathcal{P},S}(\Omega \times [0,T]; L(H)) \quad \text{and} \quad \left(\sum_{i=1}^\infty |C_i(t,\omega)|^2_{L(H)}\right)^{1/2} \le M_C, \quad \mathbb{P}-a.s. \text{ for a.e. } t \in (0,T)$$

for a suitable positive constant M_C.

On S and P_T we will need to play with two different sets of assumptions. We introduce both of them here

A4) $S \in L^1_{\mathcal{P},S}([0,T]; L^\infty(\Omega;\Sigma^+(H)))$ and $P_T \in L^\infty_S(\Omega,\mathcal{F}_T;\Sigma^+(H))$.

A5) $S \in L^2_{\mathcal{P}}(\Omega\times[0,T];\Sigma_2^+(H))$ and $P_T \in L^2(\Omega,\mathcal{F}_T;\Sigma_2^+(H))$.

We introduce, for later use, the Yosida approximants of the unbounded operator A letting:

$$A_h = AJ(h,A) \quad \text{where} \quad J(h,A) = h(hI-A)^{-1}, \quad h : 1,2,\ldots.$$

We denote by M_A a positive constant such that:

$$\sup_{t\in[0,T]} |e^{tA_h}|_{L(H)} \le M_A \quad \forall h \in \mathbb{N} \qquad \text{and} \qquad \sup_{t\in[0,T]} |e^{tA}|_{L(H)} \le M_A \tag{2.3}$$

Remark 2.2. If we set $\beta_i(t) := (f_i, W(t))_\Xi$ then $\{\beta_i : i \in \mathbb{N}\}$ is a family of independent standard (real valued) brownian motions. Moreover the term $C(t)y(t)dW(t)$ can be rewritten as $\sum_{i=1}^\infty C_i(t)y(t)\,d\beta_i(t)$.

Remark 2.3. In Section 9 we show that Assumptions A1)-A5) are satisfied by a general class of controlled stochastic delay equations. In Section 10 we point out that for stochastic controlled partial differential equations assumptions A1)-A4) are satisfied while A5) typically fails. We also notice that when H is finite dimensional A5) and A4) reduce to the requirements $S \in L^2_{\mathcal{P},S}([0,T]; L^\infty(\Omega;\Sigma^+(H)))$ and $P_T \in L^\infty_S(\Omega,\mathcal{F}_T;\Sigma^+(H))$ which slightly generalize the assumptions in [20] and [12], [13], [14] where S is uniformly bounded.

Remark 2.4. The fact that in the previous assumptions measurability and predictability has always been required with respect to the filtration $\{\mathcal{F}_t : t \ge 0\}$ generated by the noise $\{W_t : t \ge 0\}$ is not restrictive. Such a condition can in fact be easily weakened by the following standard procedure.

Let $\widehat{\Xi} \supset \Xi$ be a larger separable Hilbert space and let $\{\widehat{W}_t : t \ge 0\}$ be a cylindrical Wiener process with values in $\widehat{\Xi}$. Moreover let $\{\widehat{f}_i : i \in \mathbb{N}\}$ an orthonormal basis in $\widehat{\Xi}$ with $\{\widehat{f}_i : i \in \mathbb{N}\} \supset \{f_i : i \in \mathbb{N}\}$. Finally let $\widehat{C}_i = C_i$ if $f_i \in \Xi$, $\widehat{C}_i = 0$ if $f_i \notin \Xi$ and $\widehat{C} = \sum_{i=1}^\infty \widehat{C}_i(\cdot,\widehat{f}_i)_{\widehat{\Xi}}$. If now we replace Ξ by $\widehat{\Xi}$, W by $\widehat{W}$ and C by $\widehat{C}$ equation (2.1) is unchanged while in all the assumptions filtration $\{\mathcal{F}_t : t \ge 0\}$ can be replaced by filtration $\{\widehat{\mathcal{F}}_t : t \ge 0\}$ generated by $\widehat{W}$. In addition, in order to allow $\mathcal{F}_0$ to be non trivial there are no difficulties in letting the noise W to be defined in $[-\rho,+\infty[$, for some $\rho > 0$, instead that in $[0,+\infty[$.

3. The state equation

This section is devoted to the state equation (2.1). We recall the well known notion of *mild solution*

Definition 3.1. Given $x \in H$ and $u \in L^2_{\mathcal{P}}(\Omega\times[t,T];U)$, a mild solution of equation (2.1) is a process $y \in L^2_{\mathcal{P}}(\Omega\times[t,T];H)$ such that, almost surely in $\Omega\times[t,T]$:

$$y(s) = e^{(s-t)A}x + \int_t^s e^{(s-\sigma)A}\left[A_\sharp(\sigma)y(\sigma) + B(\sigma)u(\sigma)\right] d\sigma + \int_t^s e^{(s-\sigma)A}C(\sigma)y(\sigma)\,dW(\sigma)$$

The following existence and uniqueness result is now well known.

Theorem 3.2. *Assume A1)-A3). Given any $x \in H$ and $u \in L^2_{\mathcal{P}}(\Omega\times[t,T];U))$ problem (2.1) has a unique mild solution $y \in C_{\mathcal{P}}([t,T];L^2(\Omega;H))$. Moreover*

$$\sup_{s\in[t,T]} \mathbb{E}|y(s)|^2 \le C_2\Big[|x|^2 + \mathbb{E}\int_t^T |u(s)|^2\,ds\Big] \tag{3.1}$$

for a suitable constant C_2 depending on T, M_B, M_C $M_{A_\sharp}$ and M_A.

Finally if $p > 2$ and

$$\mathbb{E}\Big(\int_t^T |u(s)|^2\,ds\Big)^{\frac{p}{2}} < \infty$$

then we have that $y \in L^p_{\mathcal{P}}(\Omega; C([t,T];H))$ and:

$$\mathbb{E}\sup_{s\in[t,T]} |y(s)|^p \le C_p\Big[|x|^p + \mathbb{E}\Big(\int_t^T |u(s)|^2\,ds\Big)^{\frac{p}{2}}\Big] \tag{3.2}$$

for some positive constant C_p depending on p, T, M_B, M_C, M_A and $M_{A_\sharp}$.

Proof. The argument is identical to the one included in [5][Theorem 7.4] and [7][Proposition 3.2]. The only difference is that here the operators B and C are stochastic processes. Anyway, thanks to their boundedness stated in hypotheses 2.1, one can proceed exactly as in the above mentioned papers. □

To stress dependence on the initial data and on the control we will, when necessary, denote the above solution by $y^{t,x,u}$.

For all $x \in H$ and $u \in L^p_{\mathcal{P}}(\Omega; L^2([t,T];U))$, $p \geq 2$ we also introduce the following family of approximating problems, $h \in \mathbb{N}$:

$$\begin{cases} dy_h(s) = (A_h y_h(s) + A_\sharp(s) y_h(s) + B(s)u(s))\, dt + C(s) y_h(s)\, dW(s) & s \in [t,T] \\ y(t) = x \end{cases} \tag{3.3}$$

It is well known (see [5]) that, under the same hypotheses of Theorem 3.2, problem (3.3) has, for every $h \in \mathbb{N}$, a unique *classical* solution $y_h \in L^p_{\mathcal{P}}(\Omega; C([t,T];H))$ that, when necessary, we will denote by $y_h^{t,x,u}$.

The following stability result for the approximated problems holds:

Theorem 3.3. *Assume that $x_h \to x$ in H and $u_h \to u$ in $L^p_{\mathcal{P}}(\Omega; L^2([t,T];U))$, as $h \to \infty$. If $p = 2$, $y_h^{t,x_h,u_h} \to y^{t,x,u}$ in $C_{\mathcal{P}}([t,T]; L^2(\Omega;H))$. If $p > 2$, $y_h^{t,x_h,u_h} \to y^{t,x,u}$ in $L^p_{\mathcal{P}}(\Omega; C([t,T];H))$.*

Proof. The proof consists in a straightforward application of the parameter depending contraction argument (see, for instance, [25][Theorem 10.1]). The case with $p = 2$ is treated also in [23][Theorem 1.1]. For the case $p > 2$ it is enough to proceed as in [7][Proposition 3.2]. □

4. Backward stochastic equations: stability with respect to approximations

In this section we prove, for later use, a result on the stability of a generic backward stochastic equation with value in an real and separable Hilbert space K and lipschitz non-linearity. Beside the same hypotheses on the noise introduced in the previous section, we are given:

(i) a positive number $T > 0$;
(ii) an unbounded operator $G : D(G) \subset K \to K$ and a sequence of bounded operators $G_h : K \to K$:
(iii) a map $\psi : [0,T] \times \Omega \times K \times L_2(\Xi, K) \to K$
(iv) a final data $\eta \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}; K)$.

We assume the following:

Hypothesis 4.1.

(1) *G generates a C_0-semigroup $\{e^{tG} : t \geq 0\}$ in K.*
(2) *There exist a constant M_G such that:*

$$\sup_{t\in[0,T]} |e^{tG_h}|_{L(H)} \leq M_G \quad \forall h \in \mathbb{N} \qquad \textit{and} \qquad \sup_{t\in[0,T]} |e^{tG}|_{L(H)} \leq M_G. \tag{4.1}$$

(3) $\sup_{t\in[0,T]} \left| e^{tG_h} x - e^{tG} x \right| \to 0$ *for all $x \in K$.*
(4) *ψ is measurable from $\mathcal{P} \otimes \mathcal{B}(K) \otimes \mathcal{B}(L_2(\Xi, K))$ to $\mathcal{B}(K)$ and $\mathbb{E} \int_0^T |\psi(s,0,0)|_K^2\, ds < +\infty$*
(5) *There exists a constant M_ψ such that, $\mathbb{P}$ almost surely for almost every $t \in [0,T]$ the following holds for all $Y_1, Y_2 \in K$, $Z_1, Z_2 \in L_2(\Xi, K)$*

$$|\psi(t, Y_1, Z_1) - \psi(t, Y_2, Z_2)|_K \leq M_\psi \left(|Y_1 - Y_2|_K + |Z_1 - Z_2|_{L_2(\Xi,K)} \right) \tag{4.2}$$

We consider the following backward stochastic equation:

$$\begin{cases} dY(s) = -GY(s)\, ds - \psi(s, Y(s), Z(s))\, ds - Z(s)\, dW(s) & s \in [0,T] \\ Y(T) = \eta \end{cases} \tag{4.3}$$

and the following sequence of approximating problems

$$\begin{cases} dY_h(s) = -G_h Y_h(s)\, ds - \psi(s, Y_h(s), Z_h(s))\, ds - Z_h(s)\, dW(s) & s \in [0y, T] \\ Y_h(T) = \eta \end{cases} \tag{4.4}$$

Definition 4.2. A mild solution of equation (4.3) is a couple of predictable processes (Y, Z) such that Y belongs to $L^2_{\mathcal{P}}(\Omega, C([0,T];K))$, Z belongs to $L^2_{\mathcal{P}}(\Omega \times [0,T]; L_2(\Xi;K))$ and they verify for all $t \in [0,T]$:

$$Y(t) = e^{(T-t)G}\eta + \int_t^T e^{(s-t)G} \psi(s, Y(s), Z(s))\, ds + \int_t^T e^{(s-t)G} Z(s)\, dW(s) \qquad \mathbb{P}-\text{a.s.} \tag{4.5}$$

An identical definition is given for a mild solution of equation (4.4).

Remark 4.3. Being G_h bounded it is immediate to check that the couple (Y_h, Z_h) is a mild solution of equation (4.4) if and only if it is a classical solution of (4.4) that is it verifies, for all $t \in [0,T]$,

$$Y_h(t) = \eta_h + \int_t^T (G_h Y_h(s) + \psi(s, Y_h(s), Z_h(s))\, ds + \int_t^T Z_h(s)\, dW(s) \qquad \mathbb{P}-\text{a.s.} \tag{4.6}$$

The following result will be used in several occasions in the sequel. As far as the existence and uniqueness part is concerned, is very similar to the one included in [10] (except from the fact that we obtain a more regular solution). On the contrary the part dealing with stability with respect to approximations is new.

Theorem 4.4. *Under Hypothesis 4.1 problem* (4.3) *has a unique mild solution* (Y, Z). *Moreover* $\forall h \in \mathbb{N}$, *problem* (4.4) *has a unique classical (equivalently mild) solution* (Y_h, Z_h).

Finally:

$$\lim_{h\to\infty} \mathbb{E}\Big(\sup_{t\in[0,T]} |Y_h(t) - Y(t)|_K^2 \Big) = 0, \qquad \lim_{h\to\infty} \mathbb{E}\int_0^T |Z_h(s) - Z(s)|^2_{L_2(\Xi;K)} ds = 0 \tag{4.7}$$

Proof. *Part I. Existence and uniqueness for a simplified equation*

We consider the simplified equation:

$$Y(t) \;=\; e^{(T-t)G}\eta + \int_t^T e^{(s-t)G}F(s)\,ds + \int_t^T e^{(s-t)G}Z(s)\,dW(s) \qquad t\in[0,T] \tag{4.8}$$

with $F \in L^2_{\mathcal{P}}(\Omega\times[0,T];K)$. In [10] [Proposition 2.1] it is shown that the above equation admits a unique solution $(Y,Z) \in L^2_{\mathcal{P}}(\Omega\times[0,T];K)\times L^2_{\mathcal{P}}(\Omega\times[0,T];L_2(\Xi;K))$ given explicitly by

$$Y(t) = e^{(T-t)G}(\mathbb{E}^{\mathcal{F}_t}\eta) + \int_t^T e^{(s-t)G}(\mathbb{E}^{\mathcal{F}_t}F(s))\,ds \tag{4.9}$$

$$Z(t) = -e^{(T-t)G}V(t) - \int_t^T e^{(s-t)G}L(t,s)\,ds \tag{4.10}$$

where V and L verify,

$$\mathbb{E}^{\mathcal{F}_t}\eta = \eta - \int_t^T V(\sigma)\,dW(\sigma), \qquad 0\le t\le T \tag{4.11}$$

$$\mathbb{E}^{\mathcal{F}_t}F(s) = F(s) - \int_t^s L(\sigma,s)\,dW(\sigma), \qquad 0\le t\le s\le T \tag{4.12}$$

(existence and uniqueness of V and L is given by Kunita-Watanabe martingale representation result applied in the Hilbert space K, see again [10]).

We now estimate such a solution in a suitable norm. For every $\beta > 0$:

$$\mathbb{E}\sup_{t\in[0,T]} e^{2\beta t}|Y(t)|_K^2 \le 2M_G^2\Big[\mathbb{E}\sup_{t\in[0,T]} e^{2\beta t}\Big(\int_t^T \mathbb{E}^{\mathcal{F}_t}|F(\sigma)|_K\,d\sigma\Big)^2 + \mathbb{E}\sup_{t\in[0,T]} e^{2\beta t}|\mathbb{E}^{\mathcal{F}_t}\eta|_K^2\Big]$$

Since:

$$\Big(\int_t^T |F(\sigma)|_K\,d\sigma\Big)^2 \le \int_t^T e^{-2\beta s}\,ds\int_t^T e^{2\beta s}|F(s)|_K^2\,ds \le \frac{e^{-2\beta t}}{2\beta}\int_t^T e^{2\beta s}|F(s)|_K^2\,ds$$

One gets that, thanks to Jensen and Doob inequalities:

$$\mathbb{E}\sup_{t\in[0,T]} e^{2\beta t}\Big(\int_t^T \mathbb{E}^{\mathcal{F}_t}|F(s)|_K\,ds\Big)^2 \le \mathbb{E}\sup_{t\in[0,T]}\Big(\mathbb{E}^{\mathcal{F}_t}\sup_{t\in[0,T]} e^{\beta t}\int_t^T |F(s)|_K\,ds\Big)^2$$

$$\le 4\mathbb{E}\sup_{t\in[0,T]} e^{2\beta t}\Big(\int_t^T |F(s)|_K\,ds\Big)^2 \le \frac{4}{2\beta}\mathbb{E}\int_0^T e^{2\beta s}|F(\sigma)|_K^2\,ds$$

Thus we have, using again Doob inequality:

$$\mathbb{E}\sup_{t\in[0,T]} e^{2\beta t}|Y(t)|_K^2 \le \frac{4M_G^2}{\beta}\mathbb{E}\int_0^T e^{2\beta s}|F(\sigma)|_K^2\,ds + 8M_G^2 e^{2\beta T}\mathbb{E}|\eta|_K^2 \tag{4.13}$$

As far as Z is concerned we have:

$$|Z(t)|^2_{L_2(\Xi;K)} \le 2M_G^2\Big[|V(t)|^2_{L_2(\Xi;K)} + \frac{e^{-2\beta t}}{2\beta}\int_t^T e^{2\beta s}|L(t,s)|^2_{L_2(\Xi;K)}\,ds\Big]$$

Therefore:

$$\mathbb{E}\int_0^T e^{2\beta t}|Z(t)|^2_{L_2(\Xi;K)}\,dt \le 2M_G^2\Big[\mathbb{E}\int_0^T e^{2\beta t}|V(t)|^2_{L_2(\Xi;K)}\,dt + \frac{1}{2\beta}\mathbb{E}\int_0^T\int_t^T e^{2\beta s}|L(t,s)|^2_{L_2(\Xi;K)}\,ds\,dt\Big]$$

$$\le 2M_G^2\Big[4e^{2\beta T}\mathbb{E}|\eta|_K^2 + \frac{1}{2\beta}\mathbb{E}\int_0^T e^{2\beta s}\int_0^s |L(t,s)|^2_{L_2(\Xi;K)}\,dt\,ds\Big]$$

and we can conclude

$$\mathbb{E}\int_0^T e^{2\beta t}|Z(t)|^2_{L_2(\Xi;K)}\,dt \le 2M_G^2\Big[4e^{2\beta T}\mathbb{E}|\eta|^2_K + \frac{2}{\beta}\int_0^T e^{2\beta s}\mathbb{E}|F(s)|^2_K\,ds\Big]. \tag{4.14}$$

In an identical way we can prove that for all $h\in\mathbb{N}$ there exists a unique couple of processes (Y_h, Z_h) that belongs to $L^2_{\mathcal{P}}(\Omega; C([0,T];K))\times L^2_{\mathcal{P}}(\Omega\times[0,T]; L_2(\Xi;K))$ verifying, for all $t\in[0,T]$

$$Y_h(t) \quad = \quad e^{(T-t)G_h}\eta + \int_t^T e^{(s-t)G_h}F(s)\,ds + \int_t^T e^{(s-t)G_h}Z_h(s)\,dW(s) \qquad \mathbb{P}\text{-a.s.} \tag{4.15}$$

with $F\in L^2_{\mathcal{P}}(\Omega\times[0,T];K)$.

Moreover Y_h and Z_h verify (4.13) and (4.14).

Part II. Stability with respect to approximations of the simplified equation

By (4.10), we have for a.e. $t\in[0,T]$:

$$Z_h(t)-Z(t) = -e^{(T-t)G_h}V(t) + e^{(T-t)G}V(t) - \int_t^T e^{(s-t)G_h}L(t,s)\,ds + \int_t^T e^{(s-t)G}L(t,s)\,ds \quad \mathbb{P}-\text{a.e.} \tag{4.16}$$

with $V\in L^2_{\mathcal{P}}(\Omega\times[0,T];L_2(\Xi;K))$ and $L\in L^2_{\mathcal{P}}(\Omega\times[0,T]\times[0,T];L_2(\Xi;K))$

By the Dominated Convergence Theorem, we immediately have that:

$$\lim_{h\to+\infty}\mathbb{E}\int_0^T |Z_h(t)-Z(t)|^2_{L_2(\Xi;K)}\,dt = 0 \tag{4.17}$$

Now we consider the term $Y_h - Y$. We have

$$Y_h(t)-Y(t) \quad = \quad \Big[e^{(T-t)G_h}\eta - e^{(T-t)G}\eta\Big] + \mathbb{E}^{\mathcal{F}_t}\int_t^T \Big[e^{(s-t)G_h}F(s) - e^{(s-t)G}F(s)\Big]\,ds.$$

To estimate the first term of the right-hand side we can proceed as follows:

$$\begin{aligned}
&\mathbb{E}\sup_{t\in[0,T]}\big|\mathbb{E}^{\mathcal{F}_t}[e^{(T-t)G_h}\eta - e^{(T-t)G}\eta]\big|^2_K \le \mathbb{E}\sup_{t\in[0,T]}\big(\mathbb{E}^{\mathcal{F}_t}\big|e^{(T-t)G_h}\eta - e^{(T-t)G}\eta\big|_K\big)^2\\
&\le \mathbb{E}\sup_{t\in[0,T]}\big(\mathbb{E}^{\mathcal{F}_t}\sup_{t\in[0,T]}\big|e^{(T-t)G_h}\eta - e^{(T-t)G}\eta\big|_K\big)^2 \le 4\mathbb{E}\big(\sup_{t\in[0,T]}\big|\big[e^{(T-t)G_h}\eta - e^{(T-t)G}\eta\big]\big|_K\big)^2\\
&\le 4\mathbb{E}\big(\sup_{t\in[0,T]}\big|\big[e^{(T-t)G_h}\eta - e^{(T-t)G}\eta\big]\big|^2_K\big)
\end{aligned}$$

Similarly, for the second:

$$\begin{aligned}
&\mathbb{E}\sup_{t\in[0,T]}\left|\mathbb{E}^{\mathcal{F}_t}\left[\int_t^T e^{(s-t)G_h}F(s) - e^{(s-t)G}F(s)\,ds\right]\right|^2_K \le \mathbb{E}\sup_{t\in[0,T]}\left(\mathbb{E}^{\mathcal{F}_t}\int_t^T |e^{(s-t)G_h}F(s) - e^{(s-t)G}F(s)|_K\,ds\right)^2\\
&\le \mathbb{E}\sup_{t\in[0,T]}\left(\mathbb{E}^{\mathcal{F}_t}\int_0^T\sup_{\sigma\in[0,T]}|e^{\sigma G_h}F(s) - e^{\sigma G}F(s)|_K\,ds\right)^2 \le 4\mathbb{E}\left(\int_0^T\sup_{\sigma\in[0,T]}|e^{\sigma G_h}F(s) - e^{\sigma G}F(s)|_K\,ds\right)^2\\
&\le 4T\mathbb{E}\int_0^T\sup_{\sigma\in[0,T]}|e^{\sigma G_h}F(s) - e^{\sigma G}F(s)|^2_K\,ds
\end{aligned}$$

Therefore we get that:

$$\mathbb{E}\sup_{t\in[0,T]}|Y_h(t)-Y(t)|^2_K \le 8\mathbb{E}\left(\sup_{t\in[0,T]}|e^{(T-t)G_h}\eta - e^{(T-t)G}\eta|^2_K\right) + 8T\,\mathbb{E}\int_0^T\sup_{\sigma\in[0,T]}|e^{\sigma G_h}F(s) - e^{\sigma G}F(s)|^2_K\,ds$$

By point (iii) in Hypothesis 4.1 and dominated convergence Theorem we can conclude:

$$\mathbb{E}\sup_{t\in[0,T]}|Y_h(t)-Y(t)|^2_K \to 0 \tag{4.18}$$

Part III. Conclusion

We let, for $\beta>0$ $\mathbb{K}(\beta) = L^2_{\mathcal{P}}(\Omega, C([0,T];K))\times L^2_{\mathcal{P}}(\Omega\times[0,T];L_2(\Xi;K))$ endowed with the norm (equivalent to the natural one):

$$|(Y,Z)|^2_{\mathbb{K}(\beta)} = \mathbb{E}\sup_{t\in[0,T]}e^{2\beta t}|Y(t)|^2 + \mathbb{E}\int_0^T e^{2\beta s}|Z(s)|^2 ds.$$

Moreover we define a map $\Gamma : \mathbb{K}(\beta) \to \mathbb{K}(\beta)$ and a sequence of maps $\Gamma_h : \mathbb{K}(\beta) \to \mathbb{K}(\beta)$, $h \in \mathbb{N}$ letting $\Gamma(\widehat{Y}, \widehat{Z}) = (Y, Z)$ (resp. $\Gamma_h(\widehat{Y}, \widehat{Z}) = (Y_h, Z_h)$) where (Y, Z) (resp. (Y_h, Z_h)) is the solution of equation (4.8) (resp. (4.15)) with $F(s) = \psi(s, \widehat{Y}(s), \widehat{Z}(s))$.

We notice that F belongs to $L^2_{\mathcal{P}}(\Omega \times [0, T]; K)$ thus the above definition is justified by part I of the present proof.

Moreover (4.13) and (4.14) immediately yield the following inequality holding for all $(\widehat{Y}, \widehat{Z})$, $(\widetilde{Y}, \widetilde{Z})$ in $\mathbb{K}(\beta)$

$$|\Gamma(\widehat{Y}, \widehat{Z}) - \Gamma(\widetilde{Y}, \widetilde{Z})|^2_{\mathbb{K}(\beta)} \leq \frac{4M_G^2 M_\psi^2}{\beta} |(\widehat{Y}, \widehat{Z}) - (\widetilde{Y}, \widetilde{Z})|^2_{\mathbb{K}(\beta)}$$

and an identical formula holds (with the same constant) for Γ_h.

So we can conclude that, for β large enough, Γ and Γ_h are contractions in $\mathbb{K}(\beta)$. Clearly the unique fixed point of Γ (resp. Γ_h) is the unique mild solution of equation (4.3) (resp. (4.4).

Finally by the parameter depending contraction principle, see [25][Theorem 10.1], relation (4.7) follows immediately if we prove that for all fixed $(\widehat{Y}, \widehat{Z}) \in L^2_{\mathcal{P}}(\Omega, C([0, T]; K)) \times L^2_{\mathcal{P}}(\Omega \times [0, T]; L_2(\Xi; K))$ letting $(Y, Z) = \Gamma(\widehat{Y}, \widehat{Z})$ and $(Y_h, Z_h) = \Gamma_h(\widehat{Y}, \widehat{Z})$ then

$$\mathbb{E} \sup_{t \in [0,T]} |Y(s) - Y_h(s)|^2 + \mathbb{E} \int_0^T |Z(s) - Z_h(s)|^2 ds \to 0 \quad \text{as } h \to \infty.$$

The above relation is an immediate consequence of (4.18) and (4.17) letting $F(s) = \psi(s, \widehat{Y}(s), \widehat{Z}(s))$ in part II of the present proof. $\square$

Remark 4.5. As a byproduct of the previous argument we have the following estimate for the solution (Y, Z) of equation (4.3).

$$|(Y, Z)|^2_{\mathcal{K}(\beta)} \leq \hat{C} \left[e^{2\beta T} \mathbb{E} |P_T|^2_{\Sigma_2(H)} + \frac{1}{\beta} \int_0^T e^{2\beta s} \mathbb{E} |\psi(s, 0, 0)|^2_{\Sigma_2(H)} ds \right] \tag{4.19}$$

holding for β large enough, depending on T, M_G, M_ψ, and for a suitable constant $\hat{C}$, depending on T, M_G.

To prove it just remark that, for β large enough, Γ is a 1/2 contraction in $\mathcal{K}(\beta)$. Since $(Y, Z) = lim_{n \to \infty} \Gamma^n(0, 0)$ we have $|(Y, Z)|_{\mathcal{K}(\beta)} \leq 2 |\Gamma(0, 0)|_{\mathcal{K}(\beta)}$ and the claim follows by (4.14) and (4.13).

An identical estimate holds (with the same constant) for the solution (Y_h, Z_h) of the approximating equation (4.4).

Remark 4.6. Notice that although the semigroup generated by G is not, in general, a contraction semigroup and $\psi(\cdot, 0, 0)$ is only in $L^2_{\mathcal{P}}(\Omega \times [0, T]; K))$, nevertheless Y has continuous trajectories. This is not true for standard (forward) stochastic differential equations (that is when the initial datum is specified rather the final one). For instance in Theorem 3.2 if u is in $L^2_{\mathcal{P}}(\Omega \times [0, T]; K))$ then y is only mean-square continuous.

The reason for such extra regularity of Y can be founded in relation (4.9), at least for the simplified equation. Indeed in (4.9) it is clear that Y can be represented only by conditional expectations and deterministic convolutions. In particular no stochastic convolution is involved in (4.9).

5. The Riccati Equation in the Hilbert-Schmidt case

The natural space in which the deterministic Riccati equation is studied is the space $\Sigma(H)$ that is not an Hilbert space. Thus (see the introduction) we initially consider the Riccati equation in the Hilbert space $\Sigma_2(H)$ of symmetric and Hilbert-Schmidt linear operators in H.

5.1. **The Lyapunov equation.** We start from the linear part of the Riccati equation. Namely we consider the Lyapunov equation:

$$\begin{cases} -dP(t) = (A^*P(t) + P(t)A + A_\sharp^*(t)P(t) + P(t)A_\sharp(t) + L(t))\, dt + Q(t)\, dW(t) \\ \qquad\qquad + \mathrm{Tr}[C^*(t)P(t)C(t) + C^*(t)Q(t) + Q(t)C(t)]\, dt \qquad\qquad t \in [0, T] \\ P(T) = P_T \end{cases} \tag{5.1}$$

Where, expliciting the notation with respect to the basis $\{f_i : i \in \mathbb{N}\}$ of Ξ, for all $P \in \Sigma_2(H)$, and $Q \in L_2(\Xi, \Sigma_2(H))$

$$\mathrm{Tr}\Big[C^*(t)PC(t) + C^*(t)Q + QC(t)\Big] = \sum_{i=1}^{\infty} \Big[C_i^*(t)PC_i(t) + C_i^*(t)(Qf_i) + (Qf_i)C_i(t)\Big].$$

In order to give a precise definition of the *mild solution* of equation (5.1) we introduce the family $\{e^{t\mathcal{A}} : t \geq 0\}$ of linear operators $\Sigma(H) \to \Sigma(H)$ letting:

$$e^{t\mathcal{A}}X := e^{tA^*}Xe^{tA}, \qquad t \geq 0,\ X \in \Sigma(H)$$

We notice that the above family is a semigroup of bounded operators in the sense that

$$e^{t\mathcal{A}}e^{s\mathcal{A}}X = e^{(t+s)\mathcal{A}}X, \quad X \in \Sigma(H),\ t, s \geq 0$$

but is not necessarily strongly continuous, in $\Sigma(H)$, see also [1][Chapter 1]. On the other side if we restrict it to $\Sigma_2(H)$ then it becomes a strongly continuous semigroup. Namely we have the following result (also concerning approximations) that will considerably simplify our work:

Lemma 5.1. *Under hypothesis A1) the family of linear operators $\{e^{t\mathcal{A}} : t \geq 0\}$ is a strongly continuous semigroup of bounded operators in $\Sigma_2(H)$.*

Moreover for all $X \in L_2(H)$:

$$\lim_{h\to\infty} \sup_{t\in[0,T]} |e^{tA_h^*}Xe^{tA_h} - e^{tA^*}Xe^{tA}|_{L_2(H)} = 0. \tag{5.2}$$

Proof. We prove only continuity for $t = 0$. The proof of continuity in a generic t follows by semigroup law. Moreover (5.2) is proved by an identical argument. We fix $X \in L_2(H)$ and a basis $\{e_i : i \in \mathbb{N}\}$ in H. Clearly

$$\sum_{i=1}^{\infty} |e^{tA^*}Xe^{tA}e_i - Xe_i|_H^2 \leq 2\sum_{i=1}^{\infty} |e^{tA^*}Xe^{tA}e_i - e^{tA^*}Xe_i|_H^2 + 2\sum_{i=1}^{\infty} |e^{tA^*}Xe_i - Xe_i|_H^2$$

We have to prove that both the above terms converge to 0 as $t \downarrow 0$. As far as the second is concerned being:

$$|e^{tA^*}Xe_i - Xe_i|_H^2 \leq 2(M_A^2 + 1)|Xe_i|_H^2 \quad \text{and} \quad \sum_{i=1}^{\infty} |Xe_i|_H^2 = |X|_{L_2(H)}^2 < +\infty$$

the claim follows by dominated convergence theorem. As far as the first is concerned we have:

$$\sum_{i=1}^{\infty} |e^{tA^*}Xe^{tA}e_i - e^{tA^*}Xe_i|_H^2 \leq M_A^2 \sum_{i=1}^{\infty} |Xe^{tA}e_i - Xe_i|_H^2 = M_A^2 \sum_{i=1}^{\infty} |X^*e^{tA^*}e_i - X^*e_i|_H^2$$

and the claim follows again by dominated convergence theorem. □

Let us denote by $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subset \Sigma_2(H) \to \Sigma_2(H)$ the infinitesimal generator of the semigroup $\{e^{t\mathcal{A}} : t \geq 0\}$ in $\Sigma_2(H)$. Notice that

$$(\mathcal{A}Xx, y)_H = (Xx, Ay)_H + (XAx, y)_H \qquad X \in \mathcal{D}(\mathcal{A}),\ x, y \in \mathcal{D}(A)$$

We now assume that $P_T \in L^2(\Omega, \mathcal{F}_T; \Sigma_2(H))$ and $L \in L^2_{\mathcal{P}}(\Omega \times [0, T]; \Sigma_2(H))$ and give the following definition of a mild solution (P, Q) of equation (5.1) with values in Hilbert-Schmidt case. We need also the following approximations to $\mathcal{A}$.

Definition 5.2. We define a sequence of bounded operators $\mathcal{A}_n : \Sigma_2(H) \to \Sigma_2(H)$ as follows:

$$\mathcal{A}_h X \doteq A_h^* X + XA_h, \qquad X \in \Sigma_2(H), \qquad h = 1, 2, \ldots$$

Definition 5.3. A *mild solution* of problem (5.1) is a pair of processes

$$(P, Q) \in L^2_{\mathcal{P}}(\Omega, C([0, T]; \Sigma_2(H))) \times L^2_{\mathcal{P}}(\Omega \times [0, T]; L_2(\Xi; \Sigma_2(H)))$$

that verifies for all $t \in [0, T]$:

$$\begin{aligned} P(t) =& e^{(T-t)A^*}P_T e^{(T-t)A} + \int_t^T e^{(s-t)A^*}\left[L(s) + A_\sharp^*(s)P(s) + P(s)A_\sharp(s)\right]e^{(s-t)A}\,ds \\ &+ \int_t^T e^{(s-t)A^*}\mathrm{Tr}\Big[C^*(s)P(s)C(s) + C^*(s)Q(s) + Q(s)C(s)\Big]e^{(s-t)A}\,ds+ \\ &+ \int_t^T e^{(s-t)A^*}Q(s)e^{(s-t)A}\,dW(s) \qquad \mathbb{P}-\text{a.s.} \end{aligned} \tag{5.3}$$

We also introduce the regularized versions of equation (5.1) corresponding to the ones we have introduced for the state equation. Namely we consider

$$\begin{cases} -dP_h(t) = \Big(A_h^*P_h(t) + P_h(t)A_h + A_\sharp^*(t)P_h(t) + P_h(t)A_\sharp(t) + L(t)\Big)\,dt \\ \qquad +Q_h(t)\,dW(t) + \mathrm{Tr}[C^*(t)P_h(t)C(t) + C^*(t)Q_h(t) + Q_h(t)C(t)]\,dt, \quad t \in [0, T] \\ P(T) = P_T \end{cases} \tag{5.4}$$

where A_h are the Yosida approximants of A. The definition of mild solution for the above equation is obtained from the one corresponding to equation (5.1) just by replacing A by A_h. Since A_h is bounded, mild solutions are classical solutions, i.e. they satisfies $\mathbb{P}$-a.s. for all $t \in [0,T]$:

$$P_h(t) = P_T + \int_t^T \left(A_h^* P_h(s) + P_h(s)A_h + A_\sharp^*(s)P_h(s) + P_h(s)A_\sharp(s) + L(s)\right) ds$$

$$+ \int_t^T \mathrm{Tr}[C^*(s)P_h(s)C(s) + C^*(s)Q_h(s) + Q_h(s)C(s)]\, ds + \int_t^T Q_h(s)\, dW(s)$$

Theorem 5.4. *Assume hypotheses A1)-A3). Moreover assume that $P_T \in L^2_{\mathcal{P}}(\Omega, \mathcal{F}_T; \Sigma_2(H))$ and $L \in L^2_{\mathcal{P}}(\Omega \times [0,T]; \Sigma_2(H)))$.*
Then problem (5.1) *has a unique mild solution $(P,Q) \in L^2_{\mathcal{P}}(\Omega, C([0,T]; \Sigma_2(H))) \times L^2_{\mathcal{P}}(\Omega \times [0,T]; L_2(\Xi; \Sigma_2(H)))$. Moreover $\forall h \in \mathbb{N}$, problem* (5.4) *has a unique classical solution $(P_h, Q_h) \in L^2_{\mathcal{P}}(\Omega, C([0,T]; \Sigma_2(H))) \times L^2_{\mathcal{P}}(\Omega \times [0,T]; L_2(\Xi; \Sigma_2(H)))$.*

Finally the following stability result holds:

$$\lim_{h\to\infty} \mathbb{E}\left(\sup_{t\in[0,T]} |P_h(t) - P(t)|^2_{\Sigma_2(H)}\right) = 0, \qquad \lim_{h\to\infty} \mathbb{E}\int_0^T |Q_h(s) - Q(s)|^2_{L_2(\Xi;\Sigma_2(H))} ds = 0 \tag{5.5}$$

Proof. The claim is a special case of Theorem 4.4, letting $K = \Sigma_2(H)$, $G = \mathcal{A}$, $G_h = \mathcal{A}_h$, $\eta = P_T$ and defining, for all $P \in \Sigma_2(H)$, $Q \in L_2(\Xi, \Sigma_2(H))$

$$\psi(s, P, Q) = \mathrm{Tr}[C^*(s)PC(s) + C^*(s)Q + QC(s)] + L(s) + A_\sharp^*(s)P + PA_\sharp(s).$$

We have just to check that in this specific situation Hypothesis 4.1 holds, but this is a direct consequence of Hypotheses A1)-A3) and of the fact that $L \in L^2_{\mathcal{P}}(\Omega \times [0,T]; \Sigma_2(H)))$. □

Remark 5.5. Remark 4.5 gives, in the present case, the following estimate for the solution (P,Q) of equation (5.1).

$$|(P,Q)|^2_{\mathcal{K}} \le \hat{C}\left[e^{2\beta T}\mathbb{E}|P_T|^2_{\Sigma_2(H)} + \frac{1}{\beta}\int_0^T e^{2\beta s}\mathbb{E}|L(s)|^2_{\Sigma_2(H)} ds\right] \tag{5.6}$$

holding for β large enough, depending on T, M_A, $M_{A_\sharp}$, M_C, and for a suitable constant $\hat{C}$, depending on T, M_A, $M_{A_\sharp}$.

An identical estimate holds (with the same constant) for the solution (P_h, Q_h) of the approximating equation (5.4).

The following result is a key step towards the *fundamental relation* (see Proposition 5.11). Moreover it gives useful estimates on the solution to equation (5.1).

Theorem 5.6. *Besides the hypotheses of Theorem 5.4 assume that P_T belongs to $L^\infty_S(\Omega, \mathcal{F}_T; L(H))$ and L belongs to $L^1_{\mathcal{P},S}([0,T]; L^\infty(\Omega; L(H)))$. Let (P,Q) be the unique mild solution to equation* (5.1) *and let $y^{t,x,u}$ be the mild solution to equation* (2.1)*. Then for all $t \in [0,T]$, $x \in H$, $u \in L^2_{\mathcal{P}}(\Omega \times [0,T]; U)$ it holds, $\mathbb{P}$-a.s.*

$$(P(t)x,x) = \mathbb{E}^{\mathcal{F}_t}(P_T y^{t,x,u}(T), y^{t,x,u}(T)) + \mathbb{E}^{\mathcal{F}_t}\int_t^T \left[(L(s)y^{t,x,u}(s), y^{t,x,u}(s)) - 2(P(s)B(s)u(s), y^{t,x,u}(s))\right] ds \tag{5.7}$$

Moreover, for all $t \in [0,T]$:

$$|P(t)|_{L(H)} \le C_2\left[|P_T|_{L^\infty_S(\Omega;L(H))} + \int_t^T |L(s)|_{L^\infty_S(\Omega;L(H))}\, ds\right] \qquad \mathbb{P}-a.s. \tag{5.8}$$

where C_2 is the positive constant depending only on $T, M_B, M_C, M_A, M_{A_\sharp}$ defined in (3.1)*.*

Similarly if, for all $h \in \mathbb{N}$, (P_h, Q_h) is the unique solution of problem (5.4) *and for all $t \in [0,T]$:*

$$|P_h(t)|_{L(H)} \le C_2\left[|P_T|_{L^\infty_S(\Omega;L(H))} + \int_t^T |L(s)|_{L^\infty_S(\Omega;L(H))}\, ds\right] \qquad \mathbb{P}-a.s.$$

Proof. The proof will be concluded in three steps. In the first we will prove (5.7) for $u \in L^6_{\mathcal{P}}(\Omega \times [0,T]; U)$, then we will prove estimate (5.8) and finally we will extend (5.7) to all the admissible controls.

First step. The following argument is simple but has some delicate points; thus we expose it here in all details. Let $y_h = y_h^{t,x,u}$ be the classical solution to (3.3). By Theorem 3.3 we know that $y_h \in L^6_{\mathcal{P}}(\Omega, C([t,T]; H))$ and $y_h \to y^{t,x,u}$ in $L^6_{\mathcal{P}}(\Omega, C([t,T]; H))$ as $h \to \infty$.

Let $\Psi \in C^2(H)$ with $\Psi(y) = 1$ for $|y| \le 1$, $\Psi(y) = 0$ for $|y| \ge 2$ and $\Psi(y) \in [0,1], \forall y \in H$. Differentiating by Itô rule we obtain (we consider $\Psi' \in H$, $\Psi'' \in L(H)$)

$$
\begin{aligned}
d_s\Big[\Psi\left(y_h(s)/N\right)\left(P_h(s)y_h(s), y_h(s)\right)\Big] &= N^{-1}F_N(s)ds + G_N(s)dW_s + \\
&\quad -\Psi\left(y_h(s)/N\right)\Big[(L(s)y_h(s), y_h(s))_H - 2(P_h(s)B(s)u(s), y_h(s))_H\Big]ds
\end{aligned}
\tag{5.9}
$$

Where:

$$
\begin{aligned}
F_N(s) &= \left(\Psi'(N^{-1}y_h(s)), [A_h y_h(s) + A_\sharp(s)y_h(s) + B(s)u(s)]\right)_H (P_h(s)y_h(s), y_h(s))_H \\
&\quad + 2\sum_{i=1}^{\infty}\left(\Psi'(N^{-1}y_h(s)), C_i(s)y_h(s)\right)_H (P_h(s)C_i(s)y_h(s), y_h(s))_H \\
&\quad + \frac{1}{2N}\sum_{i=1}^{\infty}\left(\Psi''(N^{-1}y_h(s))C_i(s)y_h(s), C_i(s)y_h(s)\right)_H (P_h y_h(s), y_h(s))_H \\
&\quad + \sum_{i=1}^{\infty}\left(\Psi'(N^{-1}y_h(s)), C_i(s)y_h(s)\right)_H (Q_h^i y_h(s), y_h(s))_H
\end{aligned}
$$

$$
\begin{aligned}
G_N(s)f_i &= 2\Psi(N^{-1}y_h(s))(P_h(s)C_i(s)y_h(s), y_h(s))_H - \Psi(N^{-1}y_h(s))(Q_h^i(s)y_h(s), y_h(s))_H \\
&\quad + \frac{1}{N}(P_h(s)y_h(s), y_h(s))_H(\Psi'(N^{-1}y_h(s)), C_i(s)y_h(s))_H,
\end{aligned}
$$

where $Q_h^i = Q_h f_i$, with $\{f_i\}_{i\in\mathbb{N}}$ orthonormal basis of Ξ.
As it can be easily verified $\mathbb{E}\int_t^T |F_N(s)|ds \le$ const., $\forall N \in \mathbb{N}$. Moreover, since $\Psi(N^{-1}y) = 0$ and $\Psi'(N^{-1}y) = 0$ if $|y| > 2N$ we have, for all $N \in \mathbb{N}$

$$
\sum_{i=1}^{\infty}\mathbb{E}\int_t^T |G_N(s)f_i|_H^2 ds \le c_2 N^4\left\{M_C^2\mathbb{E}\int_t^T \sup_{s\in[t,T]}\|P_h(s)\|_{L(H)}^2\, ds + \mathbb{E}\int_t^T \|Q_h(s)\|_{\Sigma_2(H)}^2\, ds\right\} < +\infty
$$

where c_2 is a positive universal constant.

Finally $(Ly_h, y_h)_H$ and $(P_h Bu, y_h)$ belong to $L^1_{\mathcal{P}}(\Omega \times [t,T], \mathbb{R})$ and $\Psi(y_h(s)/N)$ converges to 1 $\mathbb{P}$-a.s., for all s.

Thus first integrating in $[t,T]$ and then computing conditional expectation with respect to $\mathcal{F}_t$ ($\mathbb{E}^{\mathcal{F}_t}$) and finally letting $N \to 0$ relation (5.9) becomes

$$
(P_h(t)x, x)_H = \mathbb{E}^{\mathcal{F}_t}(P_T y_h(T), y_h(T))_H + \mathbb{E}^{\mathcal{F}_t}\int_t^T [(L(s)y_h(s), y_h(s))_H - 2(P_h(s)B(s)u(s), y_h(s))_H]\, ds
$$

and the claim follows letting $h \to +\infty$ thank to (3.3) and (5.4).

Second step. The following L^∞ bound for the $L(H)$ norm of the mild solution to equation (5.1) will be important in the approach to the (nonlinear) Riccati equation. In the finite dimensional case a similar result is proved, by a slightly different argument in [20]. Here there is our proof. From the *first step* we know that for all $x \in H$, $\mathbb{P}$-a.s.

$$
(P(t)x, x) = \mathbb{E}^{\mathcal{F}_t}(P_T y^{t,x,0}(T), y^{t,x,0}(T)) - \mathbb{E}^{\mathcal{F}_t}\int_t^T (L(s)y^{t,x,0}(s), y^{t,x,0}(s))ds \tag{5.10}
$$

consequently

$$
|(P(t)x, x)| \le |P_T|_{L^\infty_S(\Omega;L(H))}\mathbb{E}^{\mathcal{F}_t}|y^{t,x,0}(T)|^2 + \int_t^T |L(s)|_{L^\infty_S(\Omega;L(H))}\mathbb{E}^{\mathcal{F}_t}|y^{t,x,0}(s)|^2 ds
$$

Since $y^{t,x,0}(s)$, $s > t$ is independent on $\mathcal{F}_t$ the previous relation reads

$$
|(P(t)x, x)| \le |P_T|_{L^\infty_S(\Omega;L(H))}\mathbb{E}|y^{t,x,0}(T)|^2 + \int_t^T |L(s)|_{L^\infty_S(\Omega;L(H))}\mathbb{E}|y^{t,x,0}(s)|^2 ds
$$

and by estimate (3.1) for $u \equiv 0$

$$
|(P(t)x, x)| \le C_2|P_T|_{L^\infty_S(\Omega;L(H))} + C_2\int_t^T |L(s)|_{L^\infty_S(\Omega;L(H))}ds; \quad \forall x\in H,\ |x| \le 1,\ \mathbb{P}-\text{a.s.}
$$

and the claim holds being H separable. The same estimate holds true also for every $|(P_h(t)x, x)|$, since the constant C_2 does not depend on h.

Third step For a general $u \in L^2_{\mathcal{P}}(\Omega \times [0,T];U)$ we choose a sequence $u_m \to u$ such that u_m is bounded and $u_m \to u$ in $L^2_{\mathcal{P}}(\Omega \times [0,T];U)$. By Theorem 3.3, $y^{t,x,u_m} \to y^{t,x,u}$ in $C_{\mathcal{P}}([t,T];L^2(\Omega;H))$ and, by the *second step*, $P \in L^\infty_{\mathcal{P},\mathcal{S}}(\Omega \times [0,T]y;L(H))$. Moreover:

$$\left|\mathbb{E}^{\mathcal{F}_t} \int_t^T (L(s)y^{t,x,u_m}(s), y^{t,x,u_m}(s))_H - (L(s)y^{t,x,u}(s), y^{t,x,u}(s))_H \, ds\right|$$
$$\leq \Big[(\sup_{s\in[t,T]} \mathbb{E}|y^{t,x,u_m}(s)|^2)^{1/2} + (\sup_{s\in[t,T]} \mathbb{E}|y^{t,x,u}(s)|^2)^{1/2}\Big]\Big(\sup_{s\in[t,T]} \mathbb{E}|y^{t,x,u_m}(s) - y^{t,x,u}(s)|^2\Big)^{1/2} \int_t^T |L(s)|_{L^\infty_S(\Omega,H)} \, ds$$

Thus we can pass relation (5.7) to the limit as $m \to \infty$ obtaining the claim. □

5.2. **Existence of a unique solution for the Riccati equation and the synthesis of the optimal control.**

In this section we prove the existence of a unique mild solution for the Riccati equation :

$$\begin{cases} -dP(t) = (A^*P(t) + P(t)A + \mathrm{Tr}[C^*(t)P(t)C(t) + C^*(t)Q(t) + Q(t)C(t)])\,dt \\ \qquad -(P(t)B(t)B^*(t)P(t) - A^*_\sharp(t)P(t) - P(t)A_\sharp(t) - S(t))\,dt + Q(t)\,dW(t) \qquad t\in[0,T] \\ P(T) = P_T \end{cases} \tag{5.11}$$

under assumptions A1)-A5).

The presence of a quadratic nonlinear term imposes the following approach (classical when dealing with the Riccati equation see [23]) in solving the problem: first we will find a local solution then we will prove some a-priori estimate for the solution to guarantee the existence of a global solution. The method we use to prove the a-priori bound is based on the so called *fundamental relation*, see Proposition 5.11 and uses, in an essential way, the control-theoretic interpretation of the Riccati Equation.

We start extending the notion of *mild* solution given in section 5.1:

Definition 5.7. Fix $T_0 \in [0,T]$. A mild solution for problem (5.11), considered in $[T_0,T]$ is a pair (P,Q) with

$$P \in L^2_{\mathcal{P}}(\Omega, C([T_0,T];\Sigma_2(H))) \cap L^\infty_{\mathcal{P},S}(\Omega; C([T_0,T];\Sigma^+(H))), \quad Q \in L^2_{\mathcal{P}}(\Omega \times (T_0,T); L_2(\Xi;\Sigma_2(H)))$$

such that for all $t \in [T_0,T]$:

$$\begin{aligned} P(t) &= \int_t^T e^{(s-t)A^*}\mathrm{Tr}\Big[C^*(s)P(s)C(s) + C^*(s)Q(s) + Q(s)C(s)\Big]e^{(s-t)A}\,ds \\ &+ e^{(T-t)A^*}P_T e^{(T-t)A} + \int_t^T e^{(s-t)A^*}\left[S(s) + A^*_\sharp(s)P(s) + P(s)A_\sharp(s)\right]e^{(s-t)A}\,ds \\ &+ \int_t^T e^{(s-t)A^*}Q(s)e^{(s-t)A}\,dW(s) - \int_t^T e^{(s-t)A^*}P(s)B(s)B^*(s)P(s)e^{(s-t)A}\,ds \qquad \mathbb{P}-\text{a.s.} \end{aligned} \tag{5.12}$$

Proposition 5.8 (Local Existence)**.** *Under hypotheses A1)-A5) there exists a* $\delta \in]0,T]$ *such that problem* (5.11) *has a unique mild solution in the interval* $[T-\delta,T]$

Proof. To simplify the notation we will set

$$|P_T|_{L^\infty_S(\Omega;L(H))} = M_P \qquad |S|_{L^1_{\mathcal{P},S}([0,T];L^\infty(\Omega;L(H)))} = M_S,$$

We choose $r > C_2(M_P + M_S)$ and δ such that $C_2[M_P + r^2\delta M_B^2 + M_S] \leq r$

We define:

$$B(r) = \{P \in L^2(\Omega; C([T-\delta,T];\Sigma_2(H))) : \sup_{t\in[T-\delta,T]} |P(t,\omega)|_{L(H)} \leq r \quad \mathbb{P}-a.e.\}$$

endowed with the norm

$$|P|^2_\beta = \mathbb{E}\sup_{t\in[T-\delta,T]} e^{2\beta t}|P(t)|^2_{\Sigma_2(H)}$$

On $B(r)$ we construct the map $\Lambda : B(r) \to B(r)$ letting $\Lambda(K) = P$, where (P, Q) is the unique solution to equation (5.1) (in $[T-\delta, T]$) with $L = -KBB^*K$ that is :

$$
\begin{aligned}
P(t) = & \int_t^T e^{(s-t)A^*} \mathrm{Tr}\Big[C^*(s)P(s)C(s) + C^*(s)Q(s) + Q(s)C(s)\Big] e^{(s-t)A}\, ds \\
& + \int_t^T e^{(s-t)A^*} S(s) e^{(s-t)A}\, ds + e^{(T-t)A^*} P_T e^{(T-t)A} + \int_t^T e^{(s-t)A^*} Q(s) e^{(s-t)A}\, dW(s) \\
& - \int_t^T e^{(s-t)A^*} \left[A_\sharp^*(s)P(s) + P(s)A_\sharp(s) - K(s)B(s)B^*(s)K(s)\right] e^{(s-t)A}\, ds
\end{aligned}
$$

We claim that the map Λ is a contraction in $B(r)$.

First of all we check that it maps $B(r)$ into itself. By Theorem 5.4 (applied in $[T-\delta, T]$) we know that $\Lambda(K) \in L^2_{\mathcal{P}}(\Omega \times [T-\delta, T]; \Sigma_2(H)))$. So it is enough to show that for all $t \in [T-\delta, T]$ it holds $|\Lambda(K)(t)|_{L(H)} \le r$ $\mathbb{P}$-a.s.. Thanks to (5.8) we have that $\mathbb{P}$-a.s.:

$$
\begin{aligned}
|\Lambda(K)(t)|_{L(H)} & \le C_2[|P_T|_{L^\infty_{\mathcal{S}}(\Omega, L(H))} + \int_{T-\delta}^T (|K(s)B(s)B^*(s)K(s)|_{L^\infty_{\mathcal{S}}(\Omega, L(H))} + |S(s)|_{L^\infty_{\mathcal{S}}(\Omega, L(H))})\, ds] \\
& \le C_2[M_P + r^2\delta M_B^2 + M_S] \le r
\end{aligned}
$$

Moreover by (4.19) for all K_1 and K_2 in $B(r)$ (since (4.19) is stated in the whole $[0, T]$ we should, to be precise, extend $K_1(s) = K_2(s) = 0$ for $s < T - \delta$)

$$
|\Lambda(K_2) - \Lambda(K_1)|^2_\beta \le \frac{\hat{C}}{\beta} \int_{T-\delta}^T e^{2\beta s} \mathbb{E}|K_2BB^*K_2 - K_1BB^*K_1|^2_{\Sigma_2(H)} ds
$$

Since $|K_i|_{\Sigma_2(H)} \le r$, $i = 1, 2$, $\mathbb{P}$-a.s. for all $t \in [T-\delta, T]$ the above relation gives

$$
|\Lambda(K_2) - \Lambda(K_1)|^2_\beta \le \frac{\hat{C}}{\beta} r^2 M_B^4 \int_{T-\delta}^T e^{2\beta s} \mathbb{E}|K_2 - K_1|^2_{\Sigma_2(H)} ds
$$

Therefore, if β is large enough, Λ is a contraction in $B(r)$. If P is its unique fixed point the mild solution (P, Q) of (5.1) with $L = -PBB^*P$ is the unique mild solution of (5.11). □

Clearly local uniqueness of the solution immediately implies global uniqueness

Corollary 5.9 (Global Uniqueness). *Let (P_i, Q_i), $i = 1, 2$ be two mild solutions of the Riccati equation* (5.11) *in the interval $[T_0, T]$ for some $T_0 \in [0, T)$. Then $P_1(t) = P_2(t)$, $\mathbb{P}$-a.s. for all $t \in [T_0, T]$ and $Q_1(t) = Q_2(t)$, $\mathbb{P}$-a.s. for almost all $t \in [T_0, T]$.*

Remark 5.10. The length δ of the interval on which the mild solution of the Riccati equation exists depends only on T, M_A, $M_{A_\sharp}$, M_B, M_C, $|S|_{L^1_{\mathcal{P},S}([0,T]; L^\infty(\Omega; L(H))}$ and $|P_T|_{L^\infty_{\mathcal{S}}(\Omega; L(H))}$. Thus to extend the solution to the whole $[0, T]$ it is sufficient to establish an a-priori bounds for the $L^\infty(\Omega; L(H))$ norm of the P part of any local solution, independently on the length of the interval in which it is defined.

This will be done using the following consequence of Theorem 5.6. The next relation also has an obvious control-theoretic interpretation and will be essential in performing the synthesis of the optimal control:

Proposition 5.11 (Fundamental Relation). *Assume A1)- A5) and let (P, Q) be the mild solution of* (5.11) *in an interval $[T_0, T]$. Then, $\forall t \ge T_0$, $x \in H$ $u \in L^2_{\mathcal{P}}(\Omega \times [t, T]; U)$ it holds:*

$$
(P(t)x, x)_H = J(t, x, u) - \mathbb{E}^{\mathcal{F}_t} \int_t^T |u(s) + B^*P(s)y^{t,x,u}(s)|^2_H\, ds \tag{5.13}
$$

Proof. We start noticing that, by definition, (P, Q) is a mild solution of the Lyapunov equation (5.1) with $L = -PBB^*P + S$. Thus by (5.7):

$$
\begin{aligned}
(P(t)x, x) = & \mathbb{E}^{\mathcal{F}_t}(P_T y^{t,x,u}(T), y^{t,x,u}(T)) + \mathbb{E}^{\mathcal{F}_t} \int_t^T |\sqrt{S}(s) y^{t,x,u}(s)|^2_H ds \\
& - \mathbb{E}^{\mathcal{F}_t} \int_t^T \left[(y^{t,x,u}(s), [P(s)B(s)B^*(s)P(s) + S(s)]y^{t,x,u}(s))_H + 2(P(s)B(s)u(s), y^{t,x,u}(s))_H\right] ds.
\end{aligned}
$$

Then the claim follows just adding and subtracting $\mathbb{E}^{\mathcal{F}_t} \int_t^T |u(s)|^2_U ds$. □

Proposition 5.12 (Positivity and a-priori estimate). *Let (P, Q) be the mild solution to* (5.11) *in $[T_0, T]$, then:*

(1) *for every* $t \in [T_0, T]$ *and* $x \in H$, $(P(t)x, x)_H \geq 0 \quad \mathbb{P}-a.s.$,
(2) *for every* $t \in [T_0, T]$, $|P(t)|_{L(H)} \leq C_2[M_P + M_S] \quad \mathbb{P}-a.s.$

where C_2 *is the constant defined in Theorem 3.2.*

Proof. If we apply (5.13) to $u \equiv 0$ we obtain for all $x \in H$ with $|x|_H \leq 1$ and for all $t \in [T_0, T]$:

$$\begin{aligned}(P(t)x, x)_H &= \mathbb{E}^{\mathcal{F}_t}(P_T y^{t,x,0}(T), y^{t,x,0}(T))_H + \mathbb{E}^{\mathcal{F}_t}\int_t^T |\sqrt{S}(s) y^{t,x,0}(s)|_H^2 \, ds \\ &\leq |P_T|_{L^\infty_{\mathcal{S}}(\Omega, L(H))}\mathbb{E}|y^{t,x,0}(T)|_H^2 + \int_t^T |S(s)|_{L^\infty_{\mathcal{S}}(\Omega, L(H))}\mathbb{E}|y^{t,x,0}(s)|_H^2 \, ds\end{aligned}$$

and by (3.1):

$$(P(t)x, x)_H \leq C_2[|P_T|_{L^\infty_{\mathcal{S}}(\Omega; L(H))} + |S|_{L^1_{\mathcal{P},S}([0,T]; L^\infty(\Omega; L(H)))}] \quad \mathbb{P}-a.s. \quad \forall x : |x|_H \leq 1 \tag{5.14}$$

Then consider the following *closed loop* equation, starting at a certain instant $t \geq T_0$ with an arbitrary initial data $x \in H$:

$$\left\{\begin{array}{l} d\overline{y}(r) = [A\overline{y}(r) + A_\sharp(r)\overline{y}(r) - B(r)B^*(r)P(r)\overline{y}(r)]\, dr + C(r)\overline{y}(r)\, dW(r) \\ \overline{y}(t) = x \end{array}\right. \tag{5.15}$$

Notice that if we replace $A_\sharp$ by $A_\sharp - BB^*P$ then assumptions of Theorem 3.2 still hold. Thus there exists a unique solution $\overline{y} \in L^p_{\mathcal{P}}(\Omega, C([t, T]; H))$ for every $p \geq 2$. Applying then the fundamental relation (5.13) to $\overline{u} = -B^*P\overline{y}$ and consequently to $y^{t,x,\bar{u}} = \overline{y}$ we get

$$(P(t)x, x)_H = \mathbb{E}^{\mathcal{F}_t}(P_T\overline{y}(T), \overline{y}(T))_H + \mathbb{E}^{\mathcal{F}_t}\int_t^T [|\sqrt{S}(r)\overline{y}(r)|_H^2 + |B^*(r)P(r)\overline{y}(r)|_H^2]\, dr \tag{5.16}$$

thus $(P(t)x, x)_H \geq 0$, $\mathbb{P}$-a.s. for all $x \in H$ and this together with (5.14) gives the claim. □

We summarize the content of the section in the following result:

Theorem 5.13. *Assume A1)-A5). Problem* (5.11) *has a unique mild solution* (P, Q) *with the following regularity:* $P \in L^2_{\mathcal{P}}(\Omega; C([0, T]; \Sigma_2^+(H))) \cap L^\infty_{\mathcal{P},S}(\Omega; C([0, T]; \Sigma^+(H)))$ *and* $Q \in L^2_{\mathcal{P}}(\Omega \times [0, T]; L_2(\Xi; \Sigma_2(H)))$.

Proof. The a priori estimate in Proposition 5.12 allows us to apply the local existence result in Proposition 5.8 recursively in time intervals of fixed length (see also Remark 5.10), to obtain a global solution of equation (5.11). Indeed, let $\tilde{M}_P = C_2(M_P + M_S)$, then it is enough to choose $\tilde{r}$ such that $\tilde{r} > C_2(\tilde{M}_P + M_S)$ and $\tilde{\delta}$ such that:

$$C_2[\tilde{M}_P + \tilde{r}^2\tilde{\delta}M_B^2 + M_S] \leq \tilde{r}$$

Then we can iterate the procedure in $[T - n\tilde{\delta}, T - (n-1)\tilde{\delta}]$ for a finite number of $n \geq 1$ until we cover the whole interval $[0, T]$. □

Now we are ready to solve the finite horizon problem in a standard way:

Theorem 5.14. *Fix* $T > 0$ *and* $x \in H$. *Then:*

(1) *There exists a unique* optimal control. *That is a unique control* $\bar{u} \in L^2_{\mathcal{P}}(\Omega \times [0, T]; U)$ *such that:*

$$J(0, x, \bar{u}) = \inf_{u \in L^2_{\mathcal{P}}(\Omega \times [0,T]; U)} J(0, x, u)$$

(2) *If* $\bar{y}$ *is the mild solution of the state equation corresponding to* $\bar{u}$ *(that is the* optimal state*) then* $\bar{y}$ *satisfies the* closed loop *equation*

$$\left\{\begin{array}{l} d_s\bar{y}(s) = [A\bar{y}(s) + A_\sharp(s)\bar{y}(s) - B(s)B(s)^*P(s)\bar{y}(s)]\, ds + C(s)\bar{y}(s)\, dW(s) \\ \bar{y}(0) = x \end{array}\right. \tag{5.17}$$

(3) *The following feedback law holds* $\mathbb{P}$*-a.s. for almost every* s.

$$\bar{u}(s) = -B^*(s)P(s)\bar{y}(s). \tag{5.18}$$

(4) *The optimal cost is given by* $\mathbb{E}J(0, x, \bar{u}) = \mathbb{E}(P(0)x, x)_H$, *for all* $x \in H$.

Proof. Let (P, Q) be the unique mild solution to Riccati equation (5.11). Relation (5.13) becomes

$$J(0, x, u) = (P(0)x, x)_H + \mathbb{E}\int_0^T |u(s) + B^*P(s)y^{t,x,u}(s)|^2\, ds$$

Thus $J(0, x, u) \geq (P(0)x, x)_H$ for all $u \in L^2_{\mathcal{P}}(\Omega \times [0, T]; U)$ and the equality holds if and only if (5.18) holds, that is if and only if y solves (5.17) and $u = \bar{u}$. □

6. The general case

In order to get rid of assumption A5) we introduce the following new notion of solution:

Definition 6.1. A process $P \in L^\infty_{\mathcal{P},S}(\Omega \times [0,T]; \Sigma^+(H))$, is a *generalized solution* if there exists a sequence (S^N, P^N, Q^N) where:

(i) $S^N \in L^1_{\mathcal{P},S}([0,T]; L^\infty(\Omega; \Sigma^+(H)) \cap L^2_{\mathcal{P}}(\Omega \times [0,T]; \Sigma_2(H))$ and there exists a positive function $c \in L^1([0,T])$ such that $|S^N(t)|_{L(H)} \leq c(t)$, for all $N \in \mathbb{N}$, $\mathbb{P}$-a.s. for a.e. $t \in [0,T]$.

(ii) the pair (P^N, Q^N) is a mild solution to the Riccati equation (5.11) in the space of Hilbert Schmidt operators, with forcing term S^N and final data $P^N_T = P^N(T)$. Namely (P^N, Q^N) is the unique mild solution of:

$$\begin{cases} -dP^N(t) = (A^* P^N(t) + P^N(t) A + \mathrm{Tr}[C^*(t) P^N(t) C(t) + C^*(t) Q^N(t) + Q^N(t) C(t)])\, dt \\ \\ +(A^*_\sharp(t) P^N(t) + P^N(t) A_\sharp(t) - P^N(t) B(t) B^*(t) P^N(t) + S^N(t))\, dt + Q^N(t)\, dW(t), \quad t \in [0,T], \\ \\ P^N(T) = P^N_T \end{cases}$$

such that:

(iii) for all $x \in H$:
$$S^N(t,\omega)x \to S(t,\omega)x \text{ in } H \quad \mathbb{P} \text{ a.s. for a.e. } t \in [0,T]$$

(iv) for every $t \in [0,T]$ and for all $x \in H$:
$$P^N(t,\omega)x \to P(t,\omega)x \quad \text{in } H \qquad \mathbb{P} \text{ a.s.}$$

Remark 6.2. Although in the definition the value of $P(t)$ seem determined for a.e. t, point (iv) in the definition implies that there exists a version such that for all $t \in [0,T]$ and for all $x \in H$ the value of $P(t)x$ is determined $\mathbb{P}$-a.s.. Actually we will show extra regularity property for the generalized solution if evaluated at a vector $x \in H$.

Remark 6.3. In the previous definition only the process P in the Riccati equation is characterized. By one side this is natural for control theory since Q is nor involved in the expression for the optimal cost neither in the expression for the optimal feedback law (see Theorem 6.6). On the other side it is a general feature of backward stochastic differential equations that the martingale representation term is only an auxiliary variable that can be determined computing the joint quadratic variation between the other unknown process and the noise, see [18].

We start by showing some regularity properties of the generalized solutions:

Lemma 6.4. *Every generalized solution fulfills the fundamental relation, i.e. for all $t \in [0,T]$, for all $x \in H$ and for all $u \in L^2_{\mathcal{P}}(\Omega \times [0,T]; U)$:*

$$(P(t)x, x)_H = J(t,x,u) - \mathbb{E}^{\mathcal{F}_t} \int_t^T |u(s) + B^*(s) P(s) y(s)|^2\, ds \qquad \mathbb{P} - \ a.s. \tag{6.1}$$

Proof. At each fixed N the pair (P^N, Q^N) is the mild solution of the Riccati equation, therefore by Proposition 5.11, we have that for all $t \in [0,T]$ and $x \in H$:

$$(P^N(t)x, x)_H = \mathbb{E}^{\mathcal{F}_t}(P^N_T y(T), y(T))_H + \mathbb{E}^{\mathcal{F}_t} \int_t^T [|\sqrt{S^N}(s) y(s)|^2 + |u(s)|^2]\, ds$$
$$- \mathbb{E}^{\mathcal{F}_t} \int_t^T |u(s) + B^*(s) P^N(s) y(s)|^2\, ds \qquad \mathbb{P} - \text{a.s.}$$

Now, we have to pass to the limit as $N \to \infty$ in the identity. We notice that if we show that the right-hand side converges in mean to

$$\mathbb{E}^{\mathcal{F}_t}(P_T y(T), y(T))_H + \mathbb{E}^{\mathcal{F}_t} \int_t^T [|\sqrt{S}(s) y(s)|^2 + |u(s)|^2]\, ds - \mathbb{E}^{\mathcal{F}_t} \int_t^T |u(s) + B^*(s) P(s) y(s)|^2\, ds$$

then the proof is completed just by choosing a subsequence on which convergence occurs $\mathbb{P}$ a.s..

Coming now to the proof of convergence, considering for instance the second term, by Jensen inequality it is enough to show that

$$\lim_{N \to \infty} \mathbb{E}| \int_t^T ((S - S^N)(s) y(s), y(s))_H ds| = 0.$$

Applying a first time dominated convergence theorem we get $\mathbb{E}|((S - S^N)(t) y(t), y(t))_H| \to 0$ for all fixed $s \in [t,T]$. Then we notice that

$$\mathbb{E}|((S - S^N)(s) y(s), y(s))_H| \leq 2c(s) \mathbb{E}|y(s)|^2$$

where the map c is in $L^1([0,T])$ and the map $s \to \mathbb{E}|y(s)|^2$ is in $C([0,T];\mathbb{R})$. Thus we can apply a second time dominated convergence theorem to obtain the claim. Since the other terms can be treated in an identical way the proof is completed. □

Lemma 6.5. *Let $P(t)$ be any generalized solution, then $P(t)x \in C_{\mathcal{P}}([0,T];L^p(\Omega;H))$ for all $x \in H$ and for all $p \geq 2$.*

Proof. We have to prove that, for all $x \in H$, $\lim_{r\to t}\mathbb{E}|[P(r)-P(t)]x|_H^p = 0$ or, what is equivalent

$$\lim_{r\to t}\mathbb{E}|(P(r)x,x)_H - (P(t)x,x)_H|^p = 0 \qquad \forall x \in H$$

Let us consider the state equation corresponding to $u = 0$, then for all $x \in H$ the following holds $\mathbb{P}-$a.s.:

$$(P(t)x,x)_H = \mathbb{E}^{\mathcal{F}_t}(P_T y(T), y(T))_H + \mathbb{E}^{\mathcal{F}_t}\int_t^T |\sqrt{S}(s)y(s)|^2\,ds - \mathbb{E}^{\mathcal{F}_t}\int_t^T |B^*(s)P(s)y(s)|^2\,ds$$

We set $y^{t,x} = y^{t,x,0}$ we recall that $y^{t,x} \in L^p_{\mathcal{P}}(\Omega;C([0,T];H))$ for all $p \geq 2$ (see. [7][Proposition 3.2]). Moreover the map $(t,x) \to y^{t,x}(\cdot)$ is continuous from $[0,T]\times H$ to $L^p_{\mathcal{P}}(\Omega;C([0,T];H))$, see again [7][Proposition 3.3].

Taking all these facts into account, we have that, for all $0 \leq t \leq r \leq T$:

$$\begin{aligned}\mathbb{E}|(P(r)x,x)_H - (P(t)x,x)_H|^p &\leq c(p)\mathbb{E}|\mathbb{E}^{\mathcal{F}_r}(P_T y(T),y(T))_H - \mathbb{E}^{\mathcal{F}_t}(P_T y(T),y(T))_H|^p \\ &+ \mathbb{E}|\mathbb{E}^{\mathcal{F}_r}\int_r^T |\sqrt{S(s)}y^{r,x}(s)|_H^2\,ds - \mathbb{E}^{\mathcal{F}_t}\int_t^T |\sqrt{S(s)}y^{t,x}(s)|_H^2\,ds|^p \\ &+ \mathbb{E}|\mathbb{E}^{\mathcal{F}_r}\int_r^T |B^*(s)P(s)y^{r,x}(s)|_H^2\,ds - \mathbb{E}^{\mathcal{F}_t}\int_t^T |B^*(s)P(s)^{t,x}y(s)|_H^2\,ds|^p\end{aligned} \tag{6.2}$$

Being $P_T \in L^\infty_{\mathcal{S}}(\Omega,\mathcal{F}_T,\mathbb{P},L(H))$ and consequently $(P_T y(T),y(T))_H \in L^p(\Omega,\mathcal{F}_T,\mathbb{P},\mathbb{R})$, for all $p \in [2,\infty[$ by the Kunita-Watanabe martingale representation Theorem there exists a process Z in $L^p_{\mathcal{P}}(\Omega;L^2([0,T];\Xi^*))$ such that:

$$\left|\mathbb{E}^{\mathcal{F}_r}(P_T y(T),y(T))_H - \mathbb{E}^{\mathcal{F}_t}(P_T y(T),y(T))_H\right| = \left|\int_r^t Z(s)\,dW(s)\right|$$

Now fix $\tau \in [0,T]$ by Burkholder-Davies-Gundy inequalities and the Dominated Convergence Theorem we get that:

$$\begin{aligned}&\lim_{r\uparrow\tau,\,t\downarrow\tau}\mathbb{E}|\mathbb{E}^{\mathcal{F}_r}(P_T y(T),y(T))_H - \mathbb{E}^{\mathcal{F}_t}(P_T y(T),y(T))_H|^p = \lim_{r\downarrow\tau,\,t\uparrow\tau}\mathbb{E}\Big|\int_t^r Z(s)\,dW(s)\Big|^p \\ &\leq \lim_{r\downarrow\tau,\,t\uparrow\tau}\mathbb{E}\Big(\int_t^r |Z(s)|_{\Xi^*}^2\,ds\Big)^{p/2} = 0\end{aligned} \tag{6.3}$$

Let us consider the second term at the right hand side in (6.2):

$$\begin{aligned}&\lim_{r\downarrow\tau,\,t\uparrow\tau}\mathbb{E}|\mathbb{E}^{\mathcal{F}_r}\int_r^T |\sqrt{S(s)}y^{r,x}(s)|_H^2\,ds - \mathbb{E}^{\mathcal{F}_t}\int_t^T |\sqrt{S(s)}y^{t,x}(s)|_H^2\,ds|^p \\ &\leq \lim_{r\downarrow\tau,\,t\uparrow\tau}\mathbb{E}|\mathbb{E}^{\mathcal{F}_r}\int_r^T |\sqrt{S(s)}y^{r,x}(s)|_H^2\,ds - \mathbb{E}^{\mathcal{F}_r}\int_t^T |\sqrt{S(s)}y^{t,x}(s)|_H^2\,ds|^p \\ &+ \lim_{r\downarrow\tau,\,t\uparrow\tau}\mathbb{E}|[\mathbb{E}^{\mathcal{F}_r}-\mathbb{E}^{\mathcal{F}_t}]\int_t^T |\sqrt{S(s)}y^{t,x}(s)|_H^2\,ds|^p\end{aligned} \tag{6.4}$$

Setting $y^{r,x}(s) = x$ for $s \in [t,r]$ splitting first the above expression and then applying Jensen inequality we get:

$$\begin{aligned}&\lim_{r\downarrow\tau,\,t\uparrow\tau}\mathbb{E}|\mathbb{E}^{\mathcal{F}_r}\int_r^T |\sqrt{S(s)}y^{r,x}(s)|_H^2\,ds - \mathbb{E}^{\mathcal{F}_r}\int_r^T |\sqrt{S(s)}y^{t,x}(s)|_H^2\,ds + \mathbb{E}^{\mathcal{F}_r}\int_t^r |\sqrt{S(s)}y^{t,x}(s)|_H^2\,ds|^p \\ &\leq c(p)[\lim_{r\downarrow\tau,\,t\uparrow\tau}\mathbb{E}|\int_r^T (|\sqrt{S(s)}[y^{r,x}(s)-y^{t,x}(s)]|_H^2)\,ds|^p + \lim_{r\downarrow\tau,\,t\uparrow\tau}\mathbb{E}|\mathbb{E}^{\mathcal{F}_r}\int_t^r |\sqrt{S(s)}y^{t,x}(s)|_H^2\,ds|^p] \\ &\leq C(T,p,M_S)\lim_{r\downarrow\tau,\,t\uparrow\tau}\int_t^T \mathbb{E}\sup_{s\in[t,T]}(|[y^{r,x}(s)-y^{t,x}(s)]|_H^{2p})\,ds + C(T,p)\lim_{r\downarrow\tau,\,t\uparrow\tau}\mathbb{E}\int_t^r |\sqrt{S(s)}y^{t,x}(s)|_H^{2p}\,ds]\end{aligned}$$

Therefore by the Dominated Convergence Theorem we get that:

$$\lim_{r\downarrow\tau,\,t\uparrow\tau}\mathbb{E}|\mathbb{E}^{\mathcal{F}_r}\int_r^T |\sqrt{S(s)}y^{r,x}(s)|_H^2\,ds - \mathbb{E}^{\mathcal{F}_r}\int_t^T |\sqrt{S(s)}y^{t,x}(s)|_H^2\,ds|^p = 0$$

The second term at the right hand side of (6.4) can be treated like the term $\mathbb{E}^{\mathcal{F}_r}(P_T y(T),y(T))_H$ in (6.3). The third term in (6.2) follows identically as the second term in (6.2).

This conclude the proof of the Lemma. □

Now we can state the main result of the paper:

Theorem 6.6. *Assume that hypotheses A1)-A4) hold true.*
Then there exists a unique generalized solution of problem (5.11).

Moreover we have the following characterization of the optimal control: fix $T > 0$ and $x \in H$, then:

(1) *there exists a unique control $\overline{u} \in L^2_{\mathcal{P}}(\Omega \times [0,T]; U)$ such that:*

$$J(0,x,\overline{u}) = \inf_{u \in L^2_{\mathcal{P}}(\Omega \times [0,T];U)} J(0,x,u)$$

(2) *If $\overline{y}$ is the mild solution of the state equation corresponding to $\overline{u}$ (that is the optimal state), then $\overline{y}$ is the unique mild solution to the* closed loop *equation:*

$$\begin{cases} d\overline{y}(r) = [A\overline{y}(r) + A_\sharp(r)\overline{y}(r) - B(r)B^*(r)P(r)\overline{y}(r)]\,dr + C\overline{y}(r)\,dW(r) \\ \overline{y}(0) = x \end{cases} \tag{6.5}$$

(3) *The following feedback law holds $\mathbb{P}$-a.s. for almost every s.*

$$\overline{u}(s) = -B^*(s)P(s)\overline{y}(s). \tag{6.6}$$

(4) *The optimal cost is given by $J(0,x,\overline{u}) = (P(0)x,x)_H$.*

Proof. We divide the proof in three steps.

First step: existence of a generalized solution We fix a complete orthonormal basis $\{e_i : i \in \mathbb{N}\}$ in H and introduce, for each $N \in \mathbb{N}$, the finite dimensional projections $\Pi_N : H \to H : v \to \sum_{i=1}^N (v, e_i)_H e_i$. For each $N \in \mathbb{N}$ we define at (t,ω) fixed:

$$\Pi_N P_T(\omega)\Pi_N = P_T^N(\omega) \qquad \text{and} \qquad S^N(t,\omega) = \begin{cases} \Pi_N S(t,\omega)\Pi_N & |S(t,\omega)|_{L(H)} \le N \\ 0 & |S(t,\omega)|_{\Sigma(H)} > N \end{cases} \tag{6.7}$$

First of all we notice that, from their definition $P_T^N \in \Sigma_2^+(H)$, $\mathbb{P}$-a.s., for all $N \in \mathbb{N}$ and that for all $x \in H$, $(P_T^N x, x)_H \nearrow (P_T x, x)_H$, $\mathbb{P}$-a.s.. Again from their definition it follows that $S^N(t,\omega) \in \Sigma_2^+(H)$, for all $N \in \mathbb{N}$ and that for all $x \in H$, $(S^N(t)x,x)_H \nearrow (S(t)x,x)_H$, $\mathbb{P}$-a.s. for a.e. $t \in [0,T]$. Moreover $|S^N(t)|_{\Sigma(H)} \le |S(t)|_{\Sigma(H)}\mathbb{P}$-a.s. for a.e. $t \in [0,T]$, so in particular (i) and (iii) in definition 6.1 are verified. The pair (P_T^N, S^N) will become the data of the following approximating problems:

$$\begin{cases} -dP^N(t) = (A^*P^N(t) + P^N(t)A + A_\sharp^*(t)P^N(t) + P^N(t)A_\sharp(t) - P^N(t)B(t)B^*(t)P^N(t))\,dt \\ \quad +S^N(t)dt + \mathrm{Tr}[C^*(t)P^N(t)C(t) + C^*(t)Q^N(t) + Q^N(t)C(t)]\,dt + Q^N(t)\,dW(t) \quad t \in [0,T] \\ P^N(T) = P_T^N \end{cases} \tag{6.8}$$

We notice that the above equation satisfies the assumptions of Theorem 5.13. Thus for each fixed $N \in \mathbb{N}$ there exists a mild solution (P^N, Q^N) with $P^N \in L^2_{\mathcal{P}}(\Omega; C([0,T]; \Sigma_2^+(H))) \cap L^\infty_{\mathcal{P},S}(\Omega; C([0,T]; \Sigma^+(H)))$ and $Q^N \in L^2_{\mathcal{P}}(\Omega \times [0,T]; L_2(\Xi; \Sigma_2(H)))$. Notice that P^N has strongly continuous trajectories so the final condition $P^N(T) = P_T^N$, $\mathbb{P}$-.a.s. is attained.

Points (i)-(iii) in definition 6.1 are satisfied by construction we only have to show that (iv) holds true.

We fix $t \in [0,T]$, by Theorem 5.14:

$$(P^N(t)x,x)_H = \inf\{J^N(t,x,u) : u \in L^2_{\mathcal{P}}(\Omega \times [0,T]; U)\}$$

where

$$J^N(t,x,u) = \mathbb{E}^{\mathcal{F}_t} \int_t^T \left[\left| \sqrt{S^N(s)}y^{t,x,u}(s) \right|_H^2 + |u(s)|^2 \right] ds + \left| \sqrt{P_T^N} y^{t,x,u}(T) \right|^2$$

Clearly at each $u \in L^2_{\mathcal{P}}(\Omega \times [0,T]; U)$ fixed, we have that, for every $t \in [0,T]$ and $x \in H$ the sequence $\{J^N(t,x,u) : N \in \mathbb{N}\}$ is $\mathbb{P}$-a.s. non decreasing. Moreover it is bounded by $J(t,x,u) < +\infty$ $\mathbb{P}$-a.s.. Thus the sequence of random variables $(P^N(t)x,x)_H = \inf_{u \in L^2_{\mathcal{P}}(\Omega \times [0,T];U)} J^N(t,x,u)$ is non decreasing as well. Since it is $\mathbb{P}$-a.s. bounded it has a limit. It lasts to show that this limit is actually of the form $(P(t)x,x)_H$ with $P \in L^\infty_{\mathcal{P},S}(\Omega \times [0,T]; \Sigma^+(H))$.

Let $\mathcal{D} \doteq \{x_i\}_{i \in \mathbb{N}}$ be a dense subset of H, then we can find a subset $\Omega_0 \subset \Omega$, with $P(\Omega_0) = 1$, such that for every $x_i \in \mathcal{D}$, $\exists \lim_{N \to +\infty} (P^N(t)x_i, x_i)_H$, for every $\omega \in \Omega_0$. Thus we define the limit $\phi(t, x_i, x_i)$ as follows:

$$\phi(t,x_i,x_i) = \begin{cases} \lim_{N \to +\infty} (P^N(t)x_i,x_i)_H & \forall x_i \in \mathcal{D} & \text{if } \omega \in \Omega_0 \\ 0 & \forall x_i \in \mathcal{D} & \text{if } \omega \notin \Omega_0 \end{cases}$$

For every $\omega \in \Omega_0$ the quadratic functional $\phi(t,x_i,x_i)$ defines a continuous, positive semidefinite, quadratic form on a dense subset, indeed thanks to (5.8) one has the following uniform bound, modifying if necessary Ω_0:

$$|\phi(t,x_i,x_i)| \le \sup_{N \in \mathbb{N}} |P^N(t)|_{\Sigma(H)} |x_i|_H^2 \le C_2(M_p + M_s)|x_i|_H^2 \qquad \forall \omega \in \Omega_0$$

therefore $\phi(t, x_i, x_i)$ can be extended to the whole $H \times H$, by density. Moreover, by the Riesz Theorem we can associate to this quadratic form, a linear, bounded symmetric and positive semidefinite operator $P(t)$, such that for every $t \in [0, T]$:

$$\phi(t, x, y) = (P(t)x, y)_H \qquad \forall \omega \in \Omega_0, \qquad \forall x, y \in H$$

And the following uniform bound is valid for all $t \in [0, T]$:

$$|P(t)|_{L(H)} \leq C_2[M_P + M_S] \qquad \forall \omega \in \Omega_0 \tag{6.9}$$

The process P is by construction predictable and strongly measurable. Finally, being the operators positive and symmetric the weak convergence imply also the strong convergence, for all $t \in [0, T]$ we have that

$$|P^N(t)x - P(t)x|_H \to 0 \qquad \forall \omega \in \Omega_0, \quad \forall x \in H \tag{6.10}$$

This concludes the proof of the *first step* since (6.10) implies (iv) in the definition 6.1 .

Second step: characterization of the optimal control By Lemma 6.4 we know that any generalized solution verifies (6.1), in particular for $t = 0$ we have that:

$$J(0, x, u) = (P(0)x, x)_H + \mathbb{E} \int_0^T |u(s) + B^*(s)P(s)y^{t,x,u}(s)|^2 \, ds$$

Thus $J(0, x, u) \geq (P(0)x, x)_H$ for all $u \in L^2_{\mathcal{P}}(\Omega \times [0, T]; U)$ and the equality holds if and only if (6.6) holds, that is if and only if y solves (6.5) and $u = \overline{u}$. This completely characterize the optimal control. We notice that existence and uniqueness of a solution to the closed loop equation (6.5) is guaranteed since the assumptions of Theorem 3.2 are satisfied if $A_\sharp$ is replaced by $A_\sharp - BB^*P$.

Third step: uniqueness of the generalized solution

Let P_1 and P_2 be two generalized solutions. We choose $\bar{u} = -B^* P_1 \bar{y}$ where $\bar{y}$ solves (6.5) with P replaced by P_1. By the fundamental relation (6.1) and the characterization of the optimal control proved above we immediately have:

$$(P_1(t)x, x)_H = J(t, x, \bar{u}) = (P_2(t)x, x)_H + \mathbb{E}^{\mathcal{F}_t} \int_t^T |\bar{u}(s) + B^*(s)P_2(s)\bar{y}(s)|^2 \, ds \qquad \mathbb{P}\text{-a.s.}$$

Thus $(P_1(t)x, x)_H \geq (P_2(t)x, x)_H$ $\mathbb{P}$-a.s.. The claim follows repeating the argument choosing P_2 instead of P_1 and repeating the argument. □

Remark 6.7. The idea of regularizing the data and then defining a generalized notion of solution is rather classical in the PDE context, see for instance the definition of “strong solution” in [16] an in [1] and the references therein, although it seems the first time that is used in the context of the Riccati equation.

Remark 6.8. If assumptions A1)-A5) in Hypothesis 2.1 hold then comparing relation (6.1) and relation (5.13) we immediately deduce that generalized solutions of (5.11) are mild solutions of (5.11) (see Section 5 for definition of mild solutions) and viceversa. On the contrary, when A5) fails, generalized solutions still exist, and are unique, while mild solutions can not be defined.

7. Generalized solutions and variation of constant formula

The aim of this section is to give a further characterization of the generalized solution just defined. To this purpose we notice that the state equation defines an evolution operator in a suitable sense and we recover a variation of constant formula for the value function. The main ingredient is the fundamental relation (6.1) that, on one hand is verified by the generalized solution and on the other hand will turn out to be essential to define the evolution operator.

Definition 7.1. Assume A1)-A3) and consider the state equation starting from x at time $t \in [0, T]$ and with control $u = 0$ namely:

$$\begin{cases} dy(s) = (Ay(s) + A_\sharp(s)y(s))\, ds + C(s)y(s)\, dW(s) & s \in [t, T] \\ y(t) = x \end{cases}$$

We denote by $y^{t,x}$ its mild solution and define the family of maps $L_{t,\sigma} : L^\infty_S(\Omega, \mathcal{F}_\sigma; \Sigma(H)) \to L^\infty_S(\Omega, \mathcal{F}_t; \Sigma(H))$ for $0 \leq t \leq \sigma \leq T$, in the following way. For every $V \in L^\infty_S(\Omega, \mathcal{F}_\sigma; \Sigma(H))$ we define,

$$(L_{t,\sigma} V x, x)_H = \mathbb{E}^{\mathcal{F}_t}(V y^{t,x}(\sigma), y^{t,x}(\sigma))_H. \tag{7.1}$$

We collect some properties for the evolution operator $L_{t,\sigma}$, that can be easily deduced from its definition:

Lemma 7.2. *The family of operators $\{L_{t,\sigma} : 0 \leq t \leq \sigma \leq T\}$ has the following properties:*

(1) *for every $0 \le t \le \sigma \le T$ $L_{t,\sigma}$ is a linear and bounded operator,*

$$L_{t,\sigma} : L^\infty_S(\Omega, \mathcal{F}_\sigma; \Sigma(H)) \to L^\infty_S(\Omega, \mathcal{F}_t; \Sigma(H))$$

(2) *for every $0 \le t \le r \le \sigma \le T$ one has that:*

$$L_{t,\sigma} = L_{t,r} \circ L_{r,\sigma} \qquad \mathbb{P} - a.s.$$

(3) *fixed $V \in L^\infty_S(\Omega, \mathcal{F}_\sigma; \Sigma(H))$, $x \in H$ and σ in $[0,T]$ the map $t \to (L_{t,\sigma}Vx, x)_H$ belongs to $C_{\mathcal{P}}([0,T]; L^p(\Omega, \mathbb{R}))$, for all $p \ge 2$.*

Proof.

(1) We have that:

$$\sup_{x\in H, |x|_H \le 1} |(L_{t,\sigma}Vx, x)_H = \sup_{x\in H, |x|_H \le 1} |\mathbb{E}^{\mathcal{F}_t}(Vy^{t,x}(\sigma), y^{t,x}(\sigma))_H| \le$$
$$\le |V|_{L^\infty_S(\Omega,\mathcal{F}_\sigma;L(H))} \sup_{x\in H, |x|_H \le 1} \mathbb{E}|y^{t,x}(\sigma)|^2_H \le C_2 |V|_{L^\infty_S(\Omega,\mathcal{F}_\sigma;L(H))}$$

(2) the proof follows form the semigroup property of the solution $y^{t,x}(\sigma)$ and the property of conditional expectations with respect to the filtration $\mathcal{F}_t$.

(3) the proof is identical to the one of Lemma 6.5.

□

We notice that the fundamental relation (6.1), evaluated at $u = 0$, can be rewritten in terms of the evolution operator and reads as follows, for all $t \in [0,T]$ and all $x \in H$:

$$\begin{aligned} &(P(t)x, x)_H = \\ &= \mathbb{E}^{\mathcal{F}_t}(P_T y(T), y(T))_H + \mathbb{E}^{\mathcal{F}_t}\int_t^T |\sqrt{S}(s)y(s)|^2\, ds - \mathbb{E}^{\mathcal{F}_t}\int_t^T |B^*(s)P(s)y(s)|^2\, ds \\ &= (L_{t,T}P_T x, x)_H + \int_t^T (L_{t,s}S(s)x, x)_H\, ds - \int_t^T (L_{t,s}P(s)B(s)B^*(s)P(s)x, x)_H\, ds \quad \mathbb{P} - \text{a.s.} \end{aligned} \tag{7.2}$$

This relation suggest a new characterization for a solution of (5.11):

Definition 7.3. A *variation of constants* solution to problem (5.11) is a map $P \in L^\infty_{\mathcal{P},S}((0,T)\times\Omega; \Sigma^+(H))$ such that, $\forall x \in H$, $\forall p \ge 1$, $Px \in C_{\mathcal{P}}([0,T]; L^p(\Omega; H))$ and the following variation of constant formula is verified, $\mathbb{P}$-a.s.:

$$\begin{aligned} &(P(t)x, x)_H \\ &= (L_{t,T}P_T x, x)_H + \int_t^T (L_{t,s}S(s)x, x)_H\, ds - \int_t^T (L_{t,s}P(s)B(s)B^*(s)P(s)x, x)_H\, ds \end{aligned} \tag{7.3}$$

We can prove existence and uniqueness of such solutions:

Theorem 7.4. *Assume A1)-A4) then there exists a unique solution of problem* (5.11) *in the sense of definition 7.3.*

Proof. We already know that the generalized solution defined in the previous section verifies (7.3) and it is regular enough to be a solution in the sense of definition 7.3. It lasts to prove the uniqueness of the solution in this class. Let P_1 and P_2 be two solutions in the sense of 7.3 and denote by $\bar{P}$ their difference $\bar{P}(t) = P_1(t) - P_2(t)$. Then the following holds, for all $t \in [0,T]$ and $x \in H$, $\mathbb{P}$ − a.s:

$$(\bar{P}(t)x, x)_H = \int_t^T (L_{t,s}P_2(s)B(s)B^*(s)\bar{P}(s)x, x)_H + \int_t^T (L_{t,s}\overline{P}(s)B(s)B^*(s)P_1(s)x, x)_H\, ds$$

Therefore, for every $t \in [0,T]$ we have that:

$$\begin{aligned} &|\overline{P}(t)|_{L(H)} = \sup_{x\in H, |x|_H \le 1} (\overline{P}(t)x, x)_H \le \\ &\le \sup_{x\in H, |x|_H \le 1} \int_t^T (L_{t,s}P_2(s)B(s)B^*(s)\overline{P}(s)x, x)_H + \int_t^T (L_{t,s}\overline{P}(s)B(s)B^*(s)P_1(s)x, x)_H\, ds \le \\ &\le C\int_t^T |\overline{P}(s)|_{L(H)}\, ds \quad \mathbb{P} - \text{a.s} \end{aligned}$$

where C depends on C_2, M_B, $|P_1|_{L^\infty_{\mathcal{P},S}(\Omega\times[0,T];L(H))}$ and $|P_2|_{L^\infty_{\mathcal{P},S}(\Omega\times[0,T];L(H))}$

Applying the Gronwall Lemma to $s \to |\overline{P}(s)|_{L^\infty(\Omega,\mathcal{F}_s,L(H))}$ we get that $P_1(t) = P_2(t)$, $\mathbb{P}$-a.s. for all $t \in [0,T]$. □

Remark 7.5. Since the solution of definition 7.3 is also the unique generalized solution it is obvious that it allows the synthesis of optimal controls as in Theorem 6.6.

□

8. Non-homogeneous problem

As in [13] we consider a simple generalization of our original control problem that enlarges the set of applicability of our abstract results.

We fix $\eta \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}, H)$ and $\nu \in L^2_{\mathcal{P}}(\Omega \times [0,T], H)$ and instead of $J(0,x,u)$ we minimize

$$\begin{aligned}\widehat{J}(0,x,u) &= \mathbb{E}\int_0^T \left((S(s)(y^{0,x,u}(s)-\nu(s)),(y^{0,x,u}(s)-\nu(s)))_H + |u(s)|_U^2\right)\, ds \\ &\quad +\mathbb{E}(P_T(y^{0,x,u}(T)-\eta),(y^{0,x,u}(T)-\eta))_H\end{aligned}$$

We assume that A1)-A4) in Hypothesis 2.1 hold and let P be the unique generalized solution of the Riccati equation (5.11). Moreover (p,q) with p in $L^2_{\mathcal{P}}(\Omega, C([0,T];H))$ and q in $L^2_{\mathcal{P}}(\Omega, L^2([0,T];L_2(\Xi,H)))$ is the unique mild solution of the backward equation:

$$\left\{\begin{array}{rcl} dp(s) &=& \left(-A^*p(s) - A^*_\sharp(s)p(s) + P(s)B(s)B^*(s)p(s) - \mathrm{Tr}[C^*(s)q(s)]\right) ds \\ && -S(s)\nu(s)ds + q(s)dW(s) \qquad\qquad s\in[0,T] \\ p(T) &=& P_T\eta \end{array}\right. \tag{8.1}$$

where expliciting notation with respect to the usual basis $\{f_i : i \in \mathbb{N}\}$ in Ξ

$$\mathrm{Tr}[C^*(s)q(s)] = \sum_{i=1}^{\infty} C_i^*(s)(q(s), f_i).$$

Moreover existence and uniqueness of a mild solution to equation (8.1) is guaranteed by Theorem 4.4 whose assumptions are easily verified.

The following is the analogue of Theorem 6.6:

Theorem 8.1. *Assume that hypotheses A1)-A4) hold true, then:*

1. *there exists a unique control* $\overline{u} \in L^2_{\mathcal{P}}(\Omega \times [0,T];U)$ *such that:*
$$\widehat{J}(0,x,\overline{u}) = \inf_{u \in L^2_{\mathcal{P}}(\Omega\times[0,T];U)} \widehat{J}(0,x,u)$$
2. *If* $\overline{y}$ *is the mild solution of the state equation corresponding to* $\overline{u}$ *(that is the optimal state), then* $\overline{y}$ *is the unique mild solution to the* closed loop *equation:*
$$\left\{\begin{array}{l} d\overline{y}(r) = [A\overline{y}(r) + A_\sharp(r)\overline{y}(r) - B(r)B^*(r)P(r)\overline{y}(r) + B(r)B^*(r)p(r)]\, dr + C\overline{y}(r)\, dW(r) \\ \overline{y}(0) = x \end{array}\right. \tag{8.2}$$
3. *The following feedback law holds* $\mathbb{P}$*-a.s. for almost every* s*.*
$$\overline{u}(s) = -B^*(s)P(s)\overline{y}(s) + B^*(s)p(s). \tag{8.3}$$
4. *The optimal cost is given by*
$$\widehat{J}(0,x,\overline{u}) = (P(0)x,x)_H - 2(p(0),x)_H + \mathbb{E}(P_T\eta,\eta)_H + \mathbb{E}\int_0^T \left((S(s)\nu(s),\nu(s))_H - |B^*(s)p(s)|_H^2\right) ds.$$

Proof. Let $(p_h, q_h) \in L^2_{\mathcal{P}}(\Omega; C([0,T];H)) \times L^2_{\mathcal{P}}(\Omega; L^2([0,T];L_2(\Xi,H)))$, $h = 1,2,...$ be the unique classical solution of the backward equation:

$$\left\{\begin{array}{rcl} dp(s) &=& \left(-A^*_hp_h(s) - A^*_\sharp(s)p_h(s) + P(s)B(s)B^*(s)p_h(s) - \mathrm{Tr}[C^*(s)q_h(s)]\right) ds \\ && -S(s)\nu(s)ds + q_h(s)dW(s) \qquad\qquad s\in[0,T] \\ p(T) &=& P_T\eta \end{array}\right. \tag{8.4}$$

We proceed as in the proof of Theorem 5.6. Namely we choose $\Psi \in C^2(H)$ with $\Psi(y) = 1$ for $|y| \leq 1$, $\Psi(y) = 0$ for $|y| \geq 2$ and $\Psi(y) \in [0,1], \forall y \in H$. Then we differentiate by Itô rule $d\Psi(N^{-1}y_h(s))(p_h(s), y_h(s))_H$. We integrate in $[0,T]$, and compute the mean value. Finally we let $N \to +\infty$ to obtain:

$$\begin{aligned}\mathbb{E}(P_T\eta, y_h(T))_H &= (p(0),x)_H + \mathbb{E}\int_0^T (P(s)B^*(s)B(s)p_h(s), y_h(s))_H ds \\ &\quad +\mathbb{E}\int_0^T (u(s), B^*(s)p_h(s))_H - (S(s)\nu(s), y_h(s))_H)\, ds\end{aligned}$$

Letting $h \to \infty$ we get by Theorems 3.3 and 4.4

$$\begin{aligned}\mathbb{E}(P_T\eta, y(T))_H &= (p(0), x)_H + \mathbb{E}\int_0^T (P(s)B^*(s)B(s)p(s), y(s))_H ds \\ &+\mathbb{E}\int_0^T (u(s), B^*(s)p(s))_H - (S(s)\nu(s), y(s))_H)\, ds\end{aligned}$$

Thus by easy computations

$$\begin{aligned}\widehat{J}(0,x,u) &= \mathbb{E}\int_0^T |u(s) + B^*(s)P(s)y(s) - B^*(s)p(s)|^2\, ds \\ &+(P(0)x,x)_H - 2(p(0),x)_H + \mathbb{E}(P_T\eta,\eta)_H + \mathbb{E}\int_0^T \Big[(S(s)\nu(s),\nu(s))_H - |B^*(s)p(s)|_H\Big] ds.\end{aligned}$$

The above relation completes the proof (notice that existence and uniqueness of the mild solution of equation 8.2 can be proved exactly as for existence and uniqueness of the mild solution of equation 2.1). □

9. Example: minimal variance problem for a stochastic equation with delay and random volatility

We consider the controlled stochastic differential equation with memory effects:

$$\begin{cases} d\xi(t) = \left[\int_{-1}^0 \xi(t+\theta)\, a(d\theta) + r(t)u(t)\right] dt + \sum_{i=1}^d \sigma_i(t)\xi(t)d\beta_t^i, \quad t \in [0,T] \\ \\ \xi(0) = \mu_0, \qquad \xi(\theta) = \nu_0(\theta), \qquad \text{for a.e. } \theta \in (-1,0), \end{cases} \tag{9.1}$$

where $\mu_0 \in \mathbb{R}^n$, $\nu_0 \in L^2((-1,0);\mathbb{R}^n)$, $(\Omega, \mathcal{E}, \mathcal{P})$ is a complete probability space, $\{\beta_t^i : t \geq 0,\ i = 1, ..., d\}$ are independent standard brownian motions defined in Ω. Moreover $\mathcal{F}_t$ denotes the σ-algebra generated by $\{\beta_\sigma^i,\ \sigma \in [0,t],\ i = 1,..,d\}$ and augmented with the sets of $\mathcal{F}$ with $\mathbb{P}$-measure zero (see Remark 2.4 to see how this requirement can be relaxed, and notice that for some $i = 1, ..., d$, σ_i can be null).

We assume that a is a $L(\mathbb{R}^n, \mathbb{R}^n)$-valued finite measure on $[-1, 0]$, $r : [0,T] \times \Omega \to L(\mathbb{R}^d, \mathbb{R}^n)$ is bounded and predictable stochastic process and that $\sigma_i : [0,T]\times\Omega \to L(\mathbb{R}^n, \mathbb{R}^n)$ are bounded and predictable stochastic processes, $i = 1, ..., d$.

We also consider the following cost functional of *minimal variance* type:

$$J(0, \mu_0, \nu_0, u) = \mathbb{E}\int_0^T |u(\tau)|^2_{\mathbb{R}^d} d\tau + \mathbb{E}(k(\xi(T) - \zeta), (\xi(T) - \zeta))_{\mathbb{R}^n}$$

where $k \in L^\infty(\Omega, \mathcal{F}_T, \mathbb{P}; \Sigma^+(\mathbb{R}^n))$ and $\zeta \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}; \mathbb{R}^n)$.

Our purpose is to minimize $J(0, \mu_0, \nu_0, u)$ over all predictable controls $u : [0,T] \times \Omega \to \mathbb{R}^d$.

Following [4] and [6] we set $H = \mathbb{R}^n \times L^2((-1,0);\mathbb{R}^n)$,

$$\mathcal{D}(A) = \left\{\begin{pmatrix}\mu\\ \nu\end{pmatrix} \in H : \nu \in W^{1,2}((-1,0);\mathbb{R}^n) \text{ and } \nu(0) = \mu\right\}, \quad A\begin{pmatrix}\mu\\ \nu\end{pmatrix} = \begin{pmatrix}\int_{-1}^0 \nu(\theta)a(d\theta)\\ \frac{d\nu}{d\theta}\end{pmatrix}.$$

It is proved in [9], among other places, that A generates a strongly continuous semigroup in H (see also [6]). Moreover if we set $U = \mathbb{R}^d$ and for $t \in [0,T]$, $\mu \in \mathbb{R}^n$, $\nu \in L^2((-1,0);\mathbb{R}^n)$, $u \in \mathbb{R}^d$,

$$x = \begin{pmatrix}\mu_0\\ \nu_0\end{pmatrix}, \qquad B(t)u = \begin{pmatrix}r(t)u\\ 0\end{pmatrix}, \qquad C_i(t)\begin{pmatrix}\mu\\ \nu\end{pmatrix} = \begin{pmatrix}\sigma_i(t)\mu\\ 0\end{pmatrix},$$

$$P_T\begin{pmatrix}\mu\\ \nu\end{pmatrix} = \begin{pmatrix}k(t)\mu\\ 0\end{pmatrix}, \qquad \eta = \begin{pmatrix}\zeta\\ 0\end{pmatrix}, \qquad y(\tau) = \begin{pmatrix}\xi(\tau)\\ \xi_\tau(\cdot)\end{pmatrix}$$

where $x_\tau(\theta) = x(\theta + \tau),\ \tau \geq 0,\ \theta \in [-1, 0]$.

Then equation (9.1) is equivalent (see [4] and [6][Chapter 10]) to

$$\begin{cases} dy(\tau) = (Ay(\tau) + Bu_\tau)\ d\tau + \sum_{i=1}^d C_i(\tau)y(\tau)d\beta_\tau^i, \quad \tau \in [0,T], \\ y(0) = y_0. \end{cases}$$

Moreover the cost functional becomes:

$$\widehat{J}(0,x,u) = \mathbb{E}\int_0^T |u(\tau)|^2_{\mathbb{R}^d} ds + \mathbb{E}\left|\sqrt{P_T}(y(T) - \eta)\right|^2_H$$

Moreover it is easy to verify that Hypothesis 2.1 A1)-A5) hold. Thus Theorem 8.1 can be applied to obtain the synthesis of the optimal control. We notice that in this case the Riccati equation has a unique mild solution in the sense clarified by Definition 5.3.

Remark 9.1. We believe that the present example is interesting in his own because of its simplicity. Notice that the model is finite dimensional but the presence of a simple delay term and of the stochastic coefficient σ immediately requires to use backward stochastic Riccati equations in infinite dimensional spaces.

In addition it can be regarded as a first step towards realistic financial applications of the theory. Namely in [14] (see also [26]) the authors showed that the *mean variance hedging* problem for a (incomplete) *Black and Scholes market with stochastic volatility* can be treated as a singular linear quadratic control problem with stochastic coefficients. The solution of such problem requires to prove existence and uniqueness of the solution of a backward stochastic Riccati equation in finite dimensions. On the other side in [8] it was pointed out that *memory effects* can be introduced in the market model describing the evolution of the share prices by a delay equation. Thus the present example can be seen as a contribution towards the solution of the mean variance hedging problem for a market with stochastic volatility and memory effects. To deal with the realistic formulation of the problem it would be necessary to allow control dependent noise and singular costs. This complicates the form of the Riccati equation and requires careful mixing of the techniques developed in this paper to deal with infinite dimensional stochastic Riccati equations and of the ones developed in [14] and [24] to deal with singular control problems and control dependent noise. This will be the topis of a future work.

10. Example: Optimal control for a wave equation in random media with stochastic damping

In order to show that our general results can be applied to concrete controlled stochastic PDEs arising in applications we consider a stochastic wave equation with diffused control. We assume that the system is evolving in a random media and this influences its evolution in two ways: through a stochastic force of elastic type (the term $\sum_{i=1}^{\infty} c_i(t,\zeta)\xi(t,\zeta)d\beta_i(t)$ below) and through a stochastic damping (the term $\mu(t,\zeta)\partial_t\xi(t,\zeta)dt$ below). Notice that in this model it is natural to introduce the stochastic coefficient μ, moreover although only one coefficient is stochastic the use of backward stochastic Riccati equations is necessary to solve the optimal control problem.

We consider the state equation

$$\begin{cases} d_t\partial_t\xi(t,\zeta) = \Delta_\zeta\xi(t,\zeta)dt + b(t,\zeta)u(t,\zeta)dt + \mu(t,\zeta)\partial_t\xi(t,\zeta)dt + \displaystyle\sum_{i=1}^{\infty} c_i(t,\zeta)\xi(t,\zeta)d\beta_i(t), \quad \zeta\in\mathcal{D},\ t\in[0,T], \\ \xi(t,\zeta)=0, \qquad \zeta\in\partial\mathcal{D},\ t\in[0,T], \\ \xi(0,\zeta)=x_0(\zeta),\ \partial_t\xi(0,\zeta)=v_0(\zeta) \qquad \zeta\in\mathcal{D}, \end{cases} \tag{10.1}$$

and the cost functional

$$\begin{aligned} J(0,x,u) \;=\; & \mathbb{E}\int_0^T\int_{\mathcal{D}}\left[\kappa_1(t,\zeta)\xi^2(t,\zeta)+\kappa_2(t,\zeta)\left(\frac{\partial\xi}{\partial t}(t,\zeta)\right)^2\right]d\zeta\,dt + \mathbb{E}\int_0^T\int_{\mathcal{D}} u^2(t,\zeta)d\zeta\,dt \\ & \mathbb{E}\int_{\mathcal{D}}\left[\pi_1(\zeta)\xi^2(T,\zeta)+\pi_2(\zeta)\left(\frac{\partial\xi}{\partial t}(T,\zeta)\right)^2\right]d\zeta. \end{aligned} \tag{10.2}$$

In the above formulae $\mathcal{D}\subset\mathbb{R}^d$ is a bounded domain with regular boundary. By $\mathcal{B}(\mathcal{D})$ we denote the Borel σ-field in $\mathcal{D}$.

Moreover $\{\beta_i : i=1,2...\}$ are independent standard (real valued) brownian motions defined on $(\Omega,\mathcal{F},\mathbb{P})$. We set $\mathcal{F}_t=\sigma\{\beta_i(s): s\in[0,t],\ i=1,2...\}$ and denote by $\mathcal{P}$ the predictable σ-field in $\Omega\times[0,T]$.

On the coefficients we assume the following:

(1) μ is a bounded measurable process defined on $([0,T]\times\Omega)\times\mathcal{D}$ endowed with the σ-field $\mathcal{P}\otimes\mathcal{B}(\mathcal{D})$ with values in $\mathbb{R}^+$ (with Borel σ-field).
(2) b, κ_1, κ_2 and c_i, $i=1,2,..$ are bounded measurable maps $[0,T]\times\mathcal{D}\to\mathbb{R}$. We assume that κ_1 and κ_2 have values in $\mathbb{R}^+$.
(3) There exists a constant $M>0$ such that: $\sum_{i=1}^{\infty}|c_i(t,\zeta)|^2\le M$ for a.e. $t\in[0,T]$ and a.e. $\zeta\in\mathcal{D}$
(4) π_1 and π_2 are bounded measurable maps $\mathcal{D}\to\mathbb{R}^+$.

Following, for instance, [1] we set:

(1) $H=H_0^1(\mathcal{D})\times L^2(\mathcal{D})$, $U=L^2(\mathcal{D})$
(2) $W(t)=\sum_{i=1}^{\infty} f_i\beta_i(t)$ where $\{f_i : i=1,2,...\}$ is an orthonormal basis in an arbitrary separable real Hilbert space Ξ

(3) $\mathcal{D}(A) = \left[H^2(\mathcal{D}) \cap H^1_0(\mathcal{D})\right] \times H^1_0(\mathcal{D})$ and

$$\left(A\begin{pmatrix}\xi\\ v\end{pmatrix}\right)(\zeta) = \begin{pmatrix} v(\zeta)\\ \Delta_\zeta \xi(\zeta)\end{pmatrix}, \quad \begin{pmatrix}\xi\\ v\end{pmatrix} \in \mathcal{D}(A), \quad \left(A_\sharp(t)\begin{pmatrix}\xi\\ v\end{pmatrix}\right)(\zeta) = \begin{pmatrix} 0\\ \mu(t,\zeta)\xi(\zeta)\end{pmatrix}, \quad \begin{pmatrix}\xi\\ v\end{pmatrix} \in H$$

(4) $(B(t)u)(\zeta) = \begin{pmatrix} 0\\ b(t,\zeta)u(\zeta)\end{pmatrix}, \quad \left(C_i(t)\begin{pmatrix}\xi\\ v\end{pmatrix}\right)(\zeta) = \begin{pmatrix} 0\\ c_i(t,\zeta)\xi(\zeta)\end{pmatrix}$

(5) $\left(S(t)\begin{pmatrix}\xi\\ v\end{pmatrix}\right)(\zeta) = \begin{pmatrix} \kappa_1(t,\zeta)\xi(\zeta)\\ \kappa_2(t,\zeta)v(\zeta)\end{pmatrix}, \left(P_T(t)\begin{pmatrix}\xi\\ v\end{pmatrix}\right)(\zeta) = \begin{pmatrix} \pi_1(\zeta)\xi(\zeta)\\ \pi_2(\zeta)v(\zeta)\end{pmatrix}, \quad x = \begin{pmatrix} x_0\\ v_0\end{pmatrix}$

With this setting the state equation (10.1) is equivalent to (2.1) and the cost (10.2) is equivalent to (2.2). Moreover it is easy to verify that assumptions Hypothesis 2.1 A1)-A4) are verified. So in this case we can apply the results in Theorem 6.6 to obtain existence of a unique solution of the Riccati equation both in the "generalized" sense of Definition 6.1 and in the "variation of constants" sense of Definition 7.3. Moreover such a solution allows to perform the synthesis of the optimal control as it is stated in Theorem 6.6.

Remark 10.1. Notice that assumption A5) is in general not satisfied, take, for instance, $\kappa_2 \equiv 1$ or $\pi_2 \equiv 1$.

Acknowledgements. We wish to thank Marco Fuhrman: we are indebted to him for his help and encouragement. We also thank the referees for their useful remarks.

References

[1] A. Bensoussan, G. Da Prato, M.C. Delfour and S.K. Mitter. *Representation and Control of Infinite Dimensional Systems.* Vol. 1. Systems & Control: Foundations & Applications. Birkhuser Boston, Inc., Boston, MA, 1992.

[2] J.-M. Bismut. Linear quadratic optimal stochastic control with random coefficients. *SIAM J. Contr. Optim.* 14 (1976), 419–444.

[3] J.-M. Bismut. Contrôle des systèmes linéaires quadratiques: applications de l'intégrale stochastique. In *Séminaire de Probabilités, XII.* Lecture Notes in Math., 649, Springer, Berlin, 1978.

[4] A. Chojnowska-Michalik, Representation theorem for general stochastic delay equations, *Bull. Acad. Pol. Sci. Ser. Sci. Math.* 26 (1978), no. 7, 634–641.

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

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

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

[8] M. Fuhrman and G. Tessitore. Generalized directional gradients, backward stochastic differential equations and mild solutions of semilinear parabolic equations . Preprint, Politecnico di Milano, 2003

[9] J. Hale, Theory of functional differential equations, Appl. Math. Sci., vol. 3, Springer Verlag, 1971.

[10] Y. Hu and S. Peng. Adapted solution of a backward semilinear stochastic evolution equation. *Stochastic Anal. Appl.* 9 (1991), 445–459.

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

[12] M. Kohlmann and S. Tang. New developments in backward stochastic Riccati equations and their applications. In *Mathematical finance (Konstanz, 2000)*, Trends Math., Birkhuser, Basel, 2001.

[13] M. Kohlmann and S. Tang. Global adapted solution of one-dimensional backward stochastic Riccati equations, with application to the mean-variance hedging. *Stochastic Process. Appl.* 97 (2002), 1255–288.

[14] M. Kohlmann and S. Tang. Multidimensional backward btochastic riccati equations and applications. *SIAM J. Contr. Optim.* 41 (2003), 1696–1721.

[15] M. Kohlmann and X.Y. Zhou. Relationship between backward stochastic differential equations and stochastic controls: a linear-quadratic approach. *SIAM J. Contr. Optim.* 38 (2000), 1392–1407.

[16] A. Lunardi, *Analytic semigroups and optimal regularity in parabolic problems.* Progress in Nonlinear Differential Equations and their Applications,16. Birkhäuser Verlag, Basel, 1995.

[17] J.-P Lepeltier, J. San Martín. Existence for BSDE with superlinear-quadratic coefficient. *Stochastics Stochastic Rep.*, 63 (1998),227–240.

[18] E. Pardoux, S. Peng, *Adapted solution of a backward stochastic differential equation*, *Systems and Control Lett.* 14 (1990), 55-61.

[19] A. Pazy. *Semigroups of linear operators and applications to partial differential equations.* Springer-Verlag, New York, 1983.

[20] S. Peng. Stochastic Hamilton-Jacobi-Bellman Equations. *SIAM J. Contr. Optim.* 30 (1992), 284–304.

[21] S. Peng. Open problems on backward stochastic differential equations. In *Control of distributed parameter and stochastic systems (Hangzhou, 1998)*, Kluwer Acad. Publ., Boston, 1999.

[22] S. Tang. General linear quadratic optimal control problems with random coefficients: linear stochastic Hamilton systems and backward stochastic Riccati equations. Preprint.

[23] G.Tessitore. Some remarks on the Riccati equation arising in an optimal control problem with state- and control-dependent noise. *SIAM J. Contr. and Optim. 30 (1992), 717–744.*

[24] J. Yong, X. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations.* Springer Verlag, Berlin, New York, 1999.

[25] J. Zabczyk, *Parabolic equations on Hilbert spaces.* In: *Stochastic PDE's and Kolmogorov Equations in Infinite Dimensions*, ed. G. Da Prato, 117-213. Lecture Notes in Mathematics 1715, Springer Verlag, Berlin, New York, 1999.

[26] X. Zhou, D. Li, *Continuous time mean-variance portfolio selection: a stochastic LQ framework.* Appl. Math. Optim. 42 (2000), 19–33.