Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

A nonlinear Bismut–Elworthy formula for HJB equations with quadratic Hamiltonian in Banach spaces

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

This version is available at: https://hdl.handle.net/11585/832207 since: 2021-09-14

Terms of use:

(Article begins on next page)

A nonlinear Bismut-Elworthy formula for HJB equations with quadratic Hamiltonian in Banach spaces

Davide ADDONA[*1], Elena BANDINI[†1] and Federica MASIERO[‡1]

[1]Dipartimento di Matematica e Applicazioni, Università di Milano-Bicocca, Milano, Italy

Abstract

We consider a Backward Stochastic Differential Equation (BSDE for short) in a Markovian framework for the pair of processes (Y, Z), with generator with quadratic growth with respect to Z. The forward equation is an evolution equation in an abstract Banach space. We prove an analogue of the Bismut-Elworty formula when the diffusion operator has a pseudo-inverse not necessarily bounded and when the generator has quadratic growth with respect to Z. In particular, our model covers the case of the heat equation in space dimension greater than or equal to 2. We apply these results to solve semilinear Kolmogorov equations in Banach spaces for the unknown v, with nonlinear term with quadratic growth with respect to ∇v and final condition only bounded and continuous, and to solve stochastic optimal control problems with quadratic growth.

Keywords: Stochastic heat equation in 2 and 3 dimensions, nonlinear Bismut-Elworthy formula, quadratic Backward Stochastic Differential Equation, Hamilton Jacobi Bellman equation.

MSC 2010: 60H10; 60H30; 93E20; 35Q93.

1 Introduction

In this paper we consider Markovian BSDEs whose generator has quadratic growth with respect to Z. Our BSDE is related to a forward stochastic differential equation of the form

$$\begin{cases} dX_\tau^{t,x} = AX_\tau^{t,x}d\tau + F(X_\tau^{t,x})d\tau + (-A)^{-\alpha}dW_\tau, & \tau \in [t, T], \\ X_t^{t,x} = x \in E, \end{cases} \tag{1.1}$$

where E is a Banach space which is continuously and densely embedded in a real and separable Hilbert space H, and the diffusion operator $(-A)^{-\alpha}$ admits a pseudo-inverse not bounded. The choice of setting the problem in a Banach space is due to the fact that in many situations of interest the drift F is not well defined on the whole Hilbert space H but only on a subset: a typical example is provided by the so called polynomial nonlinearities in reaction diffusion equations, see also Example 6.11 for a concrete situation.

In this context we generalize the Bismut-Elworthy type formula introduced in [10], where the Lipschitz case was studied in the Hilbert space framework, and the forward equation had a diffusion operator with bounded inverse. We apply our results to study quadratic semilinear Kolmogorov equations of the form (1.2), and to solve stochastic optimal control problems with quadratic growth.

---

[*]davide.addona@unimib.it

[†]elena.bandini@unimib.it

[‡]federica.masiero@unimib.it

The presence of the diffusion operator $(-A)^{-\alpha}$ in (1.1) allows us to deal with stochastic heat equations in 2 and 3 space dimensions, while stochastic heat equations in one space dimension can be considered without any regularization of the white noise, that is in the case with $\alpha = 0$. We also assume that the map F in (1.1) is only dissipative, so in general is well defined only on the Banach space E and not on the whole Hilbert space H; this is a natural situation arising in many models, see e.g. [4] and [6].

The solution of equation (1.1) will be denoted by X, or also by $X^{t,x}$, to stress the dependence on the initial conditions, and the transition semigroup related to $X^{t,x}$ will be denoted by

$$P_{t,\tau}[\phi](x) := \mathbb{E}[\phi(X_\tau^{t,x})]$$

for any ϕ bounded and measurable function. At least formally, the generator of $P_{t,\tau}$ is the second order differential operator

$$(\mathscr{L}f)(x) = \frac{1}{2}(Tr((-A)^{-\alpha}(-A^*)^{-\alpha}\nabla^2 f)(x) + \langle Ax, \nabla f(x)\rangle + \langle F(x), \nabla f(x)\rangle.$$

This is the link with the solution, in mild sense, of the semilinear Kolmogorov equation in E (see e.g. [5]):

$$\begin{cases} \frac{\partial v}{\partial t}(t,x) = -\mathscr{L}v(t,x) + \psi(t,x,v(t,x),\nabla v(t,x)(-A)^{-\alpha}), & t \in [0,T], \ x \in E, \\ v(T,x) = \phi(x). \end{cases} \tag{1.2}$$

We recall that by mild solution of equation (1.2) we mean a bounded and continuous function $v : [0,T] \times H \to H$, once Gâteaux differentiable with respect to x, and satisfying the integral equality

$$v(t,x) = P_{t,T}[\phi](x) + \int_t^T P_{t,s}\left[\psi(s,\cdot,v(s,\cdot),\nabla v(s,\cdot)(-A)^{-\alpha})\right](x)\,ds, \quad t \in [0,T], \ x \in E. \tag{1.3}$$

Second order differential equations like (1.2) are a widely studied topic in the literature, see e.g. [5]. In the case of ψ only locally Lipschitz continuous, we cite [14], [20], [18] and also [19], where in particular the quadratic case is studied with datum ϕ only continuous. We also mention the monograph [3], where semilinear Kolmogorov equations related to forward equations of reaction diffusion type are studied, but requiring Lipschitz continuity of the final datum.

We will consider equation (1.2) under the assumptions that the final datum ϕ is bounded and continuous, and that ψ has quadratic growth with respect to the derivative $\nabla v(-A)^{-\alpha}$. In order to prove existence and uniqueness of a mild solution of the form (1.3) for the Kolmogorov equation (1.2), we aim at representing this mild solution in terms of a Markovian BSDE of the form

$$\begin{cases} dY_\tau = -\psi(\tau, X_\tau, Y_\tau, Z_\tau)\,d\tau + Z_\tau\,dW_\tau, \\ Y_T = \phi(X_T). \end{cases} \tag{1.4}$$

We recall that, in order to solve partial differential equations by means of BSDEs, one of the crucial tasks is the identification of Z with the derivative of Y taken in the directions of the diffusion operator. In this regard, we refer to the seminal paper [21] for the finite dimensional case, and to [11] for the infinite dimensional extension in Hilbert spaces: in both papers the driver ψ is Lipschitz continuous in Y and in Z, and ψ and ϕ are differentiable. We also mention [17], where an extension to the Banach space case is studied with the same assumptions of Lipschitz continuity and differentiability on the data. Notice that in the above mentioned papers the boundedness of the inverse of the diffusion operator is a key property in order to prove the identification formula for Z.

Another key task consists in proving a nonlinear Bismut-Elworthy formula for the BSDE (1.4). Bismut-Elworthy formulas for the transition semigroup of equations of type (1.1) with invertible diffusion operator are a classical topic in the literature, see e.g. [5], and have been considered also in the case of an operator like the one in (1.1), with pseudo-inverse which is not necessarily bounded, see [3].

According to these classical Bismut formulas, for every $0 \leq t < \tau \leq T$, $x \in H$, $h \in H$, and for every bounded and continuous real function f defined on H, one has

$$\langle \nabla_x P_{t,\tau}[f](x), h \rangle = \mathbb{E} f\left(X_\tau^{t,x}\right) U_\tau^{h,t,x}, \tag{1.5}$$

where $(G(r, X_r^{t,x})$ being the general diffusion operator)

$$U_\tau^{h,t,x} := \frac{1}{\tau - t} \int_t^\tau \langle G^{-1}(r, X_r^{t,x}) \nabla_x X_r^{t,x} h, dW_r \rangle.$$

In [10] a nonlinear Bismut-Elworthy formula for the process Y solution of the BSDE (1.4) is proved when ψ is Lipschitz continuous with respect to Z and the process X takes its values in a Hilbert space H. According to this formula, for $0 \leq t < \tau \leq T$, $x \in H$, for every direction $h \in H$,

$$\mathbb{E}\left[\nabla_x Y_\tau^{t,x} h\right] = \mathbb{E}\left[\int_\tau^T \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x} \, dr\right] + \mathbb{E}\left[\phi(X_T^{t,x}) U_T^{h,t,x}\right]. \tag{1.6}$$

Identity (1.6) is then used in [10] to solve a semilinear Kolmogorov equation of the form of (1.2). When the Hamiltonian function ψ is Lipschitz continuous with respect to the derivative of v, semilinear Kolmogorov equations of the type of (1.2) can be solved also by using the estimates coming from the classical Bismut formulas (1.5) and by a fixed point argument, see e.g. [3], [5], [13].

In the quadratic case this procedure does not work anymore since the fixed point argument cannot be applied: for this reason, nonlinear versions of Bismut-Elworthy formulas, that give an alternative way to solve equations like (1.2), are particularly interesting in such a framework. In [19], a nonlinear version of the Bismut-Elworthy formula has been provided and has been applied to semilinear Kolmogorov equations of the type of (1.2), with quadratic hamiltonian, but in a Hilbert space and with diffusion operator with bounded inverse.

In the present paper, we generalize (1.6) to the Banach space framework, where the diffusion operator $(-A)^{-\alpha}$ has unbounded pseudo-inverse operator and the coefficients are not necessarily differentiable. In this context, the nonlinear Bismut formula (1.6) has its own independent interest, and moreover it allows to solve the Kolmogorov equation with Hamiltonian function quadratic with respect to $\nabla v(-A)^{-\alpha}$. We first provide an analogous of the nonlinear Bismut formula given in [10] in the case of Banach space framework and Lipschitz continuous generator, see Theorem 4.4. We then prove a nonlinear Bismut formula in the quadratic case when ψ and ϕ are differentiable, namely Theorem 5.1; the differentiability assumptions are finally removed by means of an approximation procedure, suitably generalizing the one introduced in [23]. In order to achieve these results, by denoting by $(Y^{t,x}, Z^{t,x})$ a solution to the Markovian BSDE (1.4) and by assuming that ϕ and ψ are differentiable, the two main ingredients are the identification

$$Z_t^{t,x} = \nabla_x Y_t^{t,x}(-A)^{-\alpha}, \quad t \in [0,T], \, x \in E, \tag{1.7}$$

and an a priori estimate of $Z^{t,x}$ of the form (C being a constant depending on t, T, A, F, $\|\phi\|_\infty$)

$$|Z_t^{t,x}|_H \leq C(T - t)^{-1/2}, \tag{1.8}$$

that are the objects of Theorem 3.10 and Proposition 3.12, respectively. Both (1.7) and (1.8) are new in the Banach space framework and in the case of quadratic generator with respect to z. Concerning the identification result (1.7), notice that the standard approach with the Malliavin calculus cannot be implemented in this Banach framework, nor the variational approach developed in [17] in the Lipschitz continuous case: to deal with the quadratic case, one needs to generalize this latter variational approach and to apply it not only to the the forward process X but also to the the pair of processes (Y, Z). On the other hand, estimate (1.8) is particularly delicate due to the presence of the operator $G = (-A)^{-\alpha}$ with unbounded inverse.

Our results can be applied to a stochastic optimal control problem consisting in minimizing a cost functional of the form

$$J(t,x,u) = \mathbb{E}\Big[\int_t^T l(s, X_s^u, u_s)\, ds + \phi(X_T^u) \Big] \tag{1.9}$$

over all the admissible controls u taking values in H and not necessarily bounded. Here l has quadratic growth with respect to u, and X^u is the solution of the controlled state equation

$$\begin{cases} dX_\tau^u = AX_\tau^u d\tau + F(X_\tau^u)d\tau + Qu_\tau d\tau + (-A)^{-\alpha}dW_\tau, & \tau \in [t,T] \\ X_t^u = x, \end{cases} \tag{1.10}$$

with $Q = I$ or $Q = (-A)^{-\alpha}$. Since we work in a Banach space, we are able to treat a wider class of cost functionals which respect to those which we can consider in Hilbert spaces. Indeed, in Example 6.11 we present a control problem for a stochastic heat equation in the Banach space of continuous functions, with cost functional well defined in the Banach space of continuous functions but not in the Hilbert space of square integrable functions. The aim of this latter part of the work is to characterize the value function as the solution of the associated Hamilton Jacobi Bellman (HJB in the following) equation, and to provide a feedback law for optimal controls. If $Q = (-A)^{-\alpha}$, namely when the controls affect the system only through the noise (the so called structure condition holds true), the optimal control problem (1.9) can be completely solved, see Theorem 6.10. When $Q = I$, the optimal control problem can be completely solved by restricting ourselves to the class of more regular controls taking values in $D((-A)^{-\alpha})$, see Theorem 6.16. In the general case of $Q = I$ and H-valued controls, we are able to provide an "ε-optimal solution" of the problem in the sense that the value function can be approximated by a sequence of functions which are solutions of approximating HJB equations, and we can obtain an ε-optimal control in feedback form, see Theorem 6.26. So a contribution of this paper is also to extend the BSDE approach for stochastic optimal control problems from state equation with structure conditions to cases where the structure condition is violated.

The paper is organized as follows: in Section 2 we fix the notations and we give the results on the forward process. In Section 3 we introduce the forward backward system: here the main results are the identification (1.7) of $Z_t^{t,x}$ with $\nabla Y_t^{t,x}(-A)^{-\alpha}$, which is new in the case of ψ quadratic with respect to z and in the Banach space framework, and the a priori estimate (1.8) on Z not involving derivatives of the coefficients of the BSDE. In Section 4 we give the nonlinear Bismut formula (1.6) in the Banach space E and with ψ Lipschitz continuous with respect to z, then in Section 5 we extend formula (1.6) to the case of ψ quadratic with respect to z. In both Sections 4 and 5, the Bismut formula is applied to solve the corresponding semilinear Kolmogorov equation (1.2). Finally in Section 6 we apply the previous results to solve the stochastic optimal control problem (1.9).

2   Notations and preliminary results on the forward process

We assume that E is a real and separable Banach space which admits a Schauder basis, and that E is continuously and densely embedded in a real and separable Hilbert space H. E and H are respectively endowed with the norms $|\cdot|_E$ and $|\cdot|_H$. We fix a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ endowed with a filtration $\{\mathscr{F}_t,\ t \geq 0\}$ satisfying the usual conditions.

We list below some notations that are used in the paper. Let K be a given Banach space endowed with the norm $|\cdot|_K$. For any $p, q \in [1, \infty)$ and any $t \in [0, T]$, we set

- $L^p(0, T; K)$ the space of K-valued measurable functions defined on $[0, T]$, normed by

$$\|f\|_{L^p(0,T;K)} := \Big(\int_0^T |f_s|_K^p ds \Big)^{1/p}.$$

4

- $L^q(\Omega; L^p(0, T; K))$ the space of adapted processes $(u_s)_{s \in [0, T]}$, defined on $[0, T]$ and with values in K, normed by

$$\|u\|_{L^q(\Omega; L^p(0,T;K))} := \left(\mathbb{E}\left[\left(\int_0^T |u_s|_K^p ds \right)^{q/p} \right] \right)^{1/q}.$$

- $\mathscr{S}^p((t, T]; K)$ (resp. $\mathscr{S}^p([t, T]; K)$) the space of all adapted processes $(X_s)_{s \in [t, T]}$, continuous on $(t, T]$ (resp. on $[t, T]$) and with values in K, normed by

$$\|X\|_{\mathscr{S}^p([t,T];K)} = \|X\|_{\mathscr{S}^p((t,T];K)} := \left(\mathbb{E}\left[\sup_{s \in [t,T]} |X_s|_K^p \right] \right)^{1/p}.$$

  If $K = \mathbb{R}$ we simply write $\mathscr{S}^p([t, T])$.

- $\mathscr{M}^p([t, T]; K)$ the space of all predictable processes $(Z_s)_{s \in [t, T]}$ with values in K normed by

$$\|Z\|_{\mathscr{M}^p([t,T];K)} := \left(\mathbb{E}\left[\left(\int_t^T |Z_s|_K^2 ds \right)^{p/2} \right] \right)^{1/p}.$$

  If $K = \mathbb{R}$ we simply write $\mathscr{M}^p([t, T])$.

We denote by $L(E, K)$ the space of all bounded linear operators from E to K, endowed with the usual operator norm. E^* denotes the dual space of E, and $\langle \cdot, \cdot \rangle_{E \times E^*}$ denotes the duality between E and E^*. We say that a function $f : E \to K$ belongs to the class $\mathscr{G}^1(E, K)$ if f is continuous and Gâteaux differentiable on E and if the gradient $\nabla f : E \to L(E, K)$ is strongly continuous. If $K = \mathbb{R}$ we simply write $\mathscr{G}^1(E)$. We say that $f : [0, T] \times E \to \mathbb{R}$ is in $\mathscr{G}^{0,1}([0, T] \times E)$ if f is continuous and Gâteaux differentiable with respect to every $x \in E$ and the gradient $\nabla f : [0, T] \times E \to L(E, \mathbb{R})$ is strongly continuous. For more details on this classes of Gâteaux differentiable functions see [11, Section 2.2].

2.1 The forward equation

We are given the Markov process X in E (also denoted $X^{t,x}$ to stress the dependence on the initial conditions) solution to the equation

$$\begin{cases} dX_\tau^{t,x} = AX_\tau^{t,x}d\tau + F(X_\tau^{t,x})d\tau + (-A)^{-\alpha}dW_\tau, & \tau \in [t, T], \\ X_t^{t,x} = x \in E, \end{cases} \tag{2.1}$$

where $(W_\tau)_{\tau \in [0,T]}$ is a cylindrical Wiener process with values in H, see e.g. [6] for details on cylindrical Wiener processes in infinite dimensions. From now on $\{\mathscr{F}_\tau, \tau \geq 0\}$ will be the natural filtration generated by the Wiener process and augmented in the usual way.

We assume the following on the coefficients of equation (2.1).

Hypothesis 2.1.   *1. A is a linear operator which generates a contraction analytic semigroup $(e^{tA})_{t \geq 0}$ on the Hilbert space H and there exist $c, \omega > 0$ such that $|e^{tA}h|_H \leq ce^{-\omega t}|h|_H$ for any $h \in H$ and any $t \geq 0$. Further, the restriction of A to E generates a contraction C_0 (or analytic) semigroup on E.*

     2. The stochastic convolution

$$w^A(s, t) := \int_s^t e^{(t-u)A}(-A)^{-\alpha}dW_u, \quad 0 \leq s < t \leq T,$$

admits an E-continuous version, and, for any $p \geq 2$, $\mathbb{E}[\sup_{t \in [0,T]} |w^A(t)|_E^p] < +\infty$ (when $s = 0$ we write $w^A(t)$ instead of $w^A(0, t)$).

3. $F : D(F) \subset H \to H$ is a measurable and dissipative map, and $E \subseteq D(F)$.

4. The restriction F_E of F to E is a map from E to E which is measurable and dissipative (where no confusion is possible, we simply write F instead of F_E). $F \in \mathscr{G}^1(H, H)$ and F_E is Fréchet differentiable. Further, there exist $a, c, \gamma > 0$, $m \in \mathbb{N}$ and for any $z \in E$ an element $z^* \in \partial |z|_E$, such that, for any $x \in E$, $h \in H$,

$$|F_E(x)|_E \le c(1 + |x|_E^{2m+1}),$$
$$\|\nabla F(x)\|_{\mathscr{L}(E)} \le c(1 + |x|_E^{2m}),$$
$$\langle F(x+z) - F(x), z^* \rangle_{E \times E^*} \le -a|z|_E^{2m+1} + c(1 + |x|_E^{\gamma}),$$
$$|\nabla F(x)h|_H \le c\left(1 + |x|_E^{2m}\right)|h|_H.$$

5. $\alpha \in (0, 1/2)$.

By Hypothesis 2.1-1. and the Kuratowski theorem, see e.g. [22], Chapter I, Theorem 3.9, it follows that E is a Borel set in H.

Remark 2.2. *Since by Hypothesis 2.1-3.-4. F is differentiable and dissipative, we get*

$$1 \ge |z - \alpha DF(x)z|_E, \quad x, z \in E, \quad |z|_E = 1, \quad \alpha > 0.$$

In particular, from the Hahn-Banach theorem, there exists $z^* \in \partial |z|_E$ such that $|z - \alpha DF(x)z|_E = \langle z - \alpha DF(x)z, z^* \rangle_{E \times E^*}$, and therefore $\langle DF(x)z, z^* \rangle_{E \times E^*} \le 0$. Further, from [6, Appendix D] we have

$$D_-|x|_E y = \min\{\langle y, x^* \rangle_{E \times E^*} : x^* \in \partial |x|_E\}. \tag{2.2}$$

Remark 2.3. *Since A generates a contraction semigroup on E, then A is dissipative, and for any $x \in D(A)$ we have $\langle Ax, x^* \rangle_{E \times E^*} \le 0$, $x^* \in \partial |x|_E$, see Example D.8 in [6].*

We now give an example of spaces E and H and of an operator A satisfying Hypothesis 2.1-1.-2.

Example 2.4. *Let $d, n \in \mathbb{N}$ with $d \le 3$, $\mathscr{O} \subset \mathbb{R}^d$ be an open bounded set, $H := L^2(\mathscr{O}; \mathbb{R}^n)$ and $E := C(\overline{\mathscr{O}}; \mathbb{R}^n)$. Further, let A be the realization in H of the operator*

$$\mathscr{A} - (\rho + 1)I = (\Delta - (\rho + 1)I, \dots, \Delta - (\rho + 1)I),$$

with boundary conditions $\mathscr{B}u = 0$, where $\mathscr{B} = (\mathscr{B}_1, \dots, \mathscr{B}_n)$ and

$$\mathscr{B}_k = Id, \quad or \quad \mathscr{B}_k := \sum_{i=1}^{d} \nu_i(\xi) \frac{\partial}{\partial \xi_i}, \quad \xi \in \partial \mathscr{O}, \ k = 1, \dots, n,$$

where ν_i is the normal vector to the boundary of $\overline{\mathscr{O}}$. As shown for example in [16], A satisfies Hypothesis 2.1-1. Moreover, [3, Lemma 6.1.2] with $Q = (-A)^{-\alpha}$ shows that Hypothesis 2.1-2. is satisfied with this choice of H, E and A.

In the following proposition we collect important results on the solution of the forward equation (2.1). We recall that, given $x \in E$ and $t \in [0, T]$, a mild solution to (2.1) is an adapted process $X^{t,x} : [0, T] \times \Omega \to E$ which satisfies

$$X_\tau^{t,x} = e^{(\tau-t)A}x + \int_t^\tau e^{(\tau-s)A}F(X_s^{t,x})ds + \int_t^\tau e^{(\tau-s)A}dW_s, \quad \tau \in [t, T], \ \mathbb{P}\text{-a.s.} \tag{2.3}$$

Proposition 2.5. *Let Hypothesis 2.1 hold true. Then the following hold.*

(i) For any $x \in E$, $t \in [0, T]$, the problem (2.1) admits a unique mild solution $X^{t,x} \in \mathscr{S}^p((t, T]; E)$, for any $p \geq 1$. If A generates a strongly continuous semigroup on E, then the process $X^{t,x}$ is also continuous up to t. Moreover, there exists a positive constant c such that, for any $\tau \in [t, T]$,

$$|X^{t,x}_\tau|_E \leq e^{c\tau}|x|_E + h(t, \tau), \quad \mathbb{P}\text{-a.s.}, \tag{2.4}$$

where

$$h(t, \tau) := ce^{c(\tau-t)} \int_t^\tau \left(1 + |w^A(t, s)|_E^{2m+1}\right) ds + \sup_{s \in [t, \tau]} |w^A(t, s)|_E.$$

(ii) For any $x \in E$, $t \in [0, T]$, the mild solution $X^{t,x}$ to (2.1) is Gâteaux differentiable as a map from E to $\mathscr{S}^p([t, T]; E)$, and

$$\sup_{x \in E, \tau \in [t, T]} |\nabla_x X^{t,x}_\tau z|_E \leq |z|_E, \quad z \in E, \quad \mathbb{P}\text{-a.s.} \tag{2.5}$$

Moreover, $X^{t,x}$ is Gâteaux differentiable as a map from E to $\mathscr{S}^p([0, T]; H)$, and

$$\sup_{x \in E, \tau \in [t, T]} |\nabla_x X^{t,x}_\tau h|_H \leq |h|_H, \quad h \in H, \quad \mathbb{P}\text{-a.s.} \tag{2.6}$$

(iii) For any $x \in E$, $t \in [0, T]$ and $\tau \in [t, T]$,

$$\nabla_x X^{t,x}_\tau h = e^{(\tau-t)A}h + \int_t^\tau e^{(\tau-s)A}\nabla F(X^{t,x}_s)\nabla_x X^{t,x}_s h \, ds, \quad h \in H, \quad \mathbb{P}\text{-a.s.} \tag{2.7}$$

Proof. Item (i) can be proved arguing as in [6, Theorem 7.13].

The first part of (ii) and inequality (2.5) follow from [17, Propositions 3.10 & 3.13]. We claim that

$$\sup_{x \in E, \tau \in [t, T]} |\nabla_x X^{t,x}_\tau z|_H \leq |z|_H, \quad z \in E \quad \mathbb{P}\text{-a.s.} \tag{2.8}$$

If the claim is true, since E is densely embedded into H, by approximation we immediately deduce (2.6) for any $h \in H$. In order to prove (2.8), we consider $z \in E$ and the approximating processes $G^n_\tau z := nR(n, A)\nabla_x X^{t,x}_\tau z$, $n \in \mathbb{N}$, where $R(n, A) := (nI - A)^{-1}$. Then, $G^n_\tau z$ is a strict solution to

$$\frac{d}{dt}G^n_\tau z = AG^n_\tau z + \nabla F(X^{t,x}_\tau)G^n_\tau z, \quad \tau \in (t, T], \quad G^n_t z = nR(n, A)z.$$

The dissipativity of F and A implies $\frac{d}{d\tau}|G^n_\tau z|_H^2 \leq 0$, which gives $|G^n_\tau z|_H \leq |nR(n, A)z|_H$. Letting $n \to +\infty$ we get (2.8).

It remains to prove (iii). To this end, we recall that (see e.g. [17]), for any $x, z \in E$, the process $\nabla_x X^{t,x}_\tau z$ is a mild solution to

$$\begin{cases} d\zeta_\tau = A\zeta_\tau d\tau + \nabla F(X^{t,x}_\tau)\zeta_\tau, & \tau \in [t, T], \\ \zeta_t = z \in E, \end{cases} \tag{2.9}$$

and therefore

$$\nabla_x X^{t,x}_\tau z = e^{(\tau-t)A}z + \int_t^\tau e^{(\tau-s)A}\nabla F(X^{t,x}_s)\nabla_x X^{t,x}_s z \, ds, \quad \mathbb{P}\text{-a.s.} \tag{2.10}$$

Let $h \in H$ and let $(h_n) \subset E$ be an approximating sequence of h in H. If we replace h_n to z in (2.10), from (ii) we deduce that the left-hand side of (2.10) and the first term in the right-hand side of (2.10)

converge respectively to $\nabla_x X_\tau^{t,x} h$ and to $e^{(\tau-t)A} h$, as $n \to +\infty$. As far as the integral in the right-hand side of (2.10) is considered, with z replaced by h_n, again from (ii) we infer that

$$e^{(\tau-s)A} \nabla F(X_s^{t,x}) \nabla_x X_s^{t,x} h_n \to e^{(\tau-s)A} \nabla F(X_s^{t,x}) \nabla_x X_s^{t,x} h, \quad \mathbb{P}\text{-a.s.},$$

as $n \to +\infty$. Thanks to Hypothesis 2.1-4., and estimates (2.4), (2.6), we can apply the dominated convergence theorem and therefore

$$\int_t^\tau e^{(\tau-s)A} \nabla F(X_s^{t,x}) \nabla_x X_s^{t,x} h_n ds \to \int_t^\tau e^{(\tau-s)A} \nabla F(X_s^{t,x}) \nabla_x X_s^{t,x} h ds, \quad \mathbb{P}\text{-a.s.},$$

as $n \to +\infty$, which gives (2.7). $\qquad \square$

We conclude this section by providing both integral and pointwise estimates for $\nabla_x X^{t,x} z$ in $D((-A)^\alpha)$. These estimates will be crucial to prove the identification formula for Z and the nonlinear Bismut-Elworthy formula, see respectively Proposition 3.12 and Lemma 4.1.

Proposition 2.6. *Let Hypothesis 2.1 holds true, and let $x \in E$, $z \in H$ and $t \in [0, T]$. Then, $\nabla_x X^{t,x} z \in D((-A)^{1/2})$, a.e. in (t, T) and $\mathbb{P}$-a.s., and there exists a positive constant C such that, for any $\varepsilon \in [0, 1/2]$,*

$$\int_t^\tau |(-A)^\varepsilon \nabla_x X_s^{t,x} z|_H^2 ds \le C(\tau - t)^{1-2\varepsilon} |z|_H^2, \quad \tau \in [t, T], \quad \mathbb{P}\text{-a.s.} \tag{2.11}$$

Proof. Let $x \in E$. We prove (2.11) for $t = 0$, the case $t \in [0, T]$ can be proved by analogous computations. We first assume that $z \in E$. Let $\nabla_x X_t^x z$ be a strict solution to (2.9), otherwise we can approximate it by smooth processes, as in the proof of Proposition 2.5, item (ii). The dissipativity of F in H gives

$$\frac{d}{ds} |\nabla_x X_s^x z|_H^2 = \langle A \nabla_x X_s^x z, \nabla_x X_s^x z \rangle_H + \langle \nabla F(X_s^x) \nabla_x X_s^x z, \nabla_x X_s^x z \rangle_H \le \langle A \nabla_x X_s^x z, \nabla_x X_s^x z \rangle_H,$$

for any $s \in [0, T]$. Integrating between 0 and $\tau \in [0, T]$ we get

$$|\nabla_x X_\tau^x z|_H^2 + \int_0^\tau \langle -A \nabla_x X_s^x z, \nabla_x X_s^x z \rangle_H ds \le |z|_H^2.$$

Since $\langle -A \nabla_x X_s^x z, \nabla_x X_s^x z \rangle_H = |(-A)^{1/2} \nabla_x X_s^x z|_H^2$ for any $s \in [0, t]$, from (2.6) we deduce that $\nabla_x X_\tau^x z \in D((-A)^{1/2})$ for any $\tau \in [0, T]$. Thus (2.11) holds for $\varepsilon = 1/2$, $t = 0$ and any $z \in E$.
Let us now consider $\varepsilon \in [0, 1/2)$. From interpolation estimates (see e.g. [16, Section 2.2])

$$|(-A)^\varepsilon e^{tA} x|_H \le C_\varepsilon |x|_H^{1-2\varepsilon} |(-A)^{1/2} x|_H^{2\varepsilon}, \quad x \in D((-A)^{1/2}). \tag{2.12}$$

By replacing x by $\nabla_x X_s^x z$ in (2.12), we get

$$\begin{aligned}
\int_0^\tau |(-A)^\varepsilon \nabla_x X_s^x z|_H^2 ds &\le C_\varepsilon^2 \int_0^\tau |(-A)^{1/2} \nabla_x X_s^x z|_H^{4\varepsilon} |\nabla_x X_s^x z|_H^{2-4\varepsilon} ds \\
&\le C_\varepsilon^2 \left(\int_0^\tau |(-A)^{1/2} \nabla_x X_s^x z|_H^2 ds \right)^{2\varepsilon} \left(\int_0^\tau |\nabla_x X_s^x z|_H^2 ds \right)^{1-2\varepsilon} \le C\tau^{1-2\varepsilon} |z|_H^2,
\end{aligned}$$

with $C := \sup_{\varepsilon \in (0,1/2)} C_\varepsilon^2$. We conclude that (2.11) holds for $t = 0$, $\varepsilon \in [0, 1/2]$ and any $z \in E$.
Let us now consider $z \in H$, and let $(z_n) \subset E$ be an approximating sequence of z in H. Then, from (2.6), for any $\tau \in [0, T]$ we get

$$\nabla_x X_\tau^x z_n \to \nabla_x X_\tau^x z \quad \mathbb{P}\text{-a.s. in } H, \text{ as } n \to +\infty. \tag{2.13}$$

8

Since (2.11) holds for any $z \in E$, it follows that $((-A)^{1/2}\nabla_x X^x z_n)$ is a Cauchy sequence in $\mathscr{M}^2([0,T];H)$, and therefore there exists a process $\xi \in \mathscr{M}^2([0,T];H)$ such that $(-A)^{1/2}\nabla_x X^x z_n \to \xi$ in $\mathscr{M}^2([0,T];H)$. Since $(-A)^{-1/2}$ is a bounded operator on H, it follows that

$$\nabla_x X^x z_n = (-A)^{-1/2}(-A)^{1/2}\nabla_x X^x z_n \to (-A)^{-1/2}\xi,$$

in $\mathscr{M}^2([0,T];H)$. Therefore, also by (2.13), $(-A)^{-1/2}\xi = \nabla_x X^x z$ a.e. in $(0,T)$ and $\mathbb{P}$-a.s., which means that $\nabla_x X^x z \in D((-A)^{1/2})$ a.e. in $(0,T)$ and $\mathbb{P}$-a.s., and $(-A)^{1/2}\nabla_x X^x z = \xi$ a.e. in $(0,T)$ and $\mathbb{P}$-a.s. In particular, we get

$$\int_0^t |(-A)^{1/2}\nabla_x X_s^{t,x} z|_H^2 ds \leq C|z|_H^2.$$

Again, by applying interpolation estimates we see that (2.11) holds for $\varepsilon \in [0,1/2]$, $t = 0$ and any $z \in H$. $\qquad \square$

Proposition 2.7. *Let Hypothesis 2.1 holds true and let $x \in E$, $z \in H$ and $t \in [0,T]$. Then, for any $x \in E$ and $z \in H$,*

$$\mathbb{E}\Big[\sup_{\tau \in [t,T]} |(-A)^\alpha \nabla_x X_\tau^{t,x} z|_H \Big] \leq C|z|_H \left((\tau-t)^{-\alpha} + (\tau-t)^{1-\alpha} \left(|x|_E^{2m+1} + C_T \right) \right), \qquad (2.14)$$

and if in addition $z \in D((-A)^\alpha)$, then (2.17) gives

$$\mathbb{E}\Big[\sup_{\tau \in [t,T]} |(-A)^\alpha \nabla_x X_\tau^{t,x} z|_H \Big] \leq C \left(|(-A)^\alpha z|_H + (\tau-t)^{1-\alpha} \left(|x|_E^{2m+1} + C_T \right) |z|_H \right), \qquad (2.15)$$

where $C_T := \mathbb{E}[\sup_{\tau \in [0,T]} |w^A(\tau)|_E^{2m+1}]$.

Proof. We prove estimate (2.14), then (2.15) follows from analogous arguments. Fix $x \in E$, $z \in H$ and let us consider $t = 0$. We recall that A generates an analytic semigroup on H and therefore $e^{tA}x$ belongs to $D((-A)^k)$ for any $k \in \mathbb{N}$ and any $h \in H$, and $|(-A)^\beta e^{tA} h|_H \leq C_\beta t^{-\beta} |h|_H$ for any $\beta \geq 0$ and some positive constant C_β. This means that $\nabla_x X_\tau^x z \in D((-A)^\alpha)$ for any $\tau \in [0,T]$ and, recalling (2.7),

$$(-A)^\alpha \nabla_x X_\tau^x z = (-A)^\alpha e^{\tau A} z + \int_0^\tau (-A)^\alpha e^{(\tau-s)A} \nabla_x F(X_s^x) \nabla_x X_s^x z \, dx, \quad \mathbb{P}\text{-a.s.}$$

From Hypothesis 2.1-4. and (2.6) we deduce that

$$\begin{aligned}
|(-A)^\alpha \nabla_x X_\tau^x z|_H &\leq C_\alpha \tau^{-\alpha} |z|_H + c\, C_\alpha \int_0^\tau (\tau-s)^{-\alpha} |\nabla F(X_s^x) \nabla_x X_s^x z|_H \, ds \\
&\leq C|z|_H \Big(\tau^{-\alpha} + \tau^{1-\alpha} \Big(|x|_E^{2m+1} + \sup_{\tau \in [0,T]} |w^A(\tau)|_E^{2m+1} \Big) \Big), \quad \mathbb{P}\text{-a.s.,}
\end{aligned}$$

for some positive constant C independent of x,z. Then, for any $\tau \in (t,T]$,

$$|(-A)^\alpha \nabla_x X_\tau^{t,x} z|_H \leq C|z|_H \Big((\tau-t)^{-\alpha} + (\tau-t)^{1-\alpha} \Big(|x|_E^{2m+1} + \sup_{\tau \in [t,T]} |w^A(\tau)|_E^{2m+1} \Big) \Big), \ \mathbb{P}\text{-a.s.,} \quad (2.16)$$

for some positive constant C independent of x,z,t. Further, if $z \in D((-A)^\alpha)$, then

$$|(-A)^\alpha \nabla_x X_\tau^{t,x} z|_H \leq C \Big(|(-A)^\alpha z|_H + (\tau-t)^{1-\alpha} \Big(|x|_E^{2m+1} + \sup_{\tau \in [t,T]} |w^A(\tau)|_E^{2m+1} \Big) |z|_H \Big), \ \mathbb{P}\text{-a.s.} \quad (2.17)$$

Taking the expectation in (2.16) and (2.17) we get respectively (2.14) and (2.15). $\qquad \square$

3 The forward-backward system

We consider the following forward-backward system of stochastic differential equations (FBSDE for short) for the unknown (X, Y, Z) (also denoted by $(X^{t,x}, Y^{t,x}, Z^{t,x})$ to stress the dependence on the initial conditions t and x): for given $t \in [0, T]$ and $x \in E$,

$$\begin{cases} dX_\tau = AX_\tau d\tau + F(X_\tau) d\tau + (-A)^{-\alpha} dW_\tau, & \tau \in [t, T], \\ X_t = x, \\ dY_\tau = -\psi(\tau, X_\tau, Y_\tau, Z_\tau)\, d\tau + Z_\tau\, dW_\tau, & \tau \in [t, T], \\ Y_T = \phi(X_T). \end{cases} \tag{3.1}$$

The second equation is of backward type for the unknown (Y, Z) and depends on the Markov process X. Under suitable assumptions on the coefficients ψ (the so-called generator of the BSDE) and ϕ we look for a solution consisting of a pair of processes $(Y, Z) \in \mathscr{S}^2([t, T]) \times \mathscr{M}^2([t, T]; H)$. More precisely, we will assume that ψ is Lipschitz continuous with respect to y and locally Lipschitz continuous and with quadratic growth with respect to z, as stated below.

Hypothesis 3.1. *The functions $\phi : E \to \mathbb{R}$ and $\psi : [0, T] \times E \times \mathbb{R} \times H \to \mathbb{R}$ in* (3.1) *satisfy the following.*

(i) *ϕ is continuous, and there exists a nonnegative constant K_ϕ such that $|\phi(x)| \leq K_\phi$ for every $x \in E$.*

(ii) *ψ is measurable and, for every fixed $t \in [0, T]$, the map $\psi(t, \cdot, \cdot, \cdot) : E \times \mathbb{R} \times H \to \mathbb{R}$ is continuous. Moreover, there exist nonnegative constants L_ψ and K_ψ such that*

$$|\psi(t, x_1, y_1, z_1) - \psi(t, x_2, y_2, z_2)| \leq L_\psi \left(|x_1 - x_2|_E + |y_1 - y_2| + |z_1 - z_2|_H (1 + |z_1|_H + |z_2|_H) \right),$$
$$|\psi(t, x, 0, 0)| \leq K_\psi,$$

for every $t \in [0, T]$, $x_1, x_2 \in E$, $y_1, y_2 \in \mathbb{R}$ and $z_1, z_2 \in H$.

Theorem 3.2. *Assume that Hypotheses 2.1 and 3.1 hold true, and for any $(t, x) \in [0, T] \times E$, let $(X^{t,x}, Y^{t,x}, Z^{t,x})$ be a solution to the FBSDE* (3.1). *Then, there exists a unique solution of the Markovian BSDE in* (3.1) *such that*

$$\|Y^{t,x}\|_{\mathscr{S}^2([t,T])} + \|Z^{t,x}\|_{\mathscr{M}^2([t,T];H)} \leq C,$$

where C is a constant that may depend on T, A, F, K_ψ, L_ψ, K_ϕ.. Moreover, setting $v(t, x) := Y_t^{t,x}$,

$$Y_s^{t,x} = v(s, X_s^{t,x}), \quad \mathbb{P}\text{-a.s., } s \in [t, T], \tag{3.2}$$

and there exists a Borel function $u : [t, T] \times E \to H$ such that

$$Z_s^{t,x} = u(s, X_s^{t,x}), \quad \mathbb{P}\text{-a.s., a.e. } s \in [t, T]. \tag{3.3}$$

Proof. The first part of the result substantially follows from [15]. Identities (3.2)-(3.3) are a consequence of the Markov property of X, see for instance Theorem 4.1 in [7] or the proof of Theorem 5.1 in [12]. $\square$

We recall some further estimates for the solution (Y, Z) of the forward-backward system (3.1). In particular, $Z \in \mathscr{M}^p([t, T]; H)$, for any $p \geq 1$. The corresponding proof can be found e.g. in [19].

Proposition 3.3. *Assume that Hypotheses 2.1 and 3.1 hold true, and for any $(t, x) \in [0, T] \times E$, let $(X^{t,x}, Y^{t,x}, Z^{t,x})$ be a solution to the FBSDE* (3.1). *Then, for all $p \geq 1$,*

$$\|Y^{t,x}\|_{\mathscr{S}^p([t,T])} + \|Z^{t,x}\|_{\mathscr{M}^p([t,T];H)} \leq C,$$

where C is a constant that may depend on T, A, F, K_ψ, L_ψ, K_ϕ.

At this point, we aim at proving a stability result for the BSDE when the final datum and the generator are approximated by sequences of Fréchet differentiable functions $(\phi_n)_{n\geq1}$, $(\psi_\ell)_{\ell\geq1}$, converging pointwise respectively to ϕ and ψ, and such that, for all $t \in [0,T]$, $x, x_1, x_2 \in E$, $y_1, y_2 \in \mathbb{R}$, $z_1, z_2 \in H$,

$$|\phi_n(x)| \leq K_\phi, \quad |\psi_\ell(t,x,0,0)| \leq K_\psi, \tag{3.4}$$

$$|\psi_\ell(t,x_1,y_1,z_1) - \psi_\ell(t,x_2,y_2,z_2)| \leq L_\psi(|x_1 - x_2|_E + |y_1 - y_2| + |z_1 - z_2|_H(1 + |z_1|_H + |z_2|_H)). \tag{3.5}$$

To provide such approximations we extend the results in [23] valid for Hilbert spaces: by using Schauder basis, the approximation performed in that paper can be achieved also in Banach spaces. We start by introducing the following objects.

Definition 3.4.   *i) Denote by $(e_n)_{n\geq1}$ the normalized Schauder basis in E and by (h_n) an orthonormal basis of H. For any $n \in \mathbb{N}$, we define the projections $Q_n : H \to \mathbb{R}^n$ and $P_n : E \to \mathbb{R}^n$ as follows:*

$$Q_n z := (z_1,\ldots,z_n), \quad P_n x := (x_1,\ldots,x_n),$$

for any $z \in H$ and $x \in E$ with $z = \sum_{n=1}^\infty z_n h_n$ and $x = \sum_{n=1}^\infty x_n e_n$, $z_n, x_n \in \mathbb{R}$.

ii) We consider nonnegative smooth kernels $\vartheta \in C_c^\infty(\mathbb{R})$ and $\rho_m \in C_c^\infty(\mathbb{R}^m)$, $m \in \mathbb{N}$, such that

$$\mathrm{supp}\,(\vartheta) \subseteq \{\zeta \in \mathbb{R} : |\zeta| \leq 1\}, \quad \mathrm{supp}\,(\rho_m) \subseteq \{\xi \in \mathbb{R}^m : |\xi| \leq m^{-1}\}, \quad \|\vartheta\|_{L^1(\mathbb{R})} = \|\rho_m\|_{L^1(\mathbb{R}^m)} = 1.$$

iii) For any $n, \ell \in \mathbb{N}$, we set $\vartheta_\ell(\zeta) = \ell\vartheta(\ell\zeta)$ for any $\zeta \in \mathbb{R}$, and

$$\phi_n(x) = \int_{\mathbb{R}^n} \rho_n(\xi - P_n x)\phi\Big(\sum_{i=1}^n \xi_i e_i\Big)d\xi, \tag{3.6}$$

$$\psi_\ell(t,x,y,z) := \int_{\mathbb{R}^\ell}\int_{\mathbb{R}^\ell}\int_{\mathbb{R}} \rho_\ell\,(\xi - P_\ell x)\,\rho_\ell(\eta - Q_\ell z)\vartheta_\ell(y - \zeta)\psi\Big(t,\sum_{i=1}^\ell \xi_i e_i, \zeta, \sum_{j=1}^\ell \eta_j h_j\Big)d\zeta\,d\eta\,d\xi. \tag{3.7}$$

It is not hard to prove the following lemma.

Lemma 3.5. *Le ϕ and ψ satisfy Hypothesis 3.1. Then the following hold.*

(i) For any $n \in \mathbb{N}$, the function ϕ_n in (3.6) is Fréchet differentiable, satisfies estimate (3.4), and

$$\lim_{n\to+\infty}\phi_n(x) = \phi(x), \quad x \in E.$$

(ii) For any $\ell \in \mathbb{N}$, the function ψ_ℓ in (3.7) is Fréchet differentiable with respect to x, y, z, satisfies estimates (3.4)-(3.5), and

$$\lim_{\ell\to+\infty}\psi_\ell(t,x,y,z) = \psi(t,x,y,z), \quad (t,x,y,z) \in [0,T] \times E \times \mathbb{R} \times H.$$

We can now give a stability result for the Markovian BSDE in (3.1) related to a forward process X taking values in the Banach space E, when the final datum and the generator are approximated respectively by the sequences $(\phi_n)_{n\geq1}$ and $(\psi_\ell)_{\ell\geq1}$. Notice that, differently from [19], where the forward process X takes its values in a Hilbert space H, the final datum and the generator cannot be approximated in the norm of the uniform convergence by means of their inf-sup convolutions. However, even if we only have pointwise convergence, this turns out to be enough for our purposes.

Clearly, the following result holds true if we approximate only ψ or ϕ. We postpone its proof to Section 7.1.

Proposition 3.6. *Assume that Hypotheses 2.1 and 3.1 hold true. For any $(t, x) \in [0, T] \times E$, let (X, Y, Z) be a solution to the FBSDE (3.1). Let $(Y^{n,l}, Z^{n,l})$ be the solution of the BSDE in the forward-backward system*

$$
\begin{cases}
dX_\tau = AX_\tau d\tau + F(X_\tau) d\tau + (-A)^{-\alpha} dW_\tau, & \tau \in [t, T], \\
X_t = x, \\
dY_\tau^{n,l} = -\psi_l(\tau, X_\tau, Y_\tau^{n,l}, Z_\tau^{n,l}) \, d\tau + Z_\tau^{n,l} \, dW_\tau, & \tau \in [t, T], \\
Y_T^{n,l} = \phi_n(X_T),
\end{cases}
\tag{3.8}
$$

that is, the FBSDE (3.1) with final datum equal to ϕ_n in (3.6) in place of ϕ, and with generator ψ_l in (3.7) in place of ψ. Then, for all $p \geq 1$, the unique solution of the Markovian BSDE in (3.1) is such that

$$
\|Y - Y^{n,l}\|_{\mathscr{S}^p([t,T])} + \|Z - Z^{n,l}\|_{\mathscr{M}^p([t,T];H)} \to 0 \qquad \text{as } n, l \to \infty.
$$

We now state a result on differentiability for the solution of a Markovian BSDE with generator with quadratic growth, with respect to the initial datum x.

Proposition 3.7. *Assume that Hypotheses 2.1 and 3.1 hold true, and for any $(t, x) \in [0, T] \times E$, let $(X^{t,x}, Y^{t,x}, Z^{t,x})$ be a solution to the FBSDE (3.1). Assume moreover that ϕ is Gâteaux differentiable with bounded derivative, and that ψ is Gâteaux differentiable with respect to x, y and z. Then the triple of processes $(X^{t,x}, Y^{t,x}, Z^{t,x})$ is Gâteaux differentiable as a map from E with values in $\mathscr{S}^2((t,T]; E) \times \mathscr{S}^2([t,T]) \times \mathscr{M}^2([t,T]; H)$ and, for any $h \in E$,*

$$
\begin{cases}
-d\nabla_x Y_\tau^{t,x} h & = \nabla_x \psi(\tau, X_\tau^{t,x}, Y_\tau^{t,x}, Z_\tau^{t,x}) \nabla_x X_\tau^{t,x} h \, d\tau + \nabla_y \psi(\tau, X_\tau^{t,x}, Y_\tau^{t,x}, Z_\tau^{t,x}) \nabla_x Y_\tau^{t,x} h \, d\tau \\
& \quad + \nabla_z \psi(\tau, X_\tau^{t,x}, Y_\tau^{t,x}, Z_\tau^{t,x}) \nabla_x Z_\tau^{t,x} h \, d\tau - \nabla_x Z_\tau^{t,x} h \, dW_\tau, \quad \tau \in [t, T], \\
\nabla_x Y_T^{t,x} h & = \nabla_x \phi(X_T^{t,x}) \nabla_x X_T^{t,x} h, \\
d\nabla_x X_\tau^{t,x} h & = A\nabla_x X_\tau^{t,x} h d\tau + \nabla F(X_\tau^{t,x}) \nabla_x X_\tau^{t,x} h \, d\tau, \quad \tau \in [t, T], \\
\nabla_x X_t^{t,x} h & = h.
\end{cases}
\tag{3.9}
$$

Moreover, there exists a constant C, only dependent on T, A, F, K_ψ, L_ψ, K_ϕ, such that

$$
\mathbb{E}\left[\sup_{\tau \in [t, T]} |\nabla Y_\tau^{t,x} h|^2 + \int_t^T |\nabla_x Z_\tau^{t,x} h|_H^2 \, d\tau \right] \leq C|h|_H^2.
\tag{3.10}
$$

Proof. In the case of a Markovian BSDE with generator ψ quadratic with respect to Z and related to a forward process taking values in a Hilbert space, the result is given in Theorem 4.5 of [2]. Since in Proposition 2.5 we have proved the differentiability of $X^{t,x}$ with respect to $x \in E$, the same conclusions hold when the forward process takes values in the Banach space E, namely

$$
\mathbb{E}\left[\sup_{\tau \in [t, T]} |\nabla_x Y_s^{t,x} h|^2 + \int_t^T |\nabla_x Z_\tau^{t,x} h|_H^2 \, d\tau \right] \leq C|h|_E^2.
$$

The stronger estimate (3.10) comes from Proposition 2.5, estimate (2.6). □

3.1 Identification of Z and a priori estimates on (Y, Z)

We now prove an a priori estimate on $Z^{t,x}$ depending only on the L^∞-norm of the final datum. The novelty towards [19] is that we work in a Banach space and the pseudo-inverse of the diffusion operator is the unbounded operator $(-A)^\alpha$. In order to get this estimate and also for the subsequent results of the paper, it will be crucial to prove the identification

$$
Z_t^{t,x} = \nabla_x Y_t^{t,x} (-A)^{-\alpha},
$$

12

which is new in the Banach space framework and in the case of quadratic generator with respect to z. We have to make the following assumption:

Hypothesis 3.8. *There exists a Banach space $E_0 \subset D((-A)^\alpha)$ dense in H such that $(-A)^{-\alpha} E_0 \subset E$ and $(-A)^{-\alpha} : E_0 \to E$ is continuous.*

Remark 3.9. *Notice that, if $\mathscr{D} \subset \mathbb{R}^2$ is a bounded open domain with smooth boundary, $H = L^2(\mathscr{D})$ and A is the Laplace operator in dimension 2 with Dirichlet boundary conditions, we can take $E_0 = D((-A)^{\frac{1}{2}})$ and all the requirements of Hypothesis 3.8 are verified.*

Theorem 3.10. *Assume that Hypotheses 2.1, 3.1 and 3.8 hold true, that ϕ is Gâteaux differentiable with bounded derivative, and that ψ is Gâteaux differentiable with respect to x, y and z. For any $(t, x) \in [0, T] \times E$, let $(X^{t,x}, Y^{t,x}, Z^{t,x})$ be the solution to the FBSDE (3.1). Then the triple of processes $(X^{t,x}, Y^{t,x}, Z^{t,x})$ is Gâteaux differentiable as a map from E with values in $\mathscr{S}^2((t,T]; E) \times \mathscr{S}^2([t,T]) \times \mathscr{M}^2([t,T]; H)$. Moreover, setting $v(t, x) = Y_t^{t,x}$, then, $\mathbb{P}$-a.s.,*

$$Y_s^{t,x} = v(s, X_s^{t,x}), \quad s \in [t, T], \tag{3.11}$$

$$Z_s^{t,x} h = \nabla_x v(s, X_s^{t,x}) \nabla_x X_s^{t,x} (-A)^{-\alpha} h, \quad \text{a.e. } s \in [t, T], \ h \in E_0. \tag{3.12}$$

Proof. The key ingredient consists in generalizing the variational approach introduced in [17], by applying it both to the forward process X and to the pair of processes (Y, Z): this extension, together with the quadratic growth of ψ, makes the proof more involved with respect to the one of Theorem 3.17 in [17]. Due to the many delicate passages, the proof is quite lenghty and we postpone it to the Section 7.2. $\quad\square$

Corollary 3.11. *Under the assumptions of Theorem 3.10 we have*

$$Z_s^{t,x} h = \nabla v(s, X_s^{t,x})(-A)^{-\alpha} h, \quad h \in H, \text{ for a.e. } s \in [t, T], \ \mathbb{P}\text{-a.s.,}$$

where $\nabla_x v(s, x)(-A)^{-\alpha}$ denotes an extension of the operator $\nabla_x v(s, x)(-A)^{-\alpha} : E_0 \to \mathbb{R}$ to the whole space H. Moreover, there exists a constant C, that may depend also on $\nabla_x \phi$, $\nabla_x \psi$ and L_ψ, such that

$$|Z_s^{t,x}|_H \leq C, \quad \text{for a.e. } s \in [t, T], \ \mathbb{P}\text{-a.s.} \tag{3.13}$$

Proof. Since E_0 is dense in H, by (3.12) in Theorem 3.10, for almost every $s \in [0, T]$ and almost surely with respect to the law of X, the operator $\nabla_x v(s, x)(-A)^{-\alpha} : E_0 \to \mathbb{R}$ extends to an operator defined on the whole H, which we still denote $\nabla v(s, x)(-A)^{-\alpha}$.
Moreover, from (3.12) and by the Markov property, we get

$$Z_\sigma^{t,x} = Z_\sigma^{\sigma, X_\sigma^{t,x}} = \nabla_x Y_\sigma^{\sigma,k}\big|_{k = X_\sigma^{t,x}} (-A)^{-\alpha}, \quad \text{for a.e. } \sigma \in [0, T], \ \mathbb{P}\text{-a.s.}$$

Estimate (3.13) follows from the fact that, by (3.10), we have $\sup_\sigma |\nabla_x Y_\sigma^{\sigma,k}| \leq C$, where C is a constant that does not depend on k. $\quad\square$

Now we use the previous result to give a priori estimates on $Z^{t,x}$, and on $\nabla Y^{t,x}$. Notice that, since the diffusion operator $(-A)^{-\alpha}$ has an unbounded pseudo-inverse operator, when $\alpha > 0$ estimates (3.14)-(3.15) are completely new and cannot be deduced by analogous results with bounded diffusion operators with bounded inverse. We postpone the proof to Section 7.3.

Proposition 3.12. *Assume that Hypotheses 2.1 and 3.1 hold true, and for any $(t, x) \in [0, T] \times E$, let $(X^{t,x}, Y^{t,x}, Z^{t,x})$ be the solution to the FBSDE (3.1). Then there exists a positive constant C_T only depending on T, A, F K_ϕ, L_ψ, K_ψ such that*

$$|Z_t^{t,x} h| \leq C_T (T - t)^{-1/2} |h|_H, \quad \mathbb{P}\text{-a.s.}, \ h \in H, \tag{3.14}$$

$$|\nabla_x Y_t^{t,x} h| \leq C_T (T - t)^{-1/2-\alpha} |h|_H, \quad \mathbb{P}\text{-a.s.}, \ h \in H. \tag{3.15}$$

4 The Bismut-Elworthy formula and the semilinear Kolmogorov equation: the Lipschitz case

Recall that we deal with a process X taking values in a Banach space and solution to equation (2.1), with special diffusion operator $(-A)^{-\alpha}$ with pseudo-inverse $(-A)^\alpha$ which is not bounded.

In the present section we adequate to our framework the results in [10]. More precisely, in Subsection 4.1 we present a nonlinear version of the Bismut-Elworthy formula in the case of Lipschitz generator, which extends the one provided in [10] in the case of a process X taking values in a Hilbert space, and with a bounded diffusion operator with bounded inverse. Providing the Bismut-Elworthy formula in the case of Lipschitz generator is a fundamental step in order to obtain the analogous formula in the quadratic case. Moreover, it allows us to give an existence and uniqueness result in the Banach framework for the semilinear Kolmogorov related to the process X, and with coefficients ϕ and ψ not necessarily differentiable, see Subsection 4.2.

For $0 \leq t < s \leq T$ and $h \in H$ we define the real valued random variables

$$U_s^{h,t,x} := \frac{1}{s-t} \int_t^s \langle (-A)^\alpha \nabla_x X_r^{t,x} h, dW_r \rangle. \tag{4.1}$$

Notice that, for any $h \in H$, the process $U^{h,t,x}$ is well defined thanks to formula (2.11) in Proposition 2.6. In what follows we prove some useful estimates on the process $U^{h,t,x}$.

Lemma 4.1. *Assume that Hypotheses 2.1 hold true. For any $(t,x) \in [0,T] \times E$, let $X^{t,x}$ be the unique mild solution to (2.1). Then, for any $h \in H$ and for any $q \geq 1$,*

$$\left(\mathbb{E}[|U_s^{h,t,x}|^q] \right)^{1/q} \leq C \, (s-t)^{-(\frac{1}{2}+\alpha)} \, |h|_H, \tag{4.2}$$

and also

$$\left(\mathbb{E}\Big[\sup_{s \in [\frac{t+T}{2},T]} |U_s^{h,t,x}|^q \Big] \right)^{1/q} \leq C(T-t)^{-(\frac{1}{2}+\alpha)} |h|_H. \tag{4.3}$$

Proof. We compute

$$\mathbb{E}[|U_s^{h,t,x}|^q] = \mathbb{E}\Big[\Big| \frac{1}{s-t} \int_t^s \langle (-A)^\alpha \nabla_x X_r^{t,x} h, dW_r \rangle \Big|^q \Big] \leq \frac{1}{(s-t)^q} \mathbb{E}\Big[\Big(\int_t^s |(-A)^\alpha \nabla_x X_r^{t,x} h|^2 \, dr \Big)^{q/2} \Big]$$

$$\leq \frac{1}{(s-t)^q} C((s-t)^{1-2\alpha} |h|_H^2)^{q/2} = C(s-t)^{-q(\frac{1}{2}+\alpha)} |h|_H^q,$$

where in the latter inequality we have used formula (2.11) of Proposition 2.6 with $\varepsilon = \alpha$. Analogously, we have

$$\mathbb{E}\Big[\sup_{s \in [\frac{t+T}{2},T]} |U_s^{h,t,x}|^q \Big] \leq C \frac{1}{(T-t)^q} \mathbb{E}\Big[\Big(\int_t^T |(-A)^\alpha \nabla_x X_r^{t,x} h|^2 \, dr \Big)^{q/2} \Big] \leq C(T-t)^{-q(\frac{1}{2}+\alpha)} |h|_H^q.$$

$\square$

4.1 The Bismut formula

We can now give a version of the Bismut-Elworthy formula in the case of Lipschitz generator and in the Banach space framework. We consider only the case of final datum ϕ and generator ψ bounded with respect to x, since we aim to treat such a model in the quadratic case. We start with the case when the coefficients are also differentiable. An analogous result is proved in [10] in the Hilbert space framework

using the Malliavin calculus. Since here the process X takes its values in a Banach space, we avoid the use of the Malliavin calculus by exploiting instead techniques similar to the ones used in the proof of Theorem 3.10.

In the rest of the section we will assume the following, that substitutes Hypothesis 3.1.

Hypothesis 4.2. *The functions $\phi : E \to \mathbb{R}$ and $\psi : [0,T] \times E \times \mathbb{R} \times H \to \mathbb{R}$ in (3.1) satisfy the following.*

(i) *ϕ is continuous, and there exist a nonnegative constant K_ϕ such that $|\phi(x)| \leq K_\phi$ for every $x \in E$.*

(ii) *ψ is measurable and, for every fixed $t \in [0,T]$, the map $\psi(t, \cdot, \cdot, \cdot) : E \times \mathbb{R} \times H \to \mathbb{R}$ is continuous. Moreover, there exist nonnegative constants L_ψ and K_ψ such that*

$$|\psi(t, x_1, y_1, z_1) - \psi(t, x_2, y_2, z_2)| \leq L_\psi \left(|x_1 - x_2|_E + |y_1 - y_2| + |z_1 - z_2|_H \right),$$
$$|\psi(t, x, 0, 0)| \leq K_\psi,$$

for every $t \in [0,T]$, $x_1, x_2 \in E$, $y_1, y_2 \in \mathbb{R}$ and $z_1, z_2 \in H$.

Remark 4.3. *The results in this section can be extended to the case of data ψ and ϕ having polynomial growth with respect to x. Since in Section 5, which is the core of the paper and where the driver ψ has quadratic growth with respect to z, we are able to consider only data ψ and ϕ bounded with respect to x, we have decided to present also here only the bounded case. We also notice that when ψ has quadratic growth with respect to z, also in the Hilbert case, up to our knowledge it is not known how to consider data ϕ and ψ with polynomial growth with respect to x, see [19].*

Theorem 4.4. *Let Hypotheses 2.1 and 4.2 hold true, and for any $(t,x) \in [0,T] \times E$, let $(X^{t,x}, Y^{t,x}, Z^{t,x})$ be a solution of the forward-backward system (3.1), and let $U^{h,t,x}$ be the process defined in (4.1). Assume moreover that ϕ is Gâteaux differentiable with bounded derivative, and that ψ is Gâteaux differentiable with respect to x, y and z. Then for $t \leq s \leq T$, $x \in E$, $h \in H$,*

$$\mathbb{E}\left[\nabla_x Y_s^{t,x} h \right] = \mathbb{E}\left[\int_s^T \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x} \right) U_r^{h,t,x} \, dr \right] + \mathbb{E}\left[\phi(X_T^{t,x}) U_T^{h,t,x} \right]. \tag{4.4}$$

Proof. Let ξ be a given square integrable E_0-valued predictable process, and $X^{\varepsilon,t,x}$ be a mild solution to the equation

$$\begin{cases} dX_\tau^{\varepsilon,t,x} = AX_\tau^{\varepsilon,t,x} d\tau + F\left(X_\tau^{\varepsilon,t,x} \right) d\tau + (-A)^{-\alpha} \varepsilon \xi_\tau d\tau + (-A)^{-\alpha} dW_\tau, & \tau \in [t,T], \\ X_t^{\varepsilon,t,x} = x. \end{cases} \tag{4.5}$$

We also consider the pair of processes $(Y^{\varepsilon,t,x}, Z^{\varepsilon,t,x})$ solution to the Markovian BSDE

$$\begin{cases} -dY_\tau^{\varepsilon,t,x} = \psi(\tau, X_\tau^{\varepsilon,t,x}, Y_\tau^{\varepsilon,t,x}, Z_\tau^{\varepsilon,t,x})d\tau - Z_\tau^{\varepsilon,t,x} dW_\tau, & \tau \in [t,T], \\ Y_T^{\varepsilon,t,x} = \phi(X_T^{\varepsilon,t,x}). \end{cases} \tag{4.6}$$

Arguing similarly to the proof of Theorem 3.10, we define

$$\dot{X}_\tau := \frac{d}{d\varepsilon}_{|\varepsilon=0} X_\tau^{\varepsilon,t,x}, \ \dot{Y}_\tau := \frac{d}{d\varepsilon}_{|\varepsilon=0} Y_\tau^{\varepsilon,t,x}, \ \dot{Z}_\tau := \frac{d}{d\varepsilon}_{|\varepsilon=0} Z_\tau^{\varepsilon,t,x}, \quad \tau \in [t,T], \tag{4.7}$$

which are solution to the forward-backward system (7.14) with $s - \delta = t$. We already know (see formula (7.9) with $s - \delta = t$) that

$$\dot{X}_\tau = \int_t^\tau \nabla_x X_\tau^{\sigma, X_\sigma^{t,x}} (-A)^{-\alpha} \xi_\sigma d\sigma, \quad \tau \in [t,T], \ \mathbb{P}\text{-a.s.} \tag{4.8}$$

15

Now we want to prove a similar identification for the pair $(\dot{Y}, \dot{Z})$. To this aim, for any $\sigma \in [t,T]$, we consider the Markovian BSDE in (3.1) on the time interval $[\sigma, T]$, and with initial condition y given at time σ; from Proposition 3.7 we know that the derivative with respect to $y \in E$ in the direction $h \in E$ satisfies the following BSDE, that we write in integral form: for any $\tau \in [t,T]$, $\mathbb{P}$-a.s.,

$$
\begin{cases}
\nabla_x Y_\tau^{\sigma,y} h &= \nabla_x \phi(X_T^{\sigma,y}) \nabla_x X_T^{\sigma,x} h - \displaystyle\int_\tau^T \nabla_x Z_r^{\sigma,x} h \, dW_r + \int_\tau^T \big(\nabla_x \psi(r, X_r^{\sigma,y}, Y_r^{\sigma,y}, Z_r^{\sigma,y}) \nabla_x X_r^{\sigma,y} h \\
&\quad + \nabla_y \psi(r, X_r^{\sigma,y}, Y_r^{\sigma,y}, Z_r^{\sigma,y}) \nabla_x Y_r^{\sigma,y} h \, dr + \nabla_z \psi(r, X_r^{\sigma,y}, Y_r^{\sigma,y}, Z_r^{\sigma,y}) \nabla_x Z_r^{\sigma,y} h \big) dr.
\end{cases}
$$
(4.9)

Let us take $y = X_\sigma^{t,x}$ and $h = (-A)^{-\alpha} \xi_\sigma$ in (4.9), and let us integrate both sides with respect to $\sigma \in [t,\tau]$. By inverting the order of integration where necessary, and using the Markov property, it is immediate to get

$$
\int_t^\tau \nabla_x Y_\tau^{\sigma, X_\sigma^{t,x}} (-A)^{-\alpha} \xi_\sigma \, d\sigma
$$

$$
= \int_t^\tau \nabla_x \phi(X_T^{\sigma, X_\sigma^{t,x}}) \nabla_x X_T^{\sigma, X_\sigma^{t,x}} (-A)^{-\alpha} \xi_\sigma \, d\sigma - \int_\tau^T \Big(\int_t^\tau \nabla_x Z_r^{\sigma, X_\sigma^{t,x}} (-A)^{-\alpha} \xi_\sigma \, d\sigma \Big) dW_r
$$

$$
+ \int_\tau^T \Big(\int_t^\tau \nabla_x \psi \Big(r, X_r^{\sigma, X_\sigma^{t,x}}, Y_r^{\sigma, X_\sigma^{t,x}}, Z_r^{\sigma, X_\sigma^{t,x}} \Big) \nabla_x X_r^{\sigma, X_\sigma^{t,x}} (-A)^{-\alpha} \xi_\sigma \, d\sigma \Big) dr
$$

$$
+ \int_t^\tau \Big(\int_t^\tau \nabla_y \psi(r, X_r^{\sigma, X_\sigma^{t,x}}, Y_r^{\sigma, X_\sigma^{t,x}}, Z_r^{\sigma, X_\sigma^{t,x}}) \nabla_x Y_r^{\sigma, X_\sigma^{t,x}} (-A)^{-\alpha} \xi_\sigma \, d\sigma \Big) dr
$$

$$
+ \int_t^\tau \Big(\int_t^\tau \nabla_z \psi(r, X_r^{\sigma, X_\sigma^{t,x}}, Y_r^{\sigma, X_\sigma^{t,x}}, Z_r^{\sigma, X_\sigma^{t,x}}) \nabla_x Z_r^{\sigma, X_\sigma^{t,x}} (-A)^{-\alpha} \xi_\sigma \, d\sigma \Big) dr.
$$

By (4.6) and (4.7), together with (4.8), we can conclude that

$$
\dot{Y}_\tau = \int_t^\tau \nabla_x Y_\tau^{\sigma, X_\sigma^{t,x}} (-A)^{-\alpha} \xi_\sigma d\sigma, \quad \dot{Z}_\tau = \int_t^\tau \nabla_x Z_\tau^{\sigma, X_\sigma^{t,x}} (-A)^{-\alpha} \xi_\sigma d\sigma, \quad \tau \in (t,T], \ \mathbb{P}\text{-a.s.,} \qquad (4.10)
$$

since these two pairs of processes satisfies the same BSDE. By density, arguing as in Corollary 3.11, we infer that formulas (4.8) and (4.10) hold true for any square integrable H-valued predictable process ξ. Now, let $\eta \in E$, and let us take

$$
\xi_\tau := (-A)^\alpha \nabla_x X_\tau^{t,x} \eta, \quad \tau \in (t,T]. \tag{4.11}
$$

Notice that, since $(-A)^\alpha \nabla_x X_\tau^{t,x} \eta \in D((-A)^{1/2-\alpha})$ $\mathbb{P}$-a.s., thanks to Proposition 2.6, $(-A)^\alpha \nabla_x X_\tau^{t,x} \eta \in H$ for any $\tau \in (t,T]$, $\mathbb{P}$-a.s., and so

$$
\int_t^\tau e^{(\tau - \sigma)A} (-A)^{-\alpha} (-A)^\alpha \nabla_x X_\sigma^{t,x} \eta \, d\sigma = \int_t^\tau e^{(\tau - \sigma)A} \nabla_x X_\sigma^{t,x} \eta \, d\sigma, \quad \tau \in (t,T], \ \mathbb{P}\text{-a.s.,}
$$

which belongs to E. Therefore, for all $\tau \in (t,T]$ we have $\dot{X}_\tau \in E$ $\mathbb{P}$-a.s., where $\dot{X}$ denotes the mild solution to the forward equation in (7.14) with $s - \delta = t$ with ξ given by (4.11). With this choice of ξ equalities (4.8) and (4.10) can be rewritten as

$$
\dot{X}_\tau = \int_t^\tau \nabla_x X_\tau^{\sigma, X_\sigma^{t,x}} \nabla_x X_\sigma^{t,x} \eta d\sigma = (\tau - t) \nabla_x X_\tau^{t,x} \eta, \quad \dot{Y}_\tau = \int_t^\tau \nabla_x Y_\tau^{\sigma, X_\sigma^{t,x}} \nabla_x X_\sigma^{t,x} \eta d\sigma = (\tau - t) \nabla_x Y_\tau^{t,x} \eta,
$$
(4.12)

16

$$\dot{Z}_\tau = \int_t^\tau \nabla_x Z_\tau^{\sigma, X_\sigma^{t,x}} \nabla_x X_\sigma^{t,x} \eta \, d\sigma = (\tau - t) \nabla_x Z_\tau^{t,x} \eta, \quad \tau \in [t, T], \ \mathbb{P}\text{-a.s.}$$

Let us now set

$$\dot{\psi}(\tau, t, x) := \nabla_x \psi(\tau, X_\tau^{t,x}, Y_\tau^{t,x}, Z_\tau^{t,x}) \dot{X}_\tau^{t,x} d\tau + \nabla_y \psi(\tau, X_\tau^{t,x}, Y_\tau^{t,x}, Z_\tau^{t,x}) \dot{Y}_\tau^{t,x} d\tau$$

$$+ \nabla_z \psi(\tau, X_\tau^{t,x}, Y_\tau^{t,x}, Z_\tau^{t,x}) \dot{Z}_\tau^{t,x} d\tau,$$

$$\dot{\phi}(X_T^{t,x}) := \nabla_x \phi(X_T^{t,x}) \dot{X}_T^{t,x}.$$

By (4.12), $\dot{\psi}$ and $\dot{\phi}$ can be rewritten as

$$\dot{\psi}(\tau, t, x) = (\tau - t) \left(\nabla_x \psi(\tau, X_\tau^{t,x}, Y_\tau^{t,x}, Z_\tau^{t,x}) \nabla_x X_\tau^{t,x} \eta \, d\tau + \nabla_y \psi(\tau, X_\tau^{t,x}, Y_\tau^{t,x}, Z_\tau^{t,x}) \nabla_x Y_\tau^{t,x} \eta \, d\tau \right.$$

$$\left. + \nabla_z \psi(\tau, X_\tau^{t,x}, Y_\tau^{t,x}, Z_\tau^{t,x}) \nabla_x Z_\tau^{t,x} \eta \, d\tau \right), \tag{4.13}$$

$$\dot{\phi}(X_T^{t,x}) = (T - t) \nabla_x \phi(X_T^{t,x}) \nabla_x X_T^{t,x} \eta. \tag{4.14}$$

Notice that the right-hand sides in (4.13) and in (4.14) are nothing else (modulo a renormalization) than the terms appearing in the right-hand sides of the first two equations in (3.9). Now we aim at finding an expression for $\dot{\psi}$ and $\dot{\phi}$ that does not involve the derivative of ψ, ϕ, X, Y and Z: this in turn will furnish an expression of $\nabla_x Y$ that does not involve the derivatives of ψ, ϕ, X, Y and Z, as in formula (4.4). To this end, let us consider the process

$$W_\sigma^\varepsilon = W_\sigma - \varepsilon \int_t^\sigma (-A)^\alpha \nabla_x X_r^{t,x} \eta \, dr, \qquad 0 \le t \le \sigma \le T, \tag{4.15}$$

and let us define a probability measure Q_ε such that

$$\frac{dQ_\varepsilon}{d\mathbb{P}} = \exp\left(\varepsilon \int_t^T \langle (-A)^\alpha \nabla_x X_\sigma^{t,x} \eta, \, dW_\sigma \rangle - \frac{\varepsilon^2}{2} \int_t^T \left| (-A)^\alpha \nabla_x X_\sigma^{t,x} \eta \right|_H^2 d\sigma \right).$$

By the Girsanov theorem, $(W_\sigma^\varepsilon)_{\sigma \in [t,T]}$ is a cylindrical Wiener process in H under Q_ε. Arguing as in the proof of Theorem 3.10, we also notice that the process X under Q_ε and the process X^ε under $\mathbb{P}$ have the same law. Therefore,

$$\mathbb{E}[\nabla_x Y_\tau^{t,x} \eta] = \mathbb{E}[\nabla_x \Phi(X_T^{t,x}) \nabla_x X_T^{t,x} \eta] + \mathbb{E}\left[\int_\tau^T \left(\nabla_x \psi(\sigma, X_\sigma^{t,x}, Y_\sigma^{t,x}, Z_\sigma^{t,x}) \nabla_x X_\sigma^{t,x} \eta \right. \right.$$

$$\left. \left. + \nabla_y \psi(\sigma, X_\sigma^{t,x}, Y_\sigma^{t,x}, Z_\sigma^{t,x}) \nabla_x Y_\sigma^{t,x} \eta + \nabla_z \psi(\sigma, X_\sigma^{t,x}, Y_\sigma^{t,x}, Z_\sigma^{t,x}) \nabla_x Z_\sigma^{t,x} \eta \right) d\sigma \right]$$

$$= \frac{1}{T - t} \mathbb{E}[\dot{\phi}(X_T^{t,x})] + \mathbb{E}\left[\int_\tau^T \frac{1}{\sigma - t} \dot{\psi}(\sigma, t, x) \, d\sigma \right].$$

By differentiating inside the expectation with respect to ε and changing the order of integration, we get

$$\mathbb{E}[\dot{\psi}(\sigma, t, x)] = \mathbb{E}\left[\frac{d}{d\varepsilon}_{|\varepsilon=0} \psi(\sigma, X_\sigma^{\varepsilon, t, x}, Y_\sigma^{\varepsilon, t, x}, Z_\sigma^{\varepsilon, t, x}) \right] = \frac{d}{d\varepsilon}_{|\varepsilon=0} \mathbb{E}^{\mathbb{Q}^\varepsilon}[\psi(\sigma, X_\sigma^{t,x}, Y_\sigma^{t,x}, Z_\sigma^{t,x})]$$

$$= \mathbb{E}\left[\psi(\sigma, X_\sigma^{t,x}, Y_\sigma^{t,x}, Z_\sigma^{t,x}) \int_t^\sigma \langle (-A)^\alpha \nabla_x X_r^{t,x} \eta, \, dW_r \rangle \right],$$

and so, recalling (4.1),

$$\mathbb{E}\left[\int_\tau^T \frac{1}{\sigma - t} \dot{\psi}(\sigma, t, x) \, d\sigma \right] = \mathbb{E}\left[\int_\tau^T \frac{1}{\sigma - t} \left(\int_t^\sigma \langle (-A)^\alpha \nabla_x X_r^{t,x} \eta, \, dW_r \rangle \right) \psi(\sigma, X_\sigma^{t,x}, Y_\sigma^{t,x}, Z_\sigma^{t,x}) d\sigma \right]$$

17

$$= \mathbb{E}\Big[\int_\tau^T U_\sigma^{\eta,t,x}\psi(\sigma, X_\sigma^{t,x}, Y_\sigma^{t,x}, Z_\sigma^{t,x})d\sigma\Big].$$

Similarly, $\mathbb{E}[\dot{\phi}(X_T^{t,x})] = \mathbb{E}[\phi(X_T^{t,x})U_T^{\eta,t,x}]$, and this proves (4.4) when $\eta \in E$. The general case with $\eta \in H$ follows by density, thanks to estimates (3.15) and (4.2). $\qquad\square$

In the next result we remove the differentiability assumption on ψ and ϕ in Theorem 4.4.

Theorem 4.5. *Let Hypotheses 2.1 and 4.2 hold true, and for any $(t,x) \in [0,T] \times E$, let $(X^{t,x}, Y^{t,x}, Z^{t,x})$ be a solution of the forward-backward system (3.1), and let $U^{h,t,x}$ be the process defined in (4.1). Then, for $t \le s \le T$, $x \in E$, $h \in H$, the Bismut formula given in (4.4) holds true.*

Proof. The proof follows the same lines of the one of Theorem 3.10 in [10]. The main ingredients are formula (3.12) in Theorem 3.10 and Proposition 3.6, which provide respectively the identification of Z in the Banach space case and with the diffusion operator $(-A)^{-\alpha}$, and the stability result for the BSDE in (3.1) when the generator and the final datum are approximated by (3.6)-(3.7). We underline that approximations (3.6)-(3.7) preserve the boundedness and the growth, and are only of pointwise type. Notice that in [10], the final datum and the generator are approximated by means of their inf-sup convolutions, so that the approximation is uniform. However, thanks to the aforementioned stability properties for the BSDE, our pointwise approximations (3.6)-(3.7) are sufficient to obtain the desired result. $\qquad\square$

4.2 The semilinear Kolmogorov equation

By means of Theorem 4.5, we can give an existence and uniqueness result in the Banach framework for the semilinear Kolmogorov related to the the process X, and with coefficients ϕ and ψ not necessarily differentiable, as it is assumed in [17].
Let $P_{t,\tau}$, $t \le \tau \le T$, be the transition semigroup related to the process $X^{t,x}$ solution of the forward equation (2.1), namely, for every bounded and measurable function $\varphi : E \to \mathbb{R}$, $P_{t,\tau}[\varphi](x) := \mathbb{E}\,\varphi(X_\tau^{t,x})$. We consider the following semilinear Kolmogorov equation

$$\begin{cases} \frac{\partial v}{\partial t}(t,x) = -\mathscr{L}v(t,x) + \psi\Big(t,x,v(t,x),\nabla^{(-A)^{-\alpha}}v(t,x)\Big), & t \in [0,T],\ x \in E, \\ v(T,x) = \phi(x), \end{cases} \tag{4.16}$$

where $\mathscr{L}$ is the generator of the transition semigroup $(P_{t,s})_{0 \le t \le s \le T}$, that is, at least formally,

$$(\mathscr{L}f)(x) = \frac{1}{2}(\mathrm{Tr}((-A)^{-\alpha}(-A^*)^{-\alpha}\nabla^2 f)(x) + \langle Ax, \nabla f(x)\rangle + \langle F(x), \nabla f(x)\rangle, \quad x \in E.$$

We introduce the notion of mild solution of the nonlinear Kolmogorov equation (4.16), see e.g. [11].

Definition 4.6. *A function $v : [0,T] \times E \to \mathbb{R}$ is a mild solution of the semilinear Kolmogorov equation (4.16) if $v \in \mathscr{G}^{0,1}([0,T] \times E)$, and*

$$v(t,x) = P_{t,T}[\phi](x) + \int_t^T P_{t,s}\Big[\psi(s,\cdot,v(s,\cdot),\nabla^{(-A)^{-\alpha}}v(s,\cdot))\Big](x)\,ds, \quad t \in [0,T],\ x \in E. \tag{4.17}$$

Theorem 4.7. *Let Hypotheses 2.1 and 4.2 hold true. Then the semilinear Kolmogorov equation (4.16) has a unique mild solution v given by the formula*

$$v(t,x) = Y_t^{t,x}, \qquad (t,x) \in [0,T] \times E,$$

where, for any $(t,x) \in [0,T] \times E$, $(X^{t,x}, Y^{t,x}, Z^{t,x})$ denotes the solution to the FBSDE (3.1). In addition, we have, $\mathbb{P}$-a.s.,

$$Y_s^{t,x} = v(s, X_s^{t,x}), \quad Z_s^{t,x} = \nabla_x v(s, X_s^{t,x}) \nabla_x X_s^{t,x}(-A)^{-\alpha}, \quad \text{a.e. } s \in [t,T].$$

Proof. If the data ϕ and ψ are also differentiable, the result can be proved as in [17], Theorem 6.2. When the data are not differentiable, the Bismut formula (4.4) is still true, see Theorem 4.5, and the result can be proved arguing as in [10], Theorem 4.2. $\qquad\square$

5 The Bismut-Elworthy formula and the semilinear Kolmogorov equation: the quadratic case

We are ready to state and prove the main result of the paper, which is a nonlinear Bismut-Elworthy formula as the one in Theorem 4.5, but in the case of quadratic generator. This in particular will give an existence and uniqueness result for the Kolmogorov equation (4.16) in the quadratic case and in the Banach framework, see Theorem 5.4.

Theorem 5.1. *Let Hypotheses 2.1 and 3.1 hold true. For any $(t,x) \in [0,T] \times E$, let $(X^{t,x}, Y^{t,x}, Z^{t,x})$ be the solution of the forward-backward system (3.1) and let $U^{h,t,x}$ be the process defined in (4.1). Then, for $t \le s \le T$, $x \in E$ and $h \in H$,*

$$\mathbb{E}\left[\nabla_x Y_s^{t,x} h\right] = \mathbb{E}\left[\int_s^T \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x} dr\right] + \mathbb{E}\left[\phi(X_T^{t,x}) U_T^{h,t,x}\right]. \tag{5.1}$$

Proof. We split the proof into two steps: we first prove the statement when ψ is differentiable with respect to x, y and z, and then we remove this additional assumption.

STEP 1. We start by considering ψ differentiable with respect to x, y and z. For all $n \ge 1$, let us denote by $(X^{t,x}, Y^{n,t,x}, Z^{n,t,x})$ the solution of the Markovian BSDE in (3.1) with final datum equal to ϕ_n in (3.6) in the place of ϕ:

$$\begin{cases} dY_\tau^{n,t,x} = -\psi(\tau, X_\tau^{t,x}, Y_\tau^{n,t,x}, Z_\tau^{n,t,x}) \, d\tau + Z_\tau^{n,t,x} \, dW_\tau, \quad \tau \in [t,T], \\ Y_T^{n,t,x} = \phi_n(X_T^{t,x}). \end{cases} \tag{5.2}$$

By estimate (3.13) in Corollary 3.11, for any $n \ge 1$, there exists a constant $C(n)$, depending on n, which is bounded for every n and blows up as $n \to \infty$, and such that

$$|Z_s^{n,t,x}|_H \le C(n), \quad \mathbb{P}\text{-a.s., a.e. } s \in [t,T]. \tag{5.3}$$

In particular,

$$|\psi(s,x,y,z_1) - \psi(s,x,y,z_2)| \le C(n)|z_1 - z_2|_H, \quad z_1, z_2 \in H : \ |z_i|_H \le C(n), \ i = 1, 2.$$

Therefore, the generator ψ acts as a Lipschitz generator with respect to z in the BSDE (5.2), so the Bismut-Elworthy formula stated in Theorem 4.5 holds true for the BSDE (5.2): for every $s \in [t,T]$,

$$\mathbb{E}\left[\nabla_x Y_s^{n,t,x} h\right] = \mathbb{E}\left[\int_s^T \psi\left(r, X_r^{t,x}, Y_r^{n,t,x}, Z_r^{n,t,x}\right) U_r^{h,t,x} dr\right] + \mathbb{E}\left[\phi_n(X_T^{t,x}) U_T^{h,t,x}\right]. \tag{5.4}$$

At this point we aim at taking the limit as $n \to \infty$ in (5.4).
We start by considering the right-hand side of (5.4). By the properties of the approximations $(\phi_n)_{n \ge 1}$

together with (4.2), by the dominated convergence theorem and the pointwise convergence of ϕ_n to ϕ we have

$$\mathbb{E}\Big[|[\phi_n(X_T^{t,x}) - \phi(X_T^{t,x})]U_T^{h,t,x}|\Big] \le (\mathbb{E}[|\phi_n(X_T^{t,x}) - \phi(X_T^{t,x})|^2])^{1/2}(\mathbb{E}[|U_T^{h,t,x}|^2])^{1/2}$$

$$\le C\,(T-t)^{-(1/2+\alpha)}\,(\mathbb{E}[|\phi_n(X_T^{t,x}) - \phi(X_T^{t,x})|^2])^{1/2} \to 0 \text{ as } n \to \infty.$$

Therefore,

$$\lim_{n\to\infty} \mathbb{E}\Big[\phi_n(X_T^{t,x})U_T^{h,t,x}\Big] = \mathbb{E}\Big[\phi(X_T^{t,x})U_T^{h,t,x}\Big].$$

In order to compute the limit of the remaining term in the right-hand side of (5.4), we will show that

$$\lim_{n\to\infty} \mathbb{E}\Big[\int_t^T |\psi\left(r, X_r^{t,x}, Y_r^{n,t,x}, Z_r^{n,t,x}\right)U_r^{h,t,x} - \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right)U_r^{h,t,x}|\,dr\Big] = 0.$$

We notice that

$$\mathbb{E}\Big[\int_t^T |\psi\left(r, X_r^{t,x}, Y_r^{n,t,x}, Z_r^{n,t,x}\right)U_r^{h,t,x} - \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right)U_r^{h,t,x}|\,dr\Big]$$

$$= \mathbb{E}\Big[\int_t^{\frac{t+T}{2}} |\psi\left(r, X_r^{t,x}, Y_r^{n,t,x}, Z_r^{n,t,x}\right)U_r^{h,t,x} - \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right)U_r^{h,t,x}|\,dr\Big]$$

$$+ \mathbb{E}\Big[\int_{\frac{t+T}{2}}^T |\psi\left(r, X_r^{t,x}, Y_r^{n,t,x}, Z_r^{n,t,x}\right)U_r^{h,t,x} - \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right)U_r^{h,t,x}|\,dr\Big]$$

$$=: I + II. \tag{5.5}$$

We start by estimating the term I in (5.5). We have

$$I = \mathbb{E}\Big[\int_t^{\frac{t+T}{2}} |\psi\left(r, X_r^{t,x}, Y_r^{n,t,x}, Z_r^{n,t,x}\right)U_r^{h,t,x} - \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right)U_r^{h,t,x}|\,dr\Big]$$

$$\le L_\psi \mathbb{E}\Big[\int_t^{\frac{t+T}{2}} \left(|Y_r^{n,t,x} - Y_r^{t,x}||U_r^{h,t,x}|\right)\,dr\Big]$$

$$+ L_\psi \mathbb{E}\Big[\int_t^{\frac{t+T}{2}} \left(|Z_r^{n,t,x} - Z_r^{t,x}|_H \left(1 + |Z_r^{n,t,x}|_H + |Z_r^{t,x}|_H\right)|U_r^{h,t,x}|\right)\,dr\Big]$$

$$=: I_a + I_b. \tag{5.6}$$

We recall that, by estimate (3.14) in Proposition 3.12, and since $\|\phi_n\|_\infty \le K_\phi$, there exists a constant C, not depending on n, such that

$$|Z_t^{n,t,x}|_H \le C(T-t)^{-1/2}, \quad \mathbb{P}\text{-a.s.}$$

So, since $Z_r^{n,t,x} = Z_r^{n,r,X_r^{t,x}}$ and $Z_r^{t,x} = Z_r^{r,X_r^{t,x}}$, for $r \in [t, \frac{t+T}{2}]$

$$|Z_r^{n,t,x}|_H + |Z_r^{t,x}|_H \le C \sup_{r\in[t,\frac{t+T}{2}]} (T-r)^{-1/2} \le C(T-t)^{-1/2}, \quad \mathbb{P}\text{-a.s.} \tag{5.7}$$

We only show the convergence of I_b in (5.6) since the convergence of I_a follows in a simpler way by the boundedness of $Y^{t,x}$ and of $Y^{n,t,x}$ (uniform in n), and by the convergence of $Y^{n,t,x}$ to $Y^{t,x}$ in $\mathscr{S}^p([t,T])$, $p \ge 2$. Using (5.7), Hölder inequality with $p = \frac{2}{1-\beta}$ and $q = \frac{2}{1+\beta}$, for some $2\alpha < \beta < 1$, together with estimate (4.3) in Lemma 4.1, we get

$$I_b \le C(T-t)^{-\frac{1}{2}}(T-t)^{-\frac{1}{2}\beta}\mathbb{E}\Big[\int_t^{\frac{t+T}{2}} |Z_r^{n,t,x} - Z_r^{t,x}|_H^{1-\beta}|U_r^{h,t,x}|\,dr\Big]$$

$$\leq C(T-t)^{-\frac{1}{2}}(T-t)^{-\frac{1}{2}\beta}\Big(\mathbb{E}\Big[\int_t^{\frac{t+T}{2}}|Z_r^{n,t,x}-Z_r^{t,x}|_H^2\,dr\Big]\Big)^{\frac{1-\beta}{2}}\Big(\int_t^{\frac{t+T}{2}}\mathbb{E}[|U_r^{h,t,x}|^{\frac{2}{1+\beta}}]\,dr\Big)^{\frac{1+\beta}{2}}$$

$$\leq C(T-t)^{-\frac{1}{2}(1+\beta)}\Big(\mathbb{E}\Big[\int_t^{\frac{t+T}{2}}|Z_r^{n,t,x}-Z_r^{t,x}|_H^2\,dr\Big]\Big)^{\frac{1-\beta}{2}}\Big(\int_t^{\frac{t+T}{2}}\frac{1}{(r-t)^{\frac{1+2\alpha}{1+\beta}}}\,dr\Big)^{\frac{1+\beta}{2}}$$

$$\leq C(T-t)^{-\frac{1}{2}(1+\beta)}(T-t)^{(1-\frac{1+2\alpha}{1+\beta})\frac{1+\beta}{2}}\Big(\mathbb{E}\Big[\int_t^{\frac{t+T}{2}}|Z_r^{n,t,x}-Z_r^{t,x}|_H^2\,dr\Big]\Big)^{\frac{1-\beta}{2}}$$

$$\leq C(T-t)^{-\frac{1}{2}-\alpha}\Big(\mathbb{E}\Big[\int_t^{\frac{t+T}{2}}|Z_r^{n,t,x}-Z_r^{t,x}|_H^2\,dr\Big]\Big)^{\frac{1-\beta}{2}}\to 0$$

as $n\to\infty$, since $Z^{n,t,x}\to Z^{t,x}$ in $\mathscr{M}^2([t,T];H)$.

Let us now estimate the term II in (5.5). To this end, we recall that, by Theorem 3.3, $Y^{n,t,x}$, $Y^{t,x}$ are bounded in $\mathscr{S}^p([t,T])$ and $Z^{n,t,x}$, $Z^{t,x}$ are bounded in $\mathscr{M}^{2p}([t,T];H)$, by a constant independent on n. Moreover, by Proposition 3.6, $Y^{n,t,x}$ converges to $Y^{t,x}$ in $\mathscr{S}^p([t,T])$ and $Z^{n,t,x}$ converges to $Z^{t,x}$ in $\mathscr{M}^{2p}([t,T];H)$, for any $p\geq 1$. By using again Hölder's inequality for some $p,q\geq 1$, $\frac{1}{p}+\frac{1}{q}=1$, and estimate (4.3) in Lemma 4.1, we get

$$II=\mathbb{E}\Big[\int_{\frac{t+T}{2}}^T|\psi\big(r,X_r^{t,x},Y_r^{n,t,x},Z_r^{n,t,x}\big)U_r^{h,t,x}-\psi\big(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x}\big)U_r^{h,t,x}|\,dr\Big]$$

$$\leq\mathbb{E}\Big[\sup_{s\in[\frac{t+T}{2},T]}|U_s^{h,t,x}|\int_{\frac{t+T}{2}}^T|\psi\big(r,X_r^{t,x},Y_r^{n,t,x},Z_r^{n,t,x}\big)-\psi\big(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x}\big)|\,dr\Big]$$

$$\leq\Big(\mathbb{E}\sup_{s\in[\frac{t+T}{2},T]}|U_s^{h,t,x}|^q\Big)^{\frac{1}{q}}\Big(\mathbb{E}\Big[\int_{\frac{t+T}{2}}^T|\psi\big(r,X_r^{t,x},Y_r^{n,t,x},Z_r^{n,t,x}\big)-\psi\big(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x}\big)|\,dr\Big]^p\Big)^{\frac{1}{p}}$$

$$\leq C\frac{1}{(T-t)^{\frac{1}{2}+\alpha}}\Big(\mathbb{E}\Big[\int_{\frac{t+T}{2}}^T(|Y_r^{n,t,x}-Y_r^{t,x}|+|Z_r^{n,t,x}-Z_r^{t,x}|_H(1+|Z_r^{n,t,x}|_H+|Z_r^{t,x}|_H))\,dr\Big]^p\Big)^{\frac{1}{p}}$$

$$\leq C\frac{1}{(T-t)^{\frac{1}{2}+\alpha}}\Big\{\frac{T-t}{2}\Big(\mathbb{E}\Big[\sup_{r\in[t,T]}|Y_r^{n,t,x}-Y_r^{t,x}|^p\Big]\Big)^{\frac{1}{p}}$$

$$+\Big(\mathbb{E}\Big[\Big(\int_{\frac{t+T}{2}}^T|Z_r^{n,t,x}-Z_r^{t,x}|_H^2\,dr\Big)^{\frac{p}{2}}\Big(\int_{\frac{t+T}{2}}^T\big(1+|Z_r^{n,t,x}|_H+|Z_r^{t,x}|_H\big)^2\,dr\Big)^{\frac{p}{2}}\Big]\Big)^{\frac{1}{p}}\Big\}$$

$$\leq C\frac{1}{(T-t)^{\frac{1}{2}+\alpha}}\Big\{\frac{T-t}{2}\Big(\mathbb{E}\Big[\sup_{r\in[t,T]}|Y_r^{n,t,x}-Y_r^{t,x}|^p\Big]\Big)^{\frac{1}{p}}$$

$$+\Big(\mathbb{E}\Big[\int_{\frac{t+T}{2}}^T|Z_r^{n,t,x}-Z_r^{t,x}|_H^2\,dr\Big]^p\Big)^{\frac{1}{2p}}\Big(\mathbb{E}\Big[\int_{\frac{t+T}{2}}^T\big(1+|Z_r^{n,t,x}|_H+|Z_r^{t,x}|_H\big)^2\,dr\Big]^p\Big)^{\frac{1}{2p}}\Big\}\to 0,$$

as $n\to\infty$. Collecting all the previous results, we deduce that, for every $s\in[t,T]$,

$$\lim_{n\to\infty}\mathbb{E}\big[\nabla_x Y_s^{n,t,x}h\big]=\mathbb{E}\Big[\int_s^T\psi\big(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x}\big)U_r^{h,t,x}\,dr\Big]+\mathbb{E}\Big[\phi(X_T^{t,x})U_T^{h,t,x}\Big]. \qquad (5.8)$$

In particular, by taking $s=t$ in (5.8),

$$\lim_{n\to\infty}\nabla_x Y_t^{n,t,x}h=\mathbb{E}\Big[\int_t^T\psi\big(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x}\big)U_r^{h,t,x}\,dr\Big]+\mathbb{E}\Big[\phi(X_T^{t,x})U_T^{h,t,x}\Big],$$

which shows that $\lim_{n\to\infty}\nabla_x Y_t^{n,t,x}h$ exists. Moreover, arguing as in the end of the proof of Theorem 4.1 in [19], we deduce that $\lim_{n\to\infty}\nabla_x Y_t^{n,t,x}h=\nabla_x Y_t^{t,x}h$ for all $h\in H$.

21

STEP 2. Let us now remove the differentiability assumptions on ψ. For any $k \geq 1$, let ψ_k be the function defined in (3.7). From Lemma 3.5 we know that ψ_k is differentiable and it preserves the Lipschitz constant, so that

$$|\nabla_x \psi_k|_{E^*} \leq L_\psi, \quad |\nabla_y \psi_k| \leq L_\psi.$$

Moreover, again from Lemma 3.5, $\psi_k(t,x,y,z) \to \psi(t,x,y,z)$ as $k \to +\infty$ for any $(t,x,y,z) \in [0,T] \times E \times \mathbb{R} \times H$, and for any $t \in [0,T]$, $x \in E$, $y \in \mathbb{R}$, $z_1, z_2 \in H$,

$$|\psi_k(t,x,y_1,z_1) - \psi_k(t,x,y_2,z_2)| \leq L_\psi(|y_1 - y_2| + |z_1 - z_2|_H(1 + |z_1|_H + |z_2|_H)), \tag{5.9}$$

for any $k \in \mathbb{N}$. We consider the BSDE with generator equal to ψ_k in the place of ψ:

$$\begin{cases} dY_\tau^{k,t,x} = -\psi_k(\tau, X_\tau^{t,x}, Y_\tau^{k,t,x}, Z_\tau^{k,t,x})\, d\tau + Z_\tau^{k,t,x}\, dW_\tau, \quad \tau \in [t,T], \\ Y_T^{k,t,x} = \phi(X_T^{t,x}). \end{cases} \tag{5.10}$$

By the first part of the proof, for any $k \geq 1$,

$$\mathbb{E}\left[\nabla_x Y_s^{k,t,x} h\right] = \mathbb{E}\left[\int_s^T \psi_k\left(r, X_r^{t,x}, Y_r^{k,t,x}, Z_r^{k,t,x}\right) U_r^{h,t,x}\, dr\right] + \mathbb{E}\left[\phi(X_T^{t,x}) U_T^{h,t,x}\right]. \tag{5.11}$$

We aim at taking the limit as $k \to \infty$. We start by considering the first term in the right-hand side of (5.11), and we will show that

$$\lim_{k \to \infty} \mathbb{E}\left[\int_t^T |\psi_k\left(r, X_r^{t,x}, Y_r^{k,t,x}, Z_r^{k,t,x}\right) U_r^{h,t,x} - \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x}|\, dr\right] = 0.$$

We start by splitting the integral above as follows:

$$\mathbb{E}\left[\int_t^T |\psi_k\left(r, X_r^{t,x}, Y_r^{k,t,x}, Z_r^{k,t,x}\right) U_r^{h,t,x} - \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x}|\, dr\right]$$

$$= \mathbb{E}\left[\int_t^{\frac{t+T}{2}} |\psi_k\left(r, X_r^{t,x}, Y_r^{k,t,x}, Z_r^{k,t,x}\right) U_r^{h,t,x} - \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x}|\, dr\right]$$

$$+ \mathbb{E}\left[\int_{\frac{t+T}{2}}^T |\psi_k\left(r, X_r^{t,x}, Y_r^{k,t,x}, Z_r^{k,t,x}\right) U_r^{h,t,x} - \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x}|\, dr\right]$$

$$=: I + II. \tag{5.12}$$

In order to estimate the term I in (5.12), we notice that

$$I \leq \mathbb{E}\left[\int_t^{\frac{t+T}{2}} |\psi_k\left(r, X_r^{t,x}, Y_r^{k,t,x}, Z_r^{k,t,x}\right) U_r^{h,t,x} - \psi_k\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x}|\, dr\right]$$

$$+ \mathbb{E}\left[\int_t^{\frac{t+T}{2}} |\psi_k\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x} - \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x}|\, dr\right]$$

$$=: I_a + I_b. \tag{5.13}$$

Concerning I_a in (5.13), by (3.4)-(3.5) we can argue as for I in Step 1, and get that $I_a \to 0$ as $k \to +\infty$. Let us now consider the term I_b in (5.13). From Hypothesis 3.1 and formulas (3.4) and (3.5) it follows that

$$|\psi_k\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) - \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right)| \leq C(1 + |Y_r^{t,x}| + |Z_r^{t,x}|_H^2), \tag{5.14}$$

where C is a positive constant depending on L_ψ and K_ψ. Arguing as for I is Step 1, one can prove that

$$r \mapsto (1 + |Y_r^{t,x}| + |Z_r^{t,x}|_H^2)|U_r^{h,t,x}| \in L^1\Big(\Omega; L^1\Big(t, \frac{t+T}{2}; \mathbb{R}\Big)\Big).$$

Recalling that $\psi_k \to \psi$ pointwise as $k \to +\infty$, we get that $I_b \to 0$ as $k \to +\infty$ by the dominated convergence theorem.

Let us now estimate II in (5.12). To this end, we notice that

$$
\begin{aligned}
II \leq &\mathbb{E}\Big[\int_{\frac{t+T}{2}}^T |\psi_k\left(r, X_r^{t,x}, Y_r^{k,t,x}, Z_r^{k,t,x}\right) U_r^{h,t,x} - \psi_k\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x}|\, dr\Big] \\
&+ \mathbb{E}\Big[\int_{\frac{t+T}{2}}^T |\psi_k\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x} - \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x}|\, dr\Big] \\
=: &\, II_a + II_b.
\end{aligned}
\tag{5.15}
$$

Arguing as for the term II in Step 1, we deduce that $II_a \to 0$ as $k \to 0$. As far as II_b in (5.15) is considered, we get

$$II_b \leq \Big(\mathbb{E}\Big[\sup_{s \in [\frac{t+T}{2}, T]} |U_s^{h,t,x}|^q\Big]\Big)^{1/q} \Big(\mathbb{E}\Big[\int_{\frac{t+T}{2}}^T |\psi_k\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) - \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right)|\,dr\Big]^p\Big)^{1/p}.$$

Arguing as for II in Step 1 it follows that

$$r \mapsto (1 + |Y_r^{t,x}| + |Z_r^{t,x}|_H^2)|U_r^{h,t,x}| \in L^p\Big(\Omega; L^1\Big(\frac{t+T}{2}, T; \mathbb{R}\Big)\Big).$$

Since ψ_k pointwise converges to ψ, we can again apply the dominated convergence theorem which gives $II_b \to 0$ as $k \to +\infty$. We can thus conclude that, for every $s \in [t, T]$,

$$\lim_{k \to \infty} \mathbb{E}\left[\nabla_x Y_s^{k,t,x} h\right] = \mathbb{E}\Big[\int_s^T \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x}\, dr\Big] + \mathbb{E}\left[\phi(X_T^{t,x}) U_T^{h,t,x}\right].$$

As in the end of Step 1, arguing as at the end of Theorem 4.1 in [19] we can show that, for any $s \in [t, T]$, $\lim_{k \to \infty} \mathbb{E}\left[\nabla_x Y_s^{k,t,x} h\right] = \mathbb{E}\left[\nabla_x Y_s^{t,x} h\right]$. $\qquad\square$

We now state two corollaries: the former is about integral estimates of $\nabla_x Y^{t,x}$, the latter is about the identification of $\nabla_x Y^{t,x}$ with $Z^{t,x}$ without differentiability assumptions. Notice that, by means of the Bismut formula (5.1), we can also recover estimate (3.15) on $\nabla_x Y^{t,x}$.

Corollary 5.2. *Let* $(t, x) \in [0, T] \times E$. *Under the assumptions of Theorem* 5.1, *the process* $\nabla_x Y^{t,x}$ *belongs to* $\mathscr{M}^2([t, T])$, *and there exists a constant* C *depending only on* L_ψ, K_ψ, K_ϕ *such that*

$$\mathbb{E}\Big[\int_t^T |\nabla_x Y_s^{t,x}|^2\, ds\Big] \leq C(T - t)^{-2\alpha}. \tag{5.16}$$

Proof. Integrating (5.1) between t and T we get

$$
\begin{aligned}
\int_t^T |\mathbb{E}[\nabla_x Y_s^{t,x}]|^2\, ds &= \int_t^T \Big|\mathbb{E}\Big[\int_s^T \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x}\, dr\Big] + \mathbb{E}\left[\phi(X_T^{t,x}) U_T^{h,t,x}\right]\Big|^2\, ds \\
&\leq C \int_t^T \Big|\mathbb{E}\Big[\int_s^T \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x}\, dr\Big]\Big|^2\, ds + \int_t^T \Big|\mathbb{E}\left[\phi(X_T^{t,x}) U_T^{h,t,x}\right]\Big|^2\, ds
\end{aligned}
$$

23

$$=: I + II.$$

We have

$$II \leq \int_t^T \|\phi\|_\infty \frac{1}{(T-t)^{1+2\alpha}} \, dr = C(T-t)^{-2\alpha}.$$

For what concerns I, we split it as

$$I = C\Big(\int_{\frac{t+T}{2}}^T \Big| \mathbb{E}\Big[\int_s^T \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x} \, dr \Big] \Big|^2 ds$$
$$+ \int_t^{\frac{t+T}{2}} \Big| \mathbb{E}\Big[\int_s^T \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x} \, dr \Big] \Big|^2 ds \Big)$$
$$=: I_a + I_b.$$

From (4.3) and Proposition 3.3 we have

$$I_a \leq C \int_{\frac{t+T}{2}}^T \Big(\mathbb{E}\Big[\sup_{r \in [\frac{t+T}{2}, T]} |U_r^{h,t,x}| \int_s^T |\psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right)| \, dr \Big] \Big)^2 ds$$
$$\leq C \int_{\frac{t+T}{2}}^T \mathbb{E}\Big[\sup_{r \in [\frac{t+T}{2}, T]} |U_r^{h,t,x}|^2 \Big] \mathbb{E}\Big[\Big(\int_s^T \left(1 + |Y_r^{t,x}| + |Z_r^{t,x}|_H^2\right) \, dr \Big)^2 \Big] ds$$
$$\leq C(T-t)^{-1-2\alpha} \int_{\frac{t+T}{2}}^T \mathbb{E}\Big[\Big(\int_s^T \left(1 + |Y_r^{t,x}| + |Z_r^{t,x}|_H^2\right) \, dr \Big)^2 \Big] ds \leq C(T-t)^{-2\alpha}.$$

On the other hand, we consider the function under the integral sign in I_b and we split it as follows:

$$\mathbb{E}\Big[\int_s^T \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x} \, dr \Big]$$
$$= \mathbb{E}\Big[\int_s^{\frac{t+T}{2}} \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x} \, dr \Big] + \mathbb{E}\Big[\int_{\frac{t+T}{2}}^T \psi\left(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) U_r^{h,t,x} \, dr \Big]$$
$$=: I_b' + I_b''.$$

We argue as in the proof of Theorem 5.1, Step 1. In particular, arguing as for the estimate of I we infer that

$$|I_b'| \leq C(T-t)^{-1/2-\alpha}$$

for some positive constant C. On the other hand, as far as I_b'' is considered, arguing as in the estimate of II, we get that

$$|I_b''| \leq C(T-t)^{-1/2-\alpha}$$

for some positive constant C. Hence,

$$I_b \leq C \int_t^{\frac{t+T}{2}} (T-t)^{-1-2\alpha} ds = C(T-t)^{-2\alpha},$$

and this concludes the proof. $\qquad\square$

In the following we prove that the identification of Z with the directional derivative of Y remains true also when ϕ and ψ are not differentiable.

Corollary 5.3. *Under the assumptions of Theorem 5.1, for every $(t,x) \in [0,T] \times E$,*

$$Z_t^{t,x} = \nabla_x Y_t^{t,x} (-A)^{-\alpha}. \tag{5.17}$$

Proof. Let ϕ and ψ be respectively approximated by ϕ_n and ψ_n in (3.6) and (3.7), and let $(Y^{n,t,x}, Z^{n,t,x})$ be the solution of the BSDE with final datum ϕ_n and generator ψ_n. By Theorem 4.7 we already know that $Z_t^{n,t,x} = \nabla_x Y_t^{n,t,x} (-A)^{-\alpha}$. On the other hand, we have shown in Theorem 5.1 that $x \mapsto Y_\tau^{t,x} = v(\tau, X_\tau^{t,x})$ is differentiable and that $\nabla_x Y_\tau^{n,t,x} \to \nabla_x Y_\tau^{t,x}$, $dt \otimes d\mathbb{P}$ a.e. and a.s., as $n \to \infty$. Moreover, by computing the joint quadratic variation between the process $Y_\tau^{n,t,x}$ and $\int_t^\tau \xi_s \, dW_s$, $t \leq \tau \leq T$, $\xi \in \mathscr{M}^2([t,T];H)$, it turns out that

$$\int_t^\tau \nabla v^n(s, X_s^{t,x})(-A)^{-\alpha} \xi_s \, ds = \int_t^\tau Z_s^{n,t,x} \xi_s \, ds, \ \mathbb{P}\text{-a.s., a.e. } \tau \in [t,T],$$

where we have set $v^n(\tau, X_\tau^{t,x}) := Y_\tau^{n,t,x}$. By taking a subsequence (that for simplicity we call again n) and letting $n \to \infty$ in both sides, from Proposition 3.6 we get

$$\int_t^\tau \nabla v(s, X_s^{t,x})(-A)^{-\alpha} \xi_s \, ds = \int_t^\tau Z_s^{t,x} \xi_s \, ds, \text{ a.e. } \tau \in [t,T], \ \mathbb{P}\text{-a.s.,}$$

which gives formula (5.17). $\qquad\square$

Using Theorem 3.10, we can give an existence and uniqueness result for the Kolmogorov equation (4.16) and we can provide a Feynman-Kac formula in the quadratic case and in the Banach framework.

Theorem 5.4. *Let Hypotheses 2.1 and 3.1 hold true. Then there exists a unique mild solution $v(t,x)$ of the semilinear Kolmogorov equation (4.16) given by the formula*

$$v(t,x) = Y_t^{t,x},$$

where $(X^{t,x}, Y^{t,x}, Z^{t,x})$ is the solution to the FBSDE (3.1), and $\mathbb{P}$-a.s.,

$$Y_s^{t,x} = v(s, X_s^{t,x}), \quad Z_s^{t,x} = \nabla_x v(s, X_s^{t,x}) \nabla_x X_s^{t,x} (-A)^{-\alpha}, \text{ a.e. } s \in [t,T].$$

In particular,

$$|v(t,x)| \leq C, \qquad |\nabla_x v(t,x)| \leq C \, (T-t)^{-(\frac{1}{2}+\alpha)}.$$

If in addition ϕ is Gâteaux differentiable with bounded derivative, and ψ is Gâteaux differentiable with respect to x, y and z, then

$$|Z_s^{t,x}|_H \leq C.$$

Proof. For the first part without differentiability assumptions on ϕ and ψ, it is enough to apply Theorem 5.1 and Corollary 5.3 to get existence of the solution, as well as the estimate for v. The uniqueness follows from the uniqueness of the solution of the related BSDE. The estimate for $\nabla_x v(t,x)$ is a direct consequence of Proposition 3.12. The second part of the result can be proved in a standard way by means of Proposition 3.7 and the identification of Z proved in Theorem 3.10, see e.g. the proof of Theorem 6.2 in [11]. $\qquad\square$

6 A quadratic optimal control problem

In this section we deal with the controlled state equation

$$\begin{cases} dX^u_\tau = AX^u_\tau d\tau + F(X^u_\tau)d\tau + Qu_\tau d\tau + (-A)^{-\alpha}dW_\tau, & \tau \in [t,T], \\ X^u_t = x \in E, \end{cases} \tag{6.1}$$

where $Q = (-A)^{-\alpha}$ or $Q = I$, and u is the control process belonging to a suitable space $\mathscr{U}$ of H-valued functions. We will study the optimal control problem associated to equation (6.1) with cost functional $J : [0,T] \times E \times \mathscr{U} \to \mathbb{R}$ defined by

$$J(t,x,u) := \mathbb{E}\Big[\int_t^T \ell(s, X^u_s, u_s)\, ds \Big] + \mathbb{E}[\Phi(X^u_T)], \tag{6.2}$$

that we are going to minimize over all admissible controls. We define the value function of the optimal control problem as

$$V(t,x) := \inf_{u \in \mathscr{U}} J(t,x,u), \quad x \in H,\ t \in [0,T]. \tag{6.3}$$

For any $p \geq 1$, we introduce the spaces of admissible control processes

$$\mathscr{U}_p := \big\{ u \in L^2(\Omega; L^p(0,T;H)) : u \text{ is adapted} \big\},$$
$$\mathscr{U}_p^\alpha := \big\{ u \in L^2(\Omega; L^p(0,T;D((-A)^\alpha))) : u \text{ is adapted} \big\},$$

where $D((-A)^\alpha))$ is endowed with the norm

$$|x|_\alpha := |x|_H + |(-A)^\alpha x|_H.$$

We first prove some results about well posedness of the controlled equation (6.1). The main novelty towards Section 2 and the known results in the literature is that the controls u are not necessarily bounded, together with the fact that X evolves in a Banach space E.

Beside Hypothesis 2.1 we assume the following.

Hypothesis 6.1. *There exists $\beta > 0$ such that $D((-A)^\beta) \subset E$ with continuous embedding.*

Remark 6.2. *Let A be an operator satisfying Hypothesis 2.1-(i). If Hypothesis 6.1 holds true, then we have the following.*

(i) *For any $t > 0$ and $h \in H$, $e^{tA}h \in E$ and there exists a positive constant c such that*

$$|e^{tA}h|_E \leq ct^{-\beta}|h|_H. \tag{6.4}$$

(ii) *For any $t > 0$ and $h \in H$, there exists a positive constant c such that*

$$|e^{tA}(-A)^{-\alpha}h|_E \leq ct^{(-\beta+\alpha)\wedge 0}|h|_H. \tag{6.5}$$

(iii) *For any $t > 0$ and $h \in D((-A)^\alpha)$, there exists a positive constant c such that*

$$|e^{tA}h|_E = |e^{tA}(-A)^{-\alpha}(-A)^\alpha h|_E \leq ct^{(-\beta+\alpha)\wedge 0}|(-A)^\alpha h|_H \leq ct^{(-\beta+\alpha)\wedge 0}|h|_\alpha.$$

Remark 6.3. *Hypothesis 6.1 may be replaced by the weaker condition in Remark 6.2-(i). However, this condition would not imply Remark 6.2-(ii)-(iii).*

Example 6.4. *Let $\mathscr{D} \subset \mathbb{R}^2$ be a bounded domain with smooth boundary. Set $H = L^2(\mathscr{D})$, $E = C(\overline{\mathscr{D}})$, and let A be the Laplace operator with Dirichlet boundary conditions. Then, Hypothesis 6.1 is satisfied with $\beta > 1/2$.*

We will deal with mild solutions to (6.1), namely adapted processes $X^{t,x,u} : [t,T] \times \Omega \to E$ such that

$$X_\tau^{t,x,u} = e^{\tau A}x + \int_t^\tau e^{(\tau-s)A}F(X_s^{t,x,u})ds + \int_t^\tau e^{(\tau-s)A}Qu_s ds + \int_t^\tau e^{(\tau-s)A}(-A)^{-\alpha}dW_s, \qquad (6.6)$$

for any $\tau \in [t,T]$, $\mathbb{P}$-a.s. For any $t \in [0,T]$, $u \in \mathscr{U}$, we set

$$I^u(t,\tau) := \int_t^\tau e^{(\tau-s)A}Qu_s\, ds, \quad \tau \in [t,T]. \qquad (6.7)$$

Lemma 6.5. *Let A be an operator satisfying Hypothesis 2.1-(i), and assume that Hypothesis 6.1 holds true for some positive constant β. Let $p \geq 1$, and set p' be the conjugate exponent of p, i.e., $p^{-1} + (p')^{-1} = 1$. Then the following hold.*

(i) *Case $Q = (-A)^{-\alpha}$ and $p'[(\beta - \alpha) \vee 0] < 1$.*

For any $u \in \mathscr{U}_p$, $I^u(t,\tau) \in E$ for any $\tau \in [t,T]$, $\mathbb{P}$-a.s., and there exists a positive constant $c_{\alpha,\beta,p,T}$ such that

$$|I^u(t,\tau)|_E \leq c_{\alpha,\beta,p,T}\|u\|_{L^p(0,T;H)}, \quad \tau \in [t,T], \ \mathbb{P}\text{-a.s.} \qquad (6.8)$$

(ii) *Case $Q = I$ and $p'[(\beta - \alpha) \vee 0] < 1$.*

For any $u \in \mathscr{U}_p^\alpha$, $I^u(t,\tau) \in E$ for any $\tau \in [t,T]$, $\mathbb{P}$-a.s., and there exists a positive constant $c_{\alpha,\beta,p,T}$ such that

$$|I^u(t,\tau)|_E \leq c_{\alpha,\beta,p,T}\|u\|_{L^p(0,T;D((-A)^\alpha))}, \quad \tau \in [t,T], \ \mathbb{P}\text{-a.s.}$$

(iii) *Case $Q = I$ and $p'\beta < 1$.*

For any $u \in \mathscr{U}_p$, $I^u(t,\tau) \in E$ for any $\tau \in [t,T]$, $\mathbb{P}$-a.s., and $I^u(t,\tau)$ satisfies estimate (6.8) for some positive constant $c_{\alpha,\beta,p,T}$.

Proof. Let us prove item (i), items (ii) and (iii) follow from similar arguments. From Hypothesis 6.1, we have

$$|e^{(\tau-s)A}(-A)^{-\alpha}u_s|_E \leq c(\tau-s)^{(-\beta+\alpha)\wedge 0}|u_s|_H \quad \text{a.e. } s \in (t,\tau), \mathbb{P}\text{-a.s.}$$

Therefore,

$$\left|\int_t^\tau e^{(\tau-s)A}(-A)^{-\alpha}u_s ds\right|_E \leq \int_t^\tau \left|e^{(\tau-s)A}(-A)^{-\alpha}u_s\right|_E ds \leq c\int_t^\tau (\tau-s)^{(-\beta+\alpha)\wedge 0}|u_s|_H ds$$

$$\leq c\Big(\int_t^\tau (\tau-s)^{[(-\beta+\alpha)\wedge 0]p'}ds\Big)^{1/p'}\|u\|_{L^p(0,T;H)}$$

$$\leq c(\tau-t)^{(-\beta+\alpha)\wedge 0+1/p'}\|u\|_{L^p(0,T;H)}, \quad \mathbb{P}\text{-a.s.}$$

$\square$

Thanks to Lemma 6.5, arguing as in [6, Theorem 7.11] we deduce the following result, which is the counterpart of Proposition 2.5-(i) for the controlled equation.

Proposition 6.6. *Let Hypothesis 2.1 holds true, and assume that Hypothesis 6.1 holds true for some positive constant β. Let $t \in [0, T]$, $p \geq 1$, and set p' be the conjugate exponent of p. Then the following hold.*

(i) *Case $Q = (-A)^{-\alpha}$, $p'[(\beta - \alpha) \vee 0] < 1$.*

For any $x \in E$ and $u \in \mathscr{U}_p$, there exists a unique mild solution $X_\tau^{t,x,u}$ to (6.1) belonging to $\mathscr{S}^2((t, T]; E)$. Moreover, there exists a positive constant c such that, for any $\tau \in [t, T]$,

$$|X_\tau^{t,x,u}|_E \leq c\Big(|x|_E + \|u\|_{L^p(t,T;H)}^{2m+1} + \sup_{\tau \in [t,T]} |w^A(t,\tau)|_E^{2m+1}\Big), \quad \mathbb{P}\text{-a.s.} \tag{6.9}$$

(ii) *Case $Q = I$, $p'[(\beta - \alpha) \vee 0] < 1$.*

For any $x \in E$ and $u \in \mathscr{U}_p^\alpha$, there exists a unique mild solution $X_\tau^{t,x,u}$ to (6.1) belonging to $\mathscr{S}^2((t, T]; E)$. Moreover, there exists a positive constant c such that, for any $\tau \in [t, T]$,

$$|X_\tau^{t,x,u}|_E \leq c\Big(|x|_E + \|u\|_{L^p(t,T;D((-A)^\alpha))}^{2m+1} + \sup_{\tau \in [t,T]} |w^A(t,\tau)|_E^{2m+1}\Big), \quad \mathbb{P}\text{-a.s.} \tag{6.10}$$

(iii) *Case $Q = I$, $p'\beta < 1$.*

For any $x \in E$ and $u \in \mathscr{U}_p$, there exists a unique mild solution $X_\tau^{t,x,u}$ to (6.1) belonging to $\mathscr{S}^2((t, T]; E)$. Moreover, there exists a positive constant c such that, for any $\tau \in [t, T]$,

$$|X_\tau^{t,x,u}|_E \leq c\Big(|x|_E + \|u\|_{L^p(t,T;H)}^{2m+1} + \sup_{\tau \in [t,T]} |w^A(t,\tau)|_E^{2m+1}\Big), \quad \mathbb{P}\text{-a.s.} \tag{6.11}$$

Proof. We show item (i), the proof of items (ii) and (iii) being analogous. Since by Lemma 6.5 the convolution defined in (6.7) is a well defined E-valued process for any $u \in \mathscr{U}_p$, it is possible to argue as in [6, Theorem 7.11]. Therefore, by applying the fixed point theorem we infer that for any $t \in [0, T]$, $x \in E$ and $u \in \mathscr{U}_p$, there exists a unique mild solution $X^{\alpha,t,x,u}$ to (6.1) with F replaced by its Yosida approximations F_α, $\alpha > 0$, such that $X^{\alpha,t,x,u}$ satisfies (6.9). Further, the sequence $\{X^{\alpha,t,x,u}\}_{\alpha>0}$ converges as $\alpha \to 0$ to the mild solution $X^{t,x,u}$ to (6.1). In particular, estimate (6.9) holds true also for $X^{t,x,u}$. □

6.1 The structure condition: the case $Q = (-A)^{-\alpha}$

In this section we deal with control processes $u \in \mathscr{U}_2$, and with the controlled equation

$$\begin{cases} dX_\tau^u = AX_\tau^u d\tau + F(X_\tau^u)d\tau + (-A)^{-\alpha}u_\tau d\tau + (-A)^{-\alpha}dW_\tau, & \tau \in [t, T], \\ X_t^u = x \in E, \end{cases} \tag{6.12}$$

satisfying the so called structure condition: the control affects the system only through the noise.

We make the following assumptions on the cost functional (6.2).

Hypothesis 6.7. *Let $\phi : E \to \mathbb{R}$ and $\ell : [0, T] \times E \times H \to \mathbb{R}$ be two measurable functions satisfying the following properties.*

(i) *ϕ is continuous and bounded.*

(ii) *For all $t \in [0, T]$, $u \in H$, the function $x \mapsto \ell(t, x, u)$ is bounded and continuous from E onto $\mathbb{R}$. For all $t \in [0, T]$, $x \in E$, the function $u \mapsto \ell(t, x, u)$ is continuous from H onto $\mathbb{R}$. Further, there exist c, C, R positive constants such that, for all $t \in [0, T]$, $x \in E$, $u \in H$,*

$$0 \leq \ell(t, x, u) \leq c(1 + |u|_H)^2, \tag{6.13}$$

$$\ell(t, x, u) \geq C|u|_H^2, \quad |u|_H \geq R. \tag{6.14}$$

(iii) *There exists a positive constant $L > 0$ such that, for all $t \in [0,T]$, $x_1, x_2 \in E$, $u \in H$,*

$$|\ell(t, x_1, u) - \ell(t, x_2, u)| \leq L|x_1 - x_2|_E,$$

Remark 6.8. *Under Hypothesis 6.7-(ii), it is easy to see that there exist c, R positive constants such that*

$$\ell(t, x, u) \geq c(|u|_H^2 - R^2), \quad t \in [0,T], \ x \in E, \ u \in H.$$

We introduce the Hamiltonian function

$$\psi(t, x, z) := \inf_{u \in H} \left\{ \ell(t, x, u) + \langle z, u \rangle_H \right\}, \quad t \in [0,T], \ x \in E, \ z \in H. \tag{6.15}$$

Arguing as in [9, Lemma 3.1] we deduce an analogous result.

Lemma 6.9. *Let Hypotheses 6.7 be satisfied. Then, the function ψ in (6.15) is Borel measurable, and there exists a positive constant C such that*

$$-C(1 + |z|_H^2) \leq \psi(t, x, z) \leq \ell(t, x, u) + |z|_H|u|_H, \quad t \in [0,T], \ x \in E, \ z, u \in H. \tag{6.16}$$

Further, if the minimum in (6.15) is attained, it is attained in a ball of radius $C(1 + |z|_H)$, i.e.,

$$\psi(t, x, z) = \inf_{u \in H, |u|_H \leq C(1+|z|_H)} \left\{ \ell(t, x, u) + \langle z, u \rangle_H \right\}, \quad t \in [0,T], \ x \in E, \ z \in H, \tag{6.17}$$

$$\psi(t, x, z) \leq \ell(t, x, u) + \langle z, u \rangle_H, \quad |u|_H \geq C(1 + |z|_H).$$

Finally, there exists a positive constant C such that, for any $x_1, x_2 \in E$, $z_1, z_2 \in H$,

$$|\psi(t, x_1, z_1) - \psi(t, x_2, z_2)| \leq C(|x_1 - x_2|_E + |z_1 - z_2|_H(1 + |z_1|_H + |z_2|_H)), \quad t \in [0,T]. \tag{6.18}$$

The HJB equation associated to the control problem (6.3), related to the controlled state equation (6.12), is given by

$$\begin{cases} \frac{\partial v}{\partial t}(t, x) = -\mathscr{L}v(t, x) + \psi\left(t, x, v(t, x), \nabla^{(-A)^{-\alpha}}v(t, x)\right), & t \in [0,T], \ x \in E, \\ v(T, x) = \phi(x), \end{cases} \tag{6.19}$$

where ψ is defined in (6.15). The HJB equation (6.19) turns out to be a semilinear Kolmogorv equation as (4.16), with ψ and ϕ satisfying Hypotehsis 3.1. So by Theorem 5.4 its mild solution can be represented in terms of the solution $(X^{t,x}, Y^{t,x}, Z^{t,x})$ of the forward-backward system (3.1).

In the following Theorem we state and prove the fundamental relation, and we characterize the optimal control with a feedback law.

Theorem 6.10. *Let Hypotheses 2.1, 6.7 hold true, and assume that Hypothesis 6.1 holds true with a constant β such that $\beta - \alpha < 1/2$. Let $X^{t,x,u}$ be the mild solution of (6.12), $V(t, x)$ be the value function of the control problem (6.3), and v be the mild solution of the HJB equation (6.19). Then, for any $(t, x) \in [0,T] \times E$ and $u \in \mathscr{U}_2$, the so called fundamental relation holds true:*

$$v(t, x) = J(t, x, u) + \mathbb{E}\left[\int_\tau^T \left(\psi(s, X_s^{t,x,u}, Z_s^{t,x}) - \ell(s, X_s^{t,x,u}, u_s) - Z_s^{t,x}u_s \right) ds \right].$$

In particular, $v(t, x) \leq V(t, x)$, for all $(t, x) \in [0,T] \times E$. Moreover, if there exists a measurable function $\gamma : [0,T] \times E \times H \to H$ satisfying

$$\psi(t, x, z) = \ell(t, x, \gamma(t, x, z)) + \langle z, \gamma(t, x, z) \rangle_H, \quad t \in [0,T], \ x \in E, \ z \in H,$$

then

$$v(t, x) = V(t, x)$$

and, thanks to (6.17), *the process* $\bar{u}$ *defined by*

$$\bar{u}_s := \gamma(s, X_s^{x,\bar{u}}, \nabla_x v(s, X_s^{x,\bar{u}}) \nabla_x X_s^{x,\bar{u}} (-A)^{-\alpha}) \quad \text{for-a.e. } s \in (0, T), \quad \mathbb{P}\text{-a.s.},$$

belongs to $\mathscr{U}_2$ *and it is optimal.*

Proof. The proof is standard and follows the same lines of [9, Proposition 4.1]. We notice that, by Proposition 6.6-(*i*), problem (6.12) admits a unique mild solution $X^{t,x,u}$ for any $u \in \mathscr{U}_2$. Further, for any $u \in \mathscr{U}_2$, we introduce the family of stopping times τ_n defined by

$$\tau_n := \inf \left\{ \tau \in [t, T] : \int_t^\tau |u_s|_H^2 ds > n \right\}, \quad n \in \mathbb{N}.$$

Then we proceed as in [9, Proposition 4.1], by applying the Girsanov Theorem and using the fact that ψ satisfies Hypothesis 3.1-(ii), and that the pair of processes $(Y^{t,x}, Z^{t,x})$, solution to the Markovian BSDE in (3.1), are identified respectively with the solution v of the HJB equation (6.19) and with its directional derivative $\nabla^{(-A)^{-\alpha}} v$. Namely, by Theorems 5.1 and 5.4, $Y_s^{t,x} = v(s, X_s^{t,x})$ and $Z_s^{t,x} = \nabla_x v(s, X_s^{x,\bar{u}}) \nabla_x X_s^{x,\bar{u}} (-A)^{-\alpha}$. $\square$

Example 6.11. *We now exploit Example 2.4 with $n = 1$, namely we present an heat equation with additive noise in dimension less than or equal to three. As we will see, we consider the space of continuous functions, where the heat semigroup with Dirichlet boundary conditions turns out to be analytic. Heat equations, and in general reaction diffusion equations, arise naturally in applications from physics, chemistry and so on, and it is interesting to treat optimal control problems related to them in the Banach space of continuous functions: this allows, for instance, to control some variable of the state, say, the temperature, in specific points, see below for a mathematical formulation of this problem. Besides the interest in the related controlled problems, we are able to consider dissipative drifts with polynomial growth.*
In the complete and filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t, \ t \geq 0\}, \mathbb{P})$, *let us consider the equation*

$$\begin{cases} \frac{\partial}{\partial_\tau} X^u(\tau, \xi) = \sum_{i=1}^d \frac{\partial^2}{\partial \xi_i^2} X^u(\tau, \xi) - (X^u(\tau, \xi))^3 + Qu(\tau, \xi) + Q\dot{W}(\tau, \xi), \quad \tau \in [t, T], \ \xi \in \mathscr{O}, \\ X^u(\tau, \xi) = 0, \quad \tau \in [t, T], \ \xi \in \partial\mathscr{O}, \\ X^u(t, \xi) = x_0(\xi), \quad \xi \in \mathscr{O}, \end{cases} \quad (6.20)$$

where $\mathscr{O} \subset \mathbb{R}^d$, $d \leq 3$ *is an open bounded set, and* $\dot{W}(\tau, \xi)$ *is a space-time white noise on* $[0, T] \times \mathbb{R}^d$. *The process* $u(\tau, \xi)$ *is a predictable process such that* $u(\tau, \cdot) \in L^2(\mathscr{O}; \mathbb{R})$. *We denote by A the realization of the Laplacian with Dirichlet boundary conditions and we consider* $Q = (-A)^{-\alpha}$. *As already noticed in Example 2.4, the operator A and Q satisfy Hypothesis 2.1.*
Let $H := L^2(\mathscr{O}; \mathbb{R})$ *and* $E := C(\overline{\mathscr{O}}; \mathbb{R})$. *If* $x_0 \in E$, *equation* (6.20) *can be reformulated in an abstract way in the Banach space E as*

$$\begin{cases} dX_\tau^{u,t,x} = AX_\tau^{u,t,x} d\tau + F(X_\tau^{u,t,x}) d\tau + (-A)^{-\alpha} u_\tau d\tau + (-A)^{-\alpha} dW_\tau, \quad \tau \in [t, T], \\ X_t^{u,t,x} = x_0 \in E, \end{cases} \quad (6.21)$$

where $F(X_\tau^{u,t,x}) = -(X_\tau^{u,t,x})^3$, *and it satisfies Hypothesis 2.1, points 3 and 4. Clearly F is not well defined on the whole Hilbert space H but it is well defined in E.*
We then introduce the cost functional, that has to be minimized over all admissible controls $\mathscr{U}$:

$$J(t, x, u) = \mathbb{E}\left[\int_t^T \ell(s, X_s^{u,t,x}, u_s) ds\right] + \mathbb{E}\left[\phi(X_T^{u,t,x})\right] \quad (6.22)$$

with

$$\ell(s, z, a) := \int_{\mathscr{O}} l_0\left(s, z\left(\xi\right)\right) \mu(d\xi) + l_1\left(\int_{\mathscr{O}} |a\left(\xi\right)|^2 d\xi\right), \tag{6.23}$$

$$\phi(z) := \max_{\xi \in \mathscr{O}} |z(\xi)| \wedge k. \tag{6.24}$$

Here $k > 0$ and μ is a finite regular measure on $\mathscr{O}$. The current cost in (6.23) is well defined on the space of continuous function on $\mathscr{O}$, but for an arbitrary μ it is not well defined on the Hilbert space of square integrable functions. Moreover, the final cost in (6.24) is well defined on the space of continuous functions, but it is not differentiable. This example shows therefore that in the Banach space framework we can treat a larger class of stochastic optimal control problems than the one covered by the previous Hilbert space theory.

Concerning the functions l_0 and l_1 appearing in (6.23), we ask that, for all $s \in [0, T]$, $\zeta \mapsto l_0(s, \zeta)$ is bounded and continuous from $\mathbb{R}$ to $\mathbb{R}$ and $l_1 : [0, +\infty) \to \mathbb{R}$ is continuous with

$$0 \leq l_1(y) \leq 1 + y \quad and \quad l_1(y) \geq Cy, \quad y \geq R^2.$$

Under these assumptions the function ℓ defined in (6.23) satisfies Hypothesis 6.7.

6.2 The case $Q = I$ with a special running cost

In the present section we deal with control processes $u \in \mathscr{U}_2$, and with the controlled equation

$$\begin{cases} dX_\tau^u = AX_\tau^u d\tau + F(X_\tau^u)d\tau + u_\tau d\tau + (-A)^{-\alpha}dW_\tau, & \tau \in [t, T], \\ X_t^u = x \in E. \end{cases} \tag{6.25}$$

The controlled equation (6.25) has a different structure towards (6.12) considered in Subsection 6.1, so the problem is different, and we need different assumptions on the cost functional (6.2).

Hypothesis 6.12. *Let $\phi : E \to \mathbb{R}$ and $\ell : [0, T] \times E \times H \to \mathbb{R} \cup \{+\infty\}$ be two measurable functions satisfying the following properties.*

(i) *ϕ is continuous and bounded.*

(ii) *For all $t \in [0, T]$, $u \in D((-A)^\alpha)$, the function $x \mapsto \ell(t, x, u)$ is bounded and continuous from E onto $\mathbb{R}$. For all $t \in [0, T]$, $x \in E$, the function $u \mapsto \ell(t, x, u)$ is continuous from $D((-A)^\alpha)$ onto $\mathbb{R}$. Further, there exists c, C, R positive constants such that, for all $t \in [0, T]$, $x \in E$ and $u \in D((-A)^\alpha)$,*

$$0 \leq \ell(t, x, u) \leq c(1 + |u|_\alpha)^2, \tag{6.26}$$

$$\ell(t, x, u) \geq C|u|_\alpha^2, \quad |u|_\alpha \geq R. \tag{6.27}$$

(iii) *There exists a positive constant $L > 0$ such that, for any $t \in [0, T]$, $u \in D((-A)^\alpha)$, $x_1, x_2 \in E$,*

$$|\ell(t, x_1, u) - \ell(t, x_2, u)| \leq L|x_1 - x_2|_E.$$

Remark 6.13. *Condition (6.27) in Hypothesis 6.12 implies that, if u does not take values in $D((-A)^\alpha)$, then $J(t, x, u) = +\infty$. In particular, $\inf_{u \in \mathscr{U}_2^\alpha} J(t, x, u) = \inf_{u \in \mathscr{U}_2} J(t, x, u)$, so we can limit ourselves to consider here the space of admissible controls $\mathscr{U}_2^\alpha$.*

Remark 6.14. *Under Hypothesis 6.12-(ii), there exist positive constants c, R such that, for any $t \in [0, T]$, $x \in E$, $u \in D((-A)^\alpha)$, we have $\ell(t, x, u) \geq c(|u|_\alpha^2 - R^2)$.*

We introduce the Hamiltonian function

$$\psi^\alpha(t,x,z) := \inf_{u \in D((-A)^\alpha)} \left\{ \ell(t,x,u) + \langle z, (-A)^\alpha u \rangle_H \right\}, \quad t \in [0,T], \ x \in E, \ z \in H. \tag{6.28}$$

Arguing again as in [9, Lemma 3.1], we infer the following properties of ψ^α.

Lemma 6.15. *Let Hypotheses 6.12 be satisfied. Then, the function ψ^α in (6.28) is Borel measurable and there exists a positive constant C such that*

$$-C(1 + |z|_H^2) \le \psi^\alpha(t,x,z) \le \ell(t,x,u) + |z|_H |u|_\alpha, \quad t \in [0,T], \ x \in E, \ z \in H, \ u \in D((-A)^\alpha). \tag{6.29}$$

Further, if the minimum in (6.28) is attained, it is attained in a ball of radius $C(1 + |z|_H)$, i.e.,

$$\psi^\alpha(t,x,z) = \inf_{u \in D((-A)^\alpha), |u|_\alpha \le C(1 + |z|_H)} \left\{ \ell(t,x,u) + z(-A)^\alpha u \right\}, \quad t \in [0,T], \ x \in E, \ z \in H. \tag{6.30}$$

Finally, for any $x_1, x_2 \in E$, $z_1, z_2 \in H$, ψ^α, there exists a positive constant C such that

$$|\psi^\alpha(t,x_1,z_1) - \psi^\alpha(t,x_2,z_2)| \le C(|x_1 - x_2|_E + |z_1 - z_2|_H (1 + |z_1|_H + |z_2|_H)), \quad t \in [0,T]. \tag{6.31}$$

The HJB equation associated to the control problem (6.3), related to the controlled state equation (6.25), is given by

$$\begin{cases} \frac{\partial v}{\partial t}(t,x) = -\mathcal{L}v(t,x) + \psi^\alpha\left(t,x,v(t,x), \nabla^{(-A)^{-\alpha}} v(t,x)\right), & t \in [0,T], \ x \in E, \\ v(T,x) = \phi(x), \end{cases} \tag{6.32}$$

where ψ^α is defined in (6.28). Again, the HJB equation (6.32) turns out to be a semilinear Kolmogorv equation as (4.16), with ψ^α and ϕ satisfying Hypotehsis 3.1. So by Theorem 5.4 its mild solution can be represented in terms of the solution $(X^{t,x}, Y^{t,x}, Z^{t,x})$ of the forward-backward system

$$\begin{cases} dX_\tau = AX_\tau d\tau + F(X_\tau)d\tau + (-A)^{-\alpha}dW_\tau, & \tau \in [t,T], \\ X_t = x, \\ dY_\tau = -\psi^\alpha(\tau, X_\tau, Y_\tau, Z_\tau)\,d\tau + Z_\tau\,dW_\tau, & \tau \in [t,T], \\ Y_T = \phi(X_T), \end{cases} \tag{6.33}$$

which is nothing else than the forward-backward system (3.1) with ψ instead of ψ^α.

As in Subsection 6.1, in the following Theorem we state and prove the fundamental relation, and we characterize the optimal control with a feedback law.

Theorem 6.16. *Let Hypotheses 2.1, 6.12 hold true, and assume that Hypothesis 6.1 holds true with a constant β such that $\beta - \alpha < 1/2$. Let $X^{t,x,u}$ be the mild solution of (6.25), $V(t,x)$ be the value function of the control problem (6.3), and v be the mild solution of the HJB equation (6.32). Then, for any $(t,x) \in [0,T] \times E$ and $u \in \mathscr{U}_2^\alpha$,*

$$v(t,x) = J(t,x,u) + \mathbb{E} \int_\tau^T \left(\psi^\alpha(s, X_s^{t,x,u}, Z_s^{t,x}) - \ell(s, X_s^{t,x,u}, u_s) - Z_s^{t,x}(-A)^\alpha u_s \right) ds.$$

In particular, $v(t,x) \le V(t,x)$, for all $t \in [0,T]$, $x \in E$. Moreover, if there exists a measurable function $\gamma^\alpha : [0,T] \times E \times H \to D((-A)^\alpha)$ satisfying

$$\psi^\alpha(t,x,z) = \ell(t,x,\gamma^\alpha(t,x,z)) + \langle z, (-A)^\alpha \gamma^\alpha(t,x,z) \rangle_H, \quad t \in [0,T], \ x \in E, \ z \in H,$$

then

$$v(t,x) = V(t,x)$$

and, thanks to (6.30), the process

$$\bar{u}_s^\alpha := \gamma^\alpha(s, X_s^{x,\bar{u}^\alpha}, \nabla_x v(s, X_s^{x,\bar{u}^\alpha}) \nabla_x X_s^{x,\bar{u}^\alpha}(-A)^{-\alpha}) \quad \text{for-a.e. } s \in (0,T), \quad \mathbb{P}\text{-a.s.,} \tag{6.34}$$

belongs to $\mathscr{U}_2^\alpha$ and it is optimal.

Proof. Notice that by Proposition 6.6-(*ii*), for any $u \in \mathscr{U}_2^\alpha$ there exists a unique mild solution $X^{t,x,u}$ to (6.25) which satisfies (6.10). The proof is similar to the one of Theorem 6.10. The main difference consists in the fact that, for any given $u \in \mathscr{U}_2^\alpha$, we introduce a family of stopping times depending on the norm $|\cdot|_\alpha$:

$$\tau_n := \inf \left\{ \tau \in [t, T] : \int_t^\tau |u_s|_\alpha^2 ds > n \right\}, \quad n \in \mathbb{N}.$$

Then, we set $u_\tau^n := u_\tau \mathbb{1}_{\tau \leq \tau_n} + u_0 \mathbb{1}_{\tau > \tau_n}$, $u_0 \in D((-A)^\alpha)$, and we introduce the process

$$W_\tau^n := W_\tau + \int_t^\tau (-A)^\alpha u_s^n ds.$$

Afterwards, we apply the Girsanov Theorem: writing $u_s = (-A)^{-\alpha}(-A)^\alpha u_s$ in (6.25), we get that X^{t,x,u^n} is mild solution to

$$\begin{cases} dX_\tau = AX_\tau d\tau + F(X_\tau)d\tau + (-A)^{-\alpha}dW_\tau^n, & \tau \in [t, T], \\ X_t = x \in E. \end{cases}$$

By (6.31) in Lemma 6.15, we see that Hypothesis 3.1-(ii) is verified by ψ^α. We conclude by arguing again as in [9, Proposition 4.1] and in Theorem 6.10. □

6.3   The case $Q = I$ with a general running cost

In this subsection we deal with the general controlled equation (6.25) under Hypothesis 6.7 on the coefficients of the cost functional, and we consider control processes $u \in \mathscr{U}_2$. Unlike the two cases just treated, in this framework the HJB equation would not have the structure of equation (4.16) since the Hamiltonian function would depend on ∇v, not only on the directional derivative $\nabla^{(-A)^{-\alpha}} v$, see e.g. [8], formula (6.67) and the discussion related to formulas (4.278)-(4.279). Up to our knowledge, when ϕ in only continuous, the well posedness of such an equation is an open problem: in [3] an equation of this type is solved in mild sense with Lipschitz type assumptions on the final datum ϕ.

For this reason, we will not end up identifying the value function (6.3) with the solution of the HJB equation, but instead we will approximate it. The following result will be used in the aforementioned approximation of the value function.

Proposition 6.17. *Assume that Hypothesis 2.1 holds true. Let $t \in [0, T]$, $x \in E$ and $u, \tilde{u} \in \mathscr{U}_2$. Then,*

$$\sup_{\tau \in [t,T]} |X_\tau^{t,x,u} - X_\tau^{t,x,\tilde{u}}|_H^2 \leq \int_t^T |u_s - \tilde{u}_s|_H^2 \, ds, \quad \mathbb{P}\text{-a.s.}, \tag{6.35}$$

where $X^{t,x,u}$ and $X^{t,x,\tilde{u}}$ are respectively the mild solutions to (6.25) with control u and $\tilde{u}$.

Proof. Let us set $L(\tau) := X_\tau^{t,x,u} - X_\tau^{t,x,\tilde{u}}$ and let us assume that $L(\tau)$ is a strict solution to

$$\begin{cases} \frac{d}{d\tau}L(\tau) = AL(\tau) + F(X_\tau^{t,x,u}) - F(X_\tau^{t,y,\tilde{u}}) + u_\tau - \tilde{u}_\tau, & \tau \in [t, T], \\ L(t) = 0, \end{cases}$$

otherwise we can use an approximation argument as in the proof of Proposition 2.5(ii). Then, the non-positivity of A, the dissipativity of F, the Cauchy-Schwartz inequality and the Young inequality give

$$\frac{1}{2}\frac{d}{ds}|L(s)|_H^2 \leq \langle u_s - \tilde{u}_s, L(s) \rangle_H \leq \frac{1}{2}|u_s - \tilde{u}_s|_H^2 + \frac{1}{2}|L(s)|_H^2, \quad s \in [t, T], \quad \mathbb{P}\text{-a.s.}$$

Integrating between t and τ and applying the Gronwall Lemma, we get

$$|L(\tau)|_H^2 \leq \int_t^\tau |u_s - \tilde{u}_s|_H^2 ds, \quad \forall \tau \in [t,T], \quad \mathbb{P}\text{-a.s.},$$

and we immediately deduce (6.35). $\qquad \square$

Thanks to Proposition 6.17 we deduce that, up to a subsequence, we can approximate $X^{t,x,u}$ in H by means of mild solutions X^{t,x,u^n} of problem (6.25), with u replaced by u^n, where $(u_n) \subset \mathscr{U}_p$ satisfies $u^n \to u$ in $\mathscr{U}_p$. In the following proposition we prove that a similar approximation holds true in E.

Proposition 6.18. *Let Hypothesis 2.1 holds true. Let $t \in [0,T]$, $p \geq 2$, and set p' be the conjugate exponent of p. Assume that Hypothesis 6.1 holds true for some positive constant β such that $p'\beta < 1$. Let $u \in \mathscr{U}_p$ and $(u^n) \subset \mathscr{U}_p$ be such that $u^n \to u$ in $\mathscr{U}_p$. Then, for any $x \in E$,*

$$\lim_{n \to +\infty} |X_\tau^{t,x,u^{k_n}} - X_\tau^{t,x,u}|_E = 0, \quad \forall \tau \in [t,T], \quad \mathbb{P}\text{-a.s.,} \tag{6.36}$$

where $(u^{k_n}) \subset (u^n)$ be such that $u_s^{k_n} \to u_s$ $\mathbb{P}$- a.s. for a.e. $s \in (t,T)$.

Proof. As usual, we limit ourselves to consider the case $t = 0$. For any $n \in \mathbb{N}$, let us set $L^n := X^{x,u^{k_n}} - X^{x,u}$, where $X^{x,u^{k_n}}$ and $X^{x,u}$ are mild solutions to (6.25) with initial datum x and control processes u^{k_n} and u, respectively. Further, let us denote by N the subset of Ω such that $\mathbb{P}(N) = 0$ and $u_s^{k_n} \to u_s$ on $\Omega \setminus N$ for a.e. $s \in (t,T)$. Then, for any $t \in [0,T]$

$$L_t^n = \int_0^t e^{(t-s)A}(F(X_s^{x,u^{k_n}}) - F(X_s^{x,u}))ds + \int_0^t e^{(t-s)A}(u_s^{k_n} - u_s)ds, \quad \mathbb{P}\text{-a.s.}$$

which gives

$$|L_t^n|_E \leq \int_0^t |e^{(t-s)A}(F(X_s^{x,u^{k_n}}) - F(X_s^{x,u}))|_E ds + \int_0^t |e^{(t-s)A}(u_s^{k_n} - u_s)|_E ds =: I_1^n(t) + I_2^n(t), \quad \mathbb{P}\text{-a.s.}$$

Let us estimate I_1^n and I_2^n separately. As far as I_1^n is concerned, from the boundedness of e^{tA} on E, Hypothesis 2.1-4. and (6.11), it follows that

$$|e^{(t-s)A}(F(X_s^{x,u^{k_n}}) - F(X_s^{x,u}))|_E < \infty, \quad s \in (0,T),$$

on $\Omega \setminus N$. Further, from (6.4) it follows that

$$|e^{(t-s)A}(F(X_s^{x,u^{k_n}}) - F(X_s^{x,u}))|_E \leq c(t-s)^{-\beta}|F(X_s^{x,u^{k_n}}) - F(X_s^{x,u})|_H,$$

on $\Omega \setminus N$, for any $s \in (0,T)$. Since F is continuous on H, from (6.35) we infer that $|F(X_s^{x,u^n}) - F(X_s^{x,u})|_H \to 0$ on $\Omega \setminus N$ as $n \to +\infty$ for any $s \in (0,T)$. The dominated convergence theorem implies that $I_1^n \to 0$ as $n \to +\infty$ on $\Omega \setminus N$.

Concerning I_2^n, from (6.4) and arguing as above we get

$$I_2^n(t) \leq c \int_0^t (t-s)^{-\beta}|u_s^{k_n} - u_s|_H ds \leq cT^{-\beta+1/p'}\|u^{k_n} - u\|_{L^p(0,T;H)} \to 0, \quad n \to +\infty,$$

on $\Omega \setminus N$. This concludes the proof. $\qquad \square$

6.3.1 The approximate optimal control problem

We will consider the Hamiltonian function ψ^α in (6.28) under Hypothesis 6.7. This prevents us to obtain directly estimates as those in Lemmas 6.9 and 6.15, since we don't have the structure condition and the assumptions on ℓ are not sufficient to bound the term $(-A)^\alpha u$. For this reason, for any $n \in \mathbb{N}$ we introduce the function $\ell_n : [0, T] \times E \times H \longrightarrow \mathbb{R}$ defined by

$$\ell_n(s, x, u) := \ell(s, x, u) + \frac{1}{n}|(-A)^\alpha u|_H^2. \tag{6.37}$$

Lemma 6.19. *Let A be an operator satisfying Hypothesis 2.1-(i). Then the function ℓ_n in (6.37) satisfies the following conditions: for any $(t, x) \in [0, T] \times H$,*

$$0 \leq \ell_n(t, x, u) \leq c_n(1 + |u|_\alpha^2) \quad u \in D((-A)^\alpha),$$
$$\exists c_n, R > 0 : \ \ell_n(t, x, u) \geq c_n|u|_\alpha^2 - cR^2, \quad u \in D((-A)^\alpha).$$

Proof. The first inequality directly comes from (6.37). On the other hand, for any $t \in [0, T]$, $x \in E$ and $u \in D((-A)^\alpha)$, by Remark 6.8 we have

$$\ell_n(t, x, u) \geq c\left(|u|_H^2 - R^2\right) + \frac{1}{n}|(-A)^\alpha u|_H^2 \geq c_n|u|_\alpha^2 - cR^2.$$

$\square$

For any $n \in \mathbb{N}$, we introduce the approximate Hamiltonian function

$$\psi_n(t, x, z) := \inf_{u \in D((-A)^\alpha)} \left\{\ell_n(t, x, u) + \langle z, (-A)^\alpha u\rangle_H\right\}, \quad t \in [0, T], \ x \in E, \ z \in H. \tag{6.38}$$

Estimates in Lemma 6.19 give the following result, which is analogous to Lemma 6.15.

Lemma 6.20. *Let Hypothesis 6.7 be satisfied, and let A be an operator satisfying Hypothesis 2.1-(i). Then, for any $n \in \mathbb{N}$, the function ψ_n in (6.38) is Borel measurable, and there exists a positive constant C_n such that*

$$-C_n(1 + |z|_H^2) \leq \psi_n(t, x, z) \leq \ell_n(t, x, u) + |z|_H|u|_\alpha, \quad t \in [0, T], \ x \in E, \ z \in H, \ u \in D((-A)^\alpha). \tag{6.39}$$

Further, if the minimum in (6.38) is attained, it is attained in a ball of radius $C_n(1 + |z|_H)$, i.e.,

$$\psi_n(t, x, z) = \inf_{u \in D((-A)^\alpha), |u|_\alpha \leq C_n(1+|z|_H)} \left\{\ell_n(t, x, u) + z(-A)^\alpha u\right\}, \quad t \in [0, T], \ x \in E, \ z \in H.$$

In particular, there exists a positive constant C_n such that, for any $x_1, x_2 \in E$, $z_1, z_2 \in H$,

$$|\psi_n(t, x_1, z_1) - \psi_n(t, x_2, z_2)| \leq C_n(|x_1 - x_2|_E + |z_1 - z_2|_H(1 + |z_1|_H + |z_2|_H)), \quad t \in [0, T]. \tag{6.40}$$

For any $n \in \mathbb{N}$, we introduce the approximate cost functional defined by

$$J_n(t, x, u) := \mathbb{E}\left[\int_t^T \ell_n(s, X_s^u, u_s)ds\right] + \mathbb{E}[\Phi(X_T^u)], \tag{6.41}$$

and the associated approximated optimal control problem

$$V_n(t, x) := \inf_{u \in \mathscr{U}_2^\alpha} J_n(t, x, u), \quad t \in [0, T], \ x \in E. \tag{6.42}$$

The HJB equation associated to the control problem (6.42), related to the controlled state equation (6.25), is given by

$$\begin{cases} \frac{\partial v}{\partial t}(t,x) = -\mathscr{L}v(t,x) + \psi^n\left(t,x,v(t,x),\nabla^{(-A)^{-\alpha}}v(t,x)\right), & t \in [0,T], \ x \in E, \\ v(T,x) = \phi(x), \end{cases} \tag{6.43}$$

where ψ^n is defined in (6.38). The HJB equation (6.43) is the analogous of (6.32) in Section 6.2. So again by Theorem 5.4, its solution can be represented in terms of the solution $(X^{t,x}, Y^{n,t,x}, Z^{nt,x})$ of the forward-backward system

$$\begin{cases} dY_\tau^{n,t,x} = -\psi_n(\tau, X_\tau^{t,x}, Z_\tau^{n,t,x})d\tau + Z_\tau^{n,t,x}dW_\tau, & \tau \in [t,T], \\ Y_T^{n,t,x} = \Phi(X_T^{t,x}), \\ dX_\tau^{t,x} = AX_\tau^{t,x}d\tau + F(X_\tau^{t,x})d\tau + (-A)^{-\alpha}dW_\tau, & \tau \in [t,T], \\ X_t^{t,x} = x \in E, \end{cases} \tag{6.44}$$

which is nothing else than the forward-backward system (6.33) with ψ^n instead of ψ^α.

We consider the following assumptions.

Hypothesis 6.21. *For any $n \in \mathbb{N}$, there exists a measurable function $\gamma_n : [0,T] \times E \times H \to D((-A)^\alpha)$ satisfying*

$$\psi_n(t,x,z) = \ell_n(t,x,\gamma_n(t,x,z)) + \langle z, (-A)^\alpha \gamma_n(t,x,z)\rangle_H, \quad t \in [0,T], \ x \in E, \ z \in H. \tag{6.45}$$

We state the analogous of Theorem 6.16 for the approximate optimal control problems (6.42). The results can be proved by verbatim repeating the proof of Theorem 6.16.

Theorem 6.22. *Let Hypotheses 2.1, 6.7 hold true, and assume that Hypothesis 6.1 holds true with $\beta < \frac{1}{2}$. Let $X^{u,t,x}$ be the solution of equation (6.25) and for any $n \in \mathbb{N}$, let V_n be the function defined in (6.42), and v_n be the mild solution of the HJB equation (6.43). Then, for any $(t,x) \in [0,T] \times E$ and $u \in \mathscr{U}_2^\alpha$,*

$$v_n(t,x) = J_n(t,x,u) + \mathbb{E}\Big[\int_\tau^T \left(\psi_n(s, X_s^{t,x,u}, Z_s^{n,t,x}) - \ell_n(s, X_s^{t,x,u}, u_s) - Z_s^{n,t,x}(-A)^\alpha u_s\right)ds\Big],$$

where $(X^{t,x}, Y^{n,t,x}, Z^{n,t,x})$ is the solution to (6.44). In particular, $v_n(t,x) \leq V_n(t,x)$, for all $(t,x) \in [0,T] \times E$. Finally, if Hypothesis 6.21 holds true, then

$$v_n(t,x) = V_n(t,x) \tag{6.46}$$

and, thanks to (6.30), the process

$$\bar{u}_s^n := \gamma_n(s, X_s^{x,\bar{u}^n}, \nabla_x v(s, X_s^{x,\bar{u}^n})\nabla_x X_s^{x,\bar{u}^n}(-A)^{-\alpha}) \quad \text{for-a.e. } s \in (0,T), \quad \mathbb{P}\text{-a.s.} \tag{6.47}$$

belongs to $\mathscr{U}_2^\alpha$ and it is optimal.

6.3.2 A characterization of the value function

In the present section we show that the value function V of the optimal control problem (6.3) can be approximated by the sequence (v_n) of mild solutions to (6.43), that are identified with the approximated value functions (V_n), see formula (6.46) in Theorem 6.22. As a byproduct, we deduce that the sequence $(\bar{u}^n)$ defined in (6.47) is a minimizing sequence for (6.3), and it is a bounded sequence in $\mathscr{U}_2$.

We start by introducing a family of processes by means of the Yosida approximations, namely for any $u \in \mathscr{U}_2$ we consider a suitable sequence $(u_k)_{k \geq 1} \subset \mathscr{U}_2^\alpha$ which converges to u in $\mathscr{U}_2$. Since $\bar{u}^n \in \mathscr{U}_2^\alpha$ for any $n \in \mathbb{N}$, this would allow to approximate $V(t,x)$ in terms of $J(t,x,\bar{u}^n)$.

Definition 6.23. *Let $u \in \mathscr{U}_2$. For any $k \in \mathbb{N}$ we denote by u_k the function defined as*

$$u_k(t, \omega) := \begin{cases} kR(k, A)\left(u(t, \omega)\right), & \text{if } u(t, \omega) \text{ is well defined}, \\ 0 & \text{otherwise}. \end{cases}$$

Lemma 6.24. *Let A be an operator satisfying Hypothesis 2.1-(i). Let $u \in \mathscr{U}_2$ and let u_k, with $k \in \mathbb{N}$, be the process introduced in Definition 6.23. Then, for any $k \in \mathbb{N}$, $\delta \in [0, 1]$ and $(t, \omega) \in [0, T] \times \Omega$,*

$$|u_k(t, \omega)|_H \leq C|u(t, \omega)|_H, \tag{6.48}$$

$$|(-A)^\delta u_k(t, \omega)|_H \leq c_\delta k^\delta |u(t, \omega)|_H, \tag{6.49}$$

for some positive constants C, c_δ not depending neither on k nor on u. In particular, $u_k \in \mathscr{U}_2^\delta$, $u_k \to u$ $\mathbb{P}$-a.s., a.e. in $[t, T]$ as $k \to +\infty$, and

$$u_k \to u \quad \text{in } \mathscr{U}_2 \quad \text{as } n \to +\infty. \tag{6.50}$$

Proof. Estimate (6.48) directly follows from the properties of $R(k, A)$. Further, the fact that $u_k \to u$ $\mathbb{P}$-a.s., a.e. in $[t, T]$ as $k \to +\infty$ follows from the properties of Yosida approximations. Then, convergence (6.50) follows from the dominated convergence theorem, Finally, it easily follows that

$$|(-A)u_k(t, \omega)|_H \leq kC|u(t, \omega)|_H,$$

for any $(t, \omega) \in [0, T] \times \Omega$ and any $k \in \mathbb{N}$, where C is the same positive constant as in (6.48). Interpolation estimates give (6.49). $\qquad \square$

Proposition 6.25. *Let Hypotheses 2.1, 6.7, 6.1 hold true. Let $u \in \mathscr{U}_2$ and let u_n, with $n \in \mathbb{N}$, be the process introduced in Definition 6.23. Let J, J_n be respectively the cost functionals in (6.2), (6.41). Then for any $(t, x) \in [0, T] \times E$ we have*

$$J(t, x, u_n) \to J(t, x, u), \quad n \to +\infty, \tag{6.51}$$

$$J_n(t, x, u_n) \to J(t, x, u), \quad n \to +\infty. \tag{6.52}$$

In particular, (6.51) implies that

$$V(t, x) = \inf_{u \in \mathscr{U}_2^\alpha} J(t, x, u), \quad (t, x) \in [0, T] \times E. \tag{6.53}$$

Proof. The last part of the statement is straightforward, after we prove the first part.
We first show (6.52). Since u_n pointwise converges to u, a.e. in $(0, T)$, $\mathbb{P}$-a.s., from (6.50) in Lemma 6.24 and Proposition 6.18 it follows that $X_\tau^{t,x,u_n} \to X_\tau^{t,x,u}$ $\mathbb{P}$-a.s. in E as $n \to +\infty$ for any $\tau \in [t, T]$. By dominated convergence theorem we deduce that

$$\mathbb{E}[\Phi(X_T^{t,x,u_n})] \to \mathbb{E}[\Phi(X_T^{t,x,u})], \quad n \to +\infty. \tag{6.54}$$

To estimate the convergence of the approximate running cost ℓ_n in (6.37), we consider separately the two terms in (6.37). We stress that

$$\mathbb{E}\Big[\int_t^T |\ell(s, X_s^{t,x,u_n}, (u_n)_s) - \ell(s, X_s^{t,x,u}, u_s)| ds \Big]$$

$$\leq \mathbb{E}\Big[\int_t^T |\ell(s, X_s^{t,x,u_n}, (u_n)_s) - \ell(s, X_s^{t,x,u}, u_s)| ds \Big] + \mathbb{E}\Big[\int_t^T |\ell(s, X_s^{t,x,u}, (u_n)_s) - \ell(s, X_s^{t,x,u}, u_s)| ds \Big].$$

Arguing as above, from Hypothesis 6.7-(iii) and (6.36) we get

$$|\ell(s, X_s^{t,x,u_n}, (u_n)_s) - \ell(s, X_s^{t,x,u}, (u_n)_s)| \leq L|X_s^{t,x,u_n} - X_s^{t,x,u}|_E \to 0, \quad n \to +\infty, \quad \mathbb{P}\text{-a.s.}$$

Further, from (6.13) in Hypothesis 6.7-(ii) and (6.48) we infer that

$$|\ell(s, X_s^{u_n}, (u_n)_s) - \ell(s, X_s^u, (u_n)_s)| \leq c(1 + |u_s|_H^2),$$

for any $s \in (t, T)$, $\mathbb{P}$-a.s. By dominated convergence theorem we get

$$\mathbb{E}\Big[\int_t^T |\ell(s, X_s^{u_n}, (u_n)_s) - \ell(s, X_s^u, (u_n)_s)| ds\Big] \to 0, \quad n \to +\infty. \tag{6.55}$$

Moreover, the continuity of ℓ with respect to u and the dominated convergence theorem give

$$\mathbb{E}\left[\int_t^T |\ell(s, X_s^{t,x,u}, (u_n)_s) - \ell(s, X_s^{t,x,u}, u_s)| ds\right] \to 0, \ n \to +\infty. \tag{6.56}$$

Finally, since $\alpha \in (0, 1/2)$, from (6.49) with $\delta = \alpha$ we have

$$\frac{1}{n}\mathbb{E}\Big[\int_t^T |(-A)^\alpha (u_n)_s|_H^2 ds\Big] \leq Cn^{2\alpha-1}\|u\|_{\mathscr{U}_2}^2 \to 0, \ n \to +\infty,$$

and this concludes the proof of (6.52).
To conclude, we notice that (6.54), (6.55) and (6.56) give (6.51). $\qquad\square$

The following theorem constitutes the main result of the section.

Theorem 6.26. *Let Hypotheses 2.1, 6.7, 6.21 hold true, and assume that Hypothesis 6.1 holds true with $\beta > \frac{1}{2}$. For any $n \in \mathbb{N}$, let $\bar{u}^n$ and v_n denote respectively the process in (6.47) and the mild solution to (6.43). Let V, J be respectively the functions in (6.3), (6.2). Then, for any $(t, x) \in [0, T] \times E$,*

$$V(t, x) = \lim_{n \to +\infty} v_n(t, x) = \lim_{n \to +\infty} J(t, x, \bar{u}^n). \tag{6.57}$$

Moreover, $(\bar{u}^n)$ is bounded in $\mathscr{U}_2$.

Proof. Let $(t, x) \in [0, T]$ and $\varepsilon > 0$. From (6.53) it follows that there exists $u_\varepsilon \in \mathscr{U}_2^\alpha$ such that $J(t, x, u_\varepsilon) \leq V(t, x) + \varepsilon$. We have

$$\frac{1}{n}\mathbb{E}\Big[\int_t^T |(-A)^\alpha (u_\varepsilon)_s|_H^2 ds\Big] \leq \frac{C}{n}\|u_\varepsilon\|_{\mathscr{U}_2^\alpha}^2 \to 0, \ n \to +\infty,$$

which gives that $J_n(t, x, u_\varepsilon) \to J(t, x, u)$ as $n \to +\infty$. It follows that

$$V(t, x) \leq V_n(t, x) = J_n(t, x, \overline{u}_n) \leq J_n(t, x, u_\varepsilon) \to J(t, x, u_\varepsilon) \leq V(t, x) + \varepsilon.$$

The arbitrariness of ε gives

$$\lim_{n \to +\infty} V_n(t, x) = V(t, x), \quad (t, x) \in [0, T] \times E. \tag{6.58}$$

Then the first equality in (6.57) follows from (6.58), recalling that, by Theorem 6.22, $V_n(t, x) = v_n(t, x)$ for any $n \in \mathbb{N}$.

On the other hand, since $V_n(t,x) = J_n(t,x,\bar{u}^n)$ for any $n \in \mathbb{N}$,

$$V(t,x) \le J(t,x,\bar{u}^n) \le J_n(t,x,\bar{u}^n) = V_n(t,x),$$

so that, taking into account (6.58), the second equality in (6.57) follows.

Finally, let us prove that $(\bar{u}^n)$ is bounded in $\mathscr{U}_2$. Assume by contradiction that there exists a subsequence $(u_{k_n}) \subset (\bar{u}^n)$ such that $\|u_{k_n}\|_{\mathscr{U}_2}^2 \ge n$ for any $n \in \mathbb{N}$. Then,

$$\begin{aligned}
n \le \|u_{k_n}\|_{\mathscr{U}_2}^2 &= \int_t^T \Big(\int_{\{|u_{k_n}(s)|_H \le R\}} |u_{k_n}(s)|^2 d\mathbb{P} \Big) ds + \int_t^T \Big(\int_{\{|u_{k_n}(s)|_H > R\}} |u_{k_n}(s)|^2 d\mathbb{P} \Big) ds \\
&\le TR^2 + \int_t^T \Big(\int_{\{|u_{k_n}|_H > R\}} |u_{k_n}(s)|^2 d\mathbb{P} \Big) ds.
\end{aligned}$$

On the other hand, since ℓ is nonnegative and satisfies (6.27),

$$\begin{aligned}
& \mathbb{E}\Big[\int_t^T \ell(s, X_s^{k_n}, u_{k_n}(s)) ds \Big] \\
&= \int_t^T \Big(\int_{\{|u_{k_n}(s)|_H \le R\}} \ell(s, X_s^{k_n}, u_{k_n}(s)) d\mathbb{P} \Big) ds + \int_t^T \Big(\int_{\{|u_{k_n}(s)|_H > R\}} \ell(s, X_s^{k_n}, u_{k_n}(s)) d\mathbb{P} \Big) ds \\
&\ge \int_t^T \Big(\int_{\{|u_{k_n}(s)|_H > R\}} |u_{k_n}(s)|^2 d\mathbb{P} \Big) ds.
\end{aligned}$$

Therefore $\mathbb{E}[\int_t^T \ell(s, X_s^{k_n}, u_{k_n}(s)) ds] \ge n - TR^2$, which contradicts (6.57). $\qquad\square$

Remark 6.27. *It is possible to get an explicit rate for the convergence on $J(t,x,\bar{u}^n)$. Indeed, computations in the proof of Theorem 6.26 give*

$$|J(t,x,\bar{u}^n) - V(t,x)| \le |V_n(t,x) - V(t,x)| \le |J_n(t,x,u_\varepsilon) - V(t,x)| \le \varepsilon + Cn^{-1}\|u_\varepsilon\|_{\mathscr{U}_2^\alpha}^2,$$

for any $n \in \mathbb{N}$ and any $\varepsilon > 0$.

7 Proofs of Section 3

7.1 Proof of Proposition 3.6

Thanks to (3.4), (3.5) and to Proposition 3.3, the pair of processes $(Y^{n,l}, Z^{n,l})$ is bounded in $\mathscr{S}^p([t,T]) \times \mathscr{M}^p([t,T];H)$, uniformly with respect to n,l. The BSDE satisfied by the pair of the difference processes $(Y^{n,l} - Y, Z^{n,l} - Z)$ is

$$\begin{cases} d(Y_\tau^{n,l} - Y_\tau) = \big(\psi(\tau, X_\tau, Y_\tau, Z_\tau) - \psi_l(\tau, X_\tau, Y_\tau^{n,l}, Z_\tau^{n,l})\big) d\tau + (Z_\tau^{n,l} - Z_\tau) dW_\tau, \quad \tau \in [t,T], \\ Y_T^{n,l} - Y_T = \phi_n(X_T) - \phi(X_T). \end{cases}$$

Writing the previous equation in the integral form, we get

$$\begin{aligned}
& Y_\tau^{n,l} - Y_\tau \\
&= \phi_n(X_T) - \phi(X_T) - \int_\tau^T (Z_s^{n,l} - Z_s) dW_s + \int_\tau^T \big(\psi(s, X_s, Y_s, Z_s) - \psi_l(s, X_s, Y_s, Z_s)\big) ds \\
&\quad + \int_\tau^T \big(\psi_l(s, X_s, Y_s, Z_s) - \psi_l(s, X_s, Y_s, Z_s^{n,l})\big) ds
\end{aligned}$$

39

$$+ \int_{\tau}^{T} \left(\psi_l(s, X_s, Y_s, Z_s^{n,l}) - \psi_l(s, X_s, Y_s^{n,l}, Z_s^{n,l}) \right) \, ds$$

$$= \phi_n(X_T) - \phi(X_T) + \int_{\tau}^{T} \left(\psi(s, X_s, Y_s, Z_s) - \psi_l(s, X_s, Y_s, Z_s) \right) \, ds$$

$$+ \int_{\tau}^{T} \frac{\psi_l(s, X_s, Y_s, Z_s) - \psi_l(s, X_s, Y_s, Z_s^{n,l})}{Z_s - Z_s^{n,l}} \left(Z_s - Z_s^{n,l} \right) \, ds - \int_{\tau}^{T} \left(Z_s^{n,l} - Z_s \right) dW_s$$

$$+ \int_{\tau}^{T} \frac{\psi_l(s, X_s, Y_s, Z_s^{n,l}) - \psi_l(s, X_s, Y_s^{n,l}, Z_s^{n,l})}{Y_s - Y_s^{n,l}} \left(Y_s - Y_s^{n,l} \right) \, ds$$

$$= \phi_n(X_T) - \phi(X_T) + \int_{\tau}^{T} \left(\psi(s, X_s, Y_s, Z_s) - \psi_l(s, X_\tau, Y_s, Z_s) \right) \, ds - \int_{\tau}^{T} \left(Z_s^{n,l} - Z_s \right) dW_s^{n,l}$$

$$+ \int_{\tau}^{T} \frac{\psi_l(s, X_s, Y_s, Z_s^{n,l}) - \psi_l(s, X_s, Y_s^{n,l}, Z_s^{n,l})}{Y_s - Y_s^{n,l}} \left(Y_s - Y_s^{n,l} \right) \, ds,$$

where in the last passage we have used that

$$W_\tau^{n,l} = W_\tau - \int_t^\tau \frac{\psi_l(s, X_s, Y_s, Z_s) - \psi_l(s, X_s, Y_s, Z_s^{n,l})}{Z_s - Z_s^{n,l}} ds, \quad \tau \geq t,$$

which, by the Girsanov Theorem (see, e.g., [6, Theorem 10.14]), is a cylindrical Wiener process under an equivalent probability measure $Q^{n,l}$. Taking the $Q^{n,l}$-conditional expectation $\mathbb{E}_{Q^{n,l}}^{\mathscr{F}_\tau}[\cdot] := \mathbb{E}_{Q^{n,l}}[\cdot | \mathscr{F}_\tau]$, we get

$$Y_\tau^{n,l} - Y_\tau = \mathbb{E}_{Q^{n,l}}^{\mathscr{F}_\tau}[\phi_n(X_T) - \phi(X_T)] + \mathbb{E}_{Q^{n,l}}^{\mathscr{F}_\tau}\left[\int_\tau^T \left(\psi(s, X_s, Y_s, Z_s) - \psi_l(s, X_s, Y_s, Z_s) \right) \, ds \right]$$

$$+ \mathbb{E}_{Q^{n,l}}^{\mathscr{F}_\tau}\left[\int_\tau^T \frac{\psi_l(s, X_s, Y_s, Z_s^{n,l}) - \psi_l(s, X_s, Y_s^{n,l}, Z_s^{n,l})}{Y_s - Y_s^{n,l}} \left(Y_s - Y_s^{n,l} \right) \, ds \right].$$

By taking the absolute value, the expectation and by applying the Gronwall lemma, we deduce that, for all $p \geq 1$, $Y^{n,l} \to Y$ in $\mathscr{S}^p([t, T])$ as $n, \ell \to \infty$, with respect to the probability measure $Q^{n,l}$ and also with respect to the original probability measure.

For what concerns the estimate of $Z - Z^{n,l}$, by applying the Itô formula to $|Y^{n,l} - Y|^2$ we get

$$\mathbb{E}\left[|Y_t^{n,l} - Y_t|^2 \right] + \mathbb{E}\left[\int_t^T |Z_\tau^{n,l} - Z_\tau|_H^2 \, d\tau \right]$$

$$= \mathbb{E}\left[|\phi_n(X_T) - \phi(X_T)|^2 \right] - 2\mathbb{E}\left[\int_t^T \left(Y_\tau^{n,l} - Y_\tau \right) \left(\psi(\tau, X_\tau, Y_\tau, Z_\tau) - \psi_l(\tau, X_\tau, Y_\tau^{n,l}, Z_\tau^{n,l}) \right) d\tau \right]$$

$$\leq \mathbb{E}\left[|\phi_n(X_T) - \phi(X_T)|^2 \right] + 2\mathbb{E}\left[\int_t^T |Y_\tau^{n,l} - Y_\tau| |\psi(\tau, X_\tau, Y_\tau, Z_\tau) - \psi_l(\tau, X_\tau, Y_\tau, Z_\tau)| \, d\tau \right]$$

$$+ 2\mathbb{E}\left[\int_t^T |Y_\tau^{n,l} - Y_\tau| |\psi_l(\tau, X_\tau, Y_\tau, Z_\tau) - \psi_l(\tau, X_\tau, Y_\tau^{n,l}, Z_\tau^{n,l})| \, d\tau \right]$$

$$\leq \mathbb{E}\left[|\phi_n(X_T) - \phi(X_T)|^2 \right] + 2\mathbb{E}\left[\sup_{\tau \in [t,T]} |Y_\tau^{n,l} - Y_\tau| \int_t^T |\psi(\tau, X_\tau, Y_\tau, Z_\tau) - \psi_l(\tau, X_\tau, Y_\tau, Z_\tau)| \, d\tau \right]$$

$$+ C\mathbb{E}\left[\sup_{\tau \in [t,T]} |Y_\tau^{n,l} - Y_\tau| \int_t^T \left(1 + |Y_\tau - Y_\tau^{n,l}| + |Z_\tau - Z_\tau^{n,l}|_H \left(1 + |Z_\tau|_H + |Z_\tau^{n,l}|_H \right) \right) d\tau \right].$$

Let us consider the right-hand side of the above inequality. Thanks to estimates (3.4), (3.5) and to the boundedness of Y and $Y^{n,l}$ in $\mathscr{S}^p([t, T]; E)$, the first two terms converge to 0 as $n, l \to \infty$ by the

dominated convergence theorem. For what concerns the third term, by applying Hôlder's inequality with p, q conjugate exponents, we get

$$\mathbb{E}\Big[\sup_{\tau \in [t,T]} |Y_\tau^{n,l} - Y_\tau| \int_t^T \big(1 + |Y_\tau - Y_\tau^{n,l}| + |Z_\tau - Z_\tau^{n,l}|_H \big(1 + |Z_\tau|_H + |Z_\tau^{n,l}|_H\big)\big) d\tau \Big]$$

$$\leq C \Big(\mathbb{E}\Big[\sup_{\tau \in [t,T]} |Y_\tau^{n,l} - Y_\tau|^p \Big] \Big)^{\frac{1}{p}} \Big(\mathbb{E}\Big[\Big(\int_t^T (1 + |Z_\tau|_H^2 + |Z_\tau^{n,l}|_H^2) d\tau \Big)^q \Big] \Big)^{\frac{1}{q}} \to 0$$

as $n, l \to \infty$. The stability result for $p = 2$ follows, and we can pass to the case of general p in a usual way. $\qquad\square$

7.2 Proof of Theorem 3.10

The differentiability properties of $(X^{t,x}, Y^{t,x}, Z^{t,x})$ and the identification formula (3.11) directly follow respectively from Proposition 3.7 and formula (3.2) in Theorem 3.2.

Let us now prove identification formula (3.12) for Z. Fix $t \in [0, T]$. By the definition of the function v, we can write

$$v\big(\tau, X_\tau^{t,x}\big) + \int_t^\tau Z_\sigma^{t,x} dW_\sigma = v(t,x) + \int_t^\tau \psi_\sigma d\sigma, \quad 0 \leq t \leq \tau \leq T, \tag{7.1}$$

where we have used the notation $\psi_\sigma := \psi(\sigma, X_\sigma^{t,x}, Y_\sigma^{t,x}, Z_\sigma^{t,x})$. Notice that towards [17] we do not have $\psi \in \mathcal{M}^2([0,T])$, but we only know that $\psi \in L^p(\Omega, L^1(0, T; \mathbb{R}))$ for any $p \geq 2$. As in [17], we define a family $\mathscr{S}$ of predictable processes with real values in the following way:

$$\mathscr{S} = \Big\{ \text{predictable processes } \eta : \text{for any } k = 0, \dots, 2^n - 1,$$

$$\eta_t 1_{\left[\frac{kT}{2^n}, \frac{(k+1)T}{2^n}\right)}(t) = \eta^k \big(W_{t_1}, \dots, W_{t_{l_k}}\big) \text{ for } 0 \leq t_1 \leq \cdots \leq t_{l_k} \leq \frac{kT}{2^n},$$

$$\eta^k \text{ bounded functions in } C^\infty(\mathbb{R}^{l_k}, \mathbb{R}) \text{ with bounded derivatives of all orders} \Big\}.$$

We will briefly write $\eta_t = \eta_t(W.)$, where by $W.$ we mean the trajectory of W up to time t.

Let us set $\xi_t := \eta_t \varsigma$ for $\varsigma \in E_0$. From now on we fix $s > t$, and $\delta > 0$, small enough such that $s - \delta > t$. We also identify H with its dual H^*, and we write ξ for ξ^*. Multiplying both sides of (7.1), with τ replaced by s, by $\int_{s-\delta}^s \xi_\sigma dW_\sigma$ and taking the expectation, we get

$$\mathbb{E}\Big[v\big(s, X_s^{t,x}\big) \int_{s-\delta}^s \xi_\sigma dW_\sigma \Big] = \mathbb{E}\Big[\int_t^s \psi_\sigma d\sigma \int_{s-\delta}^s \xi_\sigma dW_\sigma \Big] + \mathbb{E}\Big[\int_t^s Z_\sigma^{t,x} dW_\sigma \int_{s-\delta}^s \xi_\sigma dW_\sigma \Big]. \tag{7.2}$$

It is immediate that

$$\mathbb{E}\Big[\int_t^{s-\delta} \psi_\sigma d\sigma \int_{s-\delta}^s \xi_\sigma dW_\sigma \Big] = 0, \quad \mathbb{E}\Big[\int_t^s Z_\sigma^{t,x} dW_\sigma \int_{s-\delta}^s \xi_\sigma dW_\sigma \Big] = \mathbb{E}\Big[\int_{s-\delta}^s Z_\sigma^{t,x} \xi_\sigma d\sigma \Big],$$

so (7.2) simplifies in

$$\mathbb{E}\Big[v\big(s, X_s^{t,x}\big) \int_{s-\delta}^s \xi_\sigma dW_\sigma \Big] = \mathbb{E}\Big[\int_{s-\delta}^s \psi_\sigma d\sigma \int_{s-\delta}^s \xi_\sigma dW_\sigma \Big] + \mathbb{E}\Big[\int_{s-\delta}^s Z_\sigma^{t,x} \xi_\sigma d\sigma \Big].$$

By dividing both sides of the previous equality by δ and letting $\delta \to 0$, we get

$$\mathbb{E}\Big[Z_s^{t,x} \xi_s \Big] = \lim_{\delta \to 0} \frac{1}{\delta} \mathbb{E}\Big[v\big(s, X_s^{t,x}\big) \int_{s-\delta}^s \xi_\sigma dW_\sigma \Big] - \lim_{\delta \to 0} \frac{1}{\delta} \mathbb{E}\Big[\int_{s-\delta}^s \psi_\sigma d\sigma \int_{s-\delta}^s \xi_\sigma dW_\sigma \Big]. \tag{7.3}$$

We will prove that

$$\lim_{\delta \to 0} \frac{1}{\delta} \mathbb{E}\Big[v\left(s, X_s^{t,x}\right) \int_{s-\delta}^s \xi_\sigma dW_\sigma\Big] = \mathbb{E}\Big[\nabla_x\big(v(s, X_s^{t,x})\big)(-A)^{-\alpha}\xi_s\Big], \tag{7.4}$$

$$\lim_{\delta \to 0} \frac{1}{\delta} \mathbb{E}\Big[\int_{s-\delta}^s \psi_\sigma d\sigma \int_{s-\delta}^s \xi_\sigma dW_\sigma\Big] = 0. \tag{7.5}$$

If (7.4) and (7.5) hold, then, by (7.3), for every $\eta \in \mathscr{S}$, $\mathbb{E}[Z_\sigma^{t,x}\varsigma\eta_s] = \mathbb{E}[\nabla v(\sigma, X_\sigma^{t,x})(-A)^{-\alpha}\varsigma\eta_s]$ for almost every $\sigma \in [t, T]$. By the arbitrariness of η, we would have, for almost every $\sigma \in [t, T]$, $Z_\sigma^{t,x}\varsigma = \nabla_x\big(v(\sigma, X_\sigma^{t,x})\big)(-A)^{-\alpha}\varsigma$, $\mathbb{P}$-a.s. for all $\varsigma \in E_0$, and the formula (3.12) would follow.

Let us thus show that (7.4) and (7.5) hold true. We start by proving (7.4). One proceeds as in [1], following also [17]. In particular, for $0 \le t \le \sigma \le T$, we define

$$W_\sigma^\varepsilon = W_\sigma - \varepsilon \int_t^\sigma \xi_r\left(W_\cdot^\varepsilon\right) dr, \tag{7.6}$$

where $\xi_r(W_\cdot^\varepsilon)$ depends on the trajectories of $W_\cdot^\varepsilon$ up to time r, and the dependence is given by the definition of η. The process $(W_\sigma^\varepsilon)_\sigma$ is defined as the solution of (7.6), which is not considered as a stochastic differential equation, as specified in [1, p. 476]. Equation (7.6) can be solved step by step in each interval

$$\left[\frac{kT}{2^n}, \frac{(k+1)T}{2^n}\right), \quad k = 0, \ldots, 2^n - 1.$$

$(W_\sigma^\varepsilon)_\sigma$ is well defined for every $0 \le \sigma \le T$, see [17] for more details. Moreover, W_σ^ε is a function of the trajectories of W up to time σ, that is, $W_\sigma^\varepsilon = W_\sigma^\varepsilon(W_\cdot)$, and we can write

$$W_\sigma^\varepsilon = W_\sigma - \varepsilon \int_t^\sigma \xi_r\left(W_\cdot^\varepsilon\left(W_\cdot\right)\right) dr, \qquad 0 \le t \le \sigma \le T.$$

Now we define a probability measure Q_ε such that

$$\frac{dQ_\varepsilon}{d\mathbb{P}} = \exp\Big(\varepsilon \int_t^T \xi_\sigma\left(W_\cdot^\varepsilon\left(W_\cdot\right)\right) dW_\sigma - \frac{\varepsilon^2}{2} \int_t^T \left|\xi_\sigma\left(W_\cdot^\varepsilon\left(W_\cdot\right)\right)\right|^2 d\sigma\Big).$$

By the Girsanov Theorem, under Q_ε, $W_\sigma^\varepsilon = W_\sigma - \varepsilon \int_t^\sigma \xi_r(W_\cdot^\varepsilon(W_\cdot))dr$ is a cylindrical Wiener process in H. By this construction of $(W_\sigma^\varepsilon)_\sigma$, it is also clear that for every $0 \le \sigma \le T$, W_σ^ε is pathwise differentiable with respect to ε and $\frac{d}{d\varepsilon}_{|\varepsilon=0}W_\sigma^\varepsilon = -\int_t^\sigma \xi_r(W_\cdot)dr$, see also [1, p. 476].

By (7.1), the random varaible $v(s, X_s^{t,x})$ is square integrable and

$$\mathbb{E}[v^2(s, X_s^{t,x})] \le c\Big\{1 + \mathbb{E}\Big[\Big(\int_t^s \xi_\sigma dW_\sigma\Big)^2\Big] + \mathbb{E}\Big[\Big|\int_t^s \psi_\sigma d\sigma\Big|^2\Big]\Big\}$$

$$\le c\Big\{1 + \mathbb{E}\Big[\int_t^s |\xi_\sigma|_H^2 d\sigma\Big] + \mathbb{E}\Big[\Big|\int_t^s \psi_\sigma d\sigma\Big|^2\Big]\Big\} < \infty.$$

Therefore, by the Cauchy–Schwarz inequality the expectation of $v(s, X_s^{t,x})\int_{s-\delta}^s \xi_\sigma dW_\sigma$ is well defined. We claim that

$$\mathbb{E}\Big[v\left(s, X_s^{t,x}\right) \int_{s-\delta}^s \xi_\sigma dW_\sigma\Big] = \frac{d}{d\varepsilon}_{|\varepsilon=0} \mathbb{E}_{Q_\varepsilon}\left[v\left(s, X_s^{t,x}\right)\right]. \tag{7.7}$$

As a matter of fact,

$$\frac{d}{d\varepsilon}_{|\varepsilon=0} \mathbb{E}_{Q_\varepsilon}\left[v\left(s, X_s^{t,x}\right)\right] = \frac{d}{d\varepsilon}_{|\varepsilon=0} \mathbb{E}\left[v\left(s, X_s^{t,x}\right) \exp\left(\varepsilon \int_{s-\delta}^s \xi_\sigma dW_\sigma - \frac{\varepsilon^2}{2}\int_{s-\delta}^s \|\xi_\sigma\|_H^2 d\sigma\right)\right]$$

$$= \lim_{\varepsilon\to 0} \mathbb{E}\left[v\left(s, X_s^{t,x}\right) \frac{1}{\varepsilon}\left\{\exp\left(\varepsilon \int_{s-\delta}^s \xi_\sigma dW_\sigma - \frac{\varepsilon^2}{2}\int_{s-\delta}^s |\xi_\sigma|_H^2 d\sigma\right) - 1\right\}\right] = \mathbb{E}\left[v\left(s, X_s^{t,x}\right)\int_{s-\delta}^s \xi_\sigma dW_\sigma\right],$$

where in the last passage we have used the dominated convergence theorem being ξ bounded.
Now notice that, in $(\Omega, \mathscr{F}, Q_\varepsilon)$, $X^{t,x}$ is a mild solution to the equation

$$dX_\tau^{t,x} = AX_\tau^{t,x} d\tau + F\left(X_\tau^{t,x}\right) d\tau + (-A)^{-\alpha}\varepsilon\xi_\tau\left(W_\cdot^\varepsilon\right) d\tau + (-A)^{-\alpha} dW_\tau^\varepsilon, \quad \tau \in [s-\delta, T].$$

On the other hand, in $(\Omega, \mathscr{F}, \mathbb{P})$, we consider the process X^ε which is a mild solution to the equation

$$\begin{cases} dX_\tau^\varepsilon = AX_\tau^\varepsilon d\tau + F\left(X_\tau^\varepsilon\right) d\tau + (-A)^{-\alpha}\varepsilon\xi_\tau\left(W_\cdot\right) d\tau + (-A)^{-\alpha} dW_\tau, \quad \tau \in [s-\delta, T], \\ X_{s-\delta}^\varepsilon = X_{s-\delta}^{t,x}. \end{cases}$$

Then the process $X^{t,x}$ under Q_ε and the process X^ε under $\mathbb{P}$ have the same law, so (7.7) yields

$$\mathbb{E}\left[v\left(s, X_s^{t,x}\right)\int_{s-\delta}^s \xi_\sigma dW_\sigma\right] = \frac{d}{d\varepsilon}_{|\varepsilon=0} \mathbb{E}\left[v\left(s, X_s^\varepsilon\right)\right]. \tag{7.8}$$

Let us set $\dot{X}_\tau := \frac{d}{d\varepsilon}_{|\varepsilon=0} X_\tau^\varepsilon$ and $\Delta^\varepsilon X_\tau = \frac{X_\tau^\varepsilon - X_\tau}{\varepsilon}$, $\mathbb{P}$-a.s. for any $\tau \in [s-\delta, T]$. Arguing as in [17], one can prove that

$$\lim_{\varepsilon\to 0} |\Delta^\varepsilon X_\tau - \dot{X}_\tau|_E = 0, \quad \dot{X}_\tau = \int_{s-\delta}^\tau \nabla_x X_\tau^{\sigma, X_\sigma^{t,x}}(-A)^{-\alpha}\xi_\sigma d\sigma, \quad \tau \in [s-\delta, T], \quad \mathbb{P}\text{-a.s.} \tag{7.9}$$

Formula (7.9) in turn allows to show that

$$\frac{d}{d\varepsilon}_{|\varepsilon=0} \mathbb{E}\left[v\left(s, X_s^\varepsilon\right)\right] = \mathbb{E}\left[\nabla_x v\left(s, X_s^{t,x}\right)\dot{X}_s\right] = \mathbb{E}\left[\nabla_x v\left(s, X_s^{t,x}\right)\int_{s-\delta}^s \nabla_x X_\sigma^{\sigma, X_\sigma^{t,x}}(-A)^{-\alpha}\xi_\sigma d\sigma\right],$$

so that formula (7.8) gives

$$\mathbb{E}\left[v\left(s, X_s^{t,x}\right)\int_{s-\delta}^s \xi_\sigma dW_\sigma\right] = \mathbb{E}\left[\nabla_x v\left(s, X_s^{t,x}\right)\int_{s-\delta}^s \nabla_x X_s^{\sigma, X_\sigma^{t,x}}(-A)^{-\alpha}\xi_\sigma d\sigma\right]. \tag{7.10}$$

By (7.10) we have

$$\lim_{\delta\to 0} \frac{1}{\delta}\mathbb{E}\left[v\left(s, X_s^{t,x}\right)\int_{s-\delta}^s \xi_\sigma dW_\sigma\right] = \lim_{\delta\to 0} \frac{1}{\delta}\mathbb{E}\left[\nabla_x v\left(s, X_s^{t,x}\right)\int_{s-\delta}^s \nabla_x X_s^{\sigma, X_\sigma^{t,x}}(-A)^{-\alpha}\xi_\sigma d\sigma\right]$$

$$= \mathbb{E}\left[\nabla_x v\left(s, X_s^{t,x}\right)\nabla_x X_s^{s, X_s^{t,x}}(-A)^{-\alpha}\xi_s\right] = \mathbb{E}\left[\nabla_x\left(v(s, X_s^{t,x})\right)(-A)^{-\alpha}\xi_s\right]$$

so (7.4) is proved.
It remains to prove (7.5). Recalling identifications (3.2)-(3.3), we have

$$\frac{1}{\delta}\mathbb{E}\left[\int_{s-\delta}^s \psi(\sigma, X_\sigma^{t,x}, Y_\sigma^{t,x}, Z_\sigma^{t,x})d\sigma \int_{s-\delta}^s \xi_\sigma dW_\sigma\right]$$

$$= \frac{1}{\delta}\mathbb{E}\left[\int_{s-\delta}^s \psi(\sigma, X_\sigma, v(\sigma, X_\sigma), u(\sigma, X_\sigma))d\sigma \int_{s-\delta}^s \xi_\sigma dW_\sigma\right]$$

$$= \frac{1}{\delta} \frac{d}{d\varepsilon}_{|\varepsilon=0} \mathbb{E}\Big[\int_{s-\delta}^{s} \psi(\sigma, X_\sigma^\varepsilon, v(\sigma, X_\sigma^\varepsilon), u(\sigma, X_\sigma^\varepsilon)) d\sigma \Big], \tag{7.11}$$

which is the analogous of formula (7.8) with $f_\delta := \int_{s-\delta}^{\cdot} \psi_\sigma d\sigma$ in place of v. Now we notice that

$$\frac{1}{\delta} \frac{d}{d\varepsilon}_{|\varepsilon=0} \psi(\sigma, X_\sigma^\varepsilon, v(\sigma, X_\sigma^\varepsilon), u(\sigma, X_\sigma^\varepsilon)) = \nabla_x \psi_\sigma \frac{\dot{X}_\sigma}{\delta} + \nabla_y \psi_\sigma \frac{\dot{Y}_\sigma}{\delta} + \nabla_z \psi_\sigma \frac{\dot{Z}_\sigma}{\delta}, \quad \sigma \in [s-\delta, T], \tag{7.12}$$

where we have used the notation

$$(\dot{Y}, \dot{Z}) := \Big(\frac{d}{d\varepsilon}_{|\varepsilon=0} Y, \frac{d}{d\varepsilon}_{|\varepsilon=0} Z \Big) = (\nabla_x v(\sigma, X_\sigma^\varepsilon) \dot{X}, \nabla_x u(\sigma, X_\sigma^\varepsilon) \dot{X}).$$

By (7.9) and (2.5), we have

$$\frac{|\dot{X}_\tau|_E}{\delta} \le \frac{1}{\delta} \Big| \int_{s-\delta}^{\tau} \nabla_x X_\tau^{\sigma, X_\sigma^{t,x}} (-A)^{-\alpha} \xi_\sigma d\sigma \Big|_E \le C, \quad \tau \in [s-\delta, T]. \tag{7.13}$$

On the other hand, the pair of processes $(\dot{Y}, \dot{Z})$ is solution to the FBSDE

$$\begin{cases}
-d\dot{Y}_\tau &= \nabla_x \psi(\tau, X_\tau^{t,x}, Y_\tau^{t,x}, Z_\tau^{t,x}) \dot{X}_\tau d\tau + \nabla_y \psi(\tau, X_\tau^{t,x}, Y_\tau^{t,x}, Z_\tau^{t,x}) \dot{Y}_\tau d\tau \\
&\quad + \nabla_z \psi(\tau, X_\tau^{t,x}, Y_\tau^{t,x}, Z_\tau^{t,x}) \dot{Z}_\tau d\tau - \dot{Z}_\tau dW_\tau, \quad \tau \in [s-\delta, T], \\
\dot{Y}_T &= \nabla_x \Phi(X_T^{t,x}) \dot{X}_T, \\
d\dot{X}_\tau &= A \dot{X}_\tau d\tau + \nabla F(X_\tau^{t,x}) \dot{X}_\tau d\tau, \quad \tau \in [s-\delta, T], \\
\dot{X}_{s-\delta} &= 0.
\end{cases} \tag{7.14}$$

Moreover, taking into account (7.13) and the linearity of the BSDE (7.14), we get that the pair $(\dot{Y}, \dot{Z})$ satisfies the estimates

$$\sup_{\tau \in [s-\delta, T]} \frac{|\dot{Y}_\tau|}{\delta} \le C, \quad \mathbb{P}\text{-a.s.}, \tag{7.15}$$

$$\frac{1}{\delta} \mathbb{E}\Big[\int_{s-\delta}^{T} |\dot{Z}_\tau|_H^2 d\tau \Big] \le C. \tag{7.16}$$

By Hypothesis 3.1,

$$|\nabla_x \psi_\sigma|_{E^*} \le C, \quad |\nabla_y \psi_\sigma| \le C, \quad |\nabla_z \psi_\sigma|_H \le C(1 + |z|_H), \quad \sigma \in [s-\delta, T]. \tag{7.17}$$

Therefore, collecting (7.12)-(7.13), (7.15)-(7.16) and (7.17), (7.11) gives

$$\frac{1}{\delta} \mathbb{E}\Big[\int_{s-\delta}^{s} \psi(\sigma, X_\sigma^{t,x}, Y_\sigma^{t,x}, Z_\sigma^{t,x}) d\sigma \int_{s-\delta}^{s} \xi_\sigma dW_\sigma \Big]$$

$$= \frac{1}{\delta} \mathbb{E}\Big[\int_{s-\delta}^{s} \frac{d}{d\varepsilon} \psi(\sigma, X_\sigma^\varepsilon, v(\sigma, X_\sigma^\varepsilon), u(\sigma, X_\sigma^\varepsilon)) d\sigma \Big]$$

$$= \mathbb{E}\Big[\int_{s-\delta}^{s} \Big(\nabla_x \psi_\sigma \frac{\dot{X}_\sigma}{\delta} + \nabla_y \psi_\sigma \frac{\dot{Y}_\sigma}{\delta} + \nabla_z \psi_\sigma \frac{\dot{Z}_\sigma}{\delta} \Big) d\sigma \Big]$$

44

$$\leq C\delta + C\mathbb{E}\Big[\int_{s-\delta}^s |\nabla_y \psi_\sigma| \frac{|\dot{Y}_\sigma|}{\delta} d\sigma\Big] + C\mathbb{E}\Big[\int_{s-\delta}^s |\nabla_z \psi_\sigma|_H \frac{|\dot{Z}_\sigma|_H}{\delta} d\sigma\Big] \leq C\delta + \mathbb{E}\Big[\int_{s-\delta}^s |Z_\sigma|_H^2 d\sigma\Big]$$

which goes to zero as δ goes to zero. This shows that (7.5) holds true and concludes the proof. $\square$

7.3 Proof of Proposition 3.12

In the following C_T will denote a positive constant which may depend on $T, L_\psi, K_\psi, K_\phi$ but not on $\nabla_x \phi$, and that may vary from line to line. We fix $(t, x) \in [0\,T] \times E$.

We start by proving estimate (3.14). We first take ϕ and ψ differentiable. By Proposition 3.7, the triple of processes $(X^{t,x}, Y^{t,x}, Z^{t,x})$ is Gâteaux differentiable as a map from E with values in $\mathscr{S}^2((t,T]; E) \times \mathscr{S}^2([t,T]) \times \mathscr{M}^2([t,T]; H)$, and for any $h \in E_0$, the triple of processes $(\nabla_x X^{t,x}, \nabla_x Y^{t,x} h, \nabla_x Z^{t,x} h)$ is solution to (3.9), and satisfies estimate (3.10).

Let us now introduce the process

$$W_\tau^{\mathbb{Q}} := W_\tau - \int_t^\tau \nabla_z \psi(s, X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x}) ds, \quad \tau \in [t, T],$$

where $\mathbb{Q}$ is the probability measure such that $W^{\mathbb{Q}}$ is a Brownian motion in $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geq 0}, \mathbb{Q})$. Let us fix $h \in E_0$. Arguing as in [19, Proposition 3.6] it follows that

$$F_\tau^{t,x} h := e^{\int_t^\tau \nabla_y \psi(s, X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x}) ds} \nabla_x Y_\tau^{t,x} h$$
$$+ \int_t^\tau e^{\int_t^s \nabla_y \psi(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr} \nabla_x \psi(s, X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x}) \nabla_x X_s^{t,x} h ds, \quad \tau \in [t, T].$$

Therefore, $(|F_\tau^{t,x} h|^2)_{\tau \in [t,T]}$ is a $\mathbb{Q}$-submartingale, which implies, thanks to identification formula (3.12), that

$$\mathbb{E}^{\mathbb{Q}}\Big[\int_t^\tau |F_s^{t,x} h|^2 ds\Big] \geq (\tau - t)|F_t^{t,x} h|^2 = (\tau - t)|Z_t^{t,x}(-A)^\alpha h|, \quad \tau \in [t, T]. \tag{7.18}$$

Further, since ψ is differentiable and Lipschitz continuous with respect to x and y, and $\nabla_x X^{t,x}$ is bounded (see (2.6)), we deduce that

$$|F_\tau^{t,x} h|^2 \leq C_T \left(|\nabla_x Y_\tau^{t,x} h|^2 + |h|_H^2\right), \quad \tau \in [t, T], \ \mathbb{P}\text{-a.s.} \tag{7.19}$$

It remains to estimate $|\nabla_x Y_\tau^{t,x} h|$. To this aim, we recall the well-known estimate

$$\mathbb{E}^{\mathbb{Q}}\Big[\Big(\int_t^\tau |Z_s^{t,x}|_H^2 ds\Big)^{p/2}\Big] \leq C\|\Phi\|_\infty^p, \quad \tau \in [t, T], \tag{7.20}$$

for some $C > 0$ and any $p < +\infty$. Formulas (7.20), (3.12) and (2.15) give

$$\mathbb{E}^{\mathbb{Q}}\Big[\int_t^\tau |\nabla Y_s^{t,x} h|^2 ds\Big] \leq \mathbb{E}^{\mathbb{Q}}\Big[\int_t^\tau |Z_s^{t,x}|_H^2 |(-A)^\alpha \nabla X_s^{t,x} h|_H^2 ds\Big] \leq c\|\Phi\|_\infty^2 |(-A)^\alpha h|_H^2, \quad \tau \in [t, T],$$

which, together with (7.18) and (7.19), allows us to conclude that

$$|Z_t^{t,x}(-A)^\alpha h|^2 \leq \frac{C_T}{T-t}|(-A)^\alpha h|_H^2, \quad h \in E_0.$$

Let now fix $h \in H$. We notice that in this case we can write $h = (-A)^\alpha(-A)^{-\alpha} h$. Therefore,

$$|Z_t^{t,x} h|^2 = |Z_t^{t,x}(-A)^\alpha(-A)^{-\alpha} h|^2 \leq \frac{C_T}{T-t}|(-A)^\alpha(-A)^{-\alpha} h|_H^2 \leq \frac{C_T}{T-t}|h|_H^2, \quad h \in H,$$

which provides (3.14) in the case of ψ and ϕ differentiable.

Finally, the case ψ and ϕ non differentiable can be obtained by approximating ψ and ϕ with ψ_n and ϕ_n in (3.6) and (3.7), respectively. For the proof we refer to [19, Proposition 3.6].

Let us now prove estimate (3.15). Again, at first we prove the result when ψ and ϕ are differentiable and then we generalize it by approximation. Let us fix $h \in E_0$. For any $t < \eta < \tau \leq T$, the submartingale property of $(|F_s^{t,x}|_H^2)_{s \in [t,T]}$ gives

$$\mathbb{E}^{\mathbb{Q}}\Big[\int_\eta^\tau |F_s^{t,x}h|^2 ds\Big] = \int_\eta^\tau \mathbb{E}^{\mathbb{Q}}[|F_s^{t,x}h|^2]ds \geq \int_\eta^\tau \mathbb{E}^{\mathbb{Q}}[|F_\eta^{t,x}h|^2]ds = (\tau - \eta)\mathbb{E}^{\mathbb{Q}}[|F_\eta^{t,x}h|^2]. \qquad (7.21)$$

Moreover, for any $\tau \in (t, T]$ we split

$$\mathbb{E}^{\mathbb{Q}}\Big[\int_t^\tau |F_s^{t,x}h|^2 ds\Big] = \mathbb{E}^{\mathbb{Q}}\Big[\int_t^{(t+\tau)/2} |F_s^{t,x}h|^2 ds\Big] + \mathbb{E}^{\mathbb{Q}}\Big[\int_{(t+\tau)/2}^\tau |F_s^{t,x}h|^2 ds\Big] =: I_1 + I_2. \qquad (7.22)$$

Let us evaluate separately I_1 and I_2. Concerning I_1, identification formula (3.12), (7.20) and (2.14) give

$$\mathbb{E}^{\mathbb{Q}}\Big[\int_\eta^\tau |\nabla Y_s^{t,x}h|^2 ds\Big]$$
$$\leq \mathbb{E}^{\mathbb{Q}}\Big[\int_\eta^\tau |Z_s^{t,x}|_H^2 |(-A)^\alpha \nabla X_s^{t,x}h|_H^2 ds\Big]$$
$$+ C_T(\tau - t)^{-2\alpha}\Big(|x|_E^{2m+1} + \mathbb{E}^{\mathbb{Q}}\Big[\sup_{\tau \in [t,T]} |w^A(\tau)|_E^{2m+1}\Big]\Big)^2 |h|_H^2 \mathbb{E}^{\mathbb{Q}}\Big[\int_\eta^\tau |Z_s^{t,x}|_H^2 ds\Big]$$
$$\leq C_T \|\Phi\|_\infty^2 |h|_H^2 (\eta - t)^{-2\alpha}. \qquad (7.23)$$

Hence, from (7.19), (7.21) and (7.23) it follows that

$$\mathbb{E}^{\mathbb{Q}}[|F_\eta^{t,x}h|_H^2] \leq C_T \|\Phi\|_\infty^2 |h|_H^2 \Big(\frac{(\eta - t)^{-2\alpha}}{(\tau - \eta)} + 1\Big), \quad t < \eta < \tau \leq T. \qquad (7.24)$$

By applying Fubini's theorem and (7.24), we infer that

$$I_1 \leq C_T \|\Phi\|_\infty^2 |h|_H^2 \Big[(\tau - t)^{-1} \int_t^{(t+\tau)/2} (s - t)^{-2\alpha} ds + (\tau - t)^{-1} \int_t^{(t+\tau)/2} ds\Big]$$
$$= C_T |h|_H^2 [(\tau - t)^{-2\alpha} + (\tau - t)] \leq C_T |h|_H^2 (\tau - t)^{-2\alpha}, \qquad t < \eta < \tau \leq T. \qquad (7.25)$$

As far as I_2 is concerned, we take advantage from (7.19) and (7.23). Then, for $t < \eta < \tau \leq T$ we get

$$I_2 \leq C_T |h|_H^2 [(\tau - t) + (\tau - t)^{-2\alpha}] = C_T |h|_H^2 [(\tau - t)^{-2\alpha} + (\tau - t)] \leq C_T |h|_H^2 (\tau - t)^{-2\alpha}. \qquad (7.26)$$

Thus collecting (7.21), (7.25) and (7.26), we have

$$(\tau - \eta)\mathbb{E}^{\mathbb{Q}}[|F_\eta^{t,x}h|^2] \leq \mathbb{E}^{\mathbb{Q}} \int_t^\tau |F_s^{t,x}h|^2 ds \leq C_T |h|_H^2 (\tau - t)^{-2\alpha}, \quad t < \eta < \tau \leq T,$$

so that

$$\mathbb{E}^{\mathbb{Q}}[|F_\eta^{t,x}h|^2] \leq C_T \frac{1}{(\tau - \eta)(\tau - t)^{2\alpha}} |h|_H^2, \quad h \in E_0, \ t \leq \eta < \tau \leq T. \qquad (7.27)$$

Let us now fix $h \in H$, and let us consider a sequence $(h_n) \subset E_0$ such that $h_n \to h$ as $n \to +\infty$ in H. Taking (7.27) with h replaced by h_n and letting $n \to +\infty$, it follows that

$$\mathbb{E}^{\mathbb{Q}}[|F_\eta^{t,x}h|^2] \leq \frac{C_T}{(\tau - \eta)(\tau - t)^{2\alpha}} |h|_H^2, \quad h \in H, \ t \leq \eta < \tau \leq T. \qquad (7.28)$$

Inequality (3.15) follows from (7.28) by taking $\tau = T$ and $\eta = t$. $\qquad \square$

References

[1] J. M. Bismut. Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions. *Z. Wahrsch. Verw. Gebiete*, 56:469–505, 1981.

[2] P. Briand and F. Confortola. BSDEs with stochastic lipschitz condition and quadratic PDE s in Hilbert spaces. *Stochastic Process. Appl.*, 118(5):818–838, 2008.

[3] S. Cerrai. *Second order PDE's in finite and infinite dimension, A probabilistic approach*, volume 1762 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001.

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

[5] G. Da Prato and J. Zabczyk. *Second order partial differential equations in Hilbert spaces.*, volume 293 of *London Mathematical Society Note Series*. Cambridge University Press, Cambridge, 2002.

[6] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.

[7] S. El Karoui, N. Peng and M. C. Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.

[8] G. Fabbri, F. Gozzi, and A. Święch. *Stochastic optimal control in infinite dimensions: Dynamic programming and HJB equations*, with Chapter 6 by M. Fuhrman and G. Tessitore. Springer, 2017.

[9] M. Fuhrman, Y. Hu, and G. Tessitore. On a class of stochastic optimal control problems related to BSDEs with quadratic growth. *SIAM Journal on Control and Optimization.*, 45(4):1279–1296, 2006.

[10] M. Fuhrman and G. Tessitore. The Bismut-Elworthy formula for backward SDEs and applications to nonlinear Kolmogorov equations and control in infinite dimensional spaces. *Stoch. Stoch. Rep.*, 74(1-2):429–464, 2002.

[11] M. Fuhrman and G. Tessitore. Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations. *Ann. Probab.*, 30(3):1397–1465, 2002.

[12] M. Fuhrman and G. Tessitore. Generalized Directional Gradients, Backward Stochastic Differential Equations and Mild Solutions of Semilinear Parabolic Equations. *Appl Math Optim.*, 51(3):279–332, 2005.

[13] F. Gozzi. Regularity of solutions of second order Hamilton-Jacobi equations in Hilbert spaces and applications to a control problem. 20, 1995.

[14] F. Gozzi. Global regular solutions of second order Hamilton-Jacobi equations in Hilbert spaces with locally Lipschitz nonlinearities. 198, 1996.

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

[16] A. Lunardi. *Analytic semigroups and optimal regularity in parabolic problems*. Modern Birkhäuser Classics. Birkhäuser/Springer Basel AG, Basel, 1995. [2013 reprint of the 1995 original] [MR1329547].

[17] F. Masiero. Stochastic optimal control problems and parabolic equations in Banach spaces. *SIAM J. Control Optim.*, 47(1):251–300, 2008.

[18] F. Masiero. Hamilton Jacobi Bellman equations in infinite dimensions with quadratic and superquadratic Hamiltonian. *Discrete and Continuous Dynamical Systems*, 32(1A):223–263, 2012.

[19] F. Masiero. A Bismut-Elworthy formula for quadratic BSDEs. *Stochastic Process. Appl.*, 125(5):1945–1979, 2015.

[20] F. Masiero and A. Richou. HJB equations in infinite dimensions with locally lipschitz hamiltonian and unbounded terminal condition. *J. Differential Equations*, 257(6):1989–2034, 2014.

[21] É. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, volume 176 of *Lect. Notes Control Inf. Sci.*, pages 200–217. Springer, Berlin, 1992.

[22] K. R. Parthasarathy. *Probability measures on metric spaces*, volume 3 of *Probability and Mathematical Statistics*. Academic Press, Inc., New York-London, 1967.

[23] S. Peszat and J. Zabczyk. Strong Feller property and irreducibility for diffusions on Hilbert spaces. *Ann. Probab.*, 23(1):157–172, 1995.