INFINITE HORIZON BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS AND ELLIPTIC EQUATIONS IN HILBERT SPACES

BY MARCO FUHRMAN[1] AND GIANMARIO TESSITORE[2]

Politecnico di Milano and Università di Parma

Solutions of semilinear elliptic differential equations in infinite-dimensional spaces are obtained by means of forward and backward infinite-dimensional stochastic evolution equations. The backward equation is considered on an infinite time horizon and a suitable growth condition replaces the final condition. Elliptic equations are intended in a mild sense, suitable also for applications to optimal control. We finally notice that, due to the lack of smoothing properties, the elliptic partial differential equation considered here could not be treated by analytic methods.

1. Introduction. In this article we study a class of semilinear partial differential equations on a Hilbert space. We adopt a probabilistic approach, generalizing the theory started with the article by Pardoux and Peng [24] to an infinite-dimensional framework. We continue our previous works [12, 13], where the case of an equation of parabolic type was treated.

Our starting point is a stochastic evolution equation of the form

$$dX_\tau = AX_\tau \, d\tau + F(X_\tau) \, d\tau + G(X_\tau) \, dW_\tau, \qquad \tau \geq t,$$

(1.1)

$$X_t = x,$$

for a process X in a Hilbert space H, where $t \geq 0$, $x \in H$, W is a cylindrical Wiener process in another Hilbert space Ξ, A is the generator of a strongly continuous semigroup of bounded linear operators $(e^{tA})_{t \geq 0}$ in H, and F and G are functions with values in H and $L(\Xi, H)$, respectively, satisfying appropriate Lipschitz conditions. Under suitable assumptions, a unique solution $\{X(\tau, t, x), \tau \geq t\}$ exists and defines a Markov process with transition function $(P_t)_{t \geq 0}$ acting on measurable functions $\phi : H \to \mathbb{R}$ (satisfying suitable growth conditions) according to the formula

$$P_{\tau - t}[\phi](x) = \mathbb{E}\phi(X(\tau, t, x)), \qquad x \in H, \tau \geq t \geq 0.$$

The generator $\mathcal{L}$ corresponding to (P_t) is, at least formally, the operator

$$\mathcal{L}\phi(x) = \tfrac{1}{2}\operatorname{Trace}\bigl(G(x)G(x)^*\nabla^2\phi(x)\bigr) + \langle Ax, \nabla\phi(x)\rangle + \langle F(x), \nabla\phi(x)\rangle,$$

where $\nabla\phi(x) \in H^*$ denotes the Gâteaux derivative at point $x \in H$ and $\nabla^2\phi$ is the second Gâteaux derivative, identified with an element of $L(H)$. Notice that the above formula is, a priori, meaningful only if ϕ is sufficiently regular. In general, the characterization of the domain of $\mathcal{L}$ is difficult (refer to [7], [8] and [31] for a detailed exposition of these facts and related matters).

In [12] we gave an "infinite-dimensional" generalization of the results on nonlinear parabolic partial differential equations (PDEs) contained, for instance, in [23, 24, 27]. Namely we considered the nonlinear version of the Kolmogorov equation for X,

$$\frac{\partial u(t, x)}{\partial t} + \mathcal{L}u(t, x) = \psi\bigl(x, u(t, x), \nabla u(t, x)G(x)\bigr),$$

(1.2)
$$t \in [0, T], x \in H,$$

$$u(T, x) = \phi(x),$$

where $T > 0$ is fixed and $\psi : H \times \mathbb{R} \times \Xi^* \to \mathbb{R}$ and $\phi : H \to \mathbb{R}$ are given functions. Then we associate to (1.2) the backward stochastic evolution equation

$$dY_\tau = Z_\tau\, dW_\tau + \psi(X_\tau, Y_\tau, Z_\tau)\, d\tau, \qquad \tau \in [t, T],$$

(1.3)
$$Y_T = \phi(X_T),$$

where X is the solution of (1.1). Under suitable assumptions on ψ and ϕ, there exists a unique adapted process (Y, Z) in $\mathbb{R} \times \Xi^*$, a solution of (1.3). The processes X, Y, Z depend on the values of x and t, occurring as initial conditions in (1.1): We may denote them by $X(\tau, t, x)$, $Y(\tau, t, x)$, $Z(\tau, t, x)$, $\tau \in [t, T]$. Finally it turns out that, if we define $u(t, x) = Y(t, t, x)$, then the function u is the unique solution of (1.2), in a suitable mild sense.

In this article, instead of (1.2), we are concerned with the nonlinear elliptic equation

(1.4) $$\mathcal{L}u(x) = \lambda u(x) + \psi\bigl(x, u(x), \nabla u(x)G(x)\bigr), \qquad x \in H,$$

where $\psi : H \times \mathbb{R} \times \Xi^* \to \mathbb{R}$ is as before and $\lambda \in \mathbb{R}$. In this article we call (1.4) the nonlinear stationary Kolmogorov equation. Notice the occurrence of G in the nonlinear term: This does not imply any loss of generality in the nondegenerate case, that is, when G is boundedly invertible, whereas it involves a genuine restriction in the general case. Equations of the type of (1.4) have been studied by backward stochastic differential equations (BSDEs) techniques in several finite-dimensional situations (see [2, 10, 23, 25, 27]). However, none of the concepts of solution to (1.4) used in these articles seems suitable for immediate extension to the infinite-dimensional case. More precisely, to obtain classical solutions as in [2, 27],

that is, functions which are twice differentiable, we would be forced to impose heavy assumptions on the nonlinearity ψ as well as trace conditions on second derivatives. On the contrary, viscosity solutions can be obtained under much weaker assumptions on the coefficients than those we assume here (see [10, 23]). The main drawback is that, in comparison to the finite-dimensional case, very few uniqueness results are available for viscosity solutions and all of them, obtained by analytic techniques, impose strong assumptions on operator G such as nondegeneracy and finite trace conditions (see [14, 15, 18, 28, 29]). Moreover, in view of applications to optimal control theory, it is important to show the existence of ∇u, since this allows us to characterize optimal control by feedback laws. Since, in general, viscosity solutions are not differentiable, this characterization is not immediately available. However, we have to mention that the analytic approach and viscosity solutions allow us, in certain cases, to treat fully nonlinear equations (see [18, 28, 29] and references therein), while backward stochastic equations techniques are, in any case, limited to semilinear PDEs.

Developing the idea introduced in [12], we consider *mild solutions* of (1.4) in the following sense: A function $u : H \to \mathbb{R}$, Gâteaux differentiable and having polynomial growth, is a mild solution of (1.4) if the equality

$$(1.5) \quad u(x) = e^{-\lambda T} P_T[u](x) - \int_0^T e^{-\lambda \tau} P_\tau \big[\psi\big(\cdot, u(\cdot), \nabla u(\cdot) G(\cdot)\big)\big](x) \, d\tau$$

holds for all $x \in H$ and $T > 0$. To motivate this definition, consider the equation $\mathcal{L}u - \lambda u = \psi$, for u, ψ elements of a Banach space and $\mathcal{L}$ generator of a strongly continuous semigroup of bounded linear operators $(P_t)_{t \geq 0}$: If λ is sufficiently large, then

$$u = -\int_0^\infty e^{-\lambda \tau} P_\tau \psi \, d\tau,$$

and, for arbitrary $T \geq 0$, by a change of variable,

$$e^{-\lambda T} P_T u = -\int_T^\infty e^{-\lambda \tau} P_\tau \psi \, d\tau = u + \int_0^T e^{-\lambda \tau} P_\tau \psi \, d\tau.$$

We notice that formula (1.5) is meaningful provided u is only once differentiable with respect to x and, of course, provided ψ, u and ∇u satisfy appropriate measurability and growth conditions. Thus, mild solutions are, in a sense, intermediate between classical and viscosity solutions. Mild solutions of a similar type have been considered by more analytic methods in various situations (see [4, 14] and references within), but never in connection with the backward equations approach.

The main result of this article is the proof of existence and uniqueness of the mild solution u of (1.4), under the mere requirement of existence and boundedness (or growth conditions) of first derivatives of ψ; compare Theorem 6.1. We wish to stress that in no way do we impose nondegeneracy assumptions on the operator G;

this can even be equal to zero. As far as we know, in all the results that exist in the literature, whenever differentiable solutions of (1.4) are obtained, smoothing properties on the semigroup (P_t) associated to the operator $\mathcal{L}$ are also required and, consequently, nondegeneracy assumptions on G are needed; see, for example, [4] and [14]. So the results of this article represent an example in which backward equations give a genuine new contribution to the study of PDEs, and lead to results that cannot be obtained by more analytical approaches. As a general fact, it seems that infinite-dimensional PDEs offer, in comparison to finite-dimensional PDEs, many more cases in which the treatment by backward stochastic differential equations is the only one available.

The existence and uniqueness result is obtained for sufficiently large values of λ: this kind of restriction is natural and common to all the literature where the BSDE approach is used (see [2, 10, 23, 25, 27]). Conditions on λ also have to be expected so as to obtain some regularity (e.g., differentiability) for the solution to degenerate nonlinear elliptic equations of the type we are considering here; see, for instance, [19].

Moreover, to separate difficulties, we assume that F is Lipschitz (although many of the estimates are expressed in terms of its dissipativity constant). We remark that the same assumption is required in [2] and [27]. On the contrary, in [10] the assumption is replaced by the weaker requirement that F is monotone and has linear growth.

We also mention that in [2, 10, 23, 25, 27] the finite-dimensional analogue of (1.4) is studied in general domains by BSDE techniques, while here we consider only equations on the whole space. We believe that this method can give new results for elliptic equations on domains of a Hilbert space, but we do not address this problem here. Even in the linear case, such extensions, considered since the works of Gross [17] and Daleckij [5], possess special features and difficulties: see [6, 30] for recent results in the parabolic case, or Chapter 8 in [9].

We finally notice that in our article the derivatives are always understood in the sense of Gâteaux. This is important in view of applications where H is a space of summable functions and nonlinear terms are Nemytskii (evaluation) operators.

Coming now to more technical aspects of the present work we point out that the main difference between the elliptic case considered here and the parabolic case treated in [12] is that, following [2, 10, 23, 27], we have replaced the final condition for the process Y that occurred in (1.3) by an infinite horizon growth condition. Namely (see Proposition 5.1) we prove that we can find $\lambda > 0$, $\beta < 0$ and $p > 2$ with λ large enough such that for all $x \in H$ there exists a unique adapted process (Y, Z) in $\mathbb{R} \times \Xi^*$ such that

(1.6)
$$dY_\tau = \lambda Y_\tau \, d\tau + \psi(X_\tau, Y_\tau, Z_\tau) \, d\tau + Z_\tau \, dW_\tau, \qquad \tau \geq 0,$$

$$\mathbb{E} \sup_{\tau \geq 0} e^{p\beta\tau} |Y_\tau|^p + \mathbb{E} \left(\int_0^\infty e^{p\beta\tau} |Z_\tau|^2 \, d\tau \right)^{p/2} < +\infty.$$

In the above formulae, X is the solution to (1.1) starting from $x \in H$ at time $t = 0$. Moreover, the constants β and λ depend on the asymptotic behavior of X as well as on the nonlinearity ψ.

To stress dependence on x, let us denote by $\{X_\tau(x), Y_\tau(x), Z_\tau(x), \tau \geq 0\}$ the solution processes. Then, following again [10] or [27], we set

$$u(x) = Y_0(x)$$

and prove that u is a mild solution to (1.4).

In particular, to prove that u is differentiable, we have to study regular dependence of $X(x)$, $Y(x)$ and $Z(x)$ on x. Notice that we can limit ourselves to first order Gâteaux derivatives. As a matter of fact, the generality of our assumptions on A and G, the lack of classical tools such as the Kolmogorov continuity theorem and the fact that we are dealing with processes on an unbounded interval make the treatment of first derivatives already very delicate.

Another key point is the formula that identifies Z:

$$\text{(1.7)} \qquad Z_\tau(x) = \nabla u\big(X_\tau(x)\big)G\big(X_\tau(x)\big).$$

In [12], we proved the corresponding result in the parabolic case by deriving the equation for the Malliavin derivative of X, Y and Z. Here we argue as follows: We first compute the joint quadratic variation of $u(X(x))$ and W in an interval $[t, T]$ to obtain

$$\int_t^T \nabla u\big(X_\sigma(x)\big)G\big(X_\sigma(x)\big)\,d\sigma.$$

This is done by an application of the Malliavin calculus (on a finite time horizon). On the other hand, the joint quadratic variation of $Y(x)$ and W is $\int_t^T Z_\sigma(x)\,d\sigma$. Finally, Markovianity of the process X yields $Y(x) = u(X(x))$. Thus we can identify the two quadratic variations and deduce (1.7). Once (1.7) has been established, it is not difficult to verify that u is a mild solution to (1.4). Notice that, in this way, we avoid studying Malliavin differentiability of Y and Z.

Uniqueness is proved by showing that if u is any mild solution of (1.4) and we set

$$Y_\tau = u\big(X_\tau(x)\big), \qquad Z_\tau = \nabla u\big(X_\tau(x)\big)G\big(X_\tau(x)\big),$$

then (Y, Z) verifies (1.6). Then this is done again by computing the joint quadratic variation of $u(X(x))$ and W.

As in the parabolic case, it turns out that mild solutions to (1.4), together with their probabilistic representation formula, are particularly suitable for applications to optimal control of infinite-dimensional nonlinear stochastic systems. In Section 7, we consider a controlled process X^u solution of

$$\text{(1.8)} \quad \begin{aligned} dX_\tau^u &= AX_\tau^u\,d\tau + F(X_\tau^u)\,d\tau + C(X_\tau^u)u_\tau\,d\tau + G(X_\tau)\,dW_\tau, \qquad \tau \geq 0, \\ X_0^u &= x \in H, \end{aligned}$$

where this time u denotes the control process, taking values in a given subset $\mathcal{U}$ of another Hilbert space U, and C is a function with values in $L(U, H)$. The aim is to choose a control process u, within a set of admissible controls, in such a way to minimize an infinite horizon cost functional of the form

$$J(x, u) = \mathbb{E} \int_0^\infty e^{-\lambda \sigma} g(X_\sigma^u, u_\sigma) \, d\sigma,$$

where g is a given real function, λ is large enough and the control problem is understood in the usual weak sense (see [11] and Section 7). There is a vast literature on such control problems in infinite dimensions: Here we report only a couple of recent references that are most closely connected with our approach and refer the reader to the bibliographies therein. Namely, in [14] and [4], the authors provide a direct differentiable (in some sense) solution of the Hamilton–Jacobi–Bellman equation for the value function $v(x)$, $x \in H$, of the control problem, which is then used to prove that the optimal control u is related to the corresponding optimal trajectory X by a feedback law involving ∇v. As we already said, such results are obtained using the smoothing properties of the semigroup (P_t) and are therefore restricted to the case in which G is independent on x and (weakly) nondegenerate. Here we are able to remove the restriction on constancy of the coefficient G and any nondegeneracy assumption on G.

On the other hand, we have to assume that the control term is of the form

$$C(X) = G(X) R(X),$$

where R is a function with values in $L(U, \Xi)$. This structural requirement ensures that the Hamilton–Jacobi–Bellman equation for the value function v is of the form (1.2), provided we set

(1.9) $\psi_0(x, p) = \inf\{g(x, u) + pu : u \in \mathcal{U}\}, \qquad x \in H, \, p \in U^*,$

and $\psi(x, z) = -\psi_0(x, zR(x))$ for $z \in \Xi^*$.

Thus we are able to prove that, letting v denote the unique mild solution of (1.4), we have $J(x, u) \geq v(x)$ and the equality holds if and only if the feedback law

$$u_\tau \in \Gamma\big(X_\tau^u, \nabla v(X_\tau^u) G(X_\tau^u) R(X_\tau^u)\big),$$

where $\Gamma(x, p)$ is the set of minimizers in (1.9), is verified by u and X^u. Thus we can characterize optimal controls by a feedback law. We refer to Theorem 7.3 for precise statements and additional results. For a result proving, in a greater generality, only existence of "quasi-optimal" controls in the finite horizon case, see [3].

The plan of this article is as follows. In Section 2, some notation is fixed. In Section 3, existence and uniqueness of a solution to (1.6) is proved. In Section 4, (1.1) is studied with an infinite time horizon; in particular, regular dependence on x, Malliavin differentiability and asymptotic bounds are proved. In Section 5, (1.1) and (1.3) are studied as a system. In Section 6, we prove our main result on existence and uniqueness of a mild solution of (1.4), and Section 7 is devoted to applications to optimal control.

2. Notation. The norm of an element x of a Banach space E will be denoted $|x|_E$ or simply $|x|$, if no confusion is possible. If F is another Banach space, $L(E, F)$ denotes the space of bounded linear operators from E to F, endowed with the usual operator norm.

The letters Ξ, H and K will always denote Hilbert spaces. Scalar product is denoted $\langle \cdot, \cdot \rangle$, with a subscript to specify the space, if necessary. All Hilbert spaces are assumed to be real and separable. The space of Hilbert–Schmidt operators from Ξ to K is $L_2(\Xi, K)$, which is endowed with the Hilbert–Schmidt norm that makes it a separable Hilbert space.

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

1. for every $\xi \in \Xi$, $\{\langle \xi, W_t \rangle, t \geq 0\}$ is a real (continuous) Wiener process;
2. for every $\xi_1, \xi_2 \in \Xi$ and $t \geq 0$, $\mathbb{E}\left(\langle \xi_1, W_t \rangle \cdot \langle \xi_2, W_t \rangle\right) = \langle \xi_1, \xi_2 \rangle_\Xi\, t$.

We let $(\mathcal{F}_t)_{t \geq 0}$ denote, except in Section 7, the natural filtration of W, augmented with the family of $\mathbb{P}$-null sets. The filtration $(\mathcal{F}_t)$ satisfies the usual conditions. All the concepts of measurability for stochastic processes (e.g., predictability, etc.) refer to this filtration. By $\mathcal{P}$ we denote the predictable σ-algebra and by $\mathcal{B}(\Lambda)$ we denote, the Borel σ-algebra of any topological space Λ.

Next we define several classes of stochastic processes with values in a Hilbert space K:

- Expression $L_{\mathcal{P}}^2(\Omega \times \mathbb{R}_+; K)$ denotes the space of equivalence classes of processes $Y \in L^2(\Omega \times \mathbb{R}_+; K)$, admitting a predictable version. $L_{\mathcal{P}}^2(\Omega \times \mathbb{R}_+; K)$ is endowed with the norm

$$|Y|_{L_{\mathcal{P}}^2(\Omega \times \mathbb{R}_+; K)}^2 = \mathbb{E} \int_0^\infty |Y_\tau|_K^2 \, d\tau.$$

- Expression $L_{\mathcal{P}}^p(\Omega; L_\beta^q(K))$, defined for $\beta \in \mathbb{R}$ and $p, q \in [1, \infty)$, denotes the space of equivalence classes of processes $\{Y_t, t \geq 0\}$, with values in K, such that the norm

$$|Y|_{L_{\mathcal{P}}^p(\Omega; L_\beta^q(K))}^p = \mathbb{E}\left(\int_0^\infty e^{q\beta\sigma} |Y_\sigma|_K^q \, d\sigma\right)^{p/q}$$

is finite, and Y admits a predictable version.
- Variable $\mathcal{K}_\beta^p$ denotes the space $L_{\mathcal{P}}^p(\Omega; L_\beta^2(K)) \times L_{\mathcal{P}}^p(\Omega; L_\beta^2(L_2(\Xi, K)))$. The norm of an element $(Y, Z) \in \mathcal{K}_\beta^p$ is $|(Y, Z)|_{\mathcal{K}_\beta^p} = |Y|_{L_{\mathcal{P}}^p(\Omega; L_\beta^2(K))} + |Z|_{L_{\mathcal{P}}^p(\Omega; L_\beta^2(L_2(\Xi, K)))}$.
- Expression $L_{\mathcal{P}}^p(\Omega; C([0, T]; K))$, defined for $T > 0$ and $p \in [1, \infty)$, denotes the space of predictable processes $\{Y_t, t \in [0, T]\}$ with continuous paths in K,

such that the norm

$$|Y|^p_{L^p_{\mathcal{P}}(\Omega; C([0,T]; K))} = \mathbb{E} \sup_{\tau \in [0,T]} |Y_\tau|^p_K$$

is finite. Elements of $L^p_{\mathcal{P}}(\Omega; C([0, T]; K))$ are identified up to indistingui-shability.

- Expression $L^q_{\mathcal{P}}(\Omega; C_\eta(K))$, defined for $\eta \in \mathbb{R}$ and $q \in [1, \infty)$, denotes the space of predictable processes $\{Y_t, t \geq 0\}$ with continuous paths in K, such that the norm

$$|Y|^q_{L^q_{\mathcal{P}}(\Omega; C_\eta(E))} = \mathbb{E} \sup_{\tau \geq 0} e^{\eta q \tau} |Y_\tau|^q_K$$

is finite. Elements of $L^q_{\mathcal{P}}(\Omega; C_\eta(K))$ are identified up to indistinguishability.

- Finally, for $\eta \in \mathbb{R}$ and $q \in [1, \infty)$, we define $\mathcal{H}^q_\eta$ as the space $L^q_{\mathcal{P}}(\Omega; L^q_\eta(K)) \cap L^q_{\mathcal{P}}(\Omega; C_\eta(K))$, endowed with the norm

$$|Y|_{\mathcal{H}^q_\eta} = |Y|_{L^q_{\mathcal{P}}(\Omega; L^q_\eta(K))} + |Y|_{L^q_{\mathcal{P}}(\Omega; C_\eta(K))}.$$

Clearly, similar definitions and notations also apply to processes with values in other Hilbert spaces, different from K.

Given a process Ψ that belongs to $L^2_{\mathcal{P}}(\Omega \times [0, T]; L_2(\Xi, K))$ for every $T > 0$, the Itô stochastic integral $\int_0^t \Psi_\sigma \, dW_\sigma$, $t \geq 0$, can be defined; it is a K-valued martingale that belongs to $L^2_{\mathcal{P}}(\Omega; C([0, T]; K))$ for every $T > 0$.

In the rest of this section we recall notations and basic facts on a class of differentiable maps acting among Banach spaces that are particularly suitable for our purposes. This class was introduced in [12], to which we refer the reader for details and properties not proved here, although similar classes of differentiable functions were already used in this context (see, e.g., [31]).

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

The last requirement is equivalent to the fact that for every $h \in X$, the map $\nabla F(\cdot)h : X \to V$ is continuous. Note that $\nabla F : X \to L(X, V)$ is not continuous, in general, if $L(X, V)$ is endowed with the norm operator topology; clearly, if this happens, then F is Fréchet differentiable on X. It can be proved that if $F \in \mathcal{G}^1(X, V)$, then $(x, h) \mapsto \nabla F(x)h$ is continuous from $X \times X$ to V. If, in addition, G is in $\mathcal{G}^1(V, Z)$, then $G(F)$ belongs to $\mathcal{G}^1(X, Z)$ and the chain rule holds: $\nabla(G(F))(x) = \nabla G(F(x))\nabla F(x)$. In addition to the ordinary chain rule, a chain rule for the Malliavin derivative operator holds: see the proof of Proposition 4.5.

Generalization of these definitions and properties to functions depending on several variables is immediate. For a function $F : X \times Y \to V$, we denote by $\nabla_x F(x, y)$ the partial Gâteaux derivative with respect to the first argument, at

point (x, y) and in the direction $h \in X$, and we say that a mapping $F : X \times Y \to V$ belongs to the class $\mathcal{G}^{1,0}(X \times Y; V)$ if it is continuous, Gâteaux differentiable with respect to x on $X \times Y$ and $\nabla_x F : X \times Y \to L(X, V)$ is strongly continuous. Then we can prove that the mapping $(x, y, h) \mapsto \nabla_x F(x, y)h$ is continuous from $X \times Y \times X$ to V, and analogues of the previously stated chain rules hold. When F depends on additional arguments, further generalizations can be given.

To study regular dependence of solution of stochastic equations on their initial data we will use the parameter depending contraction principle, which is stated in the following proposition and proved in [31], Theorems 10.1 and 10.2.

PROPOSITION 2.1 (Parameter depending contraction principle). *Let $F : X \times Y \to X$ be a continuous mapping satisfying*

$$|F(x_1, y) - F(x_2, y)| \leq \alpha |x_1 - x_2|$$

for some $\alpha \in [0, 1)$ and every $x_1, x_2 \in X$, $y \in Y$. Let $\phi(y)$ denote the unique fixed point of the mapping $F(\cdot, y) : X \to X$. Then $\phi : Y \to X$ is continuous. If, in addition, $F \in \mathcal{G}^1(X \times Y, X)$, then $\phi \in \mathcal{G}^1(Y, X)$ and

$$\nabla \phi(y) = \nabla_x F(\phi(y), y) \nabla \phi(y) + \nabla_y F(\phi(y), y), \qquad y \in Y.$$

3. The backward equation on an infinite horizon. Let $\{W_\tau, \tau \geq 0\}$ be a cylindrical Wiener process with values in a Hilbert space Ξ, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let K be another Hilbert space and let $\Psi : \Omega \times \mathbb{R}_+ \times K \times L_2(\Xi, K) \to K$ be a function, measurable with respect to $\mathcal{P} \otimes \mathcal{B}(K) \otimes \mathcal{B}(L_2(\Xi, K))$ and $\mathcal{B}(K)$. As defined in Section 2, $\mathcal{P}$ denotes the predictable σ-algebra and $\mathcal{B}(\Lambda)$ denotes the Borel σ-algebra of any topological space Λ. In this section we study the backward equation, $\mathbb{P}$-*a.s.*,

$$
\begin{aligned}
(3.1) \quad & Y_\tau - Y_T + \int_\tau^T Z_\sigma \, dW_\sigma + \lambda \int_\tau^T Y_\sigma \, d\sigma \\
& = -\int_\tau^T \Psi(\sigma, Y_\sigma, Z_\sigma) \, d\sigma + \int_\tau^T f_\sigma \, d\sigma, \qquad 0 \leq \tau \leq T < \infty,
\end{aligned}
$$

where λ is a given real parameter and $f : \Omega \times \mathbb{R}_+ \to K$ is a predictable process with integrable paths. Notice that it follows immediately from the equation that any process Y satisfying (3.1) has a continuous modification. We assume the following.

HYPOTHESIS 3.1. *There exist $\mu \in \mathbb{R}$, $p \in [2, \infty)$ and nonnegative constants L_y, L_z such that*

$$|\Psi(t, y_1, z_1) - \Psi(t, y_2, z_2)| \leq L_y |y_1 - y_2| + L_z |z_1 - z_2|,$$

$$(3.2) \quad \langle \Psi(t, y_1, z) - \Psi(t, y_2, z), y_1 - y_2 \rangle_K \geq \mu |y_1 - y_2|^2,$$

$$\mathbb{E}\left(\int_0^\infty e^{2\beta\sigma} |\Psi(\sigma, 0, 0)|^2 \, d\sigma \right)^{p/2} < \infty$$

for every $t \in [0, T]$, $y_1, y_2 \in K$ and $z, z_1, z_2 \in L_2(\Xi, K)$.

Our aim is to prove the existence and uniqueness result in Theorem 3.7. We believe that further generalizations of Theorem 3.7 can be proved, for instance, in the case in which a subdifferential term occurs in the equation or in the case in which Ψ is not necessarily Lipschitz in Y. Such extensions could be based on the finite-dimensional results in [25] and [10] and the finite horizon results in [26] and [3]. Nevertheless, we report here a complete proof in the generality required for the applications to the nonlinear elliptic Kolmogorov equation (1.4).

We start from some a priori estimates for the solutions of (3.1). The spaces $L_{\mathcal{P}}^{p}(\Omega; L_{\beta}^{2}(H))$, defined for $\beta \in \mathbb{R}$, $p \in [1, \infty)$ and for any Hilbert space H, were introduced in Section 2.

THEOREM 3.2.   *Suppose that Ψ satisfies Hypothesis 3.1 for some $p \in (2, \infty)$ and assume that for some $\beta \in \mathbb{R}$, $\lambda \in \mathbb{R}$ there exist processes*

$$Y^{1}, Y^{2}, f^{1}, f^{2} \in L_{\mathcal{P}}^{p}(\Omega; L_{\beta}^{2}(K)), \qquad Z^{1}, Z^{2} \in L_{\mathcal{P}}^{p}(\Omega; L_{\beta}^{2}(L_{2}(\Xi, K))),$$

such that, $\mathbb{P}$-a.s., for $i = 1, 2$,

$$\text{(3.3)} \quad \begin{aligned} Y_{\tau}^{i} - Y_{T}^{i} + \int_{\tau}^{T} Z_{\sigma}^{i} \, dW_{\sigma} + \lambda \int_{\tau}^{T} Y_{\sigma}^{i} \, d\sigma \\ = -\int_{\tau}^{T} \Psi(\sigma, Y_{\sigma}^{i}, Z_{\sigma}^{i}) \, d\sigma + \int_{\tau}^{T} f_{\sigma}^{i} \, d\sigma, \qquad 0 \leq \tau \leq T < \infty. \end{aligned}$$

Then for every $\overline{\lambda} > -(\beta + \mu - L_{z}^{2}/2)$ there exists $C > 0$ such that, for $\lambda > \overline{\lambda}$,

$$\text{(3.4)} \quad \begin{aligned} (\lambda - \overline{\lambda})|Y^{1} - Y^{2}|_{L_{\mathcal{P}}^{p}(\Omega; L_{\beta}^{2}(K))} + (\lambda - \overline{\lambda})^{1/2}|Z^{1} - Z^{2}|_{L_{\mathcal{P}}^{p}(\Omega; L_{\beta}^{2}(L_{2}(\Xi, K)))} \\ + (\lambda - \overline{\lambda})^{1/2}\left(\mathbb{E} \sup_{\tau \geq 0} e^{\beta \tau p}|Y_{\tau}^{1} - Y_{\tau}^{2}|^{p}\right)^{1/p} \leq C|f^{1} - f^{2}|_{L_{\mathcal{P}}^{p}(\Omega; L_{\beta}^{2}(K))}. \end{aligned}$$

The constant C depends only on β, μ, L_{z}, p and $\overline{\lambda}$.

PROOF.   Let us set for brevity

$$\overline{Y}_{\tau} = Y_{\tau}^{1} - Y_{\tau}^{2}, \qquad \overline{Z}_{\tau} = Z_{\tau}^{1} - Z_{\tau}^{2}, \qquad \overline{f}_{\tau} = f_{\tau}^{1} - f_{\tau}^{2},$$
$$\overline{\Psi}_{\tau} = \Psi(\tau, Y_{\tau}^{1}, Z_{\tau}^{1}) - \Psi(\tau, Y_{\tau}^{2}, Z_{\tau}^{2}).$$

Applying the Itô formula to the process $e^{2\beta\tau}|\overline{Y}_{\tau}|^{2}$, $\tau \geq 0$, we obtain

$$e^{2\beta\tau}|\overline{Y}_{\tau}|^{2} - e^{2\beta T}|\overline{Y}_{T}|^{2} + \int_{\tau}^{T} e^{2\beta\sigma}[2(\beta + \lambda)|\overline{Y}_{\sigma}|^{2} + |\overline{Z}_{\sigma}|^{2}] \, d\sigma$$

$$= -2\int_{\tau}^{T} e^{2\beta\sigma} \langle \overline{Y}_{\sigma}, \overline{Z}_{\sigma} \, dW_{\sigma}\rangle + 2\int_{\tau}^{T} e^{2\beta\sigma} \langle \overline{Y}_{\sigma}, \overline{f}_{\sigma}\rangle \, d\sigma$$

$$- 2\int_{\tau}^{T} e^{2\beta\sigma} \langle \overline{Y}_{\sigma}, \overline{\Psi}_{\sigma}\rangle \, d\sigma.$$

By Hypothesis 3.1 and an elementary inequality, we have

$$2\langle \overline{Y}_\sigma, \overline{\Psi}_\sigma \rangle \geq 2\mu|\overline{Y}_\sigma|^2 - 2L_z|\overline{Y}_\sigma||\overline{Z}_\sigma| \geq (2\mu - L_z^2/\rho)|\overline{Y}_\sigma|^2 - \rho|\overline{Z}_\sigma|^2,$$

where ρ is an arbitrary number in $(0, 1]$ that will be chosen later. Substituting in the previous equation yields

$$(3.5) \quad \begin{aligned} &e^{2\beta\tau}|\overline{Y}_\tau|^2 - e^{2\beta T}|\overline{Y}_T|^2 \\ &\quad + \int_\tau^T e^{2\beta\sigma}\big[(2\beta + 2\lambda + 2\mu - L_z^2/\rho)|\overline{Y}_\sigma|^2 + (1 - \rho)|\overline{Z}_\sigma|^2\big]d\sigma \\ &\quad \leq -2\int_\tau^T e^{2\beta\sigma}\langle \overline{Y}_\sigma, \overline{Z}_\sigma\,dW_\sigma \rangle + 2\int_\tau^T e^{2\beta\sigma}\langle \overline{Y}_\sigma, \overline{f}_\sigma \rangle\,d\sigma. \end{aligned}$$

We split the rest of the proof into several steps.

STEP 1. We claim that $\mathbb{E}\sup_{\tau \geq 0} e^{\beta\tau p}|\overline{Y}_\tau|^p < \infty$.
By the inequality $2\langle h, k \rangle \leq \varepsilon|h|^2 + \frac{1}{\varepsilon}|k|^2$, for all $\varepsilon > 0$, we have

$$2\int_\tau^T e^{2\beta\sigma}\langle \overline{Y}_\sigma, \overline{f}_\sigma \rangle\,d\sigma \leq \varepsilon \int_\tau^T e^{2\beta\sigma}|\overline{Y}_\sigma|^2\,d\sigma + \frac{1}{\varepsilon}\int_\tau^T e^{2\beta\sigma}|\overline{f}_\sigma|^2\,d\sigma.$$

From (3.5), setting $\rho = 1$, it follows that

$$\begin{aligned} &e^{2\beta\tau}|\overline{Y}_\tau|^2 - e^{2\beta T}|\overline{Y}_T|^2 + \int_\tau^T e^{2\beta\sigma}(2\beta + 2\lambda + 2\mu - L_z^2 - \varepsilon)|\overline{Y}_\sigma|^2\,d\sigma \\ &\quad \leq -2\int_\tau^T e^{2\beta\sigma}\langle \overline{Y}_\sigma, \overline{Z}_\sigma\,dW_\sigma \rangle + \frac{1}{\varepsilon}\int_\tau^T e^{2\beta\sigma}|\overline{f}_\sigma|^2\,d\sigma. \end{aligned}$$

Since we assume $\lambda > \overline{\lambda}$, we have $2\beta + 2\lambda + 2\mu - L_z^2 > 0$, and taking $\varepsilon > 0$ sufficiently small, we obtain

$$(3.6) \quad \begin{aligned} &e^{2\beta\tau}|\overline{Y}_\tau|^2 - e^{2\beta T}|\overline{Y}_T|^2 \\ &\quad \leq -2\int_\tau^T e^{2\beta\sigma}\langle \overline{Y}_\sigma, \overline{Z}_\sigma\,dW_\sigma \rangle + \frac{1}{\varepsilon}\int_\tau^T e^{2\beta\sigma}|\overline{f}_\sigma|^2\,d\sigma. \end{aligned}$$

The quadratic variation of the stochastic integral in (3.6) can be estimated as

$$\begin{aligned} \left(\int_\tau^T e^{4\beta\sigma}|\overline{Y}_\sigma|^2|\overline{Z}_\sigma|^2\,d\sigma\right)^{1/2} &\leq e^{2|\beta|T}\sup_{\sigma\in[\tau,T]}|\overline{Y}_\sigma|\left(\int_\tau^T |\overline{Z}_\sigma|^2\,d\sigma\right)^{1/2} \\ &\leq \tfrac{1}{2}e^{2|\beta|T}\sup_{\sigma\in[\tau,T]}|\overline{Y}_\sigma|^2 + \tfrac{1}{2}e^{2|\beta|T}\int_\tau^T |\overline{Z}_\sigma|^2\,d\sigma. \end{aligned}$$

The right-hand side of this inequality is an integrable random variable by our assumptions [the fact that $\mathbb{E}\sup_{\sigma\in[\tau,T]}|\overline{Y}_\sigma|^2 < \infty$ follows easily from (3.3)]. Thus

the stochastic integral in (3.6) is an integrable random variable. Conditioning both sides of (3.6) to $\mathcal{F}_\tau$, we obtain

$$(3.7) \qquad e^{2\beta\tau}|\overline{Y}_\tau|^2 \le e^{2\beta T}\mathbb{E}^{\mathcal{F}_\tau}|\overline{Y}_T|^2 + \frac{1}{\varepsilon}\mathbb{E}^{\mathcal{F}_\tau}\int_\tau^T e^{2\beta\sigma}|\overline{f}_\sigma|^2\,d\sigma.$$

Since we assume that $\int_0^\infty e^{2\beta\sigma}\mathbb{E}|Y_\sigma^i|^2\,d\sigma < \infty$, we can find a sequence $T_n \to \infty$ such that $e^{2\beta T_n}\mathbb{E}|\overline{Y}_{T_n}|^2 \to 0$. Setting $T = T_n$ in (3.7) and letting $n \to \infty$, we arrive at

$$e^{2\beta\tau}|\overline{Y}_\tau|^2 \le \frac{1}{\varepsilon}\mathbb{E}^{\mathcal{F}_\tau}\int_\tau^\infty e^{2\beta\sigma}|\overline{f}_\sigma|^2\,d\sigma$$

$$\le \frac{1}{\varepsilon}\mathbb{E}^{\mathcal{F}_\tau}\int_0^\infty e^{2\beta\sigma}|\overline{f}_\sigma|^2\,d\sigma =: M(\tau).$$

Since M is a martingale, then for all $p > 2$, by Doob and Jensen inequalities, there exists $c_p > 0$ such that

$$\mathbb{E}\sup_{\tau\in[0,T]} e^{\beta\tau p}|\overline{Y}_\tau|^p \le c_p\mathbb{E}\big(|M(T)|\big)^{p/2}$$

$$\le \frac{c_p}{\varepsilon^{p/2}}\mathbb{E}\bigg(\int_0^T e^{2\beta\sigma}|\overline{f}_\sigma|^2\,d\sigma\bigg)^{p/2} < \infty \qquad \forall\,T > 0.$$

Setting $T \nearrow \infty$, the inequality

$$\mathbb{E}\sup_{\tau\ge 0} e^{\beta\tau p}|\overline{Y}_\tau|^p \le \frac{c_p}{\varepsilon^{p/2}}\mathbb{E}\bigg(\int_0^\infty e^{2\beta\sigma}|\overline{f}_\sigma|^2\,d\sigma\bigg)^{p/2} < \infty$$

follows and the proof of Step 1 is concluded.

STEP 2. We claim that

$$e^{2\beta\tau}|\overline{Y}_\tau|^2 + \int_\tau^\infty e^{2\beta\sigma}\big[(2\beta + 2\lambda + 2\mu - L_z^2/\rho)|\overline{Y}_\sigma|^2 + (1-\rho)|\overline{Z}_\sigma|^2\big]d\sigma$$

$$(3.8)$$

$$\le -2\int_\tau^\infty e^{2\beta\sigma}\langle\overline{Y}_\sigma, \overline{Z}_\sigma\,dW_\sigma\rangle + 2\int_\tau^\infty e^{2\beta\sigma}\langle\overline{Y}_\sigma, \overline{f}_\sigma\rangle\,d\sigma.$$

We have

$$\mathbb{E}\bigg(\int_0^\infty e^{4\beta\sigma}|\overline{Y}_\sigma|^2|\overline{Z}_\sigma|^2\,d\sigma\bigg)^{p/4}$$

$$\le \mathbb{E}\bigg[\sup_{\tau\ge 0} e^{\beta\tau p/2}|\overline{Y}_\tau|^{p/2}\bigg(\int_0^\infty e^{2\beta\sigma}|\overline{Z}_\sigma|^2\,d\sigma\bigg)^{p/4}\bigg]$$

$$\le \bigg\{\mathbb{E}\sup_{\tau\ge 0} e^{\beta\tau p}|\overline{Y}_\tau|^p\bigg\}^{1/2}\bigg\{\mathbb{E}\bigg(\int_0^\infty e^{2\beta\sigma}|\overline{Z}_\sigma|^2\,d\sigma\bigg)^{p/2}\bigg\}^{1/2}$$

and the right-hand side is finite by Step 1. It follows that the limit of the stochastic integral $\int_0^T e^{2\beta\sigma}\langle \overline{Y}_\sigma, \overline{Z}_\sigma\, dW_\sigma\rangle$ for $T \to \infty$ exists in $L^{p/2}(\Omega; \mathbb{R})$ and, for some constant $c_p > 0$,

$$(3.9) \quad \begin{aligned} &\mathbb{E}\left|\int_0^\infty e^{2\beta\sigma}\langle \overline{Y}_\sigma, \overline{Z}_\sigma\, dW_\sigma\rangle\right|^{p/2} \\ &\leq c_p\left\{\mathbb{E}\sup_{\tau\geq 0} e^{\beta\tau p}|\overline{Y}_\tau|^p\right\}^{1/2}\left\{\mathbb{E}\left(\int_0^\infty e^{2\beta\sigma}|\overline{Z}_\sigma|^2\, d\sigma\right)^{p/2}\right\}^{1/2}. \end{aligned}$$

Choosing a sequence $T_n \to \infty$ such that $e^{2\beta T_n}\mathbb{E}|\overline{Y}_{T_n}|^2 \to 0$, as in the previous step, the required inequality (3.8) follows from (3.5) by setting $T = T_n$ and letting $n \to \infty$. Step 2 is finished.

STEP 3. Conclusion. We set, for brevity,

$$|\overline{Z}|_{L_\beta^2(L_2(\Xi,K))} = \left(\int_0^\infty e^{2\beta\sigma}|\overline{Z}_\sigma|^2_{L_2(\Xi,K)}\, d\sigma\right)^{1/2},$$

$$|\overline{Y}|_{L_\beta^2(K)} = \left(\int_0^\infty e^{2\beta\sigma}|\overline{Y}_\sigma|^2_K\, d\sigma\right)^{1/2},$$

and we define $|\overline{f}|_{L_\beta^2(K)}$ in a similar way. Conditioning both sides of (3.8) to $\mathcal{F}_\tau$, we obtain

$$\begin{aligned} e^{2\beta\tau}|\overline{Y}_\tau|^2 &\leq 2\mathbb{E}^{\mathcal{F}_\tau}\int_\tau^\infty e^{2\beta\sigma}\langle \overline{Y}_\sigma, \overline{f}_\sigma\rangle\, d\sigma \\ &\leq 2\mathbb{E}^{\mathcal{F}_\tau}\int_0^\infty e^{2\beta\sigma}|\overline{Y}_\sigma||\overline{f}_\sigma|\, d\sigma \leq 2\mathbb{E}^{\mathcal{F}_\tau}\left(|\overline{Y}|_{L_\beta^2(K)}|\overline{f}|_{L_\beta^2(K)}\right), \end{aligned}$$

and by the Burkholder–Davis–Gundy inequalities, there exists a constant $c_p > 0$ such that

$$(3.10) \quad \mathbb{E}\sup_{\tau\geq 0} e^{\beta\tau p}|\overline{Y}_\tau|^p \leq c_p\mathbb{E}\left(|\overline{Y}|^{p/2}_{L_\beta^2(K)}|\overline{f}|^{p/2}_{L_\beta^2(K)}\right).$$

Now we consider again (3.8). Taking into account the inequality

$$\int_0^\infty e^{2\beta\sigma}\langle \overline{Y}_\sigma, \overline{f}_\sigma\rangle\, d\sigma \leq |\overline{Y}|_{L_\beta^2(K)}|\overline{f}|_{L_\beta^2(K)}$$

and choosing $\rho < 1$ so close to 1 that $2\beta + 2\mu - L_z^2/\rho > -2\overline{\lambda}$, we obtain, for some constant $c > 0$,

$$\begin{aligned} &2(\lambda - \overline{\lambda})|\overline{Y}|^2_{L_\beta^2(K)} + c|\overline{Z}|^2_{L_\beta^2(L_2(\Xi,K))} \\ &\leq 2|\overline{Y}|_{L_\beta^2(K)}|\overline{f}|_{L_\beta^2(K)} + 2\left|\int_0^\infty e^{2\beta\sigma}\langle \overline{Y}_\sigma, \overline{Z}_\sigma\, dW_\sigma\rangle\right|. \end{aligned}$$

Raising to the power $p/2$, taking expectation, and recalling (3.9) and (3.10), we obtain, for suitable constants c_i,

$$(\lambda - \overline{\lambda})^{p/2} \mathbb{E}|\overline{Y}|^p_{L^2_\beta(K)} + c_1 \mathbb{E}|\overline{Z}|^p_{L^2_\beta(L_2(\Xi,K))}$$

$$\leq c_2 \mathbb{E}\Big(|\overline{Y}|^{p/2}_{L^2_\beta(K)} |\overline{f}|^{p/2}_{L^2_\beta(K)}\Big)$$

$$+ c_3 \Big\{\mathbb{E}|\overline{Z}|^p_{L^2_\beta(L_2(\Xi,K))}\Big\}^{1/2} \Big\{\mathbb{E}\Big(|\overline{Y}|^{p/2}_{L^2_\beta(K)} |\overline{f}|^{p/2}_{L^2_\beta(K)}\Big)\Big\}^{1/2}$$

$$\leq c_2 \mathbb{E}\Big(|\overline{Y}|^{p/2}_{L^2_\beta(K)} |\overline{f}|^{p/2}_{L^2_\beta(K)}\Big)$$

$$+ \varepsilon \mathbb{E}|\overline{Z}|^p_{L^2_\beta(L_2(\Xi,K))} + (c_3^2/(4\varepsilon)) \mathbb{E}\Big(|\overline{Y}|^{p/2}_{L^2_\beta(K)} |\overline{f}|^{p/2}_{L^2_\beta(K)}\Big)$$

for every $\varepsilon > 0$. Choosing ε sufficiently small and using the Cauchy–Schwarz inequality, we obtain, for some $c > 0$,

$$(\lambda - \overline{\lambda})^{p/2} \mathbb{E}|\overline{Y}|^p_{L^2_\beta(K)} + c\mathbb{E}|\overline{Z}|^p_{L^2_\beta(L_2(\Xi,K))}$$

$$\leq c\Big\{\mathbb{E}|\overline{Y}|^p_{L^2_\beta(K)}\Big\}^{1/2} \Big\{\mathbb{E}|\overline{f}|^p_{L^2_\beta(K)}\Big\}^{1/2}.$$

It follows that

$$(\lambda - \overline{\lambda})^{1/2} |\overline{Y}|_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(K))} + c|\overline{Z}|_{L^p_{\mathcal{P}}(L^2_\beta(L_2(\Xi,K)))}$$

$$\leq c|\overline{Y}|^{1/2}_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(K))} |\overline{f}|^{1/2}_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(K))},$$

and the conclusion (3.4) follows immediately by taking into account (3.10) once more. $\square$

In the linear case $\Psi = 0$ we immediately obtain:

COROLLARY 3.3. *Assume that for some $\beta \in \mathbb{R}$, $p \in (2,\infty)$, $\lambda \in \mathbb{R}$ there exist processes $Y, f \in L^p_{\mathcal{P}}(\Omega; L^2_\beta(K))$ and $Z \in L^p_{\mathcal{P}}(\Omega; L^2_\beta(L_2(\Xi, K)))$ such that, $\mathbb{P}$-a.s.,*

$$(3.11) \qquad Y_\tau - Y_T + \int_\tau^T Z_\sigma \, dW_\sigma + \lambda \int_\tau^T Y_\sigma \, d\sigma = \int_\tau^T f_\sigma \, d\sigma,$$

$$0 \leq \tau \leq T < \infty.$$

Then for every $\overline{\lambda} > -\beta$, there exists a constant $C > 0$ (depending only on β, p and $\overline{\lambda}$) such that, for $\lambda > \overline{\lambda}$,

$$(\lambda - \overline{\lambda})|Y|_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(K))} + (\lambda - \overline{\lambda})^{1/2} |Z|_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(L_2(\Xi,K)))}$$

$$+ (\lambda - \overline{\lambda})^{1/2} \Big(\mathbb{E}\sup_{\tau \geq 0} e^{\beta\tau p} |Y_\tau|^p\Big)^{1/p} \leq C|f|_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(K))}.$$

The next step toward Theorem 3.7 consists of proving that the solution exists for large values of λ (see Proposition 3.6). We start with some lemmas.

LEMMA 3.4.   *Suppose that $p \in [2, \infty)$, and let $\kappa > 0$ and $f \in L^p_{\mathcal{P}}(\Omega; L^2_\kappa(K))$. Then there exists a unique pair (Y, Z) such that*

(3.12)
$$Y \in L^p_{\mathcal{P}}(\Omega; L^2_\kappa(K)),$$
$$Z \in L^p_{\mathcal{P}}\big(\Omega; L^2_\kappa(L_2(\Xi, K))\big)$$

that satisfies the equation, $\mathbb{P}$-a.s.,

$$Y_\tau - Y_T + \int_\tau^T Z_\sigma \, dW_\sigma = \int_\tau^T f_\sigma \, d\sigma, \qquad 0 \le \tau \le T < \infty.$$

PROOF.   Since $\kappa > 0$, the equation is equivalent to, $\mathbb{P}$-a.s.,

(3.13)
$$Y_\tau + \int_\tau^\infty Z_\sigma \, dW_\sigma = \int_\tau^\infty f_\sigma \, d\sigma, \qquad \tau \ge 0.$$

The assertion of the lemma, in the case $p = 2$, follows from [27], Lemma 2.1, in the finite-dimensional case ($\dim \Xi < \infty$, $\dim K < \infty$), but the arguments are the same in the infinite-dimensional case. In the general case, but with finite horizon, the result is contained in [12] and in [26].

To complete the proof, it remains to show that if Y and Z belong to $L^2_{\mathcal{P}}(\Omega; L^2_\kappa(K))$ and $L^2_{\mathcal{P}}(\Omega; L^2_\kappa(L_2(\Xi, K)))$, respectively, and if $f \in L^p_{\mathcal{P}}(\Omega; L^2_\kappa(K))$, then (3.12) holds.

Conditioning both sides of (3.13) to $\mathcal{F}_\tau$, we obtain $Y_\tau = \mathbb{E}^{\mathcal{F}_\tau} \int_\tau^\infty f_\sigma \, d\sigma$ and it follows that

$$|Y_\tau| \le \mathbb{E}^{\mathcal{F}_\tau} \int_\tau^\infty |f_\sigma| \, d\sigma \le \frac{e^{-\kappa\tau}}{\sqrt{2\kappa}} \mathbb{E}^{\mathcal{F}_\tau} \bigg(\int_\tau^\infty e^{2\kappa\sigma} |f_\sigma|^2 \, d\sigma \bigg)^{1/2}.$$

By the Burkholder–Davis–Gundy inequalities, there exists a constant $c > 0$, depending only on p and κ, such that

(3.14)
$$\mathbb{E} \sup_{\tau \ge 0} e^{\kappa\tau p} |Y_\tau|^p \le c\mathbb{E} \bigg(\int_0^\infty e^{2\kappa\sigma} |f_\sigma|^2 \, d\sigma \bigg)^{p/2}.$$

Next, applying the Itô formula to the process $e^{\kappa\tau} Y_\tau$, $\tau \ge 0$, we obtain

$$e^{\kappa\tau} Y_\tau - e^{\kappa T} Y_T + \int_\tau^T e^{\kappa\sigma} Z_\sigma \, dW_\sigma + \kappa \int_\tau^T e^{\kappa\sigma} Y_\sigma \, d\sigma$$

$$= \int_\tau^T e^{\kappa\sigma} f_\sigma \, d\sigma, \qquad 0 \le \tau \le T < \infty.$$

Again by the Burkholder–Davis–Gundy inequalities, there exists a constant $c_p > 0$, depending only on p, such that

$$\mathbb{E}\left(\int_\tau^T e^{2\kappa\sigma}|Z_\sigma|^2\,d\sigma\right)^{p/2}$$

$$\leq c_p\mathbb{E}\sup_{t\in[\tau,T]}\left|\int_\tau^t e^{\kappa\sigma}Z_\sigma\,dW_\sigma\right|^p$$

$$\leq c_p\mathbb{E}\left|2\sup_{t\in[\tau,T]}e^{\kappa t}|Y_t| + \kappa\int_\tau^T e^{\kappa\sigma}|Y_\sigma|\,d\sigma + \int_\tau^T e^{\kappa\sigma}|f_\sigma|\,d\sigma\right|^p$$

$$\leq c\mathbb{E}\sup_{t\in[\tau,T]}e^{\kappa t p}|Y_t|^p + c\mathbb{E}\left(\int_\tau^T e^{2\kappa\sigma}|f_\sigma|^2\,d\sigma\right)^{p/2},$$

where $c > 0$ denotes a constant that may depend on τ, T and κ as well. It follows from (3.14) that

$$(3.15) \qquad \mathbb{E}\left(\int_\tau^T e^{2\kappa\sigma}|Z_\sigma|^2\,d\sigma\right)^{p/2} < \infty.$$

We apply the Itô formula to the process $e^{2\kappa\tau}|Y_\tau|^2$, $\tau \geq 0$, obtaining

$$(3.16) \quad \begin{aligned} &e^{2\kappa\tau}|Y_\tau|^2 - e^{2\kappa T}|Y_T|^2 + \int_\tau^T e^{2\kappa\sigma}[2\kappa|Y_\sigma|^2 + |Z_\sigma|^2]\,d\sigma \\ &\qquad = -2\int_\tau^T e^{2\kappa\sigma}\langle Y_\sigma, Z_\sigma\,dW_\sigma\rangle + 2\int_\tau^T e^{2\kappa\sigma}\langle Y_\sigma, f_\sigma\rangle\,d\sigma. \end{aligned}$$

We estimate the right-hand side as follows. First, for every $\varepsilon > 0$,

$$2\int_\tau^T e^{2\kappa\sigma}\langle Y_\sigma, f_\sigma\rangle\,d\sigma \leq \varepsilon\int_\tau^T e^{2\kappa\sigma}|Y_\sigma|^2\,d\sigma + \frac{1}{\varepsilon}\int_\tau^T e^{2\kappa\sigma}|f_\sigma|^2\,d\sigma.$$

Next we can estimate the stochastic integral as

$$c_p\mathbb{E}\left|\int_\tau^T e^{2\kappa\sigma}\langle Y_\sigma, Z_\sigma\,dW_\sigma\rangle\right|^{p/2}$$

$$\leq \mathbb{E}\left(\int_\tau^T e^{4\kappa\sigma}|Y_\sigma|^2|Z_\sigma|^2\,d\sigma\right)^{p/4}$$

$$\leq \mathbb{E}\left[\sup_{\tau\geq 0}e^{\kappa\tau p/2}|Y_\tau|^{p/2}\left(\int_\tau^T e^{2\kappa\sigma}|Z_\sigma|^2\,d\sigma\right)^{p/4}\right]$$

$$\leq \frac{1}{2\varepsilon}\mathbb{E}\sup_{\tau\geq 0}e^{\kappa\tau p}|Y_\tau|^p + \frac{\varepsilon}{2}\mathbb{E}\left(\int_\tau^T e^{2\kappa\sigma}|Z_\sigma|^2\,d\sigma\right)^{p/2}.$$

Note that the right-hand side is finite, by (3.14) and (3.15). Raising both sides of (3.16) to the power $p/2$ and taking expectation, we obtain, for some constant

$c > 0$ independent of τ, T and ε,

$$\mathbb{E}e^{p\kappa\tau}|Y_\tau|^p + c\mathbb{E}\left(\int_\tau^T e^{2\kappa\sigma}[|Y_\sigma|^2 + |Z_\sigma|^2]\,d\sigma\right)^{p/2}$$

$$\leq c\varepsilon\mathbb{E}\left(\int_\tau^T e^{2\kappa\sigma}|Z_\sigma|^2\,d\sigma\right)^{p/2}$$

$$+ \frac{c}{\varepsilon}\mathbb{E}\sup_{\tau\geq 0}e^{\kappa\tau p}|Y_\tau|^p + c\varepsilon^{p/2}\mathbb{E}\left(\int_\tau^T e^{2\kappa\sigma}|Y_\sigma|^2\,d\sigma\right)^{p/2}$$

$$+ \frac{c}{\varepsilon^{p/2}}\mathbb{E}\left(\int_\tau^T e^{2\kappa\sigma}|f_\sigma|^2\,d\sigma\right)^{p/2} + \mathbb{E}e^{p\kappa T}|Y_T|^p.$$

Taking ε sufficiently small and recalling (3.14), we conclude that

$$\mathbb{E}\left(\int_\tau^T e^{2\kappa\sigma}[|Y_\sigma|^2 + |Z_\sigma|^2]\,d\sigma\right)^{p/2}$$

$$\leq c_1\mathbb{E}\sup_{\tau\geq 0}e^{\kappa\tau p}|Y_\tau|^p + c_2\mathbb{E}\left(\int_\tau^T e^{2\kappa\sigma}|f_\sigma|^2\,d\sigma\right)^{p/2}$$

$$\leq c_3\mathbb{E}\left(\int_0^\infty e^{2\kappa\sigma}|f_\sigma|^2\,d\sigma\right)^{p/2}$$

for constants c_i independent of τ, T. This proves that Y belongs to $L_{\mathcal{P}}^p(\Omega; L_\kappa^2(K))$ and Z belongs to $L_{\mathcal{P}}^p(\Omega; L_\kappa^2(L_2(\Xi, K)))$, and concludes the proof of the lemma. $\square$

LEMMA 3.5. *Suppose that* $p \in [2, \infty)$ *and let* $f \in L_{\mathcal{P}}^p(\Omega; L_\beta^2(K))$ *for some real number* β. *Then for every* $\lambda > -\beta$ *there exists a unique pair* (Y, Z) *such that*

(3.17) $\qquad Y \in L_{\mathcal{P}}^p(\Omega; L_\beta^2(K)), \qquad Z \in L_{\mathcal{P}}^p(\Omega; L_\beta^2(L_2(\Xi, K))),$

and satisfying the equation, $\mathbb{P}$-*a.s.,*

$$Y_\tau - Y_T + \int_\tau^T Z_\sigma\,dW_\sigma + \lambda\int_\tau^T Y_\sigma\,d\sigma = \int_\tau^T f_\sigma\,d\sigma, \qquad 0 \leq \tau \leq T < \infty.$$

PROOF. Setting $Y_\tau^\lambda = e^{-\lambda\tau}Y_\tau$, $Z_\tau^\lambda = e^{-\lambda\tau}Z_\tau$ and $f_\tau^\lambda = e^{-\lambda\tau}f_\tau$, by the Itô formula the equation of the lemma is equivalent to

$$Y_\tau^\lambda - Y_T^\lambda + \int_\tau^T Z_\sigma^\lambda\,dW_\sigma = \int_\tau^T f_\sigma^\lambda\,d\sigma.$$

Moreover, we have

$$Y, f \in L_{\mathcal{P}}^p(\Omega; L_\beta^2(K)), \qquad Z \in L_{\mathcal{P}}^p(\Omega; L_\beta^2(L_2(\Xi, K)))$$

if and only if

$$Y^\lambda, f^\lambda \in L^p_{\mathcal{P}}(\Omega; L^2_{\beta+\lambda}(K)), \qquad Z^\lambda \in L^p_{\mathcal{P}}(\Omega; L^2_{\beta+\lambda}(L_2(\Xi, K))).$$

Since $\beta + \lambda > 0$, the result follows immediately from the previous lemma. $\quad\square$

PROPOSITION 3.6. *Suppose that* Ψ *satisfies Hypothesis* 3.1 *for some* $p \in (2, \infty)$, *and assume that* $f \in L^p_{\mathcal{P}}(\Omega; L^2_\beta(K))$ *for some* $\beta \in \mathbb{R}$. *Then there exists* $\lambda_1 \in \mathbb{R}$ *such that for* $\lambda \geq \lambda_1$ *the equation* (3.1) *has a unique solution* (Y, Z) *such that*

$$Y \in L^p_{\mathcal{P}}(\Omega; L^2_\beta(K)), \qquad Z \in L^p_{\mathcal{P}}(\Omega; L^2_\beta(L_2(\Xi, K))).$$

PROOF.   Let us recall the space

$$\mathcal{K}^p_\beta = L^p_{\mathcal{P}}(\Omega; L^2_\beta(K)) \times L^p_{\mathcal{P}}(\Omega; L^2_\beta(L_2(\Xi, K))),$$

endowed with the norm $|(Y, Z)|_{\mathcal{K}^p_\beta} = |Y|_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(K))} + |Z|_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(L_2(\Xi, K)))}$ introduced in Section 2. For every λ we define a map $\Gamma : \mathcal{K}^p_\beta \to \mathcal{K}^p_\beta$, setting $(Y, Z) = \Gamma(U, V)$ if (Y, Z) is the solution of the equation, $\mathbb{P}$-*a.s.*,

$$Y_\tau - Y_T + \int_\tau^T Z_\sigma \, dW_\sigma + \lambda \int_\tau^T Y_\sigma \, d\sigma$$

$$= -\int_\tau^T \Psi(\sigma, U_\sigma, V_\sigma) \, d\sigma + \int_\tau^T f_\sigma \, d\sigma, \qquad 0 \leq \tau \leq T < \infty.$$

By the previous lemma, Γ is well defined for all sufficiently large values of λ. If, for $i = 1, 2$, $(U^i, V^i) \in \mathcal{K}^p_\beta$, $(Y^i, Z^i) = \Gamma(U^i, V^i)$, then, by Corollary 3.3, we have

$$(\lambda - \overline{\lambda})|Y^1 - Y^2|_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(K))} + (\lambda - \overline{\lambda})^{1/2}|Z^1 - Z^2|_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(L_2(\Xi, K)))}$$

$$\leq C \left\{ \mathbb{E}\left(\int_0^\infty e^{2\beta\sigma} |\Psi(\sigma, U^1_\sigma, V^1_\sigma) - \Psi(\sigma, U^2_\sigma, V^2_\sigma)|^2 \, d\sigma \right)^{p/2} \right\}^{1/p}$$

and by the Lipschitz condition on Ψ, we have, for some constant $c > 0$ independent of λ,

$$(\lambda - \overline{\lambda})|Y^1 - Y^2|_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(K))} + (\lambda - \overline{\lambda})^{1/2}|Z^1 - Z^2|_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(L_2(\Xi, K)))}$$

$$\leq c|U^1 - U^2|_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(K))} + c|V^1 - V^2|_{L^p_{\mathcal{P}}(\Omega; L^2_\beta(L_2(\Xi, K)))}.$$

This shows that Γ is a contraction in $\mathcal{K}^p_\beta$ for all λ sufficiently large. Its unique fixed point is the required solution. $\quad\square$

THEOREM 3.7.   *Suppose that Ψ satisfies Hypothesis 3.1 for some $p \in (2, \infty)$ and assume that $f \in L^p_{\mathcal{P}}(\Omega; L^2_\beta(K))$ for some $\beta \in \mathbb{R}$. Then for $\lambda > -(\beta + \mu - L^2_z/2)$, (3.1) has a unique solution (Y, Z) such that*

$$Y \in L^p_{\mathcal{P}}(\Omega; L^2_\beta(K)), \qquad Z \in L^p_{\mathcal{P}}(\Omega; L^2_\beta(L_2(\Xi, K))).$$

Moreover, for every $\overline{\lambda} > -(\beta + \mu - L^2_z/2)$ there exists $C > 0$ such that, for $\lambda > \overline{\lambda}$,

$$(\lambda - \overline{\lambda})|Y|_{L^p_{\mathcal{P}}(\Omega;L^2_\beta(K))} + (\lambda - \overline{\lambda})^{1/2}|Z|_{L^p_{\mathcal{P}}(\Omega;L^2_\beta(L_2(\Xi,K)))}$$

(3.18)
$$+ (\lambda - \overline{\lambda})^{1/2}\left(\mathbb{E}\sup_{\tau \geq 0} e^{\beta\tau p}|Y_\tau|^p\right)^{1/p}$$

$$\leq C\left\{\mathbb{E}\left(\int_0^\infty e^{2\beta\sigma}|\Psi(\sigma, 0, 0)|^2\, d\sigma\right)^{p/2}\right\}^{1/p} + C|f|_{L^p_{\mathcal{P}}(\Omega;L^2_\beta(K))}.$$

The constant C depends only on β, μ, L_z, p and $\overline{\lambda}$.

PROOF.   Let us consider again the space $\mathcal{K}^p_\beta$ used in the previous proof. We fix $\overline{\lambda} > -(\beta + \mu - L^2_z/2)$ and we define Q as the set of those real numbers $\lambda > \overline{\lambda}$ such that for every $f \in L^p_{\mathcal{P}}(\Omega; L^2_\beta(K))$ there exists a unique solution $(Y, Z) \in \mathcal{K}^p_\beta$ corresponding to λ and f. Letting C be the constant whose existence is asserted in Theorem 3.2, we also set

$$C(\lambda) = C[(\lambda - \overline{\lambda})^{-1} \vee (\lambda - \overline{\lambda})^{-1/2}].$$

We claim that if Q contains a number λ_0, then it contains every number $\lambda > \overline{\lambda}$ belonging to the interval $(\lambda_0 - C(\lambda_0)^{-1}, \lambda_0 + C(\lambda_0)^{-1})$. Indeed, for any $\lambda > \overline{\lambda}$, let us define a map $\Gamma : \mathcal{K}^p_\beta \to \mathcal{K}^p_\beta$, setting $(Y, Z) = \Gamma(U, V)$ if (Y, Z) is the solution of the equation, $\mathbb{P}$-*a.s.*, for $0 \leq \tau \leq T < \infty$,

$$Y_\tau - Y_T + \int_\tau^T Z_\sigma\, dW_\sigma + \lambda_0 \int_\tau^T Y_\sigma\, d\sigma$$

$$= -\int_\tau^T \Psi(\sigma, Y_\sigma, Z_\sigma)\, d\sigma + \int_\tau^T [(\lambda_0 - \lambda)U_\sigma + f_\sigma]\, d\sigma.$$

Thus, an element $(Y, Z) \in \mathcal{K}^p_\beta$ is a solution of (3.1) if and only if (Y, Z) is a fixed point of Γ; if, for $i = 1, 2$, $(U^i, V^i) \in \mathcal{K}^p_\beta$ and $(Y^i, Z^i) = \Gamma(U^i, V^i)$, then, by Theorem 3.2, we have

$$|(Y^1 - Y^2, Z^1 - Z^2)|_{\mathcal{K}^p_\beta} \leq C(\lambda_0)|\lambda - \lambda_0||U^1 - U^2|_{L^p_{\mathcal{P}}(\Omega;L^2_\beta(K))}$$

$$\leq C(\lambda_0)|\lambda - \lambda_0||(U^1 - U^2, V^1 - V^2)|_{\mathcal{K}^p_\beta},$$

which shows that Γ is a contraction if $C(\lambda_0)|\lambda - \lambda_0| < 1$, and the claim follows immediately from the Banach contraction principle.

Starting from this claim, we will show that Q coincides with $(\overline{\lambda}, \infty)$. If $\lambda_n \in Q$, $\lambda_\infty > \overline{\lambda}$ and $\lambda_n \to \lambda_\infty$, then for n sufficiently large, we have

$$\lambda_\infty \in \left(\lambda_n - C(\lambda_n)^{-1}, \lambda_n + C(\lambda_n)^{-1}\right)$$

and by the claim, we conclude that $\lambda_\infty \in Q$. Therefore, Q is a closed topological subspace of $(\overline{\lambda}, \infty)$. Further, invoking the claim once more, it is immediate to see that Q is also an open subspace of $(\overline{\lambda}, \infty)$. Finally Q is nonempty, since by Proposition 3.6 it contains an interval $[\lambda_1, \infty)$. We conclude that $Q = (\overline{\lambda}, \infty)$. Existence and uniqueness of the solution is now proved for every $\lambda > -(\beta + \mu - L_z^2/2)$.

The final estimate in the statement of the theorem follows from (3.4), noting that the solution corresponding to $f_\tau = \Psi(\tau, 0, 0)$ is the trivial solution $(Y, Z) = (0, 0)$.

$\square$

REMARK 3.8. It follows from (3.18) that if (Y, Z) is the mild solution to (3.1), then $Y \in L_{\mathcal{P}}^p(\Omega, C_\beta(K))$. Nevertheless uniqueness holds in the larger class $Y \in L_{\mathcal{P}}^p(\Omega, L_\beta^2(K))$.

REMARK 3.9. In this section we have allowed process Y to take values in an infinite-dimensional space. Such a generality will be needed in the sequel to treat the gradient, with respect to data, of backward equations in which process Y is real valued.

3.1. *Regular dependence on an auxiliary process.* Let us now consider a backward equation of special form, $\mathbb{P}$-*a.s.*,

$$
\begin{aligned}
&Y_\tau - Y_T + \int_\tau^T Z_\sigma \, dW_\sigma + \lambda \int_\tau^T Y_\sigma \, d\sigma \\
&\qquad = -\int_\tau^T \psi(X_\sigma, Y_\sigma, Z_\sigma) \, d\sigma, \qquad 0 \leq \tau \leq T < \infty,
\end{aligned}
$$

(3.19)

where $\psi : H \times K \times L_2(\Xi, K) \to K$ is a given measurable function, X is a predictable process with values in another Hilbert space H and λ is a real number. We want to investigate the dependence of the solution on the process X. We assume the following.

HYPOTHESIS 3.10. (i) *There exist $\mu \in \mathbb{R}$ and nonnegative constants L_y, L_z such that*

$$|\psi(x, y_1, z_1) - \psi(x, y_2, z_2)| \leq L_y |y_1 - y_2| + L_z |z_1 - z_2|,$$

$$\langle \psi(x, y_1, z) - \psi(x, y_2, z), y_1 - y_2 \rangle_K \geq \mu |y_1 - y_2|^2$$

for every $x \in H$, $y_1, y_2 \in K$ and $z, z_1, z_2 \in L_2(\Xi, K)$.

(ii) $\psi \in \mathcal{G}^1(H \times K \times L_2(\Xi, K), K)$.

(iii) *There exist $L > 0$ and $m \geq 0$ such that*

$$|\nabla_x \psi(x, y, z)h| \leq L|h|(1 + |z|)(1 + |x| + |y|)^m$$

for every $x, h \in H$, $y \in K$, $z \in L_2(\Xi, K)$.

If the process X satisfies

$$\mathbb{E}\left(\int_0^\infty e^{2\beta\sigma} |\psi(X_\sigma, 0, 0)|^2 \, d\sigma\right)^{p/2} < \infty$$

for some $p > 2$ and $\beta \in \mathbb{R}$, then it follows immediately from Theorem 3.7 that for $\lambda > -(\beta + \mu - L_z^2/2)$, (3.19) has a unique solution in the space $\mathcal{K}_\beta^p$ used in the previous section. In fact, to reach this conclusion only point (i) of Hypothesis 3.10 is needed.

Proposition 3.11 below shows that the dependence of (Y, Z) on X is regular, provided the values of the various parameters are suitably chosen (in particular, we may need larger values of λ) and X is considered as an element of appropriate spaces of processes. Process X will be taken in the spaces $L_\mathcal{P}^q(\Omega; L_\eta^q(H))$, $L_\mathcal{P}^q(\Omega; C_\eta(H))$ and $\mathcal{H}_\eta^q$; these spaces were introduced in Section 2 for every $\eta \in \mathbb{R}$ and $q \in [1, \infty)$, and for arbitrary Hilbert space H. Clearly, similar definitions and notations also apply to processes with values in other Hilbert spaces.

PROPOSITION 3.11. *Assume Hypothesis 3.10. Let $r > 2$ and $\delta < 0$ be given, and choose*

(3.20)
$$q \geq (m + 1)r, \qquad \eta > \delta/(m + 1).$$

Then the following hold:

(i) *For $X \in L_\mathcal{P}^q(\Omega; L_\eta^q(H))$ and $\lambda > -(\delta + \mu - L_z^2/2)$, (3.19) has a unique solution in $\mathcal{K}_\delta^r$ that will be denoted by $(Y_\tau(X), Z_\tau(X))$, $\tau \geq 0$.*

(ii) *The estimate*

(3.21)
$$\mathbb{E}\sup_{\tau \geq 0} |Y_\tau(X)|^r e^{r\delta\tau} + \mathbb{E}\left(\int_0^\infty e^{2\delta\sigma} |Y_\sigma(X)|^2 \, d\sigma\right)^{r/2}$$
$$+ \mathbb{E}\left(\int_0^\infty e^{2\delta\sigma} |Z_\sigma(X)|^2 \, d\sigma\right)^{r/2} \leq c\left(1 + |X|_{L_\mathcal{P}^q(\Omega; L_\eta^q(H))}^{m+1}\right)^r$$

holds for a suitable constant c. In particular, $Y(X) \in L_\mathcal{P}^r(\Omega; C_\delta(K))$.

(iii) *The map $X \to (Y(X), Z(X))$ is continuous from $L_\mathcal{P}^q(\Omega; L_\eta^q(H))$ to $\mathcal{K}_\delta^r$ and $X \to Y(X)$ is continuous from $L_\mathcal{P}^q(\Omega; L_\eta^q(H))$ to $L_\mathcal{P}^r(\Omega; C_\delta(K))$.*

(iv) *The statements of points* (i), (ii) *and* (iii) *still hold true if the space $L_\mathcal{P}^q(\Omega; L_\eta^q(H))$ is replaced by the space $L_\mathcal{P}^q(\Omega; C_\eta(H))$.*

Now suppose that $p > 2$ and $\beta < 0$ are given, and choose

(3.22)
$$q \geq (m + 1)(m + 2)p, \qquad \eta > \beta(m + 1)^{-1}(m + 2)^{-1}.$$

Then the following hold:

(v) *For* $\lambda > -(\beta + \mu - L_z^2/2)$, *the map* $X \to (Y(X), Z(X))$ *is in* $\mathcal{G}^1(L_{\mathcal{P}}^q(\Omega; C_\eta(H)), \mathcal{K}_\beta^p)$ *and the map* $X \to Y(X)$ *is in* $\mathcal{G}^1(L_{\mathcal{P}}^q(\Omega; C_\eta(H)),$ $L_{\mathcal{P}}^p(\Omega; C_\beta(H)))$.

(vi) *At every point* $X \in L_{\mathcal{P}}^q(\Omega; C_\eta(H))$, *the directional derivative process of* $(Y(X), Z(X))$ *in the direction* $N \in L_{\mathcal{P}}^q(\Omega; C_\eta(H))$, *denoted by* $(\nabla Y_\tau(X)N,$ $\nabla Z_\tau(X)N)$, $\tau \geq 0$, *is the unique solution in* $\mathcal{K}_\beta^p$ *of the backward equation,* $\mathbb{P}$*-a.s., for* $0 \leq \tau \leq T < \infty$,

$$\nabla Y_\tau(X)N - \nabla Y_T(X)N + \lambda \int_\tau^T \nabla Y_\sigma(X)N \, d\sigma + \int_\tau^T \nabla Z_\sigma(X)N \, dW_\sigma$$

$$= -\int_\tau^T \nabla_x \psi(X_\sigma, Y_\sigma(X), Z_\sigma(X)) N_\sigma \, d\sigma$$

$$-\int_\tau^T \nabla_y \psi(X_\sigma, Y_\sigma(X), Z_\sigma(X)) \nabla Y_\sigma(X)N \, d\sigma$$

$$-\int_\tau^T \nabla_z \psi(X_\sigma, Y_\sigma(X), Z_\sigma(X)) \nabla Z_\sigma(X)N \, d\sigma.$$

Moreover, $\nabla Y(X)N$ *is in* $L_{\mathcal{P}}^p(\Omega; C_\beta(H))$.
(vii) *Finally the following estimate holds*:

$$\mathbb{E} \sup_{\tau \geq 0} e^{p\beta\tau} |\nabla Y_\tau(X)N|^p + \mathbb{E} \left(\int_0^\infty e^{2\beta\sigma} |\nabla Y_\sigma(X)N|^2 \, d\sigma \right)^{p/2}$$

(3.23)
$$+ \mathbb{E} \left(\int_0^\infty e^{2\beta\sigma} |\nabla Z_\sigma(X)N|^2 \, d\sigma \right)^{p/2}$$

$$\leq c |N|_{L_{\mathcal{P}}^q(\Omega; C_\eta(H))}^p \left(1 + |X|_{L_{\mathcal{P}}^q(\Omega; C_\eta(H))}^{[(m+1)^2]} \right)^p.$$

PROOF.   It follows from Hypothesis 3.10 that

$$|\psi(x, 0, 0)| \leq c(1 + |x|)^{m+1}, \qquad x \in H.$$

Here and in the rest of this proof, c denotes a positive constant, whose value may vary from line to line. Choosing δ' such that $\delta < \delta' < \eta(m+1) \wedge 0$, we obtain

$$\mathbb{E} \left(\int_0^\infty e^{2\delta\sigma} |\psi(X_\sigma, 0, 0)|^2 \, d\sigma \right)^{r/2}$$

$$\leq c + c\mathbb{E} \left(\int_0^\infty e^{2\delta\sigma} |X_\sigma|^{2(m+1)} \, d\sigma \right)^{r/2}$$

(3.24)
$$\leq c + c\mathbb{E} \int_0^\infty e^{r\delta'\sigma} |X_\sigma|^{r(m+1)} \, d\sigma$$

$$\leq c + c\left(\mathbb{E}\int_0^\infty e^{q\eta\sigma}|X_\sigma|^q\,d\sigma\right)^{r(m+1)/q}$$

$$= c\left[1 + |X_\sigma|_{L_{\mathcal{P}}^q(\Omega;L_\eta^q(H))}^{r(m+1)}\right].$$

Now existence of a unique solution in $\mathcal{K}_\delta^r$ of (3.19), for $\lambda > -(\delta + \mu - L_z^2/2)$, and the estimate (3.21) follows from Theorem 3.7.

To prove continuous dependence stated in point (iii), let us first note that Hypothesis 3.10 also yields the inequality

$$|\psi(x,y,z)| \leq c(1 + |x|^{m+1} + |y| + |z|), \qquad x \in H,\ y \in K,\ z \in L_2(\Xi, K).$$

By estimates analogous to (3.24), we can prove that the map $(X, Y, Z) \mapsto \psi(X, Y, Z)$, which is a Nemytskii (or superposition) operator, is well defined and bounded from $L_{\mathcal{P}}^q(\Omega; L_\eta^q(H)) \times L_{\mathcal{P}}^r(\Omega; L_\delta^2(K)) \times L_{\mathcal{P}}^r(\Omega; L_\delta^2(L_2(\Xi, K)))$ to $L_{\mathcal{P}}^r(\Omega; L_\delta^2(K))$. Continuity of this map follows in a similar way by adapting the classical argument that proves continuity of Nemytskii operators in this framework (see, e.g., [1]).

Coming back to the proof of point (iii), we take $X^1, X^2 \in L_{\mathcal{P}}^q(\Omega; L_\eta^q(H))$ and let (Y^1, Z^1) and (Y^2, Z^2) be the corresponding solutions. Then $(\overline{Y}, \overline{Z}) := (Y^1 - Y^2, Z^1 - Z^2)$ solves the equation

$$\overline{Y}_\tau - \overline{Y}_T + \int_\tau^T \overline{Z}_\sigma\,dW_\sigma + \lambda \int_\tau^T \overline{Y}_\sigma\,d\sigma$$

$$= -\int_\tau^T \left[\psi(X_\sigma^1, Y_\sigma^1, Z_\sigma^1) - \psi(X_\sigma^2, Y_\sigma^1 - \overline{Y}_\sigma, Z_\sigma^1 - \overline{Z}_\sigma)\right]d\sigma.$$

The estimate of Theorem 3.7 gives

$$|\overline{Y}|_{L_{\mathcal{P}}^r(\Omega;C_\delta(K))}^r + |\overline{Y}|_{L_{\mathcal{P}}^r(\Omega;L_\delta^2(K))}^r + |\overline{Z}|_{L_{\mathcal{P}}^r(\Omega;L_\delta^2(L_2(\Xi,K)))}^r$$

$$\leq c\mathbb{E}\left(\int_0^\infty e^{2\delta\sigma}|\psi(X_\sigma^1, Y_\sigma^1, Z_\sigma^1) - \psi(X_\sigma^2, Y_\sigma^1, Z_\sigma^1)|\,d\sigma\right)^{r/2}.$$

The right-hand side of this inequality can be made arbitrarily small provided $|X^1 - X^2|_{L_{\mathcal{P}}^q(\Omega;L_\delta^q(K))}$ is chosen sufficiently small, due to the continuity of the map $(X, Y, Z) \mapsto \psi(X, Y, Z)$ introduced above.

Point (iv) follows trivially from the previous ones, since $L_{\mathcal{P}}^q(\Omega; C_\eta(H)) \subset L_{\mathcal{P}}^q(\Omega; L_{\eta-\varepsilon}^q(H))$ for every $\varepsilon > 0$.

Now we address points (v)–(vii). We choose $r = (m+2)p$ and $\delta = \beta/(m+2)$. Since $r > p$ and $\beta < \delta$, therefore $\mathcal{K}_\delta^r \subset \mathcal{K}_\beta^p$, $L_{\mathcal{P}}^r(\Omega; C_\delta(H)) \subset L_{\mathcal{P}}^p(\Omega; C_\beta(H))$, so that existence, uniqueness and continuity with respect to $X \in L_{\mathcal{P}}^q(\Omega; C_\eta(H))$ of a solution $(Y(X), Z(X)) \in \mathcal{K}_\beta^p$ with $Y(X) \in L_{\mathcal{P}}^p(\Omega; C_\beta(H))$ follow from the

previous points. Before proceeding, we prove the inequality

$$\mathbb{E}\left(\int_0^\infty e^{2\beta\sigma}|\nabla_x\psi(\sigma,X_\sigma,Y_\sigma,Z_\sigma)N_\sigma|^2\,d\sigma\right)^{p/2}$$

$$(3.25)\qquad \le c\big(1+|Z|_{L_{\mathcal{P}}^r(\Omega;L_\delta^2(L_2(\Xi,K)))}\big)^p$$

$$\times\big(1+|X|_{L_{\mathcal{P}}^q(\Omega;C_\eta(H))}+|Y|_{L_{\mathcal{P}}^r(\Omega;C_\delta(K))}\big)^{mp}|N|_{L_{\mathcal{P}}^q(\Omega;C_\eta(H))}^p.$$

Using Hypothesis 3.10(iii), an elementary inequality and the Hölder inequality with conjugate exponents r/p and $r/(r-p)$, we obtain

$$\mathbb{E}\left(\int_0^\infty e^{2\beta\sigma}|\nabla_x\psi(\sigma,X_\sigma,Y_\sigma,Z_\sigma)N_\sigma|^2\,d\sigma\right)^{p/2}$$

$$\le c\mathbb{E}\left(\int_0^\infty e^{2\beta\sigma}(1+|Z_\sigma|)^2(1+|X_\sigma|+|Y_\sigma|)^{2m}|N_\sigma|^2\,d\sigma\right)^{p/2}$$

$$(3.26)\qquad \le c\mathbb{E}\left(\sup_{\tau\ge0}e^{p(\beta-\delta)\tau}(1+|X_\tau|+|Y_\tau|)^{pm}|N_\tau|^p\right.$$

$$\left.\times\left(\int_0^\infty e^{2\delta\sigma}(1+|Z_\sigma|)^2\,d\sigma\right)^{p/2}\right)$$

$$\le cI_1^{p/r}I_2^{(r-p)/r},$$

where

$$I_1=\mathbb{E}\left(\int_0^\infty e^{2\delta\sigma}(1+|Z_\sigma|)^2\,d\sigma\right)^{r/2},$$

$$I_2=\mathbb{E}\left(\sup_{\tau\ge0}e^{pr(\beta-\delta)/(r-p)\tau}(1+|X_\tau|+|Y_\tau|)^{pmr/(r-p)}|N_\tau|^{pr/(r-p)}\right).$$

Assuming for the moment that $m>0$, we write

$$\exp\left(\frac{pr(\beta-\delta)}{r-p}\tau\right)=\exp\left(\frac{\delta rpm}{r-p}\tau\right)\exp\left(\frac{pr(\beta-\delta(1+m))}{r-p}\tau\right)$$

and use the Hölder inequality again, with conjugate exponents $(r-p)/(pm)$ and $(r-p)/(r-p(m+1))$, to obtain

$$I_2\le I_{21}^{pm/(r-p)}I_{22}^{(r-p(m+1))/(r-p)},$$

where

$$I_{21}=\mathbb{E}\left(\sup_{\tau\ge0}e^{\delta r\tau}(1+|X_\tau|+|Y_\tau|)^r\right),$$

$$I_{22}=\mathbb{E}\left(\sup_{\tau\ge0}e^{pr(\beta-\delta(m+1))/(r-p(m+1))\tau}|N_\tau|^{pr/(r-p(m+1))}\right).$$

Taking into account that $\delta < 0$, we have

$$I_1 \le c\big(1 + |Z|_{L^r_{\mathscr{P}}(\Omega; L^2_\delta(L_2(\Xi, K)))}\big)^r,$$

$$I_{21} \le c\big(1 + |X|_{L^r_{\mathscr{P}}(\Omega; C_\delta(H))} + |Y|_{L^r_{\mathscr{P}}(\Omega; C_\delta(K))}\big)^r$$

$$\le c\big(1 + |X|_{L^q_{\mathscr{P}}(\Omega; C_\eta(H))} + |Y|_{L^r_{\mathscr{P}}(\Omega; C_\delta(K))}\big)^r,$$

$$I_{22} = \mathbb{E}\left(\sup_{\tau \ge 0} e^{p\delta(m+2)}|N_\tau|^{p(m+2)}\right) = |N|^{p(m+2)}_{L^{p(m+2)}_{\mathscr{P}}(\Omega; C_\delta(H))}$$

$$\le |N|^{p(m+2)}_{L^q_{\mathscr{P}}(\Omega; C_\eta(H))} = |N|^r_{L^q_{\mathscr{P}}(\Omega; C_\eta(H))}.$$

Substituting into the previous inequalities yields (3.25). The proof of (3.25) in the case $m = 0$ is even easier.

By similar passages one can prove more, namely that the Nemytskii operator $(X, N, Y, Z) \to \nabla_x \psi(X, Y, Z)N$ is bounded and continuous from the space

$$K^\# := L^q_{\mathscr{P}}(\Omega; C_\eta(H)) \times L^q_{\mathscr{P}}(\Omega; C_\eta(H))$$

$$\times L^r_{\mathscr{P}}(\Omega; C_\delta(K)) \times L^r_{\mathscr{P}}(\Omega; L^2_\delta(L_2(\Xi, K)))$$

to $L^p_{\mathscr{P}}(\Omega; L^2_\beta(K))$.

It is convenient now to introduce another backward stochastic equation; we will eventually show that it is satisfied by the derivatives of (Y, Z) with respect to X. For any $(X, N, Y, Z) \in K^\#$, we look for $(\widehat{Y}(X, N, Y, Z), \widehat{Z}(X, N, Y, Z)) = (\widehat{Y}, \widehat{Z}) \in \mathcal{K}^p_\beta$ solving

$$\widehat{Y}_\tau - \widehat{Y}_T + \lambda \int_\tau^T \widehat{Y}_\sigma \, d\sigma + \int_\tau^T \widehat{Z}_\sigma \, dW_\sigma$$

$$(3.27) \qquad = -\int_\tau^T \nabla_x \psi(X_\sigma, Y_\sigma, Z_\sigma)N_\sigma \, d\sigma$$

$$-\int_\tau^T \nabla_y \psi(X_\sigma, Y_\sigma, Z_\sigma)\widehat{Y}_\sigma \, d\sigma - \int_\tau^T \nabla_z \psi(X_\sigma, Y_\sigma, Z_\sigma)\widehat{Z}_\sigma \, d\sigma.$$

Hypothesis 3.10(iii) implies that $\nabla_y \psi$ and $\nabla_z \psi$ are bounded and that

$$\langle \nabla_y \psi(x, y, z)k, k \rangle \ge \mu|k|^2, \qquad x \in H, y, k \in K, z \in L_2(\Xi, K).$$

Together with (3.25) this shows that Theorem 3.7 applies to (3.27) and yields existence and uniqueness of a solution in $\mathcal{K}^p_\beta$ for $\lambda > -(\beta + \mu - L_z^2/2)$. Moreover, we have $\widehat{Y} \in L^p_{\mathscr{P}}(\Omega; C_\beta(H))$ and the following estimate holds:

$$\mathbb{E} \sup_{\tau \ge 0} e^{p\beta\tau}|\widehat{Y}_\tau|^p + \mathbb{E}\left(\int_0^\infty e^{2\beta\sigma}|\widehat{Y}_\sigma|^2 \, d\sigma\right)^{p/2} + \mathbb{E}\left(\int_0^\infty e^{2\beta\sigma}|\widehat{Z}_\sigma|^2 \, d\sigma\right)^{p/2}$$

$$(3.28) \qquad \le c\big(1 + |Z|_{L^r_{\mathscr{P}}(\Omega; L^2_\delta(L_2(\Xi, K)))}\big)^p$$

$$\times \big(1 + |X|_{L^q_{\mathscr{P}}(\Omega; C_\eta(H))} + |Y|_{L^r_{\mathscr{P}}(\Omega; C_\delta(K))}\big)^{mp}|N|^p_{L^q_{\mathscr{P}}(\Omega; C_\eta(H))}.$$

The continuity of the map $(X, N, Y, Z) \to (\widehat{Y}(X, N, Y, Z), \widehat{Z}(X, N, Y, Z))$ from $K^{\#}$ to $\mathcal{K}_{\beta}^{p}$ and the continuity of the map $(X, N, Y, Z) \to \widehat{Y}(X, N, Y, Z)$ from $K^{\#}$ to $L_{\mathcal{P}}^{p}(\Omega; C_{\beta}(H))$ can be verified directly as in point (iii) above.

It remains to prove that if $X, N \in L_{\mathcal{P}}^{q}(\Omega; C_{\eta}(H))$, then the directional derivative of the process $(Y(X), Z(X))$ in the direction N is given by $(\widehat{Y}(X, N, Y(X), Z(X)), \widehat{Z}(X, N, Y(X), Z(X)))$. Let us define

$$\overline{Y}^{\varepsilon} := \frac{1}{\varepsilon}[Y(X + \varepsilon N) - Y(X)] - \widehat{Y}(X, N, Y(X), Z(X)),$$

$$\overline{Z}^{\varepsilon} := \frac{1}{\varepsilon}[Z(X + \varepsilon N) - Z(X)] - \widehat{Z}(X, N, Y(X), Z(X)).$$

We will prove that $(\overline{Y}^{\varepsilon}, \overline{Z}^{\varepsilon}) \to 0$ in $\mathcal{K}_{\beta}^{p}$ and $\overline{Y}^{\varepsilon} \to 0$ in $L_{\mathcal{P}}^{p}(\Omega; C_{\beta}(H))$ for $\varepsilon \to 0$. For short, we let $Y = Y(X)$, $Z = Z(X)$, $Y^{\varepsilon} = Y(X + \varepsilon N)$, $Z^{\varepsilon} = Z(X + \varepsilon N)$, $\widehat{Y} = \widehat{Y}(X, N, Y(X), Z(X))$ and $\widehat{Z} = \widehat{Z}(X, N, Y(X), Z(X))$. Then $(\overline{Y}^{\varepsilon}, \overline{Z}^{\varepsilon})$ is a solution of

$$\overline{Y}_{\tau}^{\varepsilon} - \overline{Y}_{T}^{\varepsilon} + \lambda \int_{\tau}^{T} \overline{Y}_{\sigma}^{\varepsilon} \, d\sigma + \int_{\tau}^{T} \overline{Z}_{\sigma}^{\varepsilon} \, d\sigma = -\int_{\tau}^{T} \nu^{\varepsilon}(\sigma) \, d\sigma,$$

where $\nu^{\varepsilon} = \nu_{1}^{\varepsilon} + \nu_{2}^{\varepsilon}$ and

$$\nu_{1}^{\varepsilon}(\sigma) = \frac{1}{\varepsilon}[\psi(X_{\sigma} + \varepsilon N_{\sigma}, Y_{\sigma}^{\varepsilon}, Z_{\sigma}^{\varepsilon}) - \psi(X_{\sigma}, Y_{\sigma}^{\varepsilon}, Z_{\sigma}^{\varepsilon})] - \nabla_{x} \psi(X_{\sigma}, Y_{\sigma}, Z_{\sigma}) N_{\sigma},$$

$$\nu_{2}^{\varepsilon}(\sigma) = \frac{1}{\varepsilon}[\psi(X_{\sigma}, Y_{\sigma}^{\varepsilon}, Z_{\sigma}^{\varepsilon}) - \psi(X_{\sigma}, Y_{\sigma}, Z_{\sigma})]$$
$$- \nabla_{y} \psi(X_{\sigma}, Y_{\sigma}, Z_{\sigma}) \widehat{Y}_{\sigma} - \nabla_{z} \psi(X_{\sigma}, Y_{\sigma}, Z_{\sigma}) \widehat{Z}_{\sigma}.$$

Writing

$$\psi(X_{\sigma} + \varepsilon N_{\sigma}, Y_{\sigma}^{\varepsilon}, Z_{\sigma}^{\varepsilon}) - \psi(X_{\sigma}, Y_{\sigma}^{\varepsilon}, Z_{\sigma}^{\varepsilon}) = \int_{0}^{1} \frac{d}{d\zeta} \psi(X_{\sigma} + \varepsilon \zeta N_{\sigma}, Y_{\sigma}^{\varepsilon}, Z_{\sigma}^{\varepsilon}) \, d\zeta$$

gives

$$\begin{aligned}
(3.29) \qquad \nu_{1}^{\varepsilon}(\sigma) &= \int_{0}^{1} \nabla_{x} \psi(X_{\sigma} + \varepsilon \zeta N_{\sigma}, Y_{\sigma}^{\varepsilon}, Z_{\sigma}^{\varepsilon}) N_{\sigma} \, d\zeta \\
&\quad - \int_{0}^{1} \nabla_{x} \psi(\sigma, X_{\sigma}, Y_{\sigma}, Z_{\sigma}) N_{\sigma} \, d\zeta.
\end{aligned}$$

Similarly, starting from

$$\psi(X_{\sigma}, Y_{\sigma}^{\varepsilon}, Z_{\sigma}^{\varepsilon}) - \psi(X_{\sigma}, Y_{\sigma}, Z_{\sigma})$$
$$= \int_{0}^{1} \frac{d}{d\zeta} \psi(X_{\sigma}, Y_{\sigma} + \zeta(Y_{\sigma}^{\varepsilon} - Y_{\sigma}), Z_{\sigma} + \zeta(Z_{\sigma}^{\varepsilon} - Z_{\sigma})) \, d\zeta,$$

evaluating the integrals and rearranging terms, we conclude that

(3.30)
$$\overline{Y}_\tau^\varepsilon - \overline{Y}_T^\varepsilon + \lambda \int_\tau^T \overline{Y}_\sigma^\varepsilon \, d\sigma + \int_\tau^T \overline{Z}_\sigma^\varepsilon \, d\sigma$$
$$= -\int_\tau^T \left[v_1^\varepsilon(\sigma) + v_3^\varepsilon(\sigma) + \psi_1^\varepsilon(\sigma)\overline{Y}_\sigma^\varepsilon + \psi_2^\varepsilon(\sigma)\overline{Z}_\sigma^\varepsilon \right] d\sigma,$$

where

$$v_3^\varepsilon(\sigma) = \int_0^1 \left[\nabla_y \psi(X_\sigma, Y_\sigma + \zeta(Y_\sigma^\varepsilon - Y_\sigma), Z_\sigma + \zeta(Z_\sigma^\varepsilon - Z_\sigma)) \right.$$
$$\left. - \nabla_y \psi(X_\sigma, Y_\sigma, Z_\sigma) \right] \widehat{Y}_\sigma \, d\zeta$$
$$+ \int_0^1 \left[\nabla_z \psi(X_\sigma, Y_\sigma + \zeta(Y_\sigma^\varepsilon - Y_\sigma), Z_\sigma + \zeta(Z_\sigma^\varepsilon - Z_\sigma)) \right.$$
$$\left. - \nabla_z \psi(X_\sigma, Y_\sigma, Z_\sigma) \right] \widehat{Z}_\sigma \, d\zeta,$$

$$\psi_1^\varepsilon(\sigma) = \int_0^1 \nabla_y \psi(X_\sigma, Y_\sigma + \zeta(Y_\sigma^\varepsilon - Y_\sigma), Z_\sigma + \zeta(Z_\sigma^\varepsilon - Z_\sigma)) \, d\zeta,$$

$$\psi_2^\varepsilon(\sigma) = \int_0^1 \nabla_z \psi(X_\sigma, Y_\sigma + \zeta(Y_\sigma^\varepsilon - Y_\sigma), Z_\sigma + \zeta(Z_\sigma^\varepsilon - Z_\sigma)) \, d\zeta.$$

Theorem 3.7 applies to the backward equation (3.30) and gives, in particular, the estimate, for $\lambda > -(\beta + \mu - L_z^2/2)$,

$$\mathbb{E} \sup_{\tau \geq 0} e^{p\beta\tau} |\overline{Y}_\tau^\varepsilon|^p + \mathbb{E}\left(\int_0^\infty e^{2\beta\sigma} |\overline{Y}_\sigma^\varepsilon|^2 \, d\sigma \right)^{p/2} + \mathbb{E}\left(\int_0^\infty e^{2\beta\sigma} |\overline{Z}_\sigma^\varepsilon|^2 \, d\sigma \right)^{p/2}$$
$$\leq c\mathbb{E}\left(\int_0^\infty e^{2\beta\sigma} |v_1^\varepsilon(\sigma)|^2 \, d\sigma \right)^{p/2} + c\mathbb{E}\left(\int_0^\infty e^{2\beta\sigma} |v_3^\varepsilon(\sigma)|^2 \, d\sigma \right)^{p/2}.$$

Now we check that the right-hand side tends to 0 as $\varepsilon \to 0$. For the term containing v_3^ε, this follows from the dominated convergence theorem since we have $|v_3^\varepsilon(\sigma)| \leq c|\widehat{Y}_\sigma| + c|\widehat{Z}_\sigma|$.

To treat the other term, we first define, for all $x, g, n \in H$, $y \in K$ and $z \in L_2(\Xi, K)$, the function $\chi(x, g, n, y, z) = \int_0^1 \nabla_x \psi(x + \zeta g, y, z) n \, d\zeta$, and note that from (3.29) it follows that v_1^ε can be written as

$$v_1^\varepsilon(\sigma) = \chi(X_\sigma, \varepsilon N_\sigma, N_\sigma, Y_\sigma^\varepsilon, Z_\sigma^\varepsilon) - \chi(X_\sigma, 0, N_\sigma, Y_\sigma, Z_\sigma).$$

Let us consider the Nemytskii operator $(X, M, N, Y, Z) \to \chi(X, M, N, Y, Z)$ associated to χ. We can show that it is a bounded and continuous mapping from the space

$$L_{\mathcal{P}}^q(\Omega; C_\eta(H)) \times L_{\mathcal{P}}^q(\Omega; C_\eta(H)) \times L_{\mathcal{P}}^q(\Omega; C_\eta(H))$$
$$\times L_{\mathcal{P}}^r(\Omega; C_\delta(K)) \times L_{\mathcal{P}}^r(\Omega; L_\delta^2(L_2(\Xi, K)))$$

to $L_{\mathcal{P}}^p(\Omega; L_\beta^2(K))$. The proof of this fact is based on the continuity of χ and the estimate $|\chi(x, g, n, y, z)| \leq L|n|(1 + |z|)(1 + |x| + |g| + |y|)^m$, which follow from Hypothesis 3.10, and it is obtained in the same way as for the operator $(X, N, Y, Z) \to \nabla_x \psi(X, Y, Z)N$ introduced before. The convergence $(X, \varepsilon N, N, Y^\varepsilon, Z^\varepsilon) \to (X, 0, N, Y, Z)$ in the appropriate space follows from point (iii) proved above and implies

$$\mathbb{E}\left(\int_0^\infty e^{2\beta\sigma}|v_1^\varepsilon(\sigma)|^2\,d\sigma\right)^{p/2} = |v_1^\varepsilon|_{L_{\mathcal{P}}^p(\Omega; L_\beta^2(K))}^{p/2} \to 0.$$

Finally, (3.23) follows plugging (3.21) into (3.28). □

REMARK 3.12.   If, in addition to Hypothesis 3.10, we suppose that $\psi(\cdot, 0, 0)$ is bounded [i.e., $\sup_{x \in H}|\psi(x, 0, 0)| < \infty$], then, with identical proof, points (i)–(iv) in Proposition 3.11 can be improved by dropping the limitation imposed by (3.20) on the choice of q and η. More precisely, for arbitrary $r > 2$, $\delta < 0$, $q \in [1, \infty)$ and $\eta \in \mathbb{R}$, the statements of points (i)–(iv) remain true with the estimate (3.21) replaced by

$$
\begin{aligned}
(3.31) \quad & \mathbb{E}\sup_{\tau \geq 0}|Y_\tau(X)|^r e^{r\delta\tau} + \mathbb{E}\left(\int_0^\infty e^{2\delta\sigma}|Y_\sigma(X)|^2\,d\sigma\right)^{r/2} \\
& + \mathbb{E}\left(\int_0^\infty e^{2\delta\sigma}|Z_\sigma(X)|^2\,d\sigma\right)^{r/2} \leq c,
\end{aligned}
$$

where c is a constant independent from the process X.

REMARK 3.13.   Now let us assume that $\psi(\cdot, 0, 0)$ is bounded and that Hypothesis 3.10 holds with $m = 0$. Then the restriction (3.22) can be weakened and points (v)–(vii) in Proposition 3.11 can be improved as follows. If, given $p > 2$ and $\beta < 0$, we choose q and η satisfying

$$q > p, \qquad \eta > \beta,$$

then the statements of points (v)–(vii) remain true with the estimate (3.23) replaced by

$$
\begin{aligned}
(3.32) \quad & \mathbb{E}\sup_{\tau \geq 0} e^{p\beta\tau}|\nabla Y_\tau(X)N|^p + \mathbb{E}\left(\int_0^\infty e^{2\beta\sigma}|\nabla Y_\sigma(X)N|^2\,d\sigma\right)^{p/2} \\
& + \mathbb{E}\left(\int_0^\infty e^{2\beta\sigma}|\nabla Z_\sigma(X)N|^2\,d\sigma\right)^{p/2} \leq c|N|_{L_{\mathcal{P}}^q(\Omega; C_\eta(H))}^p.
\end{aligned}
$$

Indeed, let us take $r > 2$ so large and $\delta < 0$ so small that

$$r > p, \qquad \frac{pr}{r - p} \leq q, \qquad \beta < \delta, \qquad \beta - \delta < \eta.$$

Then from (3.26) and (3.31) it follows that

$$\mathbb{E}\left(\int_0^\infty e^{2\beta\sigma}|\nabla_x\psi(\sigma, X_\sigma, Y_\sigma, Z_\sigma)N_\sigma|^2\, d\sigma\right)^{p/2}$$

$$\leq c\left[\mathbb{E}\left(\sup_{\tau\geq 0} e^{pr(\beta-\delta)/(r-p)\tau}|N_\tau|^{pr/(r-p)}\right)\right]^{(r-p)/r} \leq c|N|^p_{L^q_{\mathcal{P}}(\Omega; C_\eta(H))}.$$

Starting from this inequality, which improves (3.25), the proof of points (v)–(vii) in Proposition 3.11 can be repeated with minor changes and (3.32) follows as well.

4. The forward equation. As in the previous sections, we denote by $\{W_\tau, \tau\geq 0\}$ a cylindrical Wiener process with values in a Hilbert space Ξ, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Now we consider the Itô stochastic differential equation for an unknown process $\{X_\tau, \tau\geq 0\}$ with values in a Hilbert space H:

$$(4.1) \qquad \begin{aligned} X_\tau &= e^{\tau A}x + \int_0^\tau e^{(\tau-\sigma)A}F(X_\sigma)\, d\sigma \\ &\quad + \int_0^\tau e^{(\tau-\sigma)A}G(X_\sigma)\, dW_\sigma, \qquad \tau\geq 0. \end{aligned}$$

(See [22] as a reference starting study on forward SDEs in Hilbert spaces.)

We will first consider solvability of this equation for τ varying in an interval $[0, T]$ and later for $\tau\in\mathbb{R}_+$. In both cases our assumptions will be the following:

HYPOTHESIS 4.1. (i) *The operator A is the generator of a strongly continuous semigroup e^{tA}, $t\geq 0$, in the Hilbert space H. We denote by M, a two constants such that $|e^{tA}|\leq Me^{at}$ for $t\geq 0$.*

(ii) *The mapping $F: H\to H$ satisfies, for some constant $L > 0$,*

$$|F(x) - F(y)|\leq L\,|x - y|, \qquad x, y\in H.$$

(iii) *The variable G denotes a mapping from H to $L(\Xi, H)$ such that for every $\xi\in\Xi$ the map $G(\cdot)\xi: H\to H$ is measurable, $e^{tA}G(x)\in L_2(\Xi, H)$ for every $t > 0$ and $x\in H$, and*

$$(4.2) \qquad |e^{tA}G(x)|_{L_2(\Xi, H)}\leq Lt^{-\gamma}e^{at}(1 + |x|),$$

$$|e^{tA}G(x) - e^{tA}G(y)|_{L_2(\Xi, H)}\leq Lt^{-\gamma}e^{at}|x - y|, \qquad t > 0, x, y\in H,$$

$$(4.3) \qquad |G(x)|_{L(\Xi, H)}\leq L(1 + |x|), \qquad x\in H,$$

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

(iv) *For every $t > 0$, we have $F(\cdot)\in\mathcal{G}^1(H, H)$ and $e^{tA}G(\cdot)\in\mathcal{G}^1(H, L_2(\Xi, H))$.*

We start by recalling a well-known result on solvability of (4.1) on a bounded interval; see, for example, [7].

PROPOSITION 4.2. *Under the assumptions of Hypothesis* 4.1 [*only points* (i)–(iii) *are needed*], *for every* $q \in [2, \infty)$ *and* $T > 0$ *there exists a unique process* $X \in L_{\mathcal{P}}^q(\Omega; C([0, T]; H))$ *solution of* (4.1). *Moreover,*

$$\text{(4.4)} \qquad \mathbb{E} \sup_{\tau \in [0,T]} |X_\tau|^q \le C(1 + |x|)^q$$

for some constant C *depending only on* q, γ, T, L, a *and* M.

More generally, in the following discussion we need to consider a stochastic equation on an arbitrary interval $[t, T] \subset [0, T]$:

$$
\begin{aligned}
\text{(4.5)} \qquad X_\tau = {}& e^{(\tau-t)A}x + \int_t^\tau e^{(\tau-\sigma)A} F(X_\sigma)\, d\sigma \\
& + \int_t^\tau e^{(\tau-\sigma)A} G(X_\sigma)\, dW_\sigma, \qquad \tau \in [t, T].
\end{aligned}
$$

We set $X_\tau = x$ for $\tau \in [0, t)$ and we denote by $X(\tau, t, x)$, $\tau \in [0, T]$, the solution.

Our next aim is to prove Proposition 4.5, which is basic for the proof of one of our main results, Theorem 6.1. To this end, we need to recall the following results from [12]. The first result deals with regularity of $\{X(\tau, t, x), \tau \in [0, T]\}$ with respect to x; the second with its regularity in the sense of the Malliavin calculus.

PROPOSITION 4.3. *Assume Hypothesis* 4.1. *Then, for every* $q \in [2, \infty)$ *and* $T > 0$, *the following properties hold.*

(i) *The map* $(t, x) \mapsto X(\cdot, t, x)$ *belongs to* $\mathcal{G}^{0,1}([0, T] \times H, L_{\mathcal{P}}^q(\Omega; C([0, T]; H)))$.

(ii) *Denoting by* $\nabla_x X$ *the partial Gâteaux derivative, for every direction* $h \in H$, *the directional derivative process* $\nabla_x X(\tau, t, x)h$, $\tau \in [0, T]$, *solves,* $\mathbb{P}$-*a.s., the equation*

$$
\begin{aligned}
\nabla_x X(\tau, t, x)h = {}& e^{(\tau-t)A}h + \int_t^\tau e^{(\tau-\sigma)A} \nabla_x F\big(\sigma, X(\sigma, t, x)\big) \\
& \qquad\qquad \times \nabla_x X(\sigma, t, x)h\, d\sigma \\
\text{(4.6)} \qquad & + \int_t^\tau \nabla_x\big(e^{(\tau-\sigma)A} G(\sigma, X(\sigma, t, x))\big) \\
& \qquad\qquad \times \nabla_x X(\sigma, t, x)h\, dW_\sigma, \qquad \tau \in [t, T], \\
\nabla_x X(\tau, t, x)h = {}& h, \qquad\qquad\qquad\qquad\qquad\qquad\quad \tau \in [0, t).
\end{aligned}
$$

(iii) *Finally,* $|\nabla_x X(\tau, t, x)h|_{L_{\mathcal{P}}^q(\Omega; C([0,T]; H))} \le c|h|$ *for some constant* c.

To state the following result, we need to recall some basic definitions from the Malliavin calculus. We refer the reader to [20] for a detailed exposition; [16] treats the extensions to Hilbert space-valued random variables and processes.

For every $h \in L^2([0, T]; \Xi)$ we denote by $W(h)$ the integral $\int_0^T \langle h(t), dW(t) \rangle_\Xi$. Given a Hilbert space K, let S_K be the set of K-valued random variables F of the form

$$F = \sum_{j=1}^m f_j\big(W(h_1), \ldots, W(h_n)\big) e_j,$$

where $h_1, \ldots, h_n \in L^2([0, T]; \Xi)$, $\{e_j\}$ is a basis of K and $f_1, \ldots, f_m$ are infinitely differentiable functions $\mathbb{R}^n \to \mathbb{R}$ bounded together with all their derivatives. The Malliavin derivative DF of $F \in S_K$ is defined as the process $D_s F$, $s \in [0, T]$,

$$D_s F = \sum_{j=1}^m \sum_{k=1}^n \partial_k f_j\big(W(h_1), \ldots, W(h_n)\big) e_j \otimes h_k(s),$$

with values in $L_2(\Xi, K)$. By ∂_k we denote the partial derivatives with respect to the kth variable and by $e_j \otimes h_k(s)$ denote the operator $u \mapsto e_j \langle h_k(s), u \rangle_\Xi$. It is known that the operator $D : S_K \subset L^2(\Omega; K) \to L^2(\Omega \times [0, T]; L_2(\Xi; K))$ is closable. We denote by $\mathbb{D}^{1,2}(K)$ the domain of its closure, and use the same letter to denote D and its closure:

$$D : \mathbb{D}^{1,2}(K) \subset L^2(\Omega; K) \to L^2\big(\Omega \times [0, T]; L_2(\Xi; K)\big).$$

The adjoint operator of D,

$$\delta : \mathrm{dom}(\delta) \subset L^2\big(\Omega \times [0, T]; L_2(\Xi; K)\big) \to L^2(\Omega; K),$$

is called the Skorohod integral. It is known that $\mathrm{dom}(\delta)$ contains $L_{\mathcal{P}}^2(\Omega \times [0, T];$ $L_2(\Xi; K))$ and the Skorohod integral of a process in this space coincides with the Itô integral; $\mathrm{dom}(\delta)$ also contains the class $\mathbb{L}^{1,2}(L_2(\Xi; K))$, the latter being defined as the space of processes $u \in L^2(\Omega \times [0, T]; L_2(\Xi; K))$ such that $u_r \in \mathbb{D}^{1,2}(L_2(\Xi, K))$ for a.e. $r \in [0, T]$ and there exists a measurable version of $D_s u_r$ satisfying

$$\|u\|_{\mathbb{L}^{1,2}(L_2(\Xi; K))}^2$$
$$= \|u\|_{L^2(\Omega \times [0,T]; L_2(\Xi; K))}^2 + \mathbb{E} \int_0^T \int_0^T \|D_s u_r\|_{L_2(\Xi, L_2(\Xi, K))}^2 \, dr \, ds < \infty.$$

Moreover, $\|\delta(u)\|_{L^2(\Omega; K)}^2 \le \|u\|_{\mathbb{L}^{1,2}(L_2(\Xi; K))}^2$. The definition of $\mathbb{L}^{1,2}(K)$ for an arbitrary Hilbert space K is entirely analogous.

Finally, we recall that if $F \in \mathbb{D}^{1,2}(K)$ is $\mathcal{F}_t$-adapted, then $DF = 0$ a.s. on $\Omega \times (t, T]$.

With the previous notation we have the following result, proved in [12].

PROPOSITION 4.4. *Assume Hypothesis 4.1. Then the following properties hold.*

   (i) *The process* $X = \{X(\tau, t, x), \tau \in [0, T]\}$ *belongs to* $\mathbb{L}^{1,2}(H)$.
  (ii) *The variables* $X(\tau, t, x) \in \mathbb{D}^{1,2}(H)$ *for every* $\tau \in [0, T]$.
 (iii) *For a.a.* s, τ *such that* $t \leq s \leq \tau \leq T$, *we have*

$$(4.7) \qquad D_s X(\tau, t, x) = \nabla_x X(\tau, s, X(s, t, x)) G(s, X(s, t, x)), \qquad \mathbb{P}\text{-}a.s.$$

 (iv) *For a.a.* $s \in [t, T]$, *we have*

$$(4.8) \qquad D_s X(T, t, x) = \nabla_x X(T, s, X(s, t, x)) G(s, X(s, t, x)), \qquad \mathbb{P}\text{-}a.s.$$

Now, for $\xi \in \Xi$, denote by W^ξ the real Wiener process defined by $W^\xi_\tau :=$ $\langle \xi, W_\tau \rangle$, $\tau \in [0, T]$. We also set $X_\tau = X(\tau, 0, x)$ for simplicity. Given a function $u : H \to \mathbb{R}$, we investigate the existence of the joint quadratic variation of the process $\{u(X_\tau), \tau \in [0, T]\}$ with W^ξ. As usual, this is defined for every $\tau \in [0, T]$ as the limit in probability of

$$\sum_{i=1}^{n} (u_{\tau_i} - u_{\tau_{i-1}})(W^\xi_{\tau_i} - W^\xi_{\tau_{i-1}}),$$

where $\{\tau_i\}$, $0 = \tau_0 < \tau_1 < \cdots < \tau_n = \tau$, is an arbitrary subdivision of $[0, \tau]$ whose mesh tends to 0. Existence of the joint quadratic variation is not trivial. Indeed, due to the occurrence of convolution type integrals in (4.5), it is not obvious that the process X is a semimartingale. Moreover, even in this case, the process $u(X)$ might fail to be a semimartingale if u is not twice differentiable, since the Itô formula does not apply. Nevertheless, the following result holds true. Its proof could be deduced from generalization of some results obtained in [21] to the infinite-dimensional case, but we prefer to give a simpler direct proof.

PROPOSITION 4.5. *Assume Hypothesis* 4.1 *and let* u *be a function in* $\mathcal{G}^1(H, \mathbb{R})$ *having polynomial growth together with its derivative. Then the process* $\{u(X_\tau), \tau \in [0, T]\}$ *admits a joint quadratic variation process* V *with* W^ξ, *given by*

$$V_\tau = \int_0^\tau \nabla u(X_\sigma) G(X_\sigma) \xi \, d\sigma, \qquad \tau \in [0, T].$$

PROOF. Let us denote $u_\tau = u(X_\tau)$ for simplicity. A chain rule for the class $\mathcal{G}^1$ and the Malliavin derivative operator holds; see [12] for details. It follows that, by the assumptions on u for every $\tau \in [0, T]$, we have $u_\tau \in \mathbb{D}^{1,2}(\mathbb{R})$ and $Du_\tau = \nabla u(X_\tau) D X_\tau$. Taking into account (4.7) for a.e. $s \in [0, \tau]$, we obtain

$$(4.9) \qquad D_s u_\tau \xi = \nabla u(X_\tau) \nabla_x X(\tau, s, X_s) G(X_s)\xi, \qquad \mathbb{P}\text{-}a.s.,$$

whereas $D_s u_\tau \xi = 0$ $\mathbb{P}$-*a.s. for a.e.* $s \in (\tau, T]$.

Let us now compute the joint quadratic variation of u and W^ξ. Let $t = \tau_0 < \tau_1 < \cdots < \tau_n = \tau$ be a subdivision of $[0, \tau] \subset [0, T]$. By well-known rules of Malliavin calculus (see [21], Theorem 3.2, or [16], Proposition 2.11), we have

$$\big(u_{\tau_i} - u_{\tau_{i-1}}\big)\big(W^\xi_{\tau_i} - W^\xi_{\tau_{i-1}}\big)$$

$$= \big(u_{\tau_i} - u_{\tau_{i-1}}\big) \int_{\tau_{i-1}}^{\tau_i} \langle \xi, dW_s \rangle$$

$$= \int_{\tau_{i-1}}^{\tau_i} D_s\big(u_{\tau_i} - u_{\tau_{i-1}}\big)\xi \, ds + \int_{\tau_{i-1}}^{\tau_i} \big(u_{\tau_i} - u_{\tau_{i-1}}\big)\langle \xi, \widehat{dW}_s \rangle,$$

where we use the symbol $\widehat{dW}$ to denote the Skorohod integral. We note that $D_s u_{\tau_{i-1}} = 0$ for $s > \tau_{i-1}$, so recalling (4.9) and setting $U_n(s) = \sum_{i=1}^n (u_{\tau_i} - u_{\tau_{i-1}})\mathbb{1}_{(\tau_{i-1}, \tau_i]}(s)$, we obtain

$$\sum_{i=1}^n (u_{\tau_i} - u_{\tau_{i-1}})\big(W^\xi_{\tau_i} - W^\xi_{\tau_{i-1}}\big)$$

$$= \int_0^\tau U_n(s) \, \langle \xi, \widehat{dW}_s \rangle + \sum_{i=1}^n \int_{\tau_{i-1}}^{\tau_i} \nabla u(X_{\tau_i}) \nabla_x X(\tau_i, s, X_s) G(X_s) \xi \, ds.$$

By (4.9) and the continuity properties asserted in Proposition 4.3, it is easily verified that the maps $\tau \mapsto u_\tau$ and $\tau \mapsto Du_\tau \xi$ are continuous on $[0, T]$ with values in $L^2(\Omega; \mathbb{R})$ and $L^2(\Omega \times [0, T]; \mathbb{R})$, respectively. In particular, $U_n \to 0$ in $\mathbb{L}^{1,2}(\mathbb{R})$, which implies that the Skorohod integral in the last equation tends to zero in $L^2(\Omega; \mathbb{R})$. Letting the mesh of the subdivision tend to 0, we obtain

$$\sum_{i=1}^n (u_{\tau_i} - u_{\tau_{i-1}})\big(W^\xi_{\tau_i} - W^\xi_{\tau_{i-1}}\big) \to V_\tau$$

in probability, which completes the proof of the proposition. $\square$

In the rest of this section we consider (4.1) for τ varying in $\mathbb{R}_+$. By Proposition 4.2 and the arbitrariness of T in its statement, the solution is defined for every $\tau \geq 0$. To stress dependence on the parameter $x \in H$, the solution starting from $X_0 = x$ will be denoted by $X(x)$. Notice that, with the notation previously used in this section, $X_\tau(x) = X(\tau, 0, x)$.

We recall that the spaces $L^q_{\mathcal{P}}(\Omega; L^q_\eta(H))$, $L^q_{\mathcal{P}}(\Omega; C_\eta(H))$ and $\mathcal{H}^q_\eta$ were defined for arbitrary $\eta \in \mathbb{R}$ and $q \in [1, \infty)$ in Section 2.

PROPOSITION 4.6.   *Assume that Hypothesis* 4.1 *holds. Then for all* $q \in [1, \infty)$, *there exists a constant* $\eta(q)$, *depending also on* γ, L, a *and* M, *with the following properties*:

   (i) *For all* $x \in H$, *the process* $X(x)$, *solution of* (4.1), *is in* $\mathcal{H}^q_{\eta(q)}$.

(ii) *For a suitable constant $C > 0$, we have*

$$(4.10) \quad \mathbb{E} \sup_{\tau \geq 0} e^{\eta(q)q\tau} |X_\tau|^q + \mathbb{E} \int_0^\infty e^{\eta(q)q\sigma} |X_\sigma|^q \, d\sigma \leq C(1 + |x|)^q.$$

(iii) *The map $x \mapsto X(x)$ belongs to $\mathcal{G}^1(H, \mathcal{H}_{\eta(q)}^q)$ and its derivative is uniformly bounded,*

$$(4.11) \quad |\nabla X(x)h|_{\mathcal{H}_{\eta(q)}^q} \leq C|h|, \qquad x, h \in H,$$

for a suitable constant C.

PROOF. In the following, the letters M, a, L and γ denote the constants that appear in Hypothesis 4.1. Clearly it is enough to prove the claim for q large, so we can assume that $q > (1 - 2\gamma)^{-1}$.

We define a mapping $\Phi : \mathcal{H}_\eta^q \times H \to \mathcal{H}_\eta^q$ by

$$(4.12) \quad \Phi(X, x)_\tau = e^{\tau A} x + \int_0^\tau e^{(\tau - \sigma)A} F(X_\sigma) \, d\sigma + \int_0^\tau e^{(\tau - \sigma)A} G(X_\sigma) \, dW_\sigma,$$

$$\tau \geq 0.$$

We are going to show that, provided η is suitably chosen, $\Phi(\cdot, x)$ is well defined and that it is a contraction in $\mathcal{H}_\eta^q$, uniformly in x, that is, there exists $c < 1$ such that for every $x \in H$,

$$(4.13) \quad |\Phi(X^1, x) - \Phi(X^2, x)|_{\mathcal{H}_\eta^q} \leq c |X^1 - X^2|_{\mathcal{H}_\eta^q}, \qquad X^1, X^2 \in \mathcal{H}_\eta^q.$$

For simplicity, we only treat the case $F = 0$; the general case is handled in a similar way. We use the so-called factorization method (see, e.g., [8], Theorem 5.2.5). By the assumption on q we can take $\alpha \in (0, 1)$ such that

$$\frac{1}{q} < \alpha < \frac{1}{2} - \gamma,$$

and we define

$$c_\alpha^{-1} = \int_\sigma^\tau (\tau - s)^{\alpha - 1} (s - \sigma)^{-\alpha} \, ds.$$

Then, by the stochastic Fubini theorem,

$$\Phi(X, x)_\tau = e^{\tau A} x + c_\alpha \int_0^\tau \int_\sigma^\tau (\tau - s)^{\alpha - 1} (s - \sigma)^{-\alpha}$$

$$\times e^{(\tau - s)A} e^{(s - \sigma)A} \, ds \, G(\sigma, X_\sigma) \, dW_\sigma$$

$$= e^{\tau A} x + \Phi'(X)_\tau,$$

where we set

$$\Phi'(X)_\tau = c_\alpha \int_0^\tau (\tau - s)^{\alpha - 1} e^{(\tau - s)A} Y_s \, ds,$$

$$Y_s = \int_0^s (s - \sigma)^{-\alpha} e^{(s - \sigma)A} G(\sigma, X_\sigma) \, dW_\sigma.$$

Since $|e^{\tau A}x| \leq Me^{a\tau}|x|$, the process $e^{\tau A}x$, $\tau \geq 0$, belongs to $\mathcal{H}^q_\eta$ provided $a + \eta < 0$. Next we estimate $\Phi'(X)$ as

$$|\Phi'(X)_\tau| \leq c_\alpha \int_0^\tau (\tau - s)^{\alpha-1} Me^{a(\tau-s)}|Y_s|\, ds,$$

so that

$$(4.14) \qquad e^{\eta\tau}|\Phi'(X)_\tau x| \leq c_\alpha M \int_0^\tau (\tau - s)^{\alpha-1} e^{(a+\eta)(\tau-s)} e^{\eta s}|Y_s|\, ds.$$

Applying the Young inequality for convolutions in the space $L^q(0, \infty)$, we obtain

$$\left(\int_0^\infty e^{\eta\tau q}|\Phi'(X)_\tau|^q\, d\tau\right)^{1/q} \leq c_\alpha M \left(\int_0^\infty e^{\eta s q}|Y_s|^q\, ds\right)^{1/q} \int_0^\infty s^{\alpha-1} e^{(a+\eta)s}\, ds$$

and we conclude that

$$(4.15) \quad |\Phi'(X)|_{L^q_{\mathcal{P}}(\Omega; L^q_\eta(H))} \leq c_\alpha M |Y|_{L^q_{\mathcal{P}}(\Omega; L^q_\eta(H))} \int_0^\infty s^{\alpha-1} e^{(a+\eta)s}\, ds.$$

If we start again from (4.14) and apply the Hölder inequality, setting $q' = \frac{q}{q-1}$, we obtain

$$e^{\eta\tau}|\Phi'(X)_\tau x| \leq c_\alpha M \left(\int_0^\tau e^{\eta s q}|Y_s|^q\, ds\right)^{1/q} \left(\int_0^\tau s^{(\alpha-1)q'} e^{(a+\eta)sq'}\, ds\right)^{1/q'}$$

and we conclude that

$$(4.16) \quad \begin{aligned} &|\Phi'(X)|_{L^q_{\mathcal{P}}(\Omega; C_\eta(H))} \\ &\qquad \leq c_\alpha M |Y|_{L^q_{\mathcal{P}}(\Omega; L^q_\eta(H))} \left(\int_0^\infty s^{(\alpha-1)q'} e^{(a+\eta)sq'}\, ds\right)^{1/q'}. \end{aligned}$$

By the Burkholder–Davis–Gundy inequalities, taking into account the assumption (4.2), we have, for some constant c_q depending only on q,

$$\begin{aligned} \mathbb{E}|Y_s|^q &\leq c_q \mathbb{E}\left(\int_0^s (s-\sigma)^{-2\alpha}|e^{(s-\sigma)A}G(\sigma, X_\sigma)|^2_{L_2(\Xi, H)}\, d\sigma\right)^{q/2} \\ &\leq L^q c_q \mathbb{E}\left(\int_0^s (s-\sigma)^{-2\alpha-2\gamma} e^{2a(s-\sigma)}(1+|X_\sigma|)^2\, d\sigma\right)^{q/2}. \end{aligned}$$

It follows that

$$[\mathbb{E}|Y_s|^q]^{2/q} \leq L^2 c_q^{2/q} \int_0^s (s-\sigma)^{-2\alpha-2\gamma} e^{2a(s-\sigma)}[\mathbb{E}(1+|X_\sigma|)^q]^{2/q}\, d\sigma$$

so that

$$\begin{aligned} e^{2\eta s}[\mathbb{E}|Y_s|^q]^{2/q} &\leq C_1 \int_0^s (s-\sigma)^{-2\alpha-2\gamma} e^{2(a+\eta)(s-\sigma)} e^{2\eta\sigma}\, d\sigma \\ &\qquad + C_2 \int_0^s (s-\sigma)^{-2\alpha-2\gamma} e^{2(a+\eta)(s-\sigma)} e^{2\eta\sigma}[\mathbb{E}|X_\sigma|^q]^{2/q}\, d\sigma \end{aligned}$$

for suitable constants C_1, C_2. Applying the Young inequality for convolutions in the space $L^{q/2}(0, \infty)$, we obtain

$$\left(\int_0^\infty e^{q\eta s} \mathbb{E}|Y_s|^q \, ds \right)^{2/q}$$

$$\leq C_1 \int_0^\infty s^{-2\alpha - 2\gamma} e^{2(a+\eta)s} \, ds \left(\int_0^\infty e^{q\eta s} \, ds \right)^{2/q}$$

$$+ C_2 \int_0^\infty s^{-2\alpha - 2\gamma} e^{2(a+\eta)s} \, ds \left(\int_0^\infty e^{q\eta s} \mathbb{E}|X_\sigma|^q \, ds \right)^{2/q}.$$

This shows that $|Y|_{L^q_{\mathcal{P}}(\Omega; L^q_\eta(H))}$ is finite provided we assume $\eta < 0$, $a + \eta < 0$, and so the map Φ is well defined.

If X^1, X^2, are processes belonging to $\mathcal{H}^q_\eta$, and Y^1, Y^2 are defined accordingly, entirely analogous passages show that

$$|Y^1 - Y^2|_{L^q_{\mathcal{P}}(\Omega; L^q_\eta(H))}$$

$$\leq L c_q^{1/q} |X^1 - X^2|_{L^q_{\mathcal{P}}(\Omega; L^q_\eta(H))} \left(\int_0^\infty s^{-2\alpha - 2\gamma} e^{2(a+\eta)s} \, ds \right)^{1/2}.$$

Recalling the inequalities (4.15) and (4.16), and noting that the map $Y \mapsto \Phi'(X)$ is linear, we arrive at an explicit expression for the constant c in (4.13), and it is immediate to verify that $c < 1$ provided $\eta < 0$ is chosen sufficiently large. Let us fix such a value $\eta(q)$. The statement of point (i) is a consequence of the contraction principle. The estimate (4.10) also follows from the contraction property of $\Phi(\cdot, x)$.

Now we come to the regular dependence of the solution on the initial datum. To prove that the map $x \mapsto X(x)$ belongs to $\mathcal{G}^1(H, \mathcal{H}^q_{\eta(q)})$, by the parameter depending contraction principle (Proposition 2.1), it suffices to show that $\Phi \in \mathcal{G}^1(\mathcal{H}^q_{\eta(q)} \times H, \mathcal{H}^q_{\eta(q)})$. This follows easily from the following steps (see [12], Lemma 2.1, for details).

STEP 1. The variable Φ is continuous. This follows immediately from the contraction property of $\Phi(\cdot, x)$ mentioned above and the fact that $\Phi(X, \cdot)$ is continuous from H to $\mathcal{H}^q_{\eta(q)}$, which is easy to verify.

STEP 2. We claim that the directional derivative $\nabla_X \Phi(X, x; N)$ with respect to $X \in \mathcal{H}^q_{\eta(q)}$ in the direction $N \in \mathcal{H}^q_{\eta(q)}$ is the process

$$\nabla_X \Phi(X, x; N)_\tau = \int_0^\tau e^{(\tau - \sigma)A} \nabla F(X_\sigma) N_\sigma \, d\sigma$$

$$+ \int_0^\tau \nabla(e^{(\tau - \sigma)A} G(X_\sigma)) N_\sigma \, dW_\sigma, \qquad \tau \geq 0,$$

and, moreover, the mappings $(X, x) \mapsto \nabla_X \Phi(X, x; N)$ and $N \mapsto \nabla_X \Phi(X, x; N)$ are continuous.

We limit ourselves to proving this claim in the special case $F = 0$; the general case is a straightforward extension. For fixed $x \in H$ and for all $\tau \geq 0$, we define

$$
I_\tau^\varepsilon = \frac{1}{\varepsilon} \Phi(X + \varepsilon N, x)_\tau - \frac{1}{\varepsilon} \Phi(X, x)_\tau - \int_0^\tau \nabla_x (e^{(\tau-\sigma)A} G(X_\sigma)) N_\sigma \, dW_\sigma
$$

$$
= \int_0^\tau \left(\int_0^1 \left(\nabla_x (e^{(\tau-\sigma)A} G(X_\sigma + \zeta \varepsilon N_\sigma)) N_\sigma \right. \right.
$$

$$
\left. \left. - \nabla_x (e^{(\tau-\sigma)A} G(X_\sigma)) N_\sigma \right) d\zeta \right) dW_\sigma.
$$

Using the identity $\nabla(e^{(\tau-\sigma)A} G(x)) = e^{(\tau-s)A}(e^{(s-\sigma)A} G(x))$ and applying the factorization method as in the proof of Proposition 4.6, we get, for $1/q < \alpha < 1/2 - \gamma$,

$$
|I^\varepsilon|_{\mathcal{H}_{\eta(q)}^q}^q \leq c \mathbb{E} \int_0^\infty |Y_s^\varepsilon|^q \, ds,
$$

where

$$
Y_s^\varepsilon := \int_0^s (s - \sigma)^{-\alpha} \int_0^1 \left(\nabla(e^{(s-\sigma)A} G(X_\sigma + \zeta \varepsilon N_\sigma)) N_\sigma \right.
$$

$$
\left. - \nabla(e^{(s-\sigma)A} G(X_\sigma)) N_\sigma \right) d\zeta \, dW_\sigma
$$

and c, here and in the rest of this proof, denotes a suitable constant, whose value may change from line to line. Next we obtain, by the Burkholder–Davis–Gundy inequalities,

$$
\mathbb{E}|Y_s^\varepsilon|^q \leq c \mathbb{E} \left(\int_0^s (s - \sigma)^{-2\alpha} \right.
$$

$$
\times \left| \int_0^1 \left(\nabla(e^{(s-\sigma)A} G(X_\sigma + \zeta \varepsilon N_\sigma)) N_\sigma \right. \right.
$$

$$
\left. \left. - \nabla(e^{(s-\sigma)A} G(X_\sigma)) N_\sigma \right) d\zeta \right|_{L_2(\Xi, H)}^2 d\sigma \right)^{q/2},
$$

and setting

$$
f^\varepsilon(\sigma, s, \zeta) = e^{s\eta(q)} (s - \sigma)^{-\alpha} |\nabla(e^{(s-\sigma)A} G(X_\sigma + \zeta \varepsilon N_\sigma)) N_\sigma
$$

$$
- \nabla(e^{(s-\sigma)A} G(X_\sigma)) N_\sigma|_{L_2(\Xi, H)},
$$

we arrive at the inequality

$$
\mathbb{E} \int_0^\infty e^{s\eta(q)q} |Y_s^\varepsilon|^q \, ds \leq c \int_0^\infty \mathbb{E} \left(\int_0^s \left| \int_0^1 f^\varepsilon(\sigma, s, \zeta) \, d\zeta \right|^2 d\sigma \right)^{q/2} ds.
$$

To conclude that $\mathbb{E}\int_0^\infty |Y_s^\varepsilon|^q\,ds \to 0$ as $\varepsilon \to 0$, we use the dominated convergence theorem. Since we assume that $\nabla(e^{tA}G(x))h$ is continuous in x for every $h \in H$, $t > 0$, therefore $f^\varepsilon \to 0$ pointwise. Next we note that from Hypothesis 4.1(iii) and (iv) it follows that $|\nabla(e^{tA}G(x))h|_{L_2(\Xi,H)} \le Lt^{-\gamma}e^{at}|h|$, which implies $|f^\varepsilon(\sigma,s,\zeta)| \le ce^{s\eta(q)}(s-\sigma)^{-\alpha-\gamma}e^{a(s-\sigma)}|N_\sigma|$, and it remains to show that the integral

$$\int_0^\infty \mathbb{E}\left(\int_0^s |e^{s\eta(q)}(s-\sigma)^{-\alpha-\gamma}e^{a(s-\sigma)}|N_\sigma||^2\,d\sigma\right)^{q/2}\,ds$$

is finite. This is less than or equal to

$$\int_0^\infty \left(\int_0^s (s-\sigma)^{-2\alpha-2\gamma}e^{2(a+\eta(q))(s-\sigma)}e^{2\eta(q)\sigma}[\mathbb{E}|N_\sigma|^q]^{2/q}\,d\sigma\right)^{q/2}\,ds$$

$$\le \left(\int_0^\infty s^{-2\alpha-2\gamma}e^{2(a+\eta(q))s}\,ds\right)^{q/2}\int_0^\infty e^{q\eta(q)s}\mathbb{E}|N_s|^q\,ds$$

$$\le c|N|_{\mathcal{H}_{\eta(q)}^q}^q < \infty,$$

where we have used again Young's inequality for convolution in the space $L^{q/2}(0,\infty)$.

Now the existence and the required formula for $\nabla_X\Phi(X,x;N)$ have been proved. Continuity of the mappings $(X,x) \mapsto \nabla_X\Phi(X,x;N)$ and $N \mapsto \nabla_X\Phi(X,x;N)$ can be checked in a similar way.

STEP 3.   Finally, it is clear that the directional derivative $\nabla_x\Phi(X,x;h)$ in the direction $h \in H$ is the process $\nabla_x\Phi(X,x;h)_\tau = e^{\tau A}h$, $\tau \ge 0$, and that the mappings $(X,x) \mapsto \nabla_x\Phi(X,x;h)$ and $h \mapsto \nabla_x\Phi(X,x;h)$ are continuous.

It remains to prove inequality (4.11). Recalling that $X(x)$ is a fixed point of $\Phi(\cdot,x)$, by the contraction property of Φ, we obtain, for some $c < 1$ and for every $x,y \in H$,

$$|X(x) - X(y)| \le |\Phi(X(x),x) - \Phi(X(y),x)|$$
$$+ |\Phi(X(y),x) - \Phi(X(y),y)|$$
$$\le c|X(x) - X(y)| + |\Phi(X(y),x) - \Phi(X(y),y)|.$$

Since the directional derivative process in the direction $h \in H$ is $\nabla_x\Phi(X,x;h)_\tau = e^{\tau A}h$, $\tau \ge 0$, it follows that the norm of $\nabla_x\Phi(X,x)$ is bounded by a constant c_1 independent of X and x. Then we obtain $|X(x) - X(y)| \le c_1(1-c)^{-1}|x - y|$ and the required inequality follows immediately. $\square$

REMARK 4.7.   Denoting by ∇X the Gâteaux derivative whose existence is asserted in Proposition 4.6, for every direction $h \in H$, the directional derivative

process $\nabla X_\tau(x)h$, $\tau \geq 0$, is the unique solution in $\mathcal{H}^q_{\eta(q)}$ of the equation, $\mathbb{P}$-*a.s.*,

$$\nabla X_\tau(x)h = e^{\tau A}h + \int_0^\tau e^{(\tau-\sigma)A}\nabla F(X_\sigma(x))\nabla X_\sigma(x)h\,d\sigma$$

$$+ \int_0^\tau \nabla(e^{(\tau-\sigma)A}G(X_\sigma(x)))\nabla X_\sigma(x)h\,dW_\sigma, \qquad \tau \geq 0.$$

Indeed, this follows from the parameter depending contraction principle and the explicit form of $\nabla_X\Phi$ and $\nabla_x\Phi$ found in the previous proof.

5. The forward–backward system. As usual, we denote by $\{W_\tau, \tau \geq 0\}$ a cylindrical Wiener process with values in a Hilbert space Ξ and denote by $(\mathcal{F}_\tau)$ its natural filtration, augmented in the usual way. In this section, we consider the system of stochastic differential equations, $\mathbb{P}$-*a.s.*,

$$X_\tau = e^{\tau A}x + \int_0^\tau e^{(\tau-\sigma)A}F(X_\sigma)\,d\sigma + \int_0^\tau e^{(\tau-\sigma)A}G(X_\sigma)\,dW_\sigma,$$

$$\tau \geq 0,$$

(5.1)
$$Y_\tau + \int_\tau^T Z_\sigma\,dW_\sigma + \lambda\int_\tau^T Y_\sigma\,d\sigma = -\int_\tau^T \psi(X_\sigma, Y_\sigma, Z_\sigma)\,d\sigma,$$

$$0 \leq \tau \leq T < \infty,$$

where X takes values in a Hilbert space H, Y is real valued and (accordingly) Z takes values in $L_2(\Xi, \mathbb{R})$ (which coincides with Ξ^*), $\psi : H \times \mathbb{R} \times \Xi^* \to \mathbb{R}$ is a given measurable function, x is in H and λ is a real number.

We will give existence, uniqueness and regularity results for the solution, that we will denote by $\{X_\tau(x), Y_\tau(x), Z_\tau(x), \tau \geq 0\}$ when we want to stress dependence on the parameter $x \in H$. Note that the forward equation can be solved independently. Moreover, since the filtration $(\mathcal{F}_\tau)$ is generated by the Wiener process, and since Y is adapted, it follows that $Y_0(x)$ is deterministic. In the following discussion we set $u(x) = Y_0(x)$.

The equations are the same as in the previous section, in the particular case $K = \mathbb{R}$. So, if we assume that Hypotheses 4.1 and 3.10 are verified, we can immediately describe a class of processes where the system is uniquely solvable, as follows.

For any $q \in [1, \infty)$, we choose $\eta(q)$ as in Proposition 4.6. Then, for every $x \in H$, there exists a unique solution $\{X_\tau(x), \tau \geq 0\}$ in $\mathcal{H}^q_{\eta(q)}$ of the forward equation and the map $x \mapsto X(x)$ belongs to $\mathcal{G}^1(H, \mathcal{H}^q_{\eta(q)})$.

Then we fix $p > 2$ and choose q and β satisfying

(5.2) $q \geq p(m+1)(m+2), \qquad \beta < \eta(q)(m+1)(m+2), \qquad \beta < 0.$

If we set $\widehat{\lambda} = -(\beta + \mu - L_z^2/2)$, then, according to Proposition 3.11, for every $\lambda > \widehat{\lambda}$ and for arbitrary $X \in \mathcal{H}^q_{\eta(q)}$ there exists a unique solution $(Y(X), Z(X))$

in $\mathcal{K}_\beta^p$ of the backward equation; moreover $Y(X)$ is in $L_{\mathcal{P}}^q(\Omega; C_\beta(\mathbb{R}))$, the map $X \to (Y(X), Z(X))$ belongs to $\mathcal{G}^1(\mathcal{H}_{\eta(q)}^q, \mathcal{K}_\beta^p)$ and the map $X \to Y(X)$ belongs to $\mathcal{G}^1(\mathcal{H}_{\eta(q)}^q, L_{\mathcal{P}}^q(\Omega; C_\beta(\mathbb{R})))$. Therefore, with the present notation, the solution of the backward equation in (5.1) is

$$(5.3) \qquad Y(x) = Y(X(x)), \qquad Z(x) = Z(X(x)).$$

PROPOSITION 5.1. *Assume that Hypothesis* 4.1 *holds and that ψ satisfies the conditions in Hypothesis* 3.10 *(with $K = \mathbb{R}$). For $p > 2$, β and q satisfying* (5.2), *and for every $\lambda > \widehat{\lambda} = -(\beta + \mu - L_z^2/2)$, the following hold:*

(i) *For every $x \in H$ there exists a unique solution $(X(x), Y(x), Z(x))$ of the forward–backward system* (5.1) *such that $X(x) \in \mathcal{H}_{\eta(q)}^q$ and $(Y(x), Z(x)) \in \mathcal{K}_\beta^p$. Moreover, $Y(x) \in L_{\mathcal{P}}^p(\Omega; C_\beta(\mathbb{R}))$.*

(ii) *The maps $x \to X(x)$ and $x \to (Y(x), Z(x))$ and $x \to Y(x)$ belong to the spaces $\mathcal{G}^1(H, \mathcal{H}_{\eta(q)}^q)$, $\mathcal{G}^1(H, \mathcal{K}_\beta^p)$ and $\mathcal{G}^1(H, L_{\mathcal{P}}^p(\Omega; C_\beta(\mathbb{R})))$, respectively.*

(iii) *Setting $u(x) = Y_0(x)$, we have $u \in \mathcal{G}^1(H, \mathbb{R})$, and u and ∇u have polynomial growth. More precisely, there exists a constant $C > 0$ such that*

$$|u(x)| \leq C(1 + |x|)^{m+1}, \qquad |\nabla u(x)h| \leq C|h|(1 + |x|)^{[(m+1)^2]}, \qquad x, h \in H.$$

PROOF. Point (i) is already proved, and point (ii) follows from (5.3) and the chain rule. Since the (linear) functional $Y \to Y_0$ is continuous on $L_{\mathcal{P}}^p(\Omega; C_\beta(\mathbb{R}))$, it also follows that $x \to u(x) = Y_0(x)$ is in $\mathcal{G}^1(H, \mathbb{R})$. The estimate on u is a consequence of (3.21) and (4.10). The estimate on ∇u follows from the chain rule and (3.23), (4.10) and (4.11). □

REMARK 5.2. Notice that we have shown that the system (5.1) admits a unique solution [in suitable spaces $\mathcal{H}_{\eta(q)}^q$, $\mathcal{K}_\beta^p$ with parameters satisfying $p > 2$ and condition (5.2)] for all $\lambda > \widehat{\lambda}$, where

$$(5.4) \quad \widehat{\lambda} = -\mu + L_z^2/2 - \sup\{\eta(q)(m+1)(m+2) \wedge 0 : q > 2(m+1)(m+2)\}.$$

REMARK 5.3. If, in addition to Hypothesis 3.10, we suppose that $\psi(\cdot, 0, 0)$ is bounded and satisfies Hypothesis 3.10 with $m = 0$, then the above results can be improved in the following way, according to Remarks 3.12 and 3.13. Instead of invoking (5.2), it is enough to require $q > p > 2$ and $\beta < \eta(q) \wedge 0$. Then the conclusions of Proposition 5.1 still hold for $\lambda > -(\beta + \mu - L_z^2/2)$. Thus, instead of (5.4), we have

$$(5.5) \qquad \widehat{\lambda} = -\mu + L_z^2/2 - \sup\{\eta(q) \wedge 0 : q > 2\}.$$

Moreover, we have $|u(x)| \leq C$ and $|\nabla_x u(x)h| \leq C|h|$ for all $x, h \in H$.

6. Mild solutions of the Kolmogorov nonlinear equation.   Let us consider again the forward equation

(6.1)
$$X_\tau = e^{\tau A}x + \int_0^\tau e^{(\tau-\sigma)A}F(X_\sigma)\,d\sigma$$
$$+ \int_0^\tau e^{(\tau-\sigma)A}G(X_\sigma)\,dW_\sigma, \qquad \tau \geq 0,$$

studied in the previous sections. Assuming that Hypothesis 4.1 holds and denoting $\{X_\tau(x), \tau \geq 0\}$ as the solution, we define in the usual way the transition semigroup $(P_t)_{t\geq 0}$, associated to the process X:

(6.2)
$$P_t[\phi](x) = \mathbb{E}\phi(X_t(x)), \qquad x \in H,$$

for every bounded measurable function $\phi : H \to \mathbb{R}$. By Proposition 4.2, ϕ can be taken unbounded, with polynomial growth. Formally, the generator $\mathcal{L}$ of (P_t) is the operator

$$\mathcal{L}\phi(x) = \tfrac{1}{2}\operatorname{Trace}(G(x)G(x)^*\nabla^2\phi(x)) + \langle Ax + F(x), \nabla\phi(x)\rangle.$$

In this section, we address solvability of the nonlinear stationary Kolmogorov equation

(6.3)
$$\mathcal{L}u(x) - \lambda u(x) = \psi(x, u(x), \nabla u(x)G(x)), \qquad x \in H,$$

where the function $\psi : H \times \mathbb{R} \times \Xi^* \to \mathbb{R}$ satisfies the conditions in Hypothesis 3.10 (with $K = \mathbb{R}$) and λ is a given number. Note that, for $x \in H$, $\nabla u(x)$ belongs to H^*, so that $\nabla u(x)G(x)$ is in Ξ^*.

   As it is written, (6.3) is only formal. We give the following definition of solution, already mentioned in the Introduction (recall that the class $\mathcal{G}^1$ was defined in Section 2):

DEFINITION 6.1.   We say that a function $u : H \to \mathbb{R}$ is a mild solution of the nonlinear stationary Kolmogorov equation (6.3) if the following conditions hold:

   (i) Function $u \in \mathcal{G}^1(H, \mathbb{R})$.
   (ii) For all $x \in H$, $h \in H$, we have

$$|u(x)| \leq C(1 + |x|)^C, \qquad |\nabla_x u(x)h| \leq C|h|(1 + |x|)^C$$

for some constant $C > 0$.
   (iii) The following equality holds, for every $x \in H$ and $T \geq 0$:

(6.4) $$u(x) = e^{-\lambda T}P_T[u](x) - \int_0^T e^{-\lambda\tau}P_\tau[\psi(\cdot, u(\cdot), \nabla u(\cdot)G(\cdot))](x)\,d\tau.$$

Together with (6.1), we also consider the backward equation

$$
\begin{aligned}
Y_\tau - Y_T + &\int_\tau^T Z_\sigma \, dW_\sigma + \lambda \int_\tau^T Y_\sigma \, d\sigma \\
&= -\int_\tau^T \psi(X_\sigma, Y_\sigma, Z_\sigma) \, d\sigma, \qquad 0 \le \tau \le T < \infty,
\end{aligned}
$$

(6.5)

where $\psi : H \times \mathbb{R} \times \Xi^* \to \mathbb{R}$ and λ are the same that occur in the nonlinear stationary Kolmogorov equation. Under the stated assumptions, Proposition 5.1 gives a unique solution $\{X_\tau(x), Y_\tau(x), Z_\tau(x), \tau \ge 0\}$ of the forward–backward system (6.1) and (6.5). We can now state one of our main results.

THEOREM 6.1. *Assume that Hypothesis* 4.1 *holds and that* ψ *satisfies the conditions in Hypothesis* 3.10 *(with* $K = \mathbb{R}$*). Then there exists* $\widehat{\lambda} \in \mathbb{R}$ *such that, for every* $\lambda > \widehat{\lambda}$*, the nonlinear stationary Kolmogorov equation* (6.3) *has a unique mild solution. The solution u is given by the formula*

$$
u(x) = Y_0(x),
$$

(6.6)

where $\{X_\tau(x), Y_\tau(x), Z_\tau(x), \tau \ge 0\}$ *is the solution of the backward–forward system* (6.1) *and* (6.5)*, and it satisfies*

$$
|u(x)| \le C (1 + |x|)^{m+1}, \qquad |\nabla u(x)h| \le C |h|(1 + |x|)^{[(m+1)^2]}
$$

for some constant C and every $x, h \in H$.

PROOF. We need to consider the equation [which is slightly more general than (6.1)]

$$
X_\tau = e^{(\tau - t)A}x + \int_t^\tau e^{(\tau - \sigma)A} F(X_\sigma) \, d\sigma + \int_t^\tau e^{(\tau - \sigma)A} G(X_\sigma) \, dW_\sigma
$$

(6.7)

for τ varying on an arbitrary time interval $[t, \infty) \subset [0, \infty)$. We set $X_\tau = x$ for $\tau \in [0, t)$ and we denote by $\{X(\tau, t, x), \tau \ge 0\}$ the solution, to indicate dependence on x and t. By an obvious extension of the results in the previous sections, we can solve the backward equation (6.5) with X given by (6.7); we denote the corresponding solution (Y, Z) by $\{Y(\tau, t, x), Z(\tau, t, x), \tau \ge 0\}$. Thus, $\{X(\tau, 0, x), Y(\tau, 0, x), Z(\tau, 0, x), \tau \ge 0\}$ coincides with the process $\{X_\tau(x), Y_\tau(x), Z_\tau(x), \tau \ge 0\}$ that occurs in the statement of the theorem. Note that for bounded measurable $\phi : H \to \mathbb{R}$, we have

$$
P_{\tau - t}[\phi](x) = \mathbb{E}\,\phi(X(\tau, t, x)), \qquad x \in H, \ 0 \le t \le \tau,
$$

since the coefficients of (6.7) do not depend on time.

We first prove that u, given by (6.6), is a solution. The solutions of (6.7) satisfy the well-known property, for $0 \le t \le s$, $\mathbb{P}$-*a.s.*,

$$
X(\tau, s, X(s, t, x)) = X(\tau, t, x) \qquad \text{for } \tau \in [s, \infty).
$$

Since the solution of the backward equation is uniquely determined on an interval $[s, \infty)$ by the values of the process X on the same interval, for $0 \leq t \leq s$ we have, $\mathbb{P}$-*a.s.*,

(6.8)
$$Y(\tau, s, X(s, t, x)) = Y(\tau, t, x) \qquad \text{for } \tau \in [s, \infty),$$
$$Z(\tau, s, X(s, t, x)) = Z(\tau, t, x) \qquad \text{for a.a. } \tau \in [s, \infty).$$

In particular, for every $\tau \geq 0$,

(6.9)
$$Y(\tau, \tau, X(\tau, 0, x)) = Y(\tau, 0, x), \qquad \mathbb{P}\text{-}a.s.$$

Since the coefficients of (6.7) do not depend on time, we have

$$X(\cdot, 0, x) \overset{(d)}{=} X(\cdot + t, t, x), \qquad t \geq 0,$$

where $\overset{(d)}{=}$ denotes equality in distribution [both sides of the equality are viewed as random elements with values in the space $C(\mathbb{R}_+; H)$]. As a consequence, we obtain

$$\big(Y(\cdot, 0, x), Z(\cdot, 0, x)\big) \overset{(d)}{=} \big(Y(\cdot + t, t, x), Z(\cdot + t, t, x)\big), \qquad t \geq 0,$$

where both sides of the equality are viewed as random elements with values in the space $C(\mathbb{R}_+; \mathbb{R}) \times L^2_{\text{loc}}(\mathbb{R}_+; \Xi^*)$. In particular, $Y(0, 0, x) \overset{(d)}{=} Y(t, t, x)$ and since they are both deterministic, we have

$$u(x) = Y(0, 0, x) = Y(t, t, x), \qquad x \in H, t \geq 0.$$

Denoting for simplicity

$$(X_\tau, Y_\tau, Z_\tau) = \big(X(\tau, 0, x), Y(\tau, 0, x), Z(\tau, 0, x)\big), \qquad \tau \geq 0,$$

then it follows from (6.9) and path continuity that, $\mathbb{P}$-*a.s.*,

$$u(X_\tau) = Y_\tau, \qquad \tau \geq 0.$$

It follows from the backward equation that

$$u(X_\tau) = Y_\tau$$

(6.10)
$$= Y_0 + \int_0^\tau Z_\sigma \, dW_\sigma + \lambda \int_0^\tau Y_\sigma \, d\sigma + \int_0^\tau \psi(X_\sigma, Y_\sigma, Z_\sigma) \, d\sigma,$$

$$\tau \geq 0.$$

For $\xi \in \Xi$, we denote by W^ξ the real Wiener process defined by $W_\tau^\xi := \langle \xi, W_\tau \rangle$, $\tau \geq 0$. The joint quadratic variation of the right-hand side of (6.10) with W^ξ is the process $\int_0^\tau Z_\sigma \xi \, d\sigma$, $\tau \geq 0$. By Proposition 4.5, the joint quadratic variation of the left-hand side of (6.10) with W^ξ is $\int_0^\tau \nabla u(X_\sigma) G(X_\sigma) \xi \, d\sigma$, $\tau \geq 0$. It follows that, $\mathbb{P}$-*a.s.* for a.a. $\tau \geq 0$,

$$Z_\tau = \nabla u(X_\tau) G(X_\tau).$$

Applying the Itô formula to the backward equation gives

$$e^{-\lambda\tau}Y_\tau - e^{-\lambda T}Y_T + \int_\tau^T e^{-\lambda\sigma}Z_\sigma\,dW_\sigma = -\int_\tau^T e^{-\lambda\sigma}\psi(X_\sigma,Y_\sigma,Z_\sigma)\,d\sigma,$$

$$0 \le \tau \le T < \infty,$$

and it follows that

$$\int_0^T e^{-\lambda\tau}P_\tau\big[\psi\big(\cdot,u(\cdot),\nabla u(\cdot)G(\cdot)\big)\big](x)\,d\tau$$

$$= \mathbb{E}\int_0^T e^{-\lambda\tau}\psi\big(X_\tau,u(X_\tau),\nabla u(X_\tau)G(X_\tau)\big)\,d\tau$$

$$= \mathbb{E}\int_0^T e^{-\lambda\tau}\psi(X_\tau,Y_\tau,Z_\tau)\,d\tau$$

$$= \mathbb{E}\bigg[-Y_0 + e^{-\lambda T}Y_T - \int_0^T e^{-\lambda\tau}Z_\tau\,dW_\tau\bigg]$$

$$= -u(x) + e^{-\lambda T}\mathbb{E}[u(X_T)]$$

$$= -u(x) + e^{-\lambda T}P_T[u](x).$$

This completes the proof of the existence part.

Now we prove uniqueness of the solution. Assume that u is a solution. For any $y \in H$, $0 \le \tau \le T$, we have

$$u(y) = e^{-\lambda(T-\tau)}P_{T-\tau}[u](y) - \int_0^{T-\tau} e^{-\lambda t}P_t\big[\psi\big(\cdot,u(\cdot),\nabla u(\cdot)G(\cdot)\big)\big](y)\,dt.$$

Set $y = X(\tau,0,x)$, that we denote X_τ for simplicity. By the Markov property of X, denoting by $\mathbb{E}^{\mathcal{F}_\tau}$ the conditional expectation with respect to $\mathcal{F}_\tau$, we obtain

$$u(X_\tau) = e^{-\lambda(T-\tau)}\mathbb{E}^{\mathcal{F}_\tau}u(X_T)$$

$$- \int_0^{T-\tau} e^{-\lambda t}\mathbb{E}^{\mathcal{F}_\tau}\psi\big(X_{t+\tau},u(X_{t+\tau}),\nabla u(X_{t+\tau})G(X_{t+\tau})\big)\,dt,$$

and by a change of variable, we obtain

$$e^{-\lambda\tau}u(X_\tau) = e^{-\lambda T}\mathbb{E}^{\mathcal{F}_\tau}u(X_T)$$

$$- \int_\tau^T e^{-\lambda\sigma}\mathbb{E}^{\mathcal{F}_\tau}\psi\big(X_\sigma,u(X_\sigma),\nabla u(X_\sigma)G(X_\sigma)\big)\,d\sigma.$$

Now let $T > 0$ be fixed and let us define

$$\psi_\sigma = \psi\big(X_\sigma,u(X_\sigma),\nabla u(X_\sigma)G(X_\sigma)\big), \quad \sigma \in [0,T],$$

$$\xi = e^{-\lambda T}u(X_T) - \int_0^T e^{-\lambda\sigma}\psi_\sigma\,d\sigma.$$

Then we obtain

$$e^{-\lambda\tau}u(X_\tau) = \mathbb{E}^{\mathcal{F}_\tau}\xi + \mathbb{E}^{\mathcal{F}_\tau}\int_0^\tau e^{-\lambda\sigma}\psi_\sigma\,d\sigma = \mathbb{E}^{\mathcal{F}_\tau}\xi + \int_0^\tau e^{-\lambda\sigma}\psi_\sigma\,d\sigma,$$

where the last equality holds since $\int_0^\tau e^{-\lambda\sigma}\psi_\sigma\,d\sigma$ is $\mathcal{F}_\tau$-adapted. Since we assume polynomial growth for u and ∇u, therefore ξ is square-integrable. Since $(\mathcal{F}_t)$ is generated by the Wiener process W, it follows that there exists a square-integrable, $(\mathcal{F}_t)$-predictable process $\widetilde{Z}_\tau$, $\tau \in [0, T]$, with values in Ξ^*, such that, $\mathbb{P}$-*a.s.*,

$$\mathbb{E}^{\mathcal{F}_\tau}\xi = \mathbb{E}\xi + \int_0^\tau \widetilde{Z}_\sigma\,dW\sigma, \qquad \tau \in [0, T].$$

An application of the Itô formula gives

$$(6.11) \quad u(X_\tau) = \mathbb{E}\xi + \int_0^\tau e^{\lambda\sigma}\widetilde{Z}_\sigma\,dW_\sigma + \lambda\int_0^\tau u(X_\sigma)\,d\sigma + \int_0^\tau \psi_\sigma\,d\sigma.$$

This shows that $u(X_\tau)$, $\tau \in [0, T]$, is a semimartingale. For $\xi \in \Xi$, let us define W^ξ as above and let us consider the joint quadratic variation process of W^ξ with both sides of (6.11). Applying Proposition 4.5, we obtain, $\mathbb{P}$-*a.s.*,

$$\int_0^\tau \nabla u(X_\sigma)G(X_\sigma)\xi\,d\sigma = \int_0^\tau e^{\lambda\sigma}\widetilde{Z}_\sigma\xi\,d\sigma, \qquad \tau \in [0, T], \xi \in \Xi,$$

and we deduce that $\nabla u(X_\tau)G(X_\tau) = e^{\lambda\tau}\widetilde{Z}_\tau$, $\mathbb{P}$-*a.s.* for almost all $\tau \in [0, T]$. Now setting

$$Y_\tau' = u(X_\tau), \qquad Z_\tau' = e^{\lambda\tau}\nabla u(X_\tau)G(X_\tau), \qquad \tau \geq 0,$$

it follows from (6.11) that, $\mathbb{P}$-*a.s.*,

$$Y_\tau' = Y_0' + \int_0^\tau Z_\sigma'\,dW_\sigma + \lambda\int_0^\tau Y_\sigma'\,d\sigma + \int_0^\tau \psi(X_\sigma, Y_\sigma', Z_\sigma')\,d\sigma, \qquad \tau \in [0, T].$$

Since T is arbitrary, we conclude that the process (Y', Z') is a solution of the backward equation, so that, by uniqueness, it must coincide with (Y, Z). In particular,

$$u(x) = u(X_0) = Y_0' = Y_0.$$

This concludes the proof of the theorem. $\square$

REMARK 6.2. The constant $\widehat{\lambda}$ in the statement of Theorem 6.1 can be chosen equal to (5.4).

REMARK 6.3. From Remark 5.3 it follows immediately that if, in addition to Hypotheses 4.1 and 3.10, we assume that $\psi(\cdot, 0, 0)$ is bounded and ψ satisfies Hypothesis 3.10 with $m = 0$, then $\widehat{\lambda}$ can be chosen equal to (5.5) instead of (5.4). Moreover, in this case, we have $|u(x)| \leq C$, $|\nabla u(x)h| \leq C\,|h|$ for some constant C and every $x, h \in H$.

REMARK 6.4.   The results of Sections 5 and 6 can be generalized to allow the process Y and the function u to take values in a real separable Hilbert space K.

More precisely suppose that a function $\psi : H \times K \times L_2(\Xi, K) \to K$ is given and satisfies Hypothesis 3.10. Then we can look for a solution (X, Y, Z) of the forward–backward system (5.1) with values in $H \times K \times L_2(\Xi, K)$. Proposition 5.1 still holds with identical proof relying on the results of Section 3, where we already considered Hilbert-valued BSDE.

We can then consider the nonlinear elliptic Kolmogorov system, with unknown function $u : H \to K$,

$$(6.12) \quad \mathcal{L}u_i = \lambda u_i(x) + \psi_i\big(x, u(x), \nabla u(x)G(x)\big), \qquad x \in H,\ i = 1, 2, \ldots,$$

where $u_i = \langle u, k_i \rangle$ and $\psi_i = \langle \psi, k_i \rangle$ for a fixed orthonormal basis $\{k_i : i \in \mathbb{N}\}$ of K.

We say that a function u is a mild solution of (6.12) if it belongs to $\mathcal{G}^1(H, K)$, it satisfies

$$|u(x)|_K \le C(1 + |x|_H)^C, \qquad |\nabla u(x)| \le C|h|(1 + |x|_H)^C, \qquad x, h \in H,$$

for some $C > 0$ and (6.4) holds for every $x \in H$ and $T \ge 0$. Note that the definition of $P_t[\phi]$ given in (6.2) is meaningful for a measurable function $\phi : H \to K$ that has polynomial growth.

Then the obvious analogue of Theorem 6.1 holds with the same proof.

7. Applications to optimal control.   We wish to apply the above results to perform the synthesis of the optimal control for a general nonlinear control system on an infinite time horizon. To be able to use nonsmooth feedback, we settle the problem in the framework of weak control problems (see [11]).

We fix the canonical space Ω of continuous maps $\omega : \mathbb{R}_+ \to \Xi$ and endow it with the Borel σ-field $\mathcal{E}$ and the canonical filtration $(\mathcal{F}_t)_{t \ge 0}$, where $\mathcal{F}_t$ is generated by the maps $\omega \to \omega(s)$ and $s \in [0, t]$. Again H, Ξ and U denote Hilbert spaces. For fixed $x_0 \in H$, an *admissible control system* (a.c.s.) is given by $(\mathbb{P}, W_t, u)$, where:

- $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{E})$.
- The variable $\{W_t : t \ge 0\}$ is a Ξ-valued cylindrical Wiener process relative to the filtration $(\mathcal{F}_t)$ and the probability $\mathbb{P}$.
- The variable $u \in L^2_{\mathcal{P}}(\Omega \times \mathbb{R}_+; U)$ satisfies the constraint $u_t \in \mathcal{U}$ $\mathbb{P}$-a.s. for a.a. $t \ge 0$, where $\mathcal{U}$ is a fixed bounded subset of U; let $|u| \le L_{\mathcal{U}}$ for all $u \in \mathcal{U}$ and some constant $L_{\mathcal{U}}$.

To each a.c.s. we associate the mild solution $X^u \in L^r(\Omega; C([0, T]; H))$ (for arbitrary $T > 0$ and arbitrary $r \ge 1$) of the state equation

$$(7.1) \quad \begin{aligned} dX^u_\tau &= \big(AX^u_\tau + F(X^u_\tau) + G(X^u_\tau)R(X^u_\tau)u_\tau\big)\,d\tau + G(X_\tau)\,dW_\tau, \qquad \tau \ge 0, \\ X_0 &= x_0 \in H \end{aligned}$$

and the cost:

$$(7.2) \qquad J(x_0, u) = \mathbb{E} \int_0^{+\infty} e^{-\lambda\sigma} g(X_\sigma^u, u_\sigma)\, d\sigma,$$

where $g : H \times U \to \mathbb{R}$. Our purpose is to minimize the functional J over all a.c.s. Notice the occurrence of the operator G in the control term. This special structure of the state equation is imposed by our techniques. On the contrary, the presence of the operator R allows more generality.

We define in a classical way the Hamiltonian function relative to the above problem: for all $x \in H$ and $p \in U^*$,

$$(7.3) \qquad \begin{aligned} \psi_0(x, p) &= \inf\{g(x, u) + pu : u \in \mathcal{U}\} \\ \Gamma(x, p) &= \{u \in \mathcal{U} : g(x, u) + pu = \psi_0(x, p)\}. \end{aligned}$$

We make the following assumption.

HYPOTHESIS 7.1.   *The following hold*:

1. *Variables A, F and G verify Hypothesis 4.1.*
2. *The map $R : H \to L(U, \Xi)$ enjoys the following situation zR is in $\mathcal{G}^1(H, U^*)$ for every $z \in \Xi^*$; moreover, $|R(x)|_{L(U,\Xi)} \le K_R$ and $|\nabla_x(zR(x))h|_{U^*} \le L_R|z||h|$ for suitable constants $K_R, L_R > 0$ and all $z \in \Xi^*$, $x, h \in H$.*
3. *The map $g : H \times U \to \mathbb{R}$ is continuous and satisfies $|g(x, u)| \le K_g(1 + |x|^{m_g})$ for suitable constants $K_g > 0$, $m_g \ge 0$ and all $x \in H$, $u \in \mathcal{U}$.*
4. *The variable ψ_0 belongs to $\mathcal{G}^1(H \times U^*, \mathbb{R})$ with $|\nabla_x \psi_0(x, p)h| \le L_x^0|h| \times (1 + |p|)(1 + |x|^{m_\psi})$ for suitable constants $L_x^0 > 0$, $m_\psi \ge 0$ and all $x, h \in H$ and $p \in U^*$. [Notice that by its definition $|\psi_0(x, p_1) - \psi_0(x, p_2)| \le L_\mathcal{U}|p_1 - p_1|$ for all $x \in H$, $p_1, p_2 \in U^*$.]*
5. *Finally, we fix here $p > 2$, q and β verifying (5.2) with $m = m_\psi$ and such that $q > m_g$.*

We also define

$$\psi(x, z) = -\psi_0\big(x, zR(x)\big), \qquad x \in H, z \in \Xi^*,$$

and we notice that $|\nabla_z \psi(x, z)h| \le L_z|h|$ and $|\nabla_x \psi(x, z)h| \le L_x|h|(1 + |z|) \times (1 + |x|^{m_\psi})$ for $L_z := K_R L_\mathcal{U}$ and $L_x := L_x^0(1 \vee K_R) + L_\mathcal{U} L_R$.

In the following discussion, $\eta(q)$ is the constant introduced in Proposition 4.6.

EXAMPLE 7.1.1.   If $\mathcal{U}$ is the ball $\{v \in U : |v| \le r\}$ for some fixed $r > 0$, and $g(x, u) = g_0(|u|^\alpha) + g_1(x)$ with $g_0 \in C^1(\mathbb{R}^+; \mathbb{R}^+)$ convex, $g_0'(0) > 0$, $\alpha > 1$, $g_1 \in \mathcal{G}^1(H, \mathbb{R})$ with $|\nabla g_1(x)h| \le L|h|(1 + |x|^m)$ for suitable constants $L > 0$, $m \ge 0$ and all $x, h \in H$, then the conditions on g and ψ_0 in Hypothesis 7.1 hold true. Moreover, $\psi_0(x, p)$ is Fréchet differentiable with respect to p and $\Gamma(x, p) = \{\nabla_p \psi_0(x, p)\}$ turns out to be always a singleton and a continuous function of p only.

LEMMA 7.2. *Assume that $\lambda > 0$ verifies*

(7.4)
$$\lambda > \frac{L_{\mathcal{U}} K_R m_g}{2(q - m_g)} - \eta(q) m_g.$$

Then the cost functional is well defined and $J(x_0, u) < \infty$ for all $x_0 \in H$ and all a.c.s.

PROOF. Fix an *a.c.s.* $(\mathbb{P}, W, u)$ and let X^u be the unique mild solution of (7.1) [existence and uniqueness of a solution in $L^r_{\mathcal{P}}(\Omega; C([0, T]; H))$ for arbitrary $r \geq 1$ and $T > 0$ follows from an immediate extension of Proposition 4.2]. Clearly it is enough to show that, for a suitable $C > 0$,

$$\mathbb{E}(|X^u_\tau|^{m_g}) \leq C e^{(L_{\mathcal{U}} K_R m_g (2q - 2m_g)^{-1} - \eta(q) m_g) \tau}, \qquad \tau \geq 0.$$

Note that the process $\{R(X^u_\tau)u_\tau, \tau \geq 0\}$ is bounded by $K_R L_{\mathcal{U}}$ and denote by $\rho(T)$ the Girsanov density

(7.5) $\quad \rho(T) = \exp\left(-\int_0^T \langle R(X^u_\sigma)u_\sigma, dW_\sigma\rangle_\Xi - \frac{1}{2}\int_0^T |R(X^u_\sigma)u_\sigma|^2_\Xi \, d\sigma\right).$

Let $\widetilde{\mathbb{P}}$ be the unique probability (which exists by the Kolmogorov theorem) that extends to the whole $\mathcal{E}$ the probabilities

(7.6)
$$\widetilde{\mathbb{P}}\big|_{\mathcal{F}_T} = \rho(T)\mathbb{P}\big|_{\mathcal{F}_T}.$$

We notice that under $\widetilde{\mathbb{P}}$, the process

(7.7)
$$\widetilde{W}_\tau := \int_0^\tau R(X^u_\sigma)u_\sigma \, d\sigma + W_\tau, \qquad \tau \geq 0,$$

is a cylindrical Wiener process. Thus (7.1) can be rewritten as

(7.8)
$$dX^u_\tau = \big(AX^u_\tau + F(\tau, X^u_\tau)\big)\, d\tau + G(\tau, X_\tau)\, d\widetilde{W}_\tau, \qquad \tau \geq 0,$$
$$X_0 = x_0 \in H,$$

and, by Proposition 4.6(ii), we obtain $\widetilde{\mathbb{E}}(\sup_{\tau \geq 0} e^{\eta(q)q\tau}|X^u_\tau|^q) < +\infty$. Moreover, by the Hölder inequality,

(7.9)
$$\mathbb{E}(|X^u_T|^{m_g}) = \widetilde{\mathbb{E}}(\rho^{-1}(T)|X^u_T|^{m_g})$$
$$\leq C\big(\widetilde{\mathbb{E}}(\rho^{-q/(q-m_g)}(T))\big)^{(q-m_g)/q} e^{-\eta(q)m_g T}.$$

Since

$$\rho(T)^{-1} = \exp\left(\int_0^T \langle R(X^u_\sigma)u_\sigma, d\widetilde{W}_\sigma\rangle_\Xi - \frac{1}{2}\int_0^T |R(X^u_\sigma u_\sigma)|^2_\Xi \, d\sigma\right)$$

forms an exponential $\widetilde{\mathbb{P}}$-martingale, it is easy to show that for all $r \geq 1$,

$$\big(\widetilde{\mathbb{E}}(\rho^{-r}(T))\big)^{1/r} \leq e^{(1/2)K_R L_{\mathcal{U}}(r-1)T}$$

and the claim follows from (7.9), choosing $r = q/(q - m_g)$. $\quad\square$

By Theorem 6.1, for all $\lambda > \widehat{\lambda}$ [the constant $\widehat{\lambda}$ can be chosen equal to (5.4) with $L_z = K_R L_{\mathcal{u}}$] the stationary Hamilton–Jacobi–Bellman equation relative to the above stated problem, written formally

$$(7.10) \qquad \mathcal{L}v(x) = \lambda v(x) + \psi\big(x, \nabla v(x)G(x)\big), \qquad x \in H,$$

admits a unique mild solution, in the sense of Definition 6.1.

We are in a position to prove the main result of this section:

THEOREM 7.3. *Assume Hypothesis 7.1 and suppose that λ verifies*

$$(7.11) \quad \lambda > \left(-\beta + \frac{K_R^2 L_{\mathcal{u}}^2}{2}\right) \vee \left(-\beta + \frac{K_R L_{\mathcal{u}}}{2(p-1)}\right) \vee \left(\frac{L_{\mathcal{u}} K_R m_g}{2(q - m_g)} - \eta(q)m_g\right).$$

Then the following hold:

1. *For all a.c.s., we have $J(x_0, u) \geq v(x_0)$.*
2. *The equality holds if and only if the following feedback law is verified by u and X^u:*

$$(7.12) \qquad u_\tau \in \Gamma\big(X_\tau^u, \nabla v(X_\tau^u)G(X_\tau^u)R(X_\tau^u)\big), \qquad \mathbb{P}\text{-a.s. for a.a. } \tau \geq 0.$$

3. *If $\Gamma_0(x, p) \in \Gamma(x, p)$ is a measurable selection of Γ, there exists an a.c.s. for which*

$$d\overline{X}_\tau = A\overline{X}_\tau \, d\tau + G(\overline{X}_\tau)R(\overline{X}_\tau)\Gamma_0\big(\overline{X}_\tau, \nabla v(\overline{X}_\tau)G(\overline{X}_\tau)R(\overline{X}_\tau)\big) d\tau$$

$$(7.13) \qquad + F(\overline{X}_\tau) \, d\tau + G(\overline{X}_\tau) \, dW_\tau, \qquad \tau \geq 0,$$

$$\overline{X}_0 = x_0 \in H,$$

admits a solution and if $\overline{u}_\tau = \Gamma_0(\overline{X}_\tau, \nabla v(\overline{X}_\tau)G(\overline{X}_\tau)R(\overline{X}_\tau))$, then the couple $(\overline{u}, \overline{X})$ is optimal for the control problem.

PROOF. Let $\widetilde{\mathbb{P}}$ and $\widetilde{W}$ be defined as in (7.5)–(7.7). Relative to $\widetilde{W}$, (7.1) can be written:

$$dX_\tau^u = AX_\tau^u \, d\tau + F(\tau, X_\tau^u) \, d\tau + G(\tau, X_\tau) \, d\widetilde{W}_\tau, \qquad \tau \geq 0,$$

$$X_{t_0} = x_0.$$

By Proposition 5.1, the system of infinite horizon forward–backward equations

$$\widetilde{X}_\tau(x) = e^{\tau A}x + \int_0^\tau e^{(\tau - \sigma)A} F(\widetilde{X}_\sigma(x)) \, d\sigma + \int_0^\tau e^{\sigma A} G(\widetilde{X}_\sigma(x)) \, d\widetilde{W}_\sigma,$$

$$\tau \geq 0,$$

$$(7.14) \quad \widetilde{Y}_\tau(x) - \widetilde{Y}_T(x) + \int_\tau^T \widetilde{Z}_\sigma(x) \, d\widetilde{W}_\sigma + \lambda \int_\tau^T \widetilde{Y}_\sigma(x) \, d\sigma$$

$$= -\int_\tau^T \psi\big(\widetilde{X}_\sigma(x), \widetilde{Z}_\sigma(x)\big) \, d\sigma, \qquad 0 \leq \tau \leq T,$$

admits a unique solution with $X(x) \in \mathcal{H}^q_{\eta(q)}$ and $(Y(x), Z(x)) \in \mathcal{K}^p_\beta$ (with respect to $\widetilde{\mathbb{P}}$). Moreover, by Theorem 6.1,

$$(7.15) \qquad Y_\tau(x) = v\big(X_\tau(x)\big), \qquad Z_\tau(x) = \nabla v\big(X_\tau(x)\big)G\big(X_\tau(x)\big).$$

Comparing the forward equation with the state equation [see (7.8)] and choosing $x = x_0$, we conclude that $\widetilde{X}(x_0) = X^u$. Applying the Itô formula to $e^{-\lambda\tau}\widetilde{Y}_\tau(x_0)$, rewriting and restoring the original noise W, we get

$$\begin{aligned}
\widetilde{Y}_0(x_0) &+ \int_0^T e^{-\lambda\sigma}\widetilde{Z}_\sigma(x_0)\,dW_\sigma \\
(7.16) \qquad &= -\int_0^T e^{-\lambda\sigma}\big[\psi\big(X^u_\sigma, \widetilde{Z}_\sigma(x_0)\big) + \widetilde{Z}_\sigma(x_0)R(X^u_\sigma)u_\sigma\big]\,d\sigma \\
&\quad + e^{-\lambda T}Y_T(x_0).
\end{aligned}$$

Using the definition of ψ, the identification (7.15) and taking expectation with respect to $\mathbb{P}$ (7.16) yields

$$\begin{aligned}
e^{-\lambda T}\mathbb{E}\big(Y(T, x_0)\big) - v(x_0) &= -\mathbb{E}\int_0^T \psi_0\big(X^u_\sigma, \nabla v(X^u_\sigma)G(X^u_\sigma)R(X^u_\sigma)\big)\,d\sigma \\
&\quad + \mathbb{E}\int_{t_0}^T \nabla v(X^u_\sigma)G(X^u_\sigma)R(X^u_\sigma)u_\sigma\,d\sigma.
\end{aligned}$$

By Proposition 5.1, $Y(x_0)$ is in $L^p_{\mathcal{P}}(\Omega; C_\beta(\mathbb{R}))$ and so $\widetilde{\mathbb{E}}(|Y_T(x_0)|^p) \leq C\exp(-p\beta T)$ and, proceeding as in the proof of Lemma 7.2, we get

$$\mathbb{E}\big(|Y_T(x_0)|\big) \leq Ce^{(K_R L_u(2p-2)^{-1}-\beta)T}.$$

Thus adding and subtracting $\mathbb{E}\int_0^{+\infty} e^{-\lambda\sigma}g(X^u_\sigma, u_\sigma)\,d\sigma$, which is finite by Lemma 7.2, and letting $T \to \infty$, we conclude

$$\begin{aligned}
J(x_0, u) &= v(x_0) \\
(7.17) \qquad &+ \mathbb{E}\int_0^\infty e^{-\lambda\sigma}\big[-\psi_0\big(X^u_\sigma, \nabla v(X^u_\sigma)G(X^u_\sigma)R(X^u_\sigma)\big) \\
&\qquad + \nabla_x v(X^u_\sigma)G(X^u_\sigma)R(X^u_\sigma)u_\sigma + g(X^u_\sigma, u_\sigma)\big]\,d\sigma.
\end{aligned}$$

The above equality is known as the *fundamental relation* and immediately implies that $v(x_0) \leq J(x_0, u)$ and that the equality holds if and only if (7.12) holds.

Finally the existence of a weak solution to (7.13) is again a consequence of the Girsanov theorem. Namely, let $X \in \mathcal{H}^q_{\eta(q)}$ be the mild solution of

$$\begin{aligned}
dX_\tau &= AX_\tau\,d\tau + F(X_\tau)\,d\tau + G(X_\tau)\,dW_\tau, \\
X_0 &= x_0
\end{aligned}$$

and let $\widehat{P}$ be the probability on Ω under which

$$\widehat{W}_t := -\int_0^t R(X_\sigma)\Gamma_0\big(X_\sigma, \nabla v(X_\sigma^u)G(X_\sigma)R(X_\sigma)\big)\,d\sigma + W_t$$

is a Wiener process. Then X is the mild solution of (7.13) relative to the probability $\widehat{P}$ and the Wiener process $\widehat{W}$. $\quad\square$

REMARK 7.4. If, in addition to points 1–4 of Hypothesis 7.1, we also assume that g is bounded and Lipschitz in x uniformly in $u \in \mathcal{U}$, then it is easily verified that $\psi(\cdot, 0)$ is bounded and ψ satisfies Hypothesis 3.10 with $m = 0$. Thus by Remark 5.3, the results of Theorem 7.3 can be improved in the following way.

Instead of Hypothesis 7.1 point 5, it is enough to take $q > p > 2$ and $\beta < \eta(q) \wedge 0$. Moreover, instead of (7.11), it is enough to assume

$$\lambda > -\beta + \left(\frac{K_R^2 L_{\mathcal{U}}^2}{2} \vee \frac{K_R L_{\mathcal{U}}}{2(p-1)}\right).$$

EXAMPLE 7.4.1 (The controlled heat equation). Finally we briefly show that our results can be applied to perform the synthesis of optimal controls for infinite horizon costs when the state equation is a general semilinear heat equation with multiplicative noise. Namely, we consider, for $t \geq 0$, $\xi \in [0, 1]$,

$$\frac{\partial}{\partial t} X^u(t, \xi) = \frac{\partial^2}{\partial \xi^2} X^u(t, \xi) + b\big(\xi, X^u(t, \xi)\big) + \sigma\big(\xi, X^u(t, \xi)\big)r(\xi)u(t, \xi)$$

$$+ \sigma\big(\xi, X^u(t, \xi)\big)\frac{\partial}{\partial t} \mathcal{W}(t, \xi),$$

(7.18)
$$X^u(t, 0) = X^u(t, 1) = 0,$$

$$X^u(0, \xi) = x_0(\xi),$$

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

$$(7.19) \qquad J(x_0, u) = \mathbb{E}\int_0^\infty \int_0^1 e^{-\lambda t}\big[\ell(\xi, X^u(t, \xi)) + u^2(t, \xi)\big]\,d\xi\,dt$$

that we minimize over all adapted controls u such that $|u(t, \xi)| \leq \delta$ for a given constant $\delta > 0$ and almost all $t > 0$, $\xi \in [0, 1]$.

To fit the assumptions of our abstract results, we will suppose that the functions b, σ, r and ℓ are all measurable and real-valued, and moreover:

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

$$|b(t, \xi, \eta_1) - b(t, \xi, \eta_2)| \leq L_b|\eta_2 - \eta_1|, \qquad \int_0^1 b^2(\xi, 0)\,d\xi < +\infty,$$

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

2. Function σ is defined on $[0, 1] \times \mathbb{R}$ and there exist constants L_σ and K_σ such that

$$|\sigma(\xi, \eta)| \leq K_\sigma, \qquad |\sigma(\xi, \eta_1) - \sigma(\xi, \eta_2)| \leq L_\sigma |\eta_2 - \eta_1|$$

for a.a. $\xi \in [0, 1]$, and for all $\eta, \eta_1, \eta_2 \in \mathbb{R}$. Finally, $\sigma(\xi, \cdot) \in C^1(\mathbb{R})$ for a.a. $\xi \in [0, 1]$.

3. Mapping $r : [0, 1] \to \mathbb{R}$ is bounded.

4. Function ℓ is defined on $[0, 1] \times \mathbb{R}$ and, for a.a. $\xi \in [0, 1]$, the map $\ell(\xi, \cdot)$ is in $C^1(\mathbb{R}, \mathbb{R})$. Moreover,

$$\left| \frac{\partial}{\partial \eta} \ell(\xi, \eta) \right| \leq c_1(\xi) + L_\ell |\eta|, \qquad |\ell(\xi, 0)| \leq c_0(\xi)$$

with $\int_0^1 |c_1(\xi)|^2 \, d\xi < +\infty$, $\int_0^1 |c_0(\xi)| \, d\xi < +\infty$ and $L_\ell \in \mathbb{R}$.

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

To rewrite the above problem in the abstract way stated the beginning of Section 7, we set $H = \Xi = U = L^2([0, 1])$ and $\mathcal{U} = \{u \in U : |u(\xi)| \leq \delta$ for almost all $\xi \in [0, 1]\}$. By $\{W_t : t \geq 0\}$ we denote a cylindrical Wiener process in $L^2([0, 1])$. Moreover, we define the operator A with domain

$$\mathcal{D}(A) = H^2([0, 1]) \cap H_0^1([0, 1]), \qquad (Ay)(\xi) = \frac{\partial^2}{\partial \xi^2} y(\xi) \qquad \forall \, y \in \mathcal{D}(A),$$

where $H^2([0, 1])$ and $H_0^1([0, 1])$ are the usual Sobolev spaces, and we set

$$F(x)(\xi) = b(\xi, x(\xi)), \qquad (G(x)z)(\xi) = \sigma(\xi, x(\xi))z(\xi),$$

$$(R(x)u)(\xi) = r(\xi)u(\xi),$$

$$g(x, u) = |u|_U^2 + g_1(x) = \int_0^1 \left[u^2(\xi) + \ell(\xi, x(\xi)) \right] d\xi$$

for all $x, z, u \in L^2([0, 1])$ and a.a. $\xi \in [0, 1]$.

Under the previous assumptions we know (see [8], Section 11.2.1) that A, F and G verify Hypothesis 4.1. Moreover, noticing that

$$\nabla g_1(x)h = \int_0^1 \frac{\partial}{\partial \eta} \ell(\xi, x(\xi))h(\xi) \, d\xi$$

and recalling the results in Example 7.1.1, it can be easily verified that points 1–4 in Hypothesis 7.1 are satisfied with $L_\mathcal{U} = \delta$, $K_R = |r|_{L^\infty([0,1])}$, $L_R = 0$, $m_g = 2$, $m_\psi = 1$,

$$K_g = \left(\delta^2 + |c_0|_{L^1([0,1])} + \tfrac{1}{2}|c_1|_{L^2([0,1])} \right) \vee \left(\tfrac{1}{2} + L_\ell \right),$$

$$L_x^0 = |c_1|_{L^2([0,1])} \vee L_\ell,$$

$$L_x = L_x^0 (1 \vee K_R) = \left(|c_1|_{L^2([0,1])} \vee L_\ell \right) \left(1 \vee |r|_{L^\infty([0,1])} \right),$$

$$L_z = \delta |r|_{L^\infty([0,1])}.$$

REFERENCES

[1] AMBROSETTI, A. and PRODI, G. (1995). *A Primer of Nonlinear Analysis*. Cambridge Univ. Press.

[2] BUCKDAHN, R. and PENG, S. (1999). Stationary backward stochastic differential equations and associated partial differential equations. *Probab. Theory Related Fields* **115** 383–399.

[3] BUCKDAHN, R. and RĂŞCANU, A. (2002). On the existence of stochastic optimal control of distributed state system. *Nonlinear Anal.* **52** 1153–1184.

[4] CERRAI, S. (2001). *Second Order PDE's in Finite and Infinite Dimensions. A Probabilistic Approach. Lecture Notes in Math.* **1762**. Springer, New York.

[5] DALECKIJ, YU. L. and FOMIN, S. V. (1991). *Measures and Differential Equations in Infinite-Dimensional Space*. Kluwer, Dordrecht.

[6] DA PRATO, G., GOŁDYS, B. and ZABCZYK, J. (1997). Ornstein–Uhlenbeck semigroups in open sets of Hilbert spaces. *C. R. Acad. Sci. Paris Ser. I Math.* **325** 433–438.

[7] DA PRATO, G. and ZABCZYK, J. (1992). *Stochastic Equations in Infinite Dimensions. Encyclopedia of Mathematics and Its Applications* **44**. Cambridge Univ. Press.

[8] DA PRATO, G. and ZABCZYK, J. (1996). *Ergodicity for Infinite-Dimensional Systems*. Cambridge Univ. Press.

[9] DA PRATO, G. and ZABCZYK, J. (2002). *Second Order Partial Differential Equations in Hilbert Spaces*. Cambridge Univ. Press.

[10] DARLING, R. W. R. and PARDOUX, E. (1997). Backwards SDE with random terminal time and applications to semilinear elliptic PDE. *Ann. Probab.* **25** 1135–1159.

[11] FLEMING, W. H. and SONER, H. M. (1993). *Controlled Markov Processes and Viscosity Solutions*. Springer, New York.

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

[13] FUHRMAN, M. and TESSITORE, G. (2002). The Bismut–Elworthy formula for backward SDE's and applications to nonlinear Kolmogorov equations and control in infinite dimensional spaces. *Stochastics Stochastics Rep.* **74** 429–464.

[14] GOZZI, F. and ROUY, E. (1996). Regular solutions of second-order stationary Hamilton–Jacobi equations. *J. Differential Equations* **130** 201–234.

[15] GOZZI, F., ROUY, E. and ŚWIĘCH, A. (2000). Second order Hamilton–Jacobi equations in Hilbert spaces and stochastic boundary control. *SIAM J. Control Optim.* **38** 400–430.

[16] GRORUD, A. and PARDOUX, E. (1992). Intégrales Hilbertiennes anticipantes par rapport à un processus de Wiener cylindrique et calcul stochastique associé. *Appl. Math. Optim.* **25** 31–49.

[17] GROSS, L. (1967). Potential theory on Hilbert spaces. *J. Funct. Anal.* **1** 123–181.

[18] LIONS, P. L. (1989). Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. III. Uniqueness of viscosity solutions for general second-order equations. *J. Funct. Anal.* **86** 1–18.

[19] LIONS, P. L. and SOUGANIDIS, P. E. (1993). Fully nonlinear second-order degenerate elliptic equations with large zeroth-order coefficients. *Indiana Univ. Math. J.* **42** 1525–1543.

[20] NUALART, D. (1995). *The Malliavin Calculus and Related Topics. Probability and Its Applications*. Springer, Berlin.

[21] NUALART, D. and PARDOUX, E. (1988). Stochastic calculus with anticipative integrands. *Probab. Theory Related Fields* **78** 535–581.

[22] PARDOUX, E. (1975). Equations aux dérivées partielles stochastiques non linéaires monotones. Thèse, Univ. Paris Sud-Orsay.

[23] PARDOUX, E. (1998). Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order. In *Stochastic Analysis and Related Topics* (L. Decreusefond, J. Gjerde, B. Øksendal, A. S. Üstünel, eds.). 79–127. Birkhäuser, Boston.

[24] PARDOUX, E. and PENG, S. (1992). Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Stochastic Partial Differential Equations and Their Applications. Lecture Notes in Control Inform. Sci.* **176** 200–217. Springer, Berlin.

[25] PARDOUX, E. and RĂŞCANU, A. (1998). Backward stochastic differential equations with subdifferential operators and related variational inequalities. *Stochastic Process Appl.* **76** 191–215.

[26] PARDOUX, E. and RĂŞCANU, A. (1999). Backward stochastic variational inequalities. *Stochastics Stochastics Rep.* **67** 159–167.

[27] PENG, S. (1991). Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics Stochastics Rep.* **37** 61–74.

[28] ŚWIĘCH, A. (1994). "Unbounded" second order partial differential equations in infinite-dimensional Hilbert spaces. *Comm. Partial Differential Equations* **19** 1999–2036.

[29] ŚWIĘCH, A. (1993). Viscosity solutions of fully nonlinear partial differential equations with "unbounded" terms in infinite dimensions. Ph.D. dissertation, Univ. California, Santa Barbara.

[30] TALARCZYK, A. (2000). Dirichlet problem for parabolic equations on Hilbert spaces. *Studia Math.* **141** 109–142.

[31] ZABCZYK, J. (1999). Parabolic equations on Hilbert spaces. *Stochastic PDE's and Kolmogorov Equations in Infinite Dimensions. Lecture Notes in Math.* **1715** 117–213. Springer, Berlin.

DIPARTIMENTO DI MATEMATICA
POLITECNICO DI MILANO
PIAZZA LEONARDO DA VINCI 32
20133 MILANO
ITALY
E-MAIL: marco.fuhrman@polimi.it

DIPARTIMENTO DI MATEMATICA
UNIVERSITÀ DI PARMA
VIA D'AZEGLIO 85
43100 PARMA
ITALY
E-MAIL: gianmario.tessitore@unipr.it