MATHEMATICAL CONTROL AND
RELATED FIELDS
Volume 5, Number 2, June 2015

doi:10.3934/mcrf.2015.5.xx

pp. X–XX

FEEDBACK OPTIMAL CONTROL FOR STOCHASTIC VOLTERRA EQUATIONS WITH COMPLETELY MONOTONE KERNELS

Fulvia Confortola

Politecnico di Milano, Dipartimento di Matematica F. Brioschi
Piazza Leonardo da Vinci 32, 20133 Milano, Italia

Elisa Mastrogiacomo

Universitá degli studi di Milano Bicocca
Dipartimento di Statistica e Metodi Quantitativi
Piazza dell'Ateneo Nuovo, 1 - 20126, Milano

(Communicated by Andrzej Swiech)

Abstract. In this paper we are concerned with a class of stochastic Volterra integro-differential problems with completely monotone kernels, where we assume that the noise enters the system when we introduce a control. We start by reformulating the state equation into a semilinear evolution equation which can be treated by semigroup methods. The application to optimal control provides other interesting results and requires a precise description of the properties of the generated semigroup.

The first main result of the paper is the proof of existence and uniqueness of a mild solution for the corresponding Hamilton-Jacobi-Bellman (HJB) equation. The main technical point consists in the differentiability of the BSDE associated with the reformulated equation with respect to its initial datum x.

1. Introduction.

Stochastic Volterra equations represent interesting models for stochastic dynamic systems with memory. They appear naturally in many areas of mathematics such as integral transforms, transport equations, functional differential equations and so forth, and they also appear in applications in biology, physics and finance. For a detailed exposition on applications of Volterra integral equations, we refer to [4, 29] and [42, 43], the first two dealing with deterministic equations only.

In this paper we are concerned with the following optimal control problem for an infinite dimensional stochastic integral equation of Volterra type on a separable Hilbert space H:

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{t} a(t-s)u(s)\mathrm{d}s = Au(t) + f(t,u(t)) \\ \qquad\qquad\qquad\qquad + g\,[\,r(t,u(t),\gamma(t)) + \dot{W}(t)\,], \qquad t\in[0,T] \\ u(t) = u_0(t), \qquad t\le 0. \end{cases} \tag{1}$$

In the above equation $W(t)$, $t\ge 0$ is a cylindrical Wiener process defined on a suitable probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\ge 0},\mathbb{P})$ (whose properties will be specified later) with values in a (possibly different) Hilbert space Ξ; the unknown $u(\cdot)$, representing

2010 *Mathematics Subject Classification.* 45D05, 93E20, 60H30.

Key words and phrases. Stochastic Volterra integral equation, backward stochastic differential equations, optimal control, HJB equation, feedback law.

the state of the system, is an H-valued process. Also, we model the control by the predictable process γ with values in some specified subset $\mathcal{U}$ (the set of control actions) of a third Hilbert space U.

The kernel a is completely monotonic, locally integrable and singular at 0; A is a linear operator which generates an analytical semigroup; g is bounded linear mapping from Ξ into H and r is a bounded Borel measurable mapping from $[0,T]\times H\times U$ into Ξ. We notice that the control enters the system together with the noise.

The optimal control that we wish to treat in this paper consists in minimizing a cost functional of the form

$$\mathbb{J}(u_0,\gamma)=\mathbb{E}\int_0^T l(t,u(t),\gamma(t))\mathrm{d}t+\mathbb{E}(\phi(u(T))), \tag{2}$$

where l and ϕ are given real-valued functions.

We adopt the semigroup approach based on the complete monotonicity of the kernel as been initiated in [19, 39] and recently developed for the stochastic case in [6, 8, 9]. Within this approach, equation (1) is reformulated into an abstract stochastic evolution equation without memory on a different Hilbert space X. Namely, we rewrite equation (1) as

$$\begin{cases}\mathrm{d}\mathbf{x}(t)=B\mathbf{x}(t)\mathrm{d}t+(I-B)P\,f(t,J\mathbf{x}(t))\mathrm{d}t\\ \qquad\qquad +(I-B)Pg(r(t,J\mathbf{x}(t),\gamma(t))\mathrm{d}t+\mathrm{d}W(t))\\ \mathbf{x}(0)=x.\end{cases} \tag{3}$$

Here B is the infinitesimal generator of an analytic semigroup e^{tB} on X. $P:H\to X$ is a linear mapping which acts as a sort of projection into the space X. $J:\ D(J)\subset X\to H$ is an unbounded linear functional on X, which gives a way going from $\mathbf{x}$ to the solution to problem (1). In fact, it turns out that u has the representation

$$u(t)=\begin{cases}J\mathbf{x}(t), & t>0,\\ u_0(t), & t\le 0.\end{cases}$$

For more details, we refer to the original papers [6, 32].

Further, the optimal control problem, reformulated into the state setting X, consists in minimizing the cost functional

$$\mathbb{J}(x,\gamma)=\mathbb{E}\int_0^T l(t,J\mathbf{x}(t),\gamma(t))\mathrm{d}t+\mathbb{E}\phi(J\mathbf{x}(T))$$

(where the initial condition u_0 is substituted by x and the process u is substituted by $J\mathbf{x}$). It follows that γ is an optimal control for the original Volterra equation if and only if it is an optimal control for that state equation (3).

We notice that equation (3) has unbounded coefficients. Similar stochastic problems are present in literature (see [18, 9, 16, 37, 44]), also in connection with optimal control. Usually, they arise in a wide variety of applications in physical problems, see the monograph [29, 46] or the papers [1, 47, 14] for some examples, in interacting biological populations and harvesting problems and in problems in mathematical finance. For instance, an example of physically realistic situation which we have in mind is the control of a fluid in the context of thermo-dynamic or fractional diffusion-wave equations.

Our purpose is not only to prove existence of optimal controls, but mainly to characterize them by an optimal feedback law. In other words, we wish to perform the standard program of synthesis of the optimal control that consists in the following steps: first we solve (in a suitable sense) the Hamilton-Jacobi-Bellman equation; then we prove that such a solution is the value function of the control problem and allows to construct the optimal feedback law.

We focus our attention on the following (formally written) Hamilton-Jacobi-Bellman equation ((HJB) for short)

$$\begin{cases} \frac{\partial}{\partial s} v(s,x) + \mathcal{L}_s[v(s,\cdot)](x) = \psi(s,x,\nabla v(s,x)(I-B)Pg) \\ v(T,x) = \phi(x), \end{cases} \tag{4}$$

where $\mathcal{L}_t$ is the infinitesimal generator of the Markov semigroup corresponding to the process $\mathbf{x}$

$$\mathcal{L}_t[h](x) = \frac{1}{2}\mathrm{Tr}[(I-B)Pg\nabla^2 h(x) g^* p^* (I-B)^*] \\ + \langle Bx + (I-B)Pf(t,Jx), \nabla h(x)\rangle.$$

We formulate the equation (4) in a mild sense. Setting $\{P_{s,t}[\cdot] : 0 \le s \le t\}$, to be the Markov semigroup corresponding to the process $\mathbf{x}$, we seek a function v verifying the following variation of constants formula:

$$v(s,x) = P_{s,T}[\phi](x) - \int_s^T P_{s,r}\left[\psi(r,\cdot,[\nabla v(I-B)](r,\cdot)Pg\right](x)\mathrm{d}r. \tag{5}$$

We solve this equation using a method based on a system of forward-backward stochastic differential equations. In our case, this is given by the forward equation (3) and the backward equation

$$\begin{cases} \mathrm{d}Y(s) = \psi(s,\mathbf{x}(s,t,x),Z(s))\mathrm{d}s + Z(s)\mathrm{d}W(s), \qquad s \in [t,T], \\ Y(T) = \phi(\mathbf{x}(T,s,x)) \end{cases} \tag{6}$$

where ψ is the Hamiltonian function of the control problem, defined in terms of l and r, while $\mathbf{x}(s,t,x)$ stands for the solution of equation (3) starting at time t from $x \in X$. It is classical that, under suitable assumption on l, r and ϕ, problem (6) admits a unique solution. Now if we set $v(t,x) = Y(t)$, then $v(t,x)$ is the unique mild solution to the equation (4).

We notice that the formula (5) requires v to be Gâteaux differentiable and $\nabla v(t,\mathbf{x}(t))(I-B)Pg$ to be well-defined. As we will see, to prove these facts, the crucial point is to show that we can give a meaning to $\nabla v(t,\mathbf{x}(t))(I-B)Pg$ and to identify it with the process Z coming from the BSDE (6) associated with the control problem. In fact we recall that P acts from H into the real interpolation space $X_\theta := (X, D(B))_{\theta,2}$ and it turns out that Pg does not belong to $D(B)$ but only to X_θ, for suitable $\theta \in (0,1)$. Hence we are forced to prove that the map $(t,x,h) \mapsto \nabla v(t,\mathbf{x})(I-B)^{1-\theta}h$ extends to a continuous map on $[0,T]\times X \times X$. To do that, we start by proving that this extra regularity holds, in a suitable sense, for the forward equation (3) and then it is conserved if we differentiate (in Gâteaux sense) the backward equation with respect to the process $\mathbf{x}$.

On the other hand, showing first that $\mathbf{x}(\cdot;t,x)$ is regular in Malliavin sense, we can prove that if the map $(t,x,h) \mapsto \nabla v(t,x)(I-B)^{1-\theta}h$ extends to a continuous function on $[0,T]\times X\times X$ then the processes $t \mapsto v(t,\mathbf{x}(t;s,x))$ and W admit joint quadratic variation in any interval $[s,\tau]$ and this is given by $\int_t^\tau \nabla v(r,\mathbf{x}(r;s,x))(I-B)Pg\,\mathrm{d}r$. Then we proceed exploiting the characterization of $\int_t^\tau Z(r)\mathrm{d}r$ as joint quadratic variation between Y and W in $[s,\tau]$. As a consequence, the identification between Z and $\nabla v(t,x)(I-B)P\,g$ follows by the definition of v.

Once the HJB equation has been solved, we can come back to the control problem and show that the solution of (4) is the value function, that is to say, it realizes the average minimal cost "paid" by the system starting at time t in x. More precisely, v satisfies the so-called *fundamental relation*:

$$\mathbb{J}(x,\gamma) = v(0,x) + \mathbb{E}\int_0^T [\,-\psi(s,\mathbf{x}(s),\nabla v(s,\mathbf{x}(s))(I-B)Pg) \\ + \nabla v(s,\mathbf{x}(s))(I-B)Pg\; r(s,J\mathbf{x}(s),\gamma(s)) + l(s,J\mathbf{x}(s),\gamma(s))\,]\;\mathrm{d}s. \quad (7)$$

From the last relation, we are be able to construct the optimal feedback law. In fact, equality (7) immediately implies that for every admissible control γ and any initial datum x, we have $\mathbb{J}(x,\gamma) \geq v(0,x)$ and γ is optimal if and only if the following feedback law holds:

$$\psi(t,\mathbf{x}^\gamma(t),\nabla v(t,\mathbf{x}^\gamma(t))\,(I-B)P\,g) \\ = \nabla v(t,\mathbf{x}^\gamma(t))\,(I-B)P\,g\,r(t,J\mathbf{x}^\gamma(t),\gamma(t)) + l(t,J\mathbf{x}^\gamma(t),\gamma(t))$$

where $\mathbf{x}^\gamma$ is the trajectory starting at x and corresponding to the control γ (see Corollary 3).

The present paper is a first step of our program. Indeed, we consider a stochastic optimal control problem on finite horizon and under non degeneracy assumptions on the diffusion coefficient g. Further, we suppose that r and l are Borel measurable Ξ-valued functions sufficiently smooth in order that the Hamiltonian ψ is Lipschitz continuous with respect to γ. In this way the corresponding BSDE has sublinear growth in the variable Z and can be exploited, e.g., using the techniques developed in Fuhrman and Tessitore [27] or Confortola and Briand [10].

In the present article we consider a cost with linear growth, but more general situations can be treated, for instance the case of a cost functional with quadratic growth. In such case, the hamiltonian function associated to the control problem would have a quadratic growth and the synthesis of the control problem could be obtained in similar way by using the result on BSDEs with quadratic generator.

W also stress that an optimal control problem for stochastic Volterra equations is treated in [9], where the drift term of the equation has a linear growth on the control variable, the cost functional has a quadratic growth, and the control process belongs to the class of square integrable, adapted processes with no bound assumed on it. The substantial difference, in comparison with the cited paper, consists in the fact that, at our knowledge, our paper is the first attempt to study existence and uniqueness of solutions for the (HJB) equation corresponding to the Volterra equation (1) and characterize the optimal control by a feedback law.

The paper is organized as follows: the next section is devoted to notations; in Section 3 we transpose the problem in the infinite dimensional framework; in Section 4 we establish the existence result for the uncontrolled equation, while in Section 5 we study the controlled system. In section 6 we consider the regularity of the uncontrolled solution, in particular in the sense of Malliavin, while in Section 7 we will study the BSDE associated to the problem. Finally, in Section 8 we will exploit the corresponding (HJB) equation in order to construct an optimal feedback and an optimal control (see, to this end, Section 9).

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

The letters Ξ, H, U will always denote Hilbert spaces. Scalar product is denoted $\langle \cdot, \cdot \rangle$, with a subscript to specify the space, if necessary. All Hilbert spaces are assumed to be real and separable.

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 from Ξ 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 $t \geq 0$, $\mathbb{E}(\langle \xi_1, W(t)\rangle\langle \xi_2, W(t)\rangle) = \langle \xi_1, \xi_2\rangle$.

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

In the sequel we will refer to the following class of stochastic processes with values in an Hilbert space K:

1. $L^p(\Omega; L^2(0, T; K))$ defines, for $T > 0$ and $p \geq 1$, the space of equivalence classes of progressively measurable processes $y : \Omega \times [0, T) \to K$, such that
$$|y|^p_{L^p(\Omega;L^2(0,T;K))} := \mathbb{E}\left[\int_0^T |y(s)|^2_K \mathrm{d}s\right]^{p/2} < \infty.$$
Elements of $L^p(\Omega; L^2(0, T; K))$ are identified up to modification.
2. $L^p(\Omega; C([0, T]; K))$ defines, for $T > 0$ and $p \geq 1$, the space of equivalence classes of progressively measurable processes $y : \Omega \times [0, T) \to K$, with continuous paths in K, such that the norm
$$|y|^p_{L^p(\Omega;C([0,T];K))} := \mathbb{E}\left[\sup_{t\in[0,T]} |y(t)|^p_K\right]$$
is finite. Elements of $L^p(\Omega; C([0, T]; K))$ are identified if they are indistinguishable.

We also recall notation and basic facts on a class of differentiable maps acting among Banach spaces, particularly suitable for our purposes (we refer the reader to Fuhrman and Tessitore [27] or Ladas and Lakshmikantham [35, Section 1.6] (1970) for details and properties). Let now X, Y, 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 it 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 $G(F)$ is in $\mathcal{G}^1(X, Z)$ and the chain rule holds: $\nabla(G(F))(x) = \nabla G(F(x))\nabla F(x)$. When F depends on additional arguments, the previous definitions and properties have obvious generalizations. In addition to the ordinary chain rule stated above, a chain rule for the Malliavin derivative operator holds: for the reader convenience we refer to Section 6.1 for a brief introduction to this subject.

Moreover, we assume the following.

Hypothesis 2.1. 1. The kernel $a : (0, \infty) \to \mathbb{R}$ is completely monotonic, locally integrable, with $a(0+) = +\infty$. The singularity in 0 shall satisfy some technical conditions that we make precise in Section 3.

2. $A : D(A) \subset H \to H$ is a sectorial operator in H. Thus A generates an analytic semigroup e^{tA}.
3. The process $(W(t))_{t\geq 0}$ is a cylindrical Wiener process defined on a complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with values in the Hilbert space Ξ.
4. The function $f : [0,T] \times H \to H$ is measurable, for every $t \in [0,T]$ the function $f(t,\cdot) : H \to H$ is continuously Gâteaux differentiable and there exist constants $L > 0$ and $C > 0$ such that

$$|f(t,u) - f(t,v)| \leq L_f |u - v|, \qquad t \in [0,T],\ u, v \in H;$$
$$|f(t,0)| + \|\nabla_u f(t,u)\|_{\mathcal{L}(H)} \leq C, \qquad t \in [0,T],\ u \in H$$

5. g belongs to $L_2(\Xi, H)$, that is to the space of Hilbert-Schmidt operators from Ξ to H, endowed with the Hilbert-Schmidt norm $||g||^2_{L_2(\Xi,H)} = \mathrm{Tr}(gg^*)$.
6. The function $r : [0,T] \times H \times U \to \Xi$ is Borel measurable for a.e. $t \in [0,T]$ and there exists a positive constant $b_r > 0$ such that

$$|r(t,u_1,\gamma) - r(t,u_2,\gamma)| \leq b_r |u_1 - u_2|, \qquad u_1, u_2 \in H, \gamma \in U;$$
$$|r(t,u,\gamma)|_\Xi \leq b_r, \qquad u \in H, \gamma \in U.$$

The initial condition satisfies a global exponential bound as well as a linear growth bound as $t \to 0$:

Hypothesis 2.2. 1. There exist $M_1 > 0$ and $\omega > 0$ such that $|u_0(t)| \leq M_1 e^{\omega t}$ for all $t \leq 0$;

2. There exist $M_2 > 0$ and $\tau > 0$ such that $|u_0(t) - u_0(0)| \leq M_2 |t|$ for all $t \in [-\tau, 0]$;
3. $u_0(0) \in H_\varepsilon$ for some $\varepsilon \in (0, 1/2)$, where $H_\varepsilon := (H, D(A))_{\varepsilon,2}$ denotes the real interpolation space of order ε between A and H.

Concerning the functions l and ϕ appearing in the cost functional we make the following general assumptions:

Hypothesis 2.3. 1. The functions $l : [0,T] \times H \times U \to \mathbb{R}$ and $\phi : H \to \mathbb{R}$ are Borel measurable;

2. There exist a positive constant C such that for any $\gamma \in U$ the following bound is satisfied

$$|l(t,u_1,\gamma) - l(t,u_2,\gamma)| \leq C(1 + |u_1| + |u_2|)|u_1 - u_2|, \qquad u_1, u_2 \in H, \gamma \in U$$
$$0 \leq |l(t,0,\gamma)| \leq C.$$

3. There exists $L > 0$ such that, for every $u_1, u_2 \in H$ we have

$$|\phi(u_1) - \phi(u_2)|_H \leq L_\phi |u_1 - u_2|_H.$$

Moreover, $\phi \in \mathcal{G}^1(H, \mathbb{R})$.

We consider the following notion of solution for the Volterra equation (1).

Definition 2.4. We say that a process $u = (u(t))_{t\geq 0}$ is a solution to equation (1) if u is an adapted, p-mean integrable $(p \geq 1)$ continuous H-valued predictable process and the identity

$$\int_{-\infty}^{t} a(t-s)\langle u(s),\zeta\rangle_H \mathrm{d}s = \langle \bar{u},\zeta\rangle_H + \int_0^t \langle u(s), A^\star\zeta\rangle_H \mathrm{d}s \\ + \int_0^t \langle f(s,u(s)),\zeta\rangle \mathrm{d}s + \int_0^t \langle g\, r(s,u(s),\gamma(s)),\zeta\rangle_H \mathrm{d}s + \langle gW(t),\zeta\rangle$$

holds $\mathbb{P}$-a.s. for arbitrary $t \in [0,T]$ and $\zeta \in D(A^\star)$, with $A^\star$ being the adjoint of the operator A and

$$\bar{u} = \int_{-\infty}^{0} a(-s)u_0(s)\mathrm{d}s.$$

3. **The analytical setting.** A completely monotone kernel $a : (0,\infty) \to \mathbb{R}$ is a continuous, monotone decreasing function, infinitely often derivable, such that

$$(-1)^n \frac{\mathrm{d}^n}{\mathrm{d}t^n} a(t) \geq 0, \quad t \in (0,\infty),\ n = 0,1,2,\dots$$

By Bernstein's Theorem [46, pag. 90], a is completely monotone if and only if there exists a positive measure ν on $[0,\infty)$ such that

$$a(t) = \int_{[0,\infty)} e^{-\kappa t}\nu(\mathrm{d}\kappa), \quad t > 0.$$

From the required singularity of a at $0+$ we obtain that $\nu([0,+\infty)) = a(0+) = +\infty$ while for $s > 0$ the Laplace transform $\hat{a}$ of a verifies

$$\hat{a}(s) = \int_{[0,+\infty)} \frac{1}{s+\kappa}\,\nu(\mathrm{d}\kappa) < +\infty.$$

We introduce the quantity

$$\alpha(a) = \sup\left\{\rho \in \mathbb{R} :\ \int_c^\infty s^{\rho-2}\frac{1}{\hat{a}(s)}\mathrm{d}s < \infty\right\}$$

and we make the following assumption:

Hypothesis 3.1. $\alpha(a) > 1/2$.

Remark 1. The function $a(t) = e^{-\omega t} t^{\alpha-1}$, where $\omega > 0$, is an example of completely monotone kernel, with Laplace transform $\hat{a}(s) = \Gamma[\alpha]\,(\omega+s)^{-\alpha}$; an easy computation shows that $\alpha(a) = 1-\alpha$, hence we satisfy assumption 3.1 whenever we take $\alpha \in (0,\frac{1}{2})$.

Remark 2. It is known from the theory of deterministic Volterra equations that the singularity of a helps smoothing the solution. We notice that $\alpha(a)$ is independent on the choice of $c > 0$ and this quantity describes the behavior of the kernel near 0; by this way we ensure that smoothing is sufficient to keep the stochastic term tractable.

Under the assumption of complete monotonicity of the kernel, a semigroup approach to a type of abstract integro-differential equations encountered in linear viscoelasticity was introduced in [19] and extended to the case of Hilbert space valued equations in [6]. In order to simplify the exposition we quote from [6] the main result concerning the derivation of the state equation (3).

We will see that this approach allow us to treat the case of semilinear, stochastic integral equations; we start for simplicity with the equation

$$\begin{aligned} \frac{d}{dt}\int_{-\infty}^{t} a(t-s)u(s)\,\mathrm{d}s &= Au(t)+\bar f(t), \qquad && t\in[0,T] \\ u(t) &= u_0(t), && t\le 0, \end{aligned} \tag{8}$$

where $\bar f$ belongs to $L^1(0,T;H)$ and u_0 satisfies Hypothesis 2.2. A weak solution of equation (8) is defined as follows

Definition 3.2. We say that a function $u \in L^1([0,T];H)$ is a weak solution to equation (8) if u satisfies the identity

$$\int_{-\infty}^{t} a(t-s)\langle u(s),\zeta\rangle_H \mathrm{d}s = \langle \bar u,\zeta\rangle_H + \int_0^t \langle u(s), A^\star\zeta\rangle_H \mathrm{d}s + \int_0^t \langle \bar f(s),\zeta\rangle \mathrm{d}s$$

holds for arbitrary $t\in[0,T]$ and $\zeta\in D(A^\star)$, with $A^\star$ being the adjoint of the operator A and

$$\bar u = \int_{-\infty}^{0} a(-s)u_0(s)\mathrm{d}s.$$

In order to solve (8), we start from the following identity, which follows by Bernstein's theorem:

$$\int_{-\infty}^{t} a(t-s)u(s)\,\mathrm{d}s = \int_{-\infty}^{t}\int_{[0,+\infty)} e^{-\kappa(t-s)}\,\nu(\mathrm{d}\kappa)\,u(s)\mathrm{d}s = \int_{[0,+\infty)} \mathbf{x}(t,\kappa)\,\nu(\mathrm{d}\kappa)$$

where we introduce the state variable

$$\mathbf{x}(t,\kappa) = \int_{-\infty}^{t} e^{-\kappa(t-s)}u(s)\,\mathrm{d}s. \tag{9}$$

Formal differentiation yields

$$\frac{\partial}{\partial t}\mathbf{x}(t,\kappa) = -\kappa\mathbf{x}(t,\kappa)+u(t), \tag{10}$$

while the integral equation (8) can be rewritten

$$\int_{[0,+\infty)} (-\kappa\mathbf{x}(t,\kappa)+u(t))\,\nu(\mathrm{d}\kappa) = Au(t)+\bar f(t). \tag{11}$$

Now, the idea is to use equation (10) as the state equation, with $B\mathbf{x} = -\kappa\mathbf{x}(\kappa)+u$, while (11) enters in the definition of the domain of B.

In the following we want to give a formal description of the arguments above by introducing the suitable state space X for $\mathbf{x}(t,\cdot)$ and suitable operators.

Definition 3.3. Let X denote the space of all Borel measurable functions $\mathbf{y}$: $[0,+\infty)\to H$ such that the seminorm

$$\|\tilde{\mathbf{y}}\|_X^2 := \int_{[0,+\infty)} (\kappa+1)|\mathbf{y}(\kappa)|_H^2\,\nu(\mathrm{d}\kappa)$$

is finite.

We shall identify the classes $\mathbf{y}$ with respect to equality almost everywhere in ν.

Definition 3.4. We let the operator $Q : L^1((-\infty,0];H) \to L^\infty([0,+\infty);H,\mathrm{d}\nu)$ be given by

$$\mathbf{x}(0,\kappa) = Qu_0(\kappa) = \int_{-\infty}^0 e^{-\kappa s} u_0(s)\,\mathrm{d}s. \tag{12}$$

The operator Q maps the initial value of the stochastic Volterra equation in the initial value of the abstract state equation. Different initial conditions of the Volterra equation generate different initial conditions of the state equation. It has been proved in [6, Lemma 3.16] that Q maps $L^1(-\infty,0;H) \cap L^\infty(-\infty,0;H)$ into X. In particular, Hypothesis 2.2 is necessary in order to have a greater regularity on the inial value of the state equation. In fact in this case [6, Lemma 3.16 (3)]) shows that Qu_0 belongs to X_η for $\eta \in (0,\frac{1}{2})$.

Remark 3. We stress that under our assumptions we are able to treat, for instance, initial conditions for the Volterra equation of the following form

$$u_0(t) = \begin{cases} 0, & t \in (-\infty,-\delta); \\ \bar{u} & t \in [-\delta,0] \end{cases}$$

provided $\bar{u}$ has a suitable regularity.

We introduce a rigorous definition for the leading operator of the reformulated state equation (11).

Definition 3.5. Let $\mathbf{x} \in X$ and $u \in H$. We define the operators $\tilde{J} : D(\tilde{J}) \subset X \to X$ by

$$D(\tilde{J}) := \left\{ \mathbf{x} \in X : \exists u \in H s.t. \int_{[0,\infty)} (\kappa+1)|\kappa\mathbf{x}(\kappa) - u|^2 \nu(\mathrm{d}\kappa) < +\infty \right\}$$

$$\tilde{J}\mathbf{x} := u;$$

$B : D(B) \subset X \to X$

$$D(B) := \left\{ \mathbf{x} \in D(\tilde{J}) | \tilde{J}\mathbf{x} \in D(A), \int_{[0,\infty)} -\kappa\mathbf{x}(\kappa) + \tilde{J}\mathbf{x}(\kappa)\nu(\mathrm{d}\kappa) = A\tilde{J}\mathbf{x} \right\}$$

$$(B\mathbf{x})(\kappa) := -\kappa\mathbf{x}(\kappa) + \tilde{J}\mathbf{x}$$

and $P : H \to X$

$$(Pu)(\kappa) := \frac{1}{\kappa+1} R(\hat{a}(1),A)u,$$

$R(\hat{a}(1),A)$ being the resolvent $R(\hat{a}(1),A) := (\hat{a}(1) - A)^{-1}$.

We quote from [6] the main result concerning the state space setting for stochastic Volterra equations in infinite dimensions and the properties of the linear operators introduced above.

Theorem 3.6 (State space setting)**.** *Let A, a, W, $\alpha(a)$ be as in Hypotheses 2.1 and 3.1; choose numbers $\eta \in (0,1)$, $\theta \in (0,1)$ such that*

$$\eta > \frac{1}{2}(1-\alpha(a)), \quad \theta < \frac{1}{2}(1+\alpha(a)), \quad \theta - \eta > \frac{1}{2}. \tag{13}$$

Let X, Q, B, $\tilde{J}$ be defined as in Definition 3.3, 3.4 and 3.5. Then

1) *the operator $B : D(B) \subset X \to X$ is a densely defined sectorial operator generating an analytic semigroup e^{tB};*

2) the operator $\tilde{J} : D(B) \to D(A)$ is onto and admits a unique extension J as a continuous linear operator $J : X_\eta \to H$, where X_η denotes the real interpolation space $X_\eta := (X, D(B))_{\eta,2}$;
3) the linear operator P maps H continuously into X_θ.

Proof. The assertion of the theorem can be proved by following [6, Section 3]. □

Our idea is to rewrite the semilinear inhomogeneous integral equation (8) into an abstract evolution equation on the state space X, by using the linear operators introduced above. More precisely, equation (8) can be reformulated as

$$\begin{aligned} \mathbf{x}'(t) &= B\mathbf{x}(t)\mathrm{d}t + (I - B)P\bar{f}(t), \qquad t \geq 0 \\ \mathbf{x}(0) &= Qu_0. \end{aligned} \tag{14}$$

Since B is the generator of an analytic semigroup in X, we can give a meaning to the solution of (14) in a mild sense as

$$\mathbf{x}(t) = e^{(t-s)B}\mathbf{x}(0) + \int_s^t e^{(t-\sigma)B}(I - B)P\bar{f}(\sigma)\mathrm{d}\sigma. \tag{15}$$

In the following we will explore the relation between the abstract state space and Problem (8).

Theorem 3.7. *We assume Hypotheses 2.1 (i-iii), 2.2, $\bar{f} \in L^1(0, T; H)$ and let X, Q, B, P be defined as in Definition 3.3, 3.4, 3.5 and J as in Theorem 3.6 (ii); choose numbers $\eta \in (0, 1)$, $\theta \in (0, 1)$ such that*

$$\eta > \frac{1}{2}(1 - \alpha(a)), \quad \theta < \frac{1}{2}(1 + \alpha(a)), \quad \theta - \eta > \frac{1}{2}.$$

Set $x = Qu_0$. If $\mathbf{x}$ given by (15) is the solution in mild sense of the abstract Cauchy problem (14), then the following assertions hold:

1. $\mathbf{x}(t) \in L^1_{loc}([0, \infty); X_\eta)$, thus $J\mathbf{x}(t)$ is well defined almost everywhere, and $J\mathbf{x} \in L^1_{loc}([0, \infty); H)$;
2. The function $u : [0, \infty) \to H$, defined by

$$u(t) = \begin{cases} u_0(t), & if\ t < 0, \\ J\mathbf{x}(t) & if\ t \geq 0, \end{cases} \tag{16}$$

is the unique weak solution of (8).

Proof. The result is a direct consequence of [6, Theorem 2.7]. We notice that in the aforementioned result the authors prove the existence and uniqueness of an integrated solution, which is, clearly also a weak solution (compare [6, Definition 2.6] and Definition 3.2 above). □

It is remarkable that B generates an analytic semigroup, since in this case we have at our disposal a powerful theory of optimal regularity results. In particular, besides the interpolation spaces X_θ, we may construct the extrapolation space X_{-1}, which is the completion of X with respect to the norm $\|\mathbf{x}\|_{-1} := \|(B - I)\mathbf{x}\|_X$.

Assume for simplicity that B is of negative type with growth bound $\omega_0 > 0$ (otherwise, one may consider $B - \omega_0$ instead of B in the following discussion). The semigroup e^{tB} extends to X_{-1} and the generator of this extension, that we denote B_{-1}, is the unique continuous extension of B to an isometry between X and X_{-1}. See for instance [20, Definition 5.4] for further details.

Remark 4. 1. In the sequel, we shall always denote the operator with the letter B, even in case where formally B_{-1} should be used instead. This should cause no confusion, due to the similarity of the operators.

2. We notice that the interpolation spaces $X_\gamma, \gamma \in (0,1)$ and the domains of fractional powers of B are linked by the following inclusion

$$D((-B)^{\gamma+\varepsilon}) \subset X_\gamma, \varepsilon > 0,$$

(see [17, Proposition A.15]) Hence, if x is any element of $X_\gamma, \gamma \in (0,1)$, we can find sufficiently small $\varepsilon > 0$ such that

$$\|x\|_\gamma \leq \|x\|_{D((-B)^{\gamma+\varepsilon})} = \|(I-B)^{\gamma+\varepsilon} x\|_X.$$

We will frequently make use of the above inequality in the following, especially when we will need to prove that an element of X is, instead in X_η.

3. We notice that since $B - \omega_0$ is the infinitesimal generator of the analytic semigroup of contraction $(e^{-\omega_0 t} e^{tB})_{t \geq 0}$ on X with $0 \in \rho(B - \omega_0)$ we have that $X_\theta = (D(-B)^\theta)$ for any $\alpha \in (0,1)$ and there exists $M > 0$ such that

$$\|(B-\omega_0) e^{-\omega_0 t} e^{tB}\|_{\mathcal{L}(X_\theta, X_\eta)} \leq M t^{-1+\theta-\eta}, \qquad t \geq 0$$

Moreover, the norm in X_θ is equivalent to the norm $x \mapsto \|(-B)^\theta x\|_X$ on $D((-B)^\theta)$. The above properties and estimate are well explained in [32, pgg. 23,25 Theorems 1.4.27, Lemma 1.4.15, Corollary 1.4.30(ii)]. For more detail see also [36, pages 114, 97, 120, Theorems 4.3.5, 4.2.6, Corollary 4.3.12]. We assume, for simplicity, that $\omega_0 > -1$, so that the same estimate holds with $B - I$ instead of $B - \omega_0$, i.e.

$$\|(B-I) e^{tB}\|_{\mathcal{L}(X_\theta, X_\eta)} \leq M_T t^{-1+\theta-\eta}, \qquad t \in [0,T],$$

where M_T is a positive constant depending only on T. In the next pages we will make use of this inequality frequently and of the equivalence between interpolation spaces and the domains of the fractional power of B, especially when studying the forward equation, its differentiability and the Malliavin regularity.

4. The state equation: Existence and uniqueness.

In this section, motivated by the construction in Section 3, we shall establish existence and uniqueness result for the following stochastic controlled Cauchy problem on the space X defined in Section 3:

$$\begin{cases} \mathrm{d}\mathbf{x}(t) = B\mathbf{x}(t)\mathrm{d}t + (I-B)Pf(t, J\mathbf{x}(t))\mathrm{d}t + \\ \qquad\qquad (I-B)P\, r(t, J\mathbf{x}(t), \gamma(t))\mathrm{d}t + (I-B)Pg\,\mathrm{d}W(t) \\ \mathbf{x}(s) = x. \end{cases} \tag{17}$$

for $0 \leq s \leq t \leq T$ and initial condition $x \in X_\eta$, for $\eta > \frac{1}{2}(1-\alpha(a))$. The above expression is only formal since the coefficients do not belong to the state space; however, we can give a meaning to the mild form of the equation:

Definition 4.1. We say that a continuous, X_η-valued, adapted process $\mathbf{x} = (\mathbf{x}(t))_{t \in [s,T]}$ is a (mild) solution of the state equation (17) if $\mathbb{P}$-a.s.,

$$\mathbf{x}(t) = e^{(t-s)B} x + \int_s^t e^{(t-\sigma)B}(I-B)Pf(\sigma, J\mathbf{x}(\sigma))\mathrm{d}\sigma$$
$$+ \int_s^t e^{(t-\sigma)B}(I-B)P\, r(\sigma, J\mathbf{x}(\sigma), \gamma(\sigma))\mathrm{d}\sigma + \int_s^t e^{(t-\sigma)B}(I-B)Pg\,\mathrm{d}W(\sigma).$$

Before proceeding with the existence and uniqueness result, we list relevant properties of the nonlinear term of the reformulated equation.

Remark 5. It follows directly by the properties of the nonlinear mappings f, r and the operator J that, under Hypothesis 2.1 the function $(t,x) \mapsto f(t,Jx)$ from $[0,T] \times X_\eta$ into H is measurable and it verifies the following estimates

$$|f(t,Jx) - f(t,Jy)|_H \le L_f \|J\|_{L(X_\eta;H)} \|x-y\|_\eta \qquad t \in [0,T],\ x,y \in X_\eta;$$
$$|f(t,J0)| \le C, \qquad t \in [0,T];$$
$$|r(t,Jx,\gamma) - f(t,Jy,\gamma)|_H \le b_r \|J\|_{L(X_\eta;H)} \|x-y\|_\eta \qquad t \in [0,T],\ x,y \in X_\eta, \gamma \in \mathcal{U};$$
$$|r(t,Jx,\gamma)| \le C, \qquad t \in [0,T], x \in X_\eta.$$

Moreover, for every $t \in [0,T]$, $(t,x) \mapsto f(t,J(\,\cdot\,))$ has a Gâteaux derivative at every point $x \in X_\eta$: this is given by the linear operator on X_η

$$\nabla_x(f(t,Jx))[h] = \nabla_u f(t,Jx)[Jh].$$

Finally, the function $(x,h) \to \nabla f(t,Jx)[h]$ is continuous as a map $X_\eta \times X_\eta \to \mathbb{R}$ and $\|\nabla_u f(t,Jx)\|_\eta \le C$, for $t \in [0,T]$, $x \in X_\eta$ and a suitable constant $C > 0$.

Let us state the main existence result for the solution of equation (17).

Theorem 4.2. *Under Hypothesis 2.1, 2.2, chosen η, θ, p such that*

$$\eta > \frac{1}{2}(1-\alpha(a)) \qquad \theta < \frac{1}{2}(1+\alpha(a)) \qquad \theta - \eta > \frac{1}{2} \qquad \frac{1}{p} < \theta - \eta - \frac{1}{2}$$

for an arbitrary predictable process γ with values in $\mathcal{U}$, for every $0 \le s \le t \le T$ and $x \in X_\eta$, there exists a unique adapted process $\mathbf{x} \in L^p(\Omega, C([s,T];X_\eta))$ solution of (17). Moreover, the estimate

$$\mathbb{E} \sup_{t\in[s,T]} ||\mathbf{x}(t)||_\eta^p \le C(1 + ||x||_\eta^p) \tag{18}$$

holds for some positive constant C depending on T, p and the parameters of the problem.

Proof. The proof of the above theorem proceeds, basically, on the same lines as the proof of Theorem 3.2 in [8]. First, we define a mapping $\mathcal{K}$ from $L^p(\Omega; C([0,T];X_\eta))$ to itself by the formula

$$\mathcal{K}(\mathbf{x})(t) := e^{(t-s)B}x + \Lambda(\mathbf{x})(t) + \Delta(\mathbf{x})(t) + \Gamma(t), \tag{19}$$

where the second, third and last term in the right side of (19) are given by

$$\Lambda(\mathbf{x})(t) = \int_s^t e^{(t-\tau)B}(I-B)Pf(\tau, J\mathbf{x}(\tau))\mathrm{d}\tau \tag{20}$$

$$\Delta(\mathbf{x})(t) = \int_s^t e^{(t-\tau)B}(I-B)Pg\, r(\tau, J\mathbf{x}(\tau), \gamma(\tau))\mathrm{d}\tau \tag{21}$$

$$\Gamma(t) = \int_s^t e^{(t-\tau)B}(I-B)Pg\,\mathrm{d}W(\tau) \tag{22}$$

Then, we will prove that the mapping $\mathcal{K}$ is a contraction on $L^p(\Omega; C([0,T];X_\eta))$ with respect to the equivalent norm

$$|\!|\!|\mathbf{x}|\!|\!|_\eta^p := \mathbb{E} \sup_{t\in[0,T]} e^{-\beta p t} ||\mathbf{x}(t)||_\eta^p,$$

where $\beta > 0$ will be chosen later. For simplicity we fix the initial time $s = 0$ and write $\Lambda(t), \Delta(t)$ instead of $\Lambda(\mathbf{x})(t), \Delta(\mathbf{x})(t)$.

We start by proving that Γ is a well-defined mapping on the space $L^p(\Omega; C([0,T]; X_\eta))$ and to give estimates on their norm. The proof of this part is divided in two steps and is proved combining the classical factorization method by Da Prato and Zabczyk [17, Theorem 5.9] and the result in [16, Proposition A.1.1]. According to these results and taking into account Remark 4, we notice that the thesis is equivalent to the statement: $t \mapsto (I-B)^{\eta+\varepsilon}\Gamma(t) \in L^p(\Omega; C([0,T]; X))$, for suitable $\varepsilon > 0$.

Step 1. For given $\gamma \in (0,1)$ and $\eta \in (0,1)$, the following identity holds:

$$e^{-\beta t}\,(I-B)^{\eta+\varepsilon}\Gamma(t) = c_\gamma \int_0^t e^{-\beta(t-\tau)}(t-\tau)^{\gamma-1} e^{(t-\tau)B} y_\eta(\tau)\,\mathrm{d}\tau$$

where y_η is the process

$$y_\eta(\tau) = \int_0^\tau e^{-\beta\sigma}(\tau-\sigma)^{-\gamma} e^{-\beta(\tau-\sigma)} e^{(\tau-\sigma)B}(I-B)^{\eta+\varepsilon+1} Pg\,\mathrm{d}W(\sigma). \tag{23}$$

We notice that y_η is a well-defined process with values in X. In fact, g maps Ξ into H and P maps H into X_θ for arbitrary $\theta > \frac{1+\alpha(a)}{2}$. Moreover, since the semigroup e^{tB} is analytic, $\theta > \eta$ and $\eta > \frac{1}{2}$, for each $t > 0$ the operator $e^{tB}(I-B)$ maps X_θ into X_η and $(I-B)^{\eta+\varepsilon}$ is well-defined on X_η (due to the inclusion in $D((I-B)^{\eta+\varepsilon})$). As a consequence, there exists a constant M such that for $t \in [0,T]$ the following estimates holds:

$$\begin{aligned} &\|e^{tB}(I-B)\|_{\mathcal{L}(X_\theta, X_\eta)} \le M t^{\theta-\eta-1};\\ &\|(I-B)^{\eta+\varepsilon}x\|_X \le M\|x\|_\eta, \qquad x \in X_\eta. \end{aligned}$$

We shall estimate the $L^p(\Omega; X)$-norm of this process:

$$\mathbb{E}|y_\eta(\tau)|^p = \mathbb{E}\left|\int_0^\tau e^{-\beta\sigma}(\tau-\sigma)^{-\gamma} e^{-\beta(\tau-\sigma)} e^{(\tau-\sigma)B}(I-B)^{\eta+\varepsilon+1} Pg\,\mathrm{d}W(\sigma)\right|^p.$$

Proceeding as in [17, Lemma 7.2] this leads to

$$\begin{aligned} &\mathbb{E}|y_\eta(\tau)|^p\\ &\quad \le C\left[\int_0^\tau \|e^{-\beta\sigma} e^{(\tau-\sigma)B}(I-B)^{\eta+\varepsilon+1}Pg\,(\tau-\sigma)^{-\gamma}e^{-\beta(\tau-\sigma)}\|^2_{L_2(\Xi,X)}\,\mathrm{d}\sigma\right]^{p/2}. \end{aligned}$$

hence

$$\|e^{(\tau-\sigma)B}(I-B)^{\eta+\varepsilon+1}Pg\|_{L_2(\Xi,X)} \le C(\tau-\sigma)^{\theta-1-\eta-\varepsilon}\|g\|_{L_2(\Xi,H)}$$

and the process y_η is estimated by

$$\mathbb{E}|y_\eta(\tau)|^p \le C\left(\int_0^\tau e^{-2\beta\sigma}\|g\|^2_{L_2(\Xi,H)}\,e^{-2\beta(\tau-\sigma)}\,(\tau-\sigma)^{-2(\gamma+1+\eta+\varepsilon-\theta)}\,\mathrm{d}\sigma\right)^{p/2}$$

We apply Young's inequality to get

$$||y_\eta||^p_{L^p_{\mathcal{F}}(\Omega;L^p(0,T;X))} = \left(\mathbb{E}\int_0^T |y_\eta(\tau)|^p \,\mathrm{d}\tau\right)$$
$$\leq C\left[\left(\int_0^T e^{-2\beta\sigma}\|g\|^2_{L_2(\Xi,H)}\,\mathrm{d}\sigma\right)^{2/p}\right]^{p/2}$$
$$\left(\int_0^\infty e^{-\tau}(2\beta)^{1+2\gamma-2(\theta-\eta-\varepsilon)}\tau^{-2(\gamma+1+\eta+\varepsilon-\theta)}\,\mathrm{d}\tau\right)^{p/2}$$

where the last integral is obtained by a change of variables. Hence, for any $\gamma+\varepsilon<(\theta-\eta)-1/2$ (notice that we can always choose $\gamma+\varepsilon>0$ small enough such that this holds) we obtain

$$\|y_\eta\|_{L^p(\Omega;L^p(0,T;H))} \leq C_T\,(2\beta)^{\frac{1}{2}+\gamma+\varepsilon-(\theta-\eta)}\left[\int_0^T e^{-p\beta\sigma}\|g\|^p_{L_2(\Xi,H)}\,\mathrm{d}\sigma\right]^{1/p}. \tag{24}$$

Now, taking into account the assumptions on g, we estimate the integral term above and we finally arrive at

$$\|y_\eta\|_{L^p(\Omega;L^p(0,T;X))} \leq C(2\beta)^{\frac{1}{2}+\gamma-(\theta-\eta-\varepsilon)}. \tag{25}$$

It follows that in particular $y_\eta\in L^p(0,T;X)$, $\mathbb{P}$-a.s.

Step 2. In [16, Appendix A] it is proved that for any $\gamma\in(0,1)$, p large enough such that $\gamma-\frac{1}{p}>0$, the linear operator

$$R_\gamma\phi(t) = \int_0^t (t-\sigma)^{\gamma-1}e^{(t-\sigma)B}\phi(\sigma)\,\mathrm{d}\sigma \tag{26}$$

is a bounded operator from $L^p(0,T;X)$ into $C([0,T];X)$. Using this result in Step 1. the thesis follows.

In a similar (and easier) way it is possible to show that $\Lambda(t)$ and $\Delta(t)$ belong to $L^p(\Omega,C([0,T];X_\eta))$ and moreover that the following estimates hold

$$\mathbb{E}\left(\sup_{t\in[0,T]} e^{-\beta pt}\|\Lambda(t)\|^p_\eta\right)^{1/p} \leq C(L_f,\|J\|,T)\beta^{\eta+\varepsilon-\theta}(1+|\!|\!|\mathbf{x}|\!|\!|^p_\eta) \tag{27}$$

$$\mathbb{E}\left(\sup_{t\in[0,T]} e^{-\beta pt}\|\Delta(t)\|^p_\eta\right)^{1/p} \leq C(\|g\|_{L^2(\Xi,H)},b_r,T)\beta^{\eta+\varepsilon-\theta}(1+|\!|\!|\mathbf{x}|\!|\!|^p_\eta), \tag{28}$$

where L_f, b_r are the constant in Hypothesis 2.1. Hence, we conclude that $\mathcal{K}$ maps $L^p(\Omega;C([0,T];X_\eta))$ into itself; but this follows immediately from the analyticity of the semigroup, provided that e^{tB} is extended by a constant x for $t<s$:

$$e^{(t-s)B}x = x \quad \text{for } t<s.$$

Now we claim that $\mathcal{K}$ is a contraction in $L^p(\Omega,C([0,T];X_\eta))$. Let $\mathbf{x}$, $\mathbf{y}$ in it and consider $\mathcal{K}(\mathbf{x})-\mathcal{K}(\mathbf{y})$; then

$$|\!|\!|\mathcal{K}(\mathbf{x})-\mathcal{K}(\mathbf{y})|\!|\!|^p \leq c_p\left(|\!|\!|\Lambda(\mathbf{x})-\Lambda(\mathbf{y})|\!|\!|^p + |\!|\!|\Delta(\mathbf{x})-\Delta(\mathbf{y})|\!|\!|^p\right).$$

Straightforward calculation show that the following estimates hold:

$$\|\Lambda(\mathbf{x}) - \Lambda(\mathbf{y})\|^p \le C(L_f, \|J\|, T)\beta^{\eta+\varepsilon-\theta}\|\mathbf{x} - \mathbf{y}\|$$
$$\|\Delta(\mathbf{x}) - \Delta(\mathbf{y})\|^p \le C(\|g\|_{L_2(\Xi,H)}, b_r, T)\beta^{\eta+\varepsilon-\theta}\|\mathbf{x} - \mathbf{y}\|$$

Now find β large enough such that

$$(C(L_f, \|J\|, T) + C(\|g\|_{L_2(\Xi,H)}, b_r, T))\beta^{\eta+\varepsilon-\theta} \le \delta < 1$$

(this is possible by choosing ε such that $\eta + \varepsilon - \theta$ is negative). Then, $\mathcal{K}$ becomes a contraction on the time interval $[0, T]$ and by the Banach fixed point argument we get that there exists a unique solution of the mild equation (17) on $[0, T]$.

Since the solution to (17) verifies $X = \mathcal{K}(X)$ we also deduce from the above computations that

$$\|X\|^p = \|\mathcal{K}(X)\|^p \le \delta(1 + \|X\|^p) + C(T)\|x\|_\eta^p$$

hence

$$\|X\| \le C(1 + \|x\|_\eta).$$

□

Remark 6. In the following it will be also useful to consider the uncontrolled version of equation (17), namely:

$$\begin{cases} \mathrm{d}\mathbf{x}(t) = B\mathbf{x}(t)\mathrm{d}t + (I - B)Pf(t, J\mathbf{x}(t))\mathrm{d}t + (I - B)Pg\,\mathrm{d}W(t) \\ \mathbf{x}(0) = x. \end{cases} \tag{29}$$

We will refer to (29) as the forward equation. We then notice that existence and uniqueness for the above equation can be treated in an identical way as in the proof of Theorem 4.2.

5. **The controlled stochastic Volterra equation.** As a preliminary step for the sequel, we state two results of existence and uniqueness for (a special case of) the original Volterra equation. The proofs can be found in [9, Section 2].

Proposition 1. *The linear equation*

$$\begin{gathered} \frac{d}{dt}\int_{-\infty}^t a(t-s)u(s)\,\mathrm{d}s = Au(t), \qquad t \in [0, T] \\ u(t) = 0, \qquad t \le 0. \end{gathered} \tag{30}$$

has a unique solution $u \equiv 0$.

Now we deal with existence and uniqueness of the Stochastic Volterra equation with non-homogeneous terms. The result extends Theorem 2.14 in [6], where the case $f(t) \equiv 0$ is treated, by considering the case where f is a deterministic function in $L^1(0, T; H)$ and g is as in Hypothesis 2.1.

Proposition 2. *In our assumptions, let* $x_0 \in X_\eta$ *for some* $\frac{1-\alpha(a)}{2} < \eta < \frac{1}{2}\alpha(a)$. *Given the process*

$$\mathbf{x}(t) = e^{tB}x_0 + \int_0^t e^{(t-s)B}(I - B)Pf(s)\,\mathrm{d}s + \int_0^t e^{(t-s)B}(I - B)Pg\,\mathrm{d}W(s) \tag{31}$$

we define the process

$$u(t) = \begin{cases} J\mathbf{x}(t), & t \ge 0, \\ u_0(t), & t \le 0. \end{cases} \tag{32}$$

Then $u(t)$ is a weak solution to problem

$$\begin{aligned}\frac{d}{dt}\int_{-\infty}^{t} a(t-s)u(s)\,\mathrm{d}s = Au(t) + f(t) + g\dot{W}(t), \qquad t\in[0,T]\\ u(t) = u_0(t), \qquad t\le 0.\end{aligned} \tag{33}$$

After the preparatory results stated above, here we prove that main result of existence and uniqueness of solutions of the original controlled Volterra equation (1).

Theorem 5.1. *Assume Hypothesis 2.1 and 2.2. Let γ be an admissible control and $\mathbf{x}$ be the solution to problem (3)) (associated with γ) in X_η with η satisfying the assumptions of Theorem 4.2. Then the process*

$$u(t) = \begin{cases} u_0(t), & t\le 0\\ J\mathbf{x}(t), & t\in[0,T]\end{cases} \tag{34}$$

is the unique solution of the stochastic Volterra equation

$$\begin{aligned}\frac{d}{dt}\int_{-\infty}^{t} a(t-s)u(s)\,\mathrm{d}s = Au(t) + f(t,u(t)) + g\left[r(t,u(t),\gamma(t)) + \dot{W}(t)\right], \quad t\in[0,T]\\ u(t) = u_0(t), \quad t\le 0.\end{aligned} \tag{35}$$

Proof. We propose to fulfil the following steps: first, we prove that the affine equation

$$\begin{aligned}\frac{d}{dt}\int_{-\infty}^{t} a(t-s)u(s)\,\mathrm{d}s = Au(t) + f(t,\tilde{u}(t)) + g\left[r(t,\gamma(t),\tilde{u}(t)) + \dot{W}(t)\right], \quad t\in[0,T]\\ u(t) = u_0(t), \quad t\le 0.\end{aligned} \tag{36}$$

defines a contraction mapping $\mathcal{Q}:\tilde{u}\mapsto u$ on the space $L^p(\Omega;C([0,T];H))$. Therefore, equation (35) admits a unique solution.

Then we show that the process u defined in (34) satisfies equation (35). Accordingly, by the uniqueness of the solution, the thesis of the theorem follows.

First step. We proceed by defining the mapping

$$\mathcal{Q}: L^p(\Omega;C([0,T];H)) \to L^p(\Omega;C([0,T];H))$$

where $\mathcal{Q}(\tilde{u}) = u$ is the solution of the problem

$$\begin{aligned}\frac{d}{dt}\int_{-\infty}^{t} a(t-s)u(s)\,\mathrm{d}s = Au(t) + g(t,\tilde{u}(t))\left[r(t,\tilde{u}(t),\gamma(t)) + \dot{W}(t)\right], \qquad t\in[0,T]\\ u(t) = u_0(t), \qquad t\le 0.\end{aligned} \tag{37}$$

Let $\tilde{u}_1,\tilde{u}_2$ be two processes belonging to $L^p(\Omega;C([0,T];H))$ and take $u_1 = \mathcal{Q}(\tilde{u}_1)$ and $u_2 = \mathcal{Q}(\tilde{u}_2)$. It follows from Proposition 2, that, if $\mathbf{x}_i, i=1,2$ are the processes defined as

$$\mathbf{x}_i(t) = e^{tB}x + \int_0^t e^{(t-s)B}(I-B)Pgr(s,\tilde{u}_i(s),\gamma(s))\,\mathrm{d}s + \int_0^t e^{(t-s)B}(I-B)Pg\,\mathrm{d}W(s),$$

then the processes $w_i(t)$ $(i = 1, 2)$ defined through the formula

$$w_i(t) = \begin{cases} J\mathbf{x}_i(t), & t \in [0, T] \\ u_0(t), & t \le 0 \end{cases}$$

satisfy the stochastic Volterra equations for any $t \in [0, T]$,

$$\frac{d}{dt}\int_{-\infty}^{t} a(t-s)w_i(s)\,\mathrm{d}s = Aw_i(t) + f(t, \tilde{u}_i(t)) + g\left[r(t, \tilde{u}_i(t), \gamma(t)) + \dot{W}(t)\right],$$

with initial condition

$$u(t) = u_0(t), \qquad t \le 0.$$

By the uniqueness of the stochastic homogeneous Volterra equation stated in Proposition 1, we have $w_i = u_i, i = 1, 2$. Now define $U(t) := u_1(t) - u_2(t)$. We have

$$U(t) = \begin{cases} J(\mathbf{x}_1(t) - \mathbf{x}_2(t)), & t \in [0, T] \\ 0, & t \le 0; \end{cases}$$

and

$$\mathbb{E}\sup_{t\in[0,T]} e^{-\beta p t}|U(t)|^p \le \|J\|^p_{L(X_\eta, H)}\mathbb{E}\sup_{t\in[0,T]} e^{-\beta p t}\|\mathbf{x}_1(t) - \mathbf{x}_2(t)\|^p_\eta$$

The quantity on the right hand side can be treated as in Theorem 4.2 and the claim follows.

Second step. It follows from the previous step that there exists at most a unique solution u of problem (36); hence it only remains to prove the representation formula (34) for u.

Let $\tilde{f}(t) = f(t, J\mathbf{x}(t)) + gr(t, J\mathbf{x}(t), \gamma(t))$; it is a consequence of Proposition 2 that u, defined in (34), is a weak solution of the problem

$$\begin{aligned} \frac{d}{dt}\int_{-\infty}^{t} a(t-s)u(s)\,\mathrm{d}s &= Au(t) + \tilde{f}(t) + g\dot{W}(t), \qquad t \in [0, T] \\ u(t) &= u_0(t), \qquad t \le 0, \end{aligned} \tag{38}$$

and the definition of $\tilde{f}$ implies that u is a weak solution on $[0, T]$ of

$$\frac{d}{dt}\int_{-\infty}^{t} a(t-s)u(s)\,\mathrm{d}s = Au(t) + f(t, J\mathbf{x}(t)) + g\left[r(t, J\mathbf{x}(t), \gamma(t)) + \dot{W}(t)\right], \tag{39}$$

with initial condition

$$u(t) = u_0(t), \qquad t \le 0,$$

that is problem (35). □

6. **The forward SDE.** In the following we are concerned with smoothness properties of the forward equation, i.e. of the uncontrolled state equation (29) on the time interval $[s, T]$ with initial condition $x \in X_\eta$. It will be denoted by $\mathbf{x}(t; s, x)$, to stress dependence on the initial data s and x. Also, we extend $\mathbf{x}(\cdot; s, x)$ letting $\mathbf{x}(t; s, x) = x$ for $t \in [0, s]$.

Now we consider the dependence of the process $(\mathbf{x}(t; s, x))_{t\ge 0}$ on the initial data. More precisely, we prove that $(\mathbf{x}(t; s, x))_{t\ge 0}$ depends continuously on s and x and it is also Gâteaux differentiable with respect to x. The following result rely on Proposition 2.4 in Fuhrman and Tessitore [27], where a parameter depending contraction principle is provided.

Proposition 3. *For any $p \geq 1$ the following holds.*

1. *for each $t \in [s,T]$, the map $(s,x) \mapsto \mathbf{x}(t;s,x)$ defined on $[0,T] \times X_\eta$ and with values in $L^p(\Omega, C([0,T];X_\eta))$ is continuous.*
2. *For every $s \in [0,T]$ the map $x \mapsto \mathbf{x}(t;s,x)$ has, at every point $x \in X_\eta$, a Gâteaux derivative $\nabla_x \mathbf{x}(\cdot;s,x)$. The map $(s,x,h) \mapsto \nabla_x \mathbf{x}(\cdot;s,x)[h]$ is a continuous map from $[0,T] \times X_\eta \times X_\eta \to L^p(\Omega, C([0,T];X_\eta))$ and, for every $h \in X_\eta$, the following equation holds $\mathbb{P}$-a.s.:*

$$\nabla_x \mathbf{x}(t;s,x)[h] = e^{(t-s)B}h + \int_s^t e^{(t-\tau)B}(I-B)P\,\nabla_u f(\tau, J\mathbf{x}(\tau;s,x))J\nabla_x \mathbf{x}(\tau;s,x)[h]\mathrm{d}\tau,$$

for any $t \in [s,T]$, whereas $\nabla_x \mathbf{x}(t;s,x)[h] = h$ for $t \in [0,s]$.

Proof. Point 1*: Continuity.* As before, we deal with the mappings $\mathcal{K}$, Λ, Γ defined in (19), (20), (22). We will denote $\mathcal{K}$, Λ, Γ and respectively by $\mathcal{K}(\mathbf{x};s,x)$, $\Lambda(\mathbf{x};s,x)$, $\Gamma(\cdot;s)$ in order to stress the dependence on the initial conditions s and x. Moreover we set $\mathcal{K}(\mathbf{x};s,x) = x$, $\Lambda(\cdot\,;s,x) = 0$ and $\Gamma(\mathbf{x};s) = 0$ for $t < s$ and we recall that $\mathcal{K}(\cdot;s,x)$ is a contraction, with contraction constant independent on s and x, in the space $L^p(\Omega, C([0,T];X_\eta))$ with respect to the norm

$$|\!|\!|\mathbf{x}|\!|\!|_\eta^p := \mathbb{E} \sup_{t\in[0,T]} e^{-\beta t}\|\mathbf{x}(t)\|_\eta^p.$$

By a parameter dependent contraction argument (see, for instance [27, Proposition 2.4]), the claim follows if we show that for all $\mathbf{x} \in L^p(\Omega, C([0,T];X_\eta))$ the map $t \mapsto \mathcal{K}(\mathbf{x};s,x)$ is a continuous map from $[0,T] \times X_\eta$ with values in $L^p(\Omega, C([0,T];X_\eta))$.

To this end, we introduce two sequences $\{s_n^+\}$ and $\{s_n^-\}$ such that $s_n^+ \searrow s$ and $s_n^- \nearrow s$ and we estimate the norm of $\mathcal{K}(\mathbf{x};s_n^+,x) - \mathcal{K}(\mathbf{x};s_n^-,x)$ in the space $L^p(\Omega, C([0,T];X_\eta))$. We have

$$|\!|\!|\mathcal{K}(\mathbf{x};s_n^+,x) - \mathcal{K}(\mathbf{x};s_n^-,x)|\!|\!|_\eta^p \leq \mathbb{E} \sup_{t\in[0,T]} e^{-\beta t}\|e^{(t-s_n^+)B}x - e^{(t-s_n^-)B}x\|_\eta^p$$
$$+\mathbb{E} \sup_{t\in[0,T]} e^{-\beta t}\|\Lambda(\mathbf{x};s_n^+,x)(t) - \Lambda(\mathbf{x};s_n^-,x)(t)\|_\eta^p + \mathbb{E} \sup_{t\in[0,T]} e^{-\beta t}\|\Gamma(t;s_n^+) - \Gamma(t;s_n^-)\|_\eta^p.$$

Now we focus on the third member of the above inequality: introducing a change of variables we obtain

$$\begin{aligned}
&\mathbb{E} \sup_{t\in[0,T]} e^{-\beta t}\|\Gamma(t;s_n^+) - \Gamma(t;s_n^-)\|_\eta^p \\
&\qquad \leq \mathbb{E} \sup_{t\in[s_n^-,T]} e^{-\beta t}\left\|\int_{s_n^-}^{t\wedge s_n^+} e^{(t-\tau)B}(I-B)P\,g\mathrm{d}W(\tau)\right\|_\eta^p \\
&\qquad \leq \mathbb{E} \sup_{t\in[s_n^-,s_n^+]} e^{-\beta t}\left\|\Gamma(t;s_n^-)\right\|_\eta^p \\
&\qquad \leq \mathbb{E} \sup_{t\in[0,s_n^+-s_n^-]} e^{-\beta t}\left\|\Gamma(t;0)\right\|_\eta^p \to 0,
\end{aligned}$$

where the final convergence comes as an immediate consequence of [17, Lemma 7.2] and the dominated theorem, since $\Gamma(\cdot\,;0) \in L^p(\Omega, C([0,T];X_\eta))$. Similarly, taking into account Remark 4,

$$\begin{aligned}
&\mathbb{E}\sup_{t\in[0,T]}\left\|\Lambda(\mathbf{x};s_n^+,x)(t)-\Lambda(\mathbf{x};s_n^-,x)(t)\right\|_\eta^p \\
&\quad\le \mathbb{E}\sup_{t\in[s_n^-,T]}\left\|\int_{s_n^-}^{t\wedge s_n^+} e^{(t-\tau)B}(I-B)Pf(\tau,J\mathbf{x}(\tau))\mathrm{d}\tau\right\|_\eta^p \\
&\quad\le \mathbb{E}\sup_{t\in[s_n^-,s_n^+]}\left|\int_{s_n^-}^{t}\|e^{(t-\tau)B}(I-B)\|_{\mathcal{L}(X_\theta;X_\eta)}\|P\|_{\mathcal{L}(H;X_\theta)}\|f(\tau,J\mathbf{x}(\tau))\|_H\mathrm{d}\tau\right|^p \\
&\quad\le C\,\mathbb{E}\sup_{t\in[0,T]}(1+\|\mathbf{x}(\tau)\|_\eta^p)\sup_{t\in[s_n^-,s_n^+]}\left(\int_{s_n^-}^t (t-\tau)^{\theta-\eta-1}\mathrm{d}\tau\right)^p \\
&\quad\le C(s_n^+-s_n^-)^{p(\theta-\eta)}(1+\|\!|\mathbf{x}|\!\|_\eta^p)\ \to 0.
\end{aligned}$$

Finally, if we extend $e^{(t-s)B}$ to the identity for $t<s$ we have

$$\begin{aligned}
&\mathbb{E}\sup_{t\in[0,T]}e^{-\beta t}\left\|e^{(t-s_n^+)B}x-e^{(t-s_n^-)B}x\right\|_\eta^p \\
&\qquad\qquad\qquad = \mathbb{E}\sup_{t\in[s_n^-,T]}e^{-\beta t}\left\|e^{(t-s_n^+)B}[x-e^{(s_n^+-s_n^-)B}x]\right\|_\eta^p\ \to\ 0
\end{aligned}$$

and also the map $x\mapsto\left\{t\mapsto e^{(t-s)B}x\right\}$ is clearly continuous in x uniformly in s from X_η into the space $C([0,T];X_\eta)$.

Point 2: Differentiability. Again by [27, Proposition 2.4], it is enough to show that the map $(\mathbf{x},s,x)\mapsto\mathcal{K}(\mathbf{x};s,x)$ defined on $L^p(\Omega,C([0,T];X_\eta))\times[0,T]\times X_\eta$ with values in $L^p(\Omega,C([0,T];X_\eta))$ is Gâteaux differentiable in $(\mathbf{x},x)$ and has strongly continuous derivatives.

The directional derivative $\nabla_{\mathbf{x}}\mathcal{K}(\mathbf{x};s,x)$ in the direction $\mathbf{h}\in L^p(\Omega,C([0,H];X_\eta))$ is defined as

$$\lim_{\varepsilon\to0}\frac{\mathcal{K}(\mathbf{x}+\varepsilon\mathbf{h};s,x)-\mathcal{K}(\mathbf{x};s,x)}{\varepsilon}.$$

We claim that the above limit coincides with the process

$$\nabla_{\mathbf{x}}\mathcal{K}(\mathbf{x};s,x)[\mathbf{h}](t)=\int_s^t e^{(t-\tau)B}(I-B)P\nabla_u f(\tau,J\mathbf{x}(\tau))[J\mathbf{h}(\tau)]\mathrm{d}\tau,\quad t\in[s,T],$$

whereas $\nabla_{\mathbf{x}}\mathcal{K}(\mathbf{x};s,x)[\mathbf{h}](t)=0$ when $t<s$. Moreover, the mappings $(\mathbf{x};s,x)\mapsto\nabla_{\mathbf{x}}\mathcal{K}(\mathbf{x};s,x)$ and $\mathbf{h}\mapsto\nabla_{\mathbf{x}}\mathcal{K}(\mathbf{x};s,x)[\mathbf{h}]$ are continuous. In fact, let us define the process

$$\begin{aligned}
I^\varepsilon(t):=&\frac{\mathcal{K}(\mathbf{x}+\varepsilon\mathbf{h};s,x)(t)-\mathcal{K}(\mathbf{x};s,x)(t)}{\varepsilon}\\
&-\int_s^t e^{(t-\tau)B}(I-B)P\nabla_u f(\tau,J\mathbf{x}(\tau))[J\mathbf{h}(\tau)]\mathrm{d}\tau.
\end{aligned}$$

Since we have the identity

$$\frac{\mathcal{K}(\mathbf{x}+\varepsilon\mathbf{h};s,x)(t)-\mathcal{K}(\mathbf{x};s,x)(t)}{\varepsilon}$$

$$= \int_s^t \int_0^1 \frac{1}{\varepsilon} \frac{\mathrm{d}}{\mathrm{d}\xi} e^{(t-\tau)B}(I-B)Pf(\tau\,, J\mathbf{x}(\tau) + \varepsilon\,\xi\, J\mathbf{h}(\tau))\mathrm{d}\xi\mathrm{d}\tau,$$
$$= \int_s^t \int_0^1 e^{(t-\tau)B}(I-B)P\nabla_u f(\tau, J\mathbf{x}(\tau) + \varepsilon\,\xi\, J\mathbf{h}(\tau))[J\mathbf{h}(\tau)]\mathrm{d}\xi\,\mathrm{d}\tau,$$

$I^\varepsilon(t)$ can be rewritten as

$$I^\varepsilon(t) = \int_s^t \mathrm{d}\tau \left(\int_0^1 e^{(t-\tau)B}(I-B)P\nabla_u f(\tau, J\mathbf{x}(\tau) + \varepsilon\,\xi\, J\mathbf{h}(\tau))[J\mathbf{h}(\tau)] \right.$$
$$\left. - e^{(t-\tau)B}(I-B)P\nabla_u f(\tau, J\mathbf{x}(\tau))[J\mathbf{h}(\tau)]\mathrm{d}\xi \right).$$

Moreover by the assumption on the gradient of f, for all $\varepsilon > 0$ we have

$$\mathbb{E} \sup_{t\in[s,T]} \left\| \int_s^t \mathrm{d}\tau \int_0^1 e^{(t-\tau)B}(I-B)P\nabla_u f(\tau, J\mathbf{x}(\tau) + \varepsilon\,\xi\, J\mathbf{h}(\tau))[J\mathbf{h}(\tau)]\mathrm{d}\xi \right\|_\eta^p$$
$$\leq \|\nabla_u f\|^p \|P\|^p_{L(H;X_\theta)} \|\mathbf{h}\|^p_\eta \sup_{t\in[s,T]} \left| \int_s^t \mathrm{d}\tau \|e^{(t-\tau)B}(I-B)\|_{\mathcal{L}(X_\theta;X_\eta)} \right|^p$$
$$\leq \|\nabla_u f\|^p \|P\|^p_{L(H;X_\theta)} \|\mathbf{h}\|^p_\eta (T-s)^{p(\theta-\eta-1)} < \infty$$

and $e^{(t-\tau)B}(I-B)P\,\nabla_u f(\tau, J\mathbf{x})\,J$ is continuous in $\mathbf{x}$. Therefore, by the dominated convergence theorem, we get $\|I^\varepsilon\|^p_\eta \to 0$, as $\varepsilon \to 0$ and the claim follows. Continuity of the mappings $(\mathbf{x}, s, x) \mapsto \nabla_{\mathbf{x}}\mathcal{K}(\mathbf{x}; s, x)[\mathbf{h}]$ and $\mathbf{h} \mapsto \nabla_{\mathbf{x}}\mathcal{K}(\mathbf{x}; s, x)[\mathbf{h}]$ can be proved in a similar way.

Finally, we consider the differentiability of $\mathcal{K}(\mathbf{x}; s, x)$ with respect to x. It is clear that the directional derivative $\nabla_x\mathcal{K}(\mathbf{x}; s, x)[h]$ in the direction $h \in X_\eta$ is the process given by

$$\nabla_x\mathcal{K}(\mathbf{x}; s, x)[h] = e^{(t-s)B}h, \qquad t \in [s, T],$$

whereas $\nabla_x\mathcal{K}(\mathbf{x}; s, x)[h] = h$ for $t \in [0, s]$ and the continuity of $\nabla_x\mathcal{K}(\mathbf{x}; s, x)[h]$ in all variables is immediate. □

In the rest of the section we introduce an auxiliary process that will turn to be useful when dealing with the formulation of the fundamental relation for the value function of our control problem. More precisely, for any $x \in X_\theta$ and $t \in [0, T]$ and h in the space

$$\mathcal{D} := \left\{ h \in X_{1-\theta} \subset X :\ (I-B)^{1-\theta}h \in X_\eta \right\}, \tag{40}$$

we define the process $(\Theta(t; s, x)[h])_{t\in[0,T]}$ as

$$\Theta(t; s, x)[h] := (\nabla_x\mathbf{x}(t; s, x) - e^{(t-s)B})(I-B)^{1-\theta}h, \qquad \text{for } t \in [s, T] \tag{41}$$

and $\Theta(t; s, x)[h] := 0$ whenever $t \in [0, s)$. We notice that $\Theta(\cdot; s, x)$ can be seen as a stochastic process with values into the space of linear mappings on $\mathcal{D}$. In the following we shall prove that it can be continuously extended to the whole space X (that with an abuse of notation we still denote by Θ). To this end, the first step is to verify that the domain $\mathcal{D}$ is dense in X.

Proposition 4. *The space $\mathcal{D}$ defined in (40) is dense in X.*

Proof. We recall that the linear operator B is densely defined (see Theorem (3.6)) and that $D(B)$ is contained in all real interpolation spaces X_ρ, $\rho \in (0,1)$. We notice that if $h \in D(B)$ we have

$$\|(I-B)^{1-\theta} h\|_\eta = \|(I-B)^{1-\theta+\eta} h\|_X < \infty.$$

This implies that $D(B) \subset \mathcal{D}$ so that the claim follows. □

Now we are ready to prove that $\Theta(t;s,x)$ can be extended as a linear operator from X into itself.

Proposition 5. *There exists a process* $\{\Theta(\cdot\,;s,x)[h] : h \in X; x \in X_\eta, s \in [0,T]\}$ *defined on* $\Omega \times [0,T] \to X_\eta$ *such that the following hold:*

1. *The map* $h \mapsto \Theta(\,\cdot\,;s,x)[h]$ *is linear and, on the space* $\mathcal{D} \subset X$*, has the representation given in* (41)*;*
2. *The map* $(s,x,h) \mapsto \Theta(\,\cdot\,;s,x)[h]$ *is continuous from* $[0,T] \times X_\eta \times X$ *into* $L^\infty(\Omega; C([0,T];X_\eta))$*;*
3. *There exists a positive constant* C *such that*

$$|\Theta(\,\cdot\,;s,x)[h]|_{L^\infty(\Omega;C([0,T];X_\eta))} \le C\|h\|, \qquad \text{for all } s \in [0,T], x \in X_\eta, h \in X. \tag{42}$$

Proof. For fixed $s \in [0,T]$, $x \in X_\eta$ and $h \in X$ we consider the integral equation

$$\begin{aligned}\Theta(t;s,x)[h] = &\int_s^t e^{(t-\sigma)B}(I-B)P\nabla_u f(\sigma, J\mathbf{x}(\sigma;s,x))J\Theta(\sigma;s,x)[h]\mathrm{d}\sigma \\ &+ \int_s^t e^{(t-\sigma)B}(I-B)P\nabla_u f(\sigma, J\mathbf{x}(\sigma;s,x))J(I-B)^{1-\theta}e^{(\sigma-s)B}h\mathrm{d}\sigma.\end{aligned} \tag{43}$$

Notice that

$$\begin{aligned}&\int_s^t \left\| e^{(t-\sigma)B}(I-B)P\nabla_u f(\sigma, J\mathbf{x}(\sigma;s,x))J(I-B)^{1-\theta}e^{(\sigma-s)B}h \right\|_\eta \mathrm{d}\sigma \\ &\le \int_s^t \|e^{(t-\sigma)B}(I-B)\|_{L(X_\theta;X)}\|P\|_{L(H;X_\theta)}\|\nabla_u f(\sigma, J\mathbf{x}(\sigma;s,x))\|_{L(H)} \\ &\qquad \|J\|_{\mathcal{L}(X_\eta;H)}\|(I-B)^{1-\theta}e^{(\sigma-s)B}\|_{L(X;X_\eta)}\|h\|\mathrm{d}\sigma \\ &\le C\|h\|_X \int_s^t (t-\sigma)^{\theta-1}(\sigma-s)^{\theta-1-\eta}\mathrm{d}\sigma \\ &\le C\|h\|_X (t-s)^{\theta-\eta+1}\int_0^1 (1-\sigma)^{\theta-1}\sigma^{\theta-1-\eta}\mathrm{d}\sigma \\ &= C\|h\|_X B(\theta-\eta,\theta)\end{aligned}$$

where $B(\alpha,\beta)$ is the beta-distribution of parameter $\alpha,\beta > 0$. By the above estimate we then obtain that equation (43) has $\mathbb{P}$-almost surely a unique mild solution in $C([s,T];X_\eta)$. Moreover, extending $\Theta(t;s,x)[h] = 0$, for $t < s$, we have $\Theta(\cdot;t,x)[h] \in L^\infty(\Omega;C([0,T];X_\eta))$ and $|\Theta(\cdot;s,x)|_{L^\infty(\Omega;C([0,T];X_\eta))} \le C\|h\|$. Concerning the continuity with respect to s,x and h, we can argue as in the proof of Proposition 3. Moreover, linearity is straight - forward.

Finally, we prove the representation formula (41) for $\Theta(\cdot;s,x)[k]$ when $k \in \mathcal{D}$. Setting $h = (I-B)^{1-\theta}k$, the equation satisfied by the Gâteaux derivative of $\mathbf{x}(\cdot;s,x)$

at h is given by

$$(\nabla_x \mathbf{x}(t;s,x) - e^{(t-s)B})(I-B)^{1-\theta}[k] = \int_s^t e^{(t-\tau)B}(I-B)P\,\nabla_u f(\tau, J\mathbf{x}(\tau;s,x))J\nabla_x \mathbf{x}(\tau;s,x)(I-B)^{1-\theta}[k]\mathrm{d}\tau,$$

and by adding and subtracting the term $e^{(\tau-s)B}(I-B)^{1-\theta}[k]$ suitably in the integral of the above equality we obtain

$$\begin{aligned}
&(\nabla_x \mathbf{x}(t;s,x) - e^{(t-s)B})(I-B)^{1-\theta}[k]\\
&= \int_s^t e^{(t-\tau)B}(I-B)P\,\nabla_u f(\tau, J\mathbf{x}(\tau;s,x))Je^{(\tau-s)B}(I-B)^{1-\theta}[k]\mathrm{d}\tau\\
&+ \int_s^t e^{(t-\tau)B}(I-B)P\,\nabla_u f(\tau, J\mathbf{x}(\tau;s,x))J[\nabla_x \mathbf{x}(\tau;s,x) - e^{(\tau-s)B}](I-B)^{1-\theta}[k]\mathrm{d}\tau.
\end{aligned}$$

Comparing the above equality with the equation satisfied by $\Theta(\cdot;s,x)[k]$ and applying Gronwall's lemma we conclude that $\Theta(\cdot;s,x)[k] = (\nabla_x \mathbf{x}(\tau;s,x) - e^{(\tau-s)B})(I-B)^{1-\theta}[k]$, $\mathbb{P}$-a.s. for all $t \in [s,T]$. □

6.1. **Regularity in the sense of Malliavin.** In order to state the main results concerning the Malliavin regularity of the process $\mathbf{x}$ we need to recall some basic definitions from the Malliavin calculus. We refer the reader the book [40] for a detailed exposition. The paper [30] treats the extension to Hilbert space valued random variables and processes. For every $h \in L^2(0,T;\Xi)$ we denote by $W(h)$ the integral $\int_0^T \langle h(t), \mathrm{d}W(t)\rangle_\Xi$. Given a Hilbert space K, let us denote by S_K the set of K-valued random variables F of the form

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

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 $D_\sigma F$, $\sigma \in [0,T]$,

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

with values in $L_2(\Xi,K)$; by ∂_k we denote the partial derivatives with respect to the k-th variable and by $e_j \otimes h_k(\sigma)$ the operator $u \mapsto e_j\langle h_k(\sigma), u\rangle_\Xi$. 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:

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

The adjoint operator of D,

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

is called Skorohod integral. It is known that $\mathrm{dom}(\delta)$ contains $L^2_{\mathcal{F}}(\Omega\times[0,T];L_2(\Xi;K))$ 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 $u(t) \in \mathbb{D}^{1,2}(L_2(\Xi; K))$ for a.e. $t \in [0,T]$ and there exists a measurable version $D_\sigma u(t)$ satisfying

$$\|u\|^2_{\mathbb{L}^{1,2}(L_2(\Xi;K))} = \|u\|^2_{L^2(\Omega\times[0,T];L_2(\Xi;K))} + \mathbb{E}\int_0^T \int_0^T |D_\sigma u(t)|^2_{L_2(\Xi;L_2(\Xi;K))} \mathrm{d}t\mathrm{d}\sigma < \infty.$$

Moreover,$\|\delta(u)\|^2_{L^2(\Omega;K)} \le \|u\|^2_{\mathbb{L}^{1,2}(L_2(\Xi;K))}$.

The definition of $\mathbb{L}^{1,2}(K)$ for an arbitrary Hilbert space K is entirely analogous.

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

Now let us consider again the process $\mathbf{x} = (\mathbf{x}(t;s,x))_{t\in[s,T]}$, denoted simply by $\mathbf{x}(t)$, solution of the forward equation (17), with $s \in [0,T]$, $x \in X_\eta$ fixed. We set as before $\mathbf{x}(t) = x$ for $t \in [0,s)$. We will soon prove that $\mathbf{x} \in \mathbb{L}^{1,2}(X_\eta)$. Then it is clear that the equality $D_\sigma \mathbf{x}(t) = 0$, holds, $\mathbb{P}$-a.s., for a.a. σ, s, t if $t < s$ or $\sigma > t$.

Proposition 6. *Assume Hypothesis 2.1. Let $s \in [0,T]$ and $x \in X_\eta$ be fixed. Then the following properties hold:*

1. $\mathbf{x} \in \mathbb{L}^{1,2}(X_\eta)$;
2. *For a.a. σ and t such that $s \le \sigma \le t \le T$, we have $\mathbb{P}$-a.s.*

$$D_\sigma \mathbf{x}(t) = (I-B)e^{(t-\sigma)B}P\ g + \int_\sigma^t e^{(t-r)B}(I-B)P\nabla_u f(r, J\mathbf{x}(r))JD_\sigma \mathbf{x}(r)\mathrm{d}r. \quad (44)$$

and

$$\|D_\sigma \mathbf{x}(t)\|_{L_2(\Xi,X_\eta)} \le C[(t-\sigma)^{\theta-\eta-1} + 1], \quad (45)$$

from which it follows that for every $t \in [0,T]$ we have

$$D\mathbf{x}(t) \in L^\infty(\Omega; L^2(0,T; L_2(\Xi, X_\eta))).$$

3. *There exists a version of $D\mathbf{x}$ such that for every $\sigma \in [0,T)$, $\{D_\sigma \mathbf{x}(t) :\ t \in (\sigma, T]\}$ is a predictable process in $L_2(\Xi; X_\eta)$ with continuous path, satisfying, for $p \in [2,\infty)$,*

$$\sup_{\sigma\in[0,T]} \mathbb{E}\left(\sup_{t\in(\sigma,T]} (t-\sigma)^{p(\theta-1-\eta)}\|D_\sigma \mathbf{x}(t)\|^p_{L_2(\Xi,X_\eta)}\right) \le C, \quad (46)$$

for some positive constant C depending only on p, L, T and $M := \sup_{t\in[0,T]} |e^{tB}|$. Further, for every $t \in [0,T]$ we have $\mathbf{x}(t) \in \mathbb{D}^{1,2}(X_\eta)$.

4. *For every $q \in [2,\infty)$, the map $t \mapsto D\mathbf{x}(t)$ is continuous from $[0,T]$ (hence uniformly continuous and bounded) with values in $L^p(\Omega; L^2([0,T]; L_2(\Xi; X_\eta))$.*

In order to prove this proposition we need some preparation. We start with the following lemma.

Lemma 6.1. *If $\mathbf{x} \in \mathbb{L}^{1,2}(X_\eta)$ then the random processes*

$$\int_0^t e^{(t-r)B}(I-B)P\ f(r, J\mathbf{x}(r))\mathrm{d}r \quad \textit{and} \quad \int_0^t e^{(t-r)}(I-B)Pg\mathrm{d}W(r)$$

belong to $\mathbb{L}^{1,2}(X_\eta)$ *and for a.a.* σ *and* t *with* $\sigma < t$,

$$D_\sigma \int_0^t e^{(t-r)B}(I-B)P\ f(r, J\mathbf{x}(r))\mathrm{d}r = \int_\sigma^t (I-B)P\nabla_u f(r, J\mathbf{x}(r))JD_\sigma \mathbf{x}(r)\mathrm{d}r \tag{47}$$

while

$$D_\sigma \int_0^t e^{(t-r)B}(I-B)Pg\mathrm{d}W(r) = e^{(t-r)B}(I-B)Pg \tag{48}$$

Proof. We start by proving that the process

$$I(t) := \int_0^t e^{(t-r)B}(I-B)Pf(r, J\mathbf{x}(r))\mathrm{d}r$$

belong to $\mathbb{L}^{1,2}(X_\eta)$ and that equality (47) hold. By definition, we need to prove that for a.e. $t \in [0,T]$, the random variable $I(t)$ belong to $\mathbb{D}^{1,2}(X_\eta)$ with

$$\mathbb{E}\int_0^T \int_0^T \left\| D_\sigma \int_0^t e^{(t-r)B}(I-B)Pf(r, J\mathbf{x}(r))\mathrm{d}r \right\|^2_{L_2(\Xi;H)} \mathrm{d}r\,\mathrm{d}\sigma < \infty. \tag{49}$$

To prove (47) we will use the relation between the Malliavin derivative and the Skorohod integral. In particular, we recall that $\mathbb{D}^{1,2}(X_\eta) = \mathrm{dom}(\delta^\star)$ where $\delta^\star$ is the adjoint of the Skorohod integral. Hence, (47) will be proved once we will have shown that, for any $\mathbf{y} \in L^2(\Omega \times [0,T]; X_\eta)$ the equality

$$\langle \int_0^t e^{(t-r)B}(I-B)Pf(r, J\mathbf{x}(r))\mathrm{d}r, \delta(\mathbf{y})\rangle_{L^2(\Omega;X_\eta)} = \langle \int_0^t e^{(t-r)}(I-B)PD.f(r, J\mathbf{x}(r))\mathrm{d}r, \mathbf{y}\rangle_{L^2(\Omega\times[0,T];X_\eta)}$$

holds. For fixed $\mathbf{y} \in \mathrm{dom}(\delta)$ we consider the scalar product between $I(t)$ and $\delta(\mathbf{y})$ in the space $L^2(\Omega; X_\eta)$; applying Fubini's Theorem (see [17, Theorem 4.18]) we have

$$\begin{aligned}
\langle I(t), \delta(\mathbf{y})\rangle_{L^2(\Omega;X_\eta)} &= \mathbb{E}\langle I(t), \delta(\mathbf{y})\rangle_\eta \\
&= \int_0^t \mathbb{E}\langle e^{(t-r)B}(I-B)Pf(r, J\mathbf{x}(r)), \delta(\mathbf{y})\rangle_\eta\ \mathrm{d}r \\
&= \int_0^t \langle e^{(t-r)B}(I-B)Pf(r, J\mathbf{x}(r)), \delta(\mathbf{y})\rangle_{L^2(\Omega;X_\eta)}\ \mathrm{d}r.
\end{aligned}$$

Now using the duality between the operators δ and D we obtain

$$\begin{aligned}
&\langle I(t), \delta(\mathbf{y})\rangle_{L^2(\Omega;X_\eta)} \\
&= \int_0^t \langle D\left(e^{(t-r)B}(I-B)Pf(r, J\mathbf{x}(r))\right), \mathbf{y}\rangle_{L^2(\Omega\times[0,T];L_2(\Xi,X_\eta))}\mathrm{d}r \\
&= \int_0^t \int_0^t \mathbb{E}\langle D_\sigma\left(e^{(t-r)B}(I-B)Pf(r, J\mathbf{x}(r))\right), \mathbf{y}\rangle_{L_2(\Xi;X_\eta)}\mathrm{d}\sigma\mathrm{d}r \\
&= \mathbb{E}\int_0^t \langle \int_0^t e^{(t-r)B}(I-B)PD_\sigma f(r, J\mathbf{x}(r))\mathrm{d}r, \mathbf{y}\rangle\,\mathrm{d}\sigma \\
&= \langle \int_0^T e^{(t-r)B}(I-B)PD.f(r, J\mathbf{x}(r))\mathrm{d}r, \mathbf{y}\rangle_{L^2(\Omega\times[0,T];X_\eta)}
\end{aligned}$$

Comparing the first and the last term in the above expression we conclude that $I(t) \in \mathbb{D}^{1,2}(X_\eta)$ with

$$\delta^\star I(t) = DI(t) = \int_0^t e^{(t-r)B}(I-B)PD.f(r, J\mathbf{x}(r))\mathrm{d}r.$$

Now we prove estimate (49). First we notice that

$$\begin{aligned}\mathbb{E}\int_0^T \|I(t)\|_\eta^2 \mathrm{d}t &\le \mathbb{E}\int_0^T \int_0^t \|e^{(t-r)B}(I-B)Pf(r, J\mathbf{x}(r))\|_\eta^2 \mathrm{d}r\ \mathrm{d}t \\ &\le\ L_f^2\|P\|^2_{L(H;X_\theta)}\mathbb{E}\int_0^T \int_0^t (t-r)^{2(\theta-1-\eta)}(1+\|J\|_{L(X_\eta;H)}\|\mathbf{x}(r)\|_\eta)^2 \mathrm{d}r\ \mathrm{d}t\end{aligned}$$

where L_f is the Lipschitz constant of f. The right-hand side is finite for a.a. $t \in [0,T]$; in fact, by exchanging the integrals we verify that

$$\begin{aligned}&\int_0^T \left(\mathbb{E}\int_0^t (t-r)^{2(\theta-1-\eta)}(1+\|J\|_{L(X_\eta;H)}\|\mathbf{x}(r)\|_\eta)^2 \mathrm{d}r\right)\mathrm{d}t \\ &\qquad\le \left(\int_0^T t^{-2(\theta-1-\eta)}\mathrm{d}t\right)\left(\int_0^T \mathbb{E}(1+\|J\|_{L(X_\eta;H)}\|\mathbf{x}(r)\|_\eta)^2 \mathrm{d}r\right) < \infty,\end{aligned}$$

since $\mathbf{x} \in \mathbb{L}^{1,2}(X_\eta) \subset L^2(\Omega\times[0,T]; X_\eta)$. Next, for every $t \in [0,T]$, by the chain rule for Malliavin derivative (see Fuhrman and Tessitore [27, Lemma 3.4 (*ii*)]), $D_\sigma[f(r, J\mathbf{x}(r))] = \nabla_u f(r, J\mathbf{x}(r))JD_\sigma\mathbf{x}(r)$ for a.a. $\sigma \le r$, whereas $D_\sigma[f(r, J\mathbf{x}(r))] = 0$ for a.a. $\sigma > r$, by adaptiveness. Next, recalling the assumption on $\nabla_u f$,

$$\begin{aligned}&\mathbb{E}\int_0^T \int_0^T \|D_\sigma I(t)\|^2_{L_2(\Xi,X_\eta)}\mathrm{d}t\ \mathrm{d}\sigma \\ &\le \mathbb{E}\int_0^T \int_0^T \int_\sigma^t \|e^{(t-r)B}(I-B)P\nabla_u f(r, J\mathbf{x}(r))JD_\sigma\mathbf{x}(r)\|^2_{L_2(\Xi;X_\eta)}\mathrm{d}r\ \mathrm{d}t\ \mathrm{d}\sigma \\ &\le C_{\|\nabla f\|,\|J\|,\|P\|}\mathbb{E}\int_0^T \int_0^\sigma \int_\sigma^t (t-r)^{2(\theta-1-\eta)}\|D_\sigma\mathbf{x}(r)\|^2_{L_2(\Xi;X_\eta)}\mathrm{d}r\ \mathrm{d}t\ \mathrm{d}\sigma,\end{aligned}$$

so that, by an easy application of Fubini's theorem,

$$\begin{aligned}&\mathbb{E}\int_0^T \int_0^T \|D_\sigma I(t)\|^2_{L_2(\Xi;X_\eta)}\mathrm{d}t\ \mathrm{d}\sigma \\ &\le C\,\mathbb{E}\int_0^T \int_0^t \int_\sigma^t (t-r)^{2(\theta-1-\eta)}\|D_\sigma\mathbf{x}(r)\|^2_{L_2(\Xi;X_\eta)}\mathrm{d}r\ \mathrm{d}\sigma\ \mathrm{d}t \\ &= C\,\mathbb{E}\int_0^T \int_0^t \int_0^r (t-r)^{2(\theta-1-\eta)}\|D_\sigma\mathbf{x}(r)\|^2_{L_2(\Xi,X_\eta)}\mathrm{d}\sigma\ \mathrm{d}r\ \mathrm{d}t \\ &= C\int_0^T \int_0^t (t-r)^{2(\theta-1-\eta)}\int_0^r \mathbb{E}\|D_\sigma\mathbf{x}(r)\|^2_{L_2(\Xi;X_\eta)}\mathrm{d}\sigma\ \mathrm{d}r\ \mathrm{d}t.\end{aligned}$$

Now the right hand side of the previous inequality is finite; in fact, by exchanging the integrals we verify that

$$\int_0^T \left(\int_0^t (t-r)^{2(\theta-1-\eta)}\int_0^r \mathbb{E}\|D_\sigma\mathbf{x}(r)\|^2_{L_2(\Xi;X_\eta)}\mathrm{d}\sigma\ \mathrm{d}r\right)\mathrm{d}t$$

$$\leq \left(\int_0^T t^{2(\theta-1-\eta)}\mathrm{d}t\right)\left(\int_0^T\int_0^r \mathbb{E}\|D_\sigma \mathbf{x}(r)\|^2_{L_2(\Xi;X_\eta)}\mathrm{d}\sigma\ \mathrm{d}r\right)$$
$$= \left(\int_0^T t^{2(\theta-1-\eta)}\mathrm{d}t\right)\ \|D_\sigma \mathbf{x}\|^2_{L^2(\Omega\times[0,T];L_2(\Xi;X_\eta))} < \infty,$$

since $\mathbf{x}\in\mathbb{L}^{1,2}(X_\eta)$. This proves (49).

Now we consider the Malliavin derivative of the stochastic term

$$\int_0^t e^{(t-r)B}(I-B)Pg\mathrm{d}W(r)$$

and we prove that it belongs to $\mathbb{L}^{1,2}(X_\eta)$ and satisfies (48). This will be consequence of an easy application of the following fact, proved in [27, Proposition 3.4]: if $\mathbf{y}\in\mathbb{L}^{1,2}(X_\eta)$, and for a.a. $\sigma\in[0,T]$ the process $\{D_\sigma\mathbf{y}(t): t\in[0,T]\}$ belongs to $\mathrm{dom}(\delta)$, and the map $\sigma\mapsto\delta(D_\sigma\mathbf{x})$ belongs to $L^2(\Omega\times[0,T];X_\eta)$, then $\delta(\mathbf{y})\in\mathbb{D}^{1,2}(X_\eta)$ and $D_\sigma\delta(\mathbf{y}) = \mathbf{y}(\sigma)+\delta(D_\sigma\mathbf{y})$.

We fix $t\in[0,T]$ and we define $\mathbf{y}(r) := e^{(t-r)B}(I-B)Pg$ for $t\geq r$ and $\mathbf{y}(r) = 0, t<r$. Clearly $\mathbf{y}\in\mathbb{L}^{1,2}(X_\eta)$ and $D_\sigma\mathbf{y} = 0$ for every $\sigma\in[0,T]$. This implies that $(D_\sigma\mathbf{y}(t))_{t\geq 0}\in\mathrm{dom}(\delta)$ for a.a. $\sigma\in[0,T]$ and $\sigma\mapsto\delta(D_\sigma\mathbf{y})$ belongs to $L^2(\Omega\times[0,T];X_\eta)$. On the other hand, we recall that the Skorohod and the Itô integral coincide for adapted integrand, so that

$$\delta(\mathbf{y}) = \int_0^t e^{(t-r)B}(I-B)Pg\mathrm{d}W(r).$$

Applying the result mentioned above we get $\delta(\mathbf{y})\in\mathbb{L}^{1,2}(X_\eta)$ with

$$D_\sigma\delta(\mathbf{y}) = D_\sigma\int_0^t e^{(t-r)B}(I-B)Pg\mathrm{d}W(r)$$
$$= e^{(t-r)B}(I-B)Pg$$

for $a.a.t\in[0,T]$ so that formula (48) is proved. □

Now for $\sigma\in[0,T)$ and for arbitrary predictable processes $(\mathbf{x}(t))_{t\in[\sigma,T]}$ and $(\mathbf{q}(t))_{t\in[\sigma,T]}$ with values respectively in X_η and $L_2(\Xi;X_\eta)$ we define the process $\mathcal{H}=\mathcal{H}(\mathbf{x},\mathbf{q})$ as

$$\mathcal{H}(\mathbf{x},\mathbf{q})_{\sigma t} = e^{(t-\sigma)B}(I-B)Pg + \int_\sigma^t e^{(t-r)B}(I-B)P\nabla_u f(r,J\mathbf{x}(r))J\mathbf{q}(r)\mathrm{d}r.$$

We are now ready to prove Proposition 6.

Proof. [of Proposition 6] We fix $s\in[0,T)$. Let us consider the sequence $\mathbf{x}^n$ defined as follows: $\mathbf{x}^0 = 0\in X_\eta$,

$$\begin{aligned}&\mathbf{x}^{n+1}(t)\\ &= e^{(t-s)B}x + \int_s^t e^{(t-r)B}(I-B)Pf(r,J\mathbf{x}^n(r))\mathrm{d}r + \int_s^t e^{(t-r)B}(I-B)Pg\mathrm{d}W(r)\\ &= \mathcal{K}(\mathbf{x}^n)(t),\end{aligned}$$

and $\mathbf{x}^n(t) = x$ for $t<s$, where the mapping $\mathcal{K}$ was defined in Section 4 (see (19)). It was proved in Theorem 4.2 that $\mathcal{K}$ is a contraction in $L^p_{\mathcal{F}}(\Omega;C([0,T];X_\eta))$, hence, in particular, in the space $L^2(\Omega\times[0,T];X_\eta)$. This implies that $\mathbf{x}^n$ converges to an

element $\mathbf{x}$ in this space. By Lemma 6.1, $\mathbf{x}^{n+1}$ belongs to $\mathbb{L}^{1,2}(X_\eta)$ and, for a.a. σ and t with $\sigma < t$

$$D_\sigma \mathbf{x}^{n+1}(t) = e^{(t-\sigma)B}(I-B)Pg + \int_\sigma^t e^{(t-r)B}(I-B)P\nabla_u f(r, J\mathbf{x}^n(r))JD_\sigma \mathbf{x}^n(r)\mathrm{d}r. \tag{50}$$

Hence, recalling the operator introduced above, we may write equality (50) as

$$D\mathbf{x}^{n+1} = \mathcal{H}(\mathbf{x}^n, D\mathbf{x}^n).$$

In the following we prove that

$$|||\mathcal{H}(\mathbf{x}^n, D\mathbf{x}^n)|||_\beta^2 \le \alpha |||D\mathbf{x}^n|||_\beta^2,$$

for some $\alpha \in [0,1)$ and $\beta > 0$ to be chosen later, where $|||\cdot|||_\beta$ denotes an equivalent norm in $L^2(\Omega \times [0,T] \times [0,T]; L_2(\Xi; X_\eta))$. More precisely, for $\beta > 0$, we introduce the norm

$$|||Y|||_\beta^2 := \int_0^T \int_0^T e^{-\beta(t-\sigma)} \mathbb{E}\|Y_{\sigma t}\|^2_{L_2(\Xi;X_\eta)} \mathrm{d}t\, \mathrm{d}\sigma.$$

First we estimate the term $U := \{U_{\sigma t} : 0 \le \sigma \le t \le T\}$ defined as $U_{\sigma t} := e^{(t-\sigma)B}(I-B)Pg$. We have

$$\begin{aligned}
|||U|||_\beta^2 &= \int_0^T \int_0^T e^{-\beta(t-\sigma)} \|e^{(t-\sigma)B}(I-B)Pg\|^2_{L_2(\Xi;X_\eta)} \mathrm{d}t\ \mathrm{d}\sigma \\
&\le C_{P,g} \int_0^T \int_0^t e^{-\beta(t-\sigma)}(t-\sigma)^{2(\theta-1-\eta)} \mathrm{d}\sigma\ \mathrm{d}t \\
&= C_{P,g} \int_0^T \int_0^t e^{-\beta\sigma}\sigma^{2(\theta-1-\eta)} \mathrm{d}\sigma\ \mathrm{d}t \\
&\le C_{P,g}\, T\, \frac{\Gamma(2(\theta-1-\eta))}{\beta^{2(\theta-1-\eta)+1}} < \infty,
\end{aligned}$$

where $C_{P,g}$ is a positive constant depending only on the norms of P and g. Hence $\left\{e^{(t-\sigma)B}(I-B)Pg :\ 0 \le \sigma < t \le T\right\}$ is bounded in the space $L^2(\Omega\times[0,T]\times[0,T]; L_2(\Xi; X_\eta))$. Now let us consider the norm of $\mathcal{H}(\mathbf{x}^n, D\mathbf{x}^n)$: taking into account the above inequality we get

$$\begin{aligned}
|||\mathcal{H}(\mathbf{x}^n, D\mathbf{x}^n)|||_\beta^2 &= \int_0^T \int_0^T e^{-\beta(t-\sigma)} \mathbb{E}\|\mathcal{H}(\mathbf{x}^n, D\mathbf{x}^n)_{\sigma t}\|^2_{L_2(\Xi;X_\eta)} \mathrm{d}\sigma\ \mathrm{d}t \\
\le C_{P,g}\nu(\beta) + &\int_0^T \int_0^t \int_\sigma^t e^{-\beta(t-\sigma)} \mathbb{E}\left\| e^{(t-r)B}(I-B)P\nabla_u f(r, J\mathbf{x}^b(r))JD_\sigma \mathbf{x}^n(r)\right\|_\eta^2 \mathrm{d}r\, \mathrm{d}\sigma\, \mathrm{d}t
\end{aligned}$$

where we set $\nu(\beta) := T\, \frac{\Gamma(2(\theta-1-\eta))}{\beta^{2(\theta-1-\eta)+1}}$. Now changing the order of integration, we have

$$\begin{aligned}
&|||\mathcal{H}(\mathbf{x}^n, D\mathbf{x}^n)|||_\beta^2 \\
&\le C_{P,g}\nu(\beta) + C_{P,f,J} \int_0^T \int_\sigma^T \left(\int_r^T (t-r)^{-2(\theta-1-\eta)} e^{-\beta(t-r)} \mathrm{d}t \right) \\
&\qquad e^{-\beta(r-\sigma)} \mathbb{E}\|D_\sigma \mathbf{x}^n(r)\|^2_{L_2(\Xi;X_\eta)}\, \mathrm{d}r\, \mathrm{d}\sigma \\
&\le C_{P,g}\nu(\beta) + C_{P,f,J} \int_0^T \int_\sigma^T \left(\sup_{r\in[\sigma,T]} \int_r^T (t-r)^{-2(\theta-1-\eta)} e^{-\beta(t-r)} \mathrm{d}t \right) \\
&\qquad e^{-\beta(r-\sigma)} \mathbb{E}\|D_\sigma \mathbf{x}^n(r)\|^2_{L_2(\Xi;X_\eta)} \mathrm{d}r\, \mathrm{d}\sigma,
\end{aligned}$$

where $C_{P,f,J}$ is a positive constant depending only on the norms of P, f, J. Since the supremum on the right-hand side can be estimated by $\int_0^T t^{2(\theta-1-\eta)}e^{-\beta t}\mathrm{d}t$ we obtain

$$\begin{aligned}&\|\mathcal{H}(\mathbf{x}^n, D\mathbf{x}^n)\|_\beta^2\\ \le & C_{P,g}\nu(\beta) + C_{P,f,J}\nu(\beta)\int_0^T\int_\sigma^T e^{-\beta(r-\sigma)}\mathbb{E}\|D_\sigma\mathbf{x}^n(r)\|^2_{L_2(\Xi;X_\eta)}\mathrm{d}r\ \mathrm{d}\sigma.\end{aligned}$$

Now we choose β large enough such that $\nu(\beta)(C_{P,g} + C_{P,J,f}\|D\mathbf{x}^n\|_\beta^2) \le \alpha\|D\mathbf{x}^n\|_\beta^2$, for $\alpha \in [0,1)$, so that the above inequality means that

$$\|\mathcal{H}(\mathbf{x}^n, D\mathbf{x}^n)\|_\beta^2 \le \alpha\|D\mathbf{x}^n\|_\beta^2, \tag{51}$$

with $\alpha \in [0,1)$. From (51) and from the fact that $e^{(t-\sigma)B}(I-B)Pg$ is bounded in $L^2(\Omega \times [0,T] \times [0,T]; L_2(\Xi; X_\eta))$, it follows that the sequence $\{D\mathbf{x}^n\}_{n\in\mathbb{N}}$ is also bounded in this space. Since, as mentioned before $\mathbf{x}^n$ converges to $\mathbf{x}$ in $L^2(\Omega\times[0,T]; X_\eta)$ it follows from the closedness of the operator D that $\mathbf{x}$ belongs to $\mathbb{L}^{1,2}(X_\eta)$. Thus point 1 of Proposition 6 is proved.

We now consider point 2 in Proposition 6. First, we notice that, by Lemma 6.1, we can compute the Malliavin derivative of both side of equation (17) and we obtain, for a.a. σ and t such that $\sigma < t$ the following equality $\mathbb{P}$-a.s.:

$$\begin{aligned}D_\sigma\mathbf{x}(t) = {}& e^{(t-\sigma)B}(I-B)Pg\\ &+ \int_\sigma^t e^{(t-r)B}(I-B)P\nabla_u f(r, J\mathbf{x}(r))JD_\sigma\mathbf{x}(r)\mathrm{d}r.\end{aligned} \tag{52}$$

Since, for any $\sigma < t$

$$\|e^{(t-\sigma)B}(I-B)Pg\|_{L_2(\Xi;X_\eta)} \le (t-\sigma)^{(\theta-1-\eta)}\|P\|_{L(H;X_\theta)}\|g\|_{L_2(\Xi;X_\eta)}, \tag{53}$$

by the boundedness of $\nabla_u f$ and the Gronwall's lemma it easy to deduce that

$$\|D_\sigma\mathbf{x}(t)\|_{L_2(\Xi,X_\eta)} \le C((t-\sigma)^{\theta-\eta-1}+1).$$

In particular it follows that, for every $t\in[0,T]$, the mapping $\sigma \mapsto D_\sigma\mathbf{x}(t)$ is bounded in $L^\infty(\Omega, L^2(0,T; L_2(\Xi; X_\eta)))$.

We now consider point 2 in Proposition 6, which concerns with the regularity of the trajectories of $D_\sigma\mathbf{x}$, $\sigma\in[0,T)$. We introduce the space $\mathcal{V}$ of processes $(\mathbf{q}_{\sigma t})_{0\le\sigma\le t\le T}$, such that, for every $\sigma\in[s,T)$, $(\mathbf{q}_{\sigma t})_{t\in(\sigma,T]}$ is a predictable process in X_η with continuous paths, and such that

$$\sup_{\sigma\in[0,T]}\mathbb{E}\left(\sup_{t\in(\sigma,T]} e^{-\beta(t-\sigma)p}(t-\sigma)^{p(\theta-1-\eta)}\|\mathbf{q}_{\sigma t}\|^2_{L_2(\Xi;X_\eta)}\right) < \infty.$$

Here $p\in[2,\infty)$ is fixed and $\beta > 0$ is a parameter to be chosen later. Let us consider the equation

$$\mathbf{q}_{\sigma t} = e^{(t-\sigma)B}(I-B)Pg + \int_\sigma^t e^{(t-r)B}(I-B)P\nabla_u f(r, J\mathbf{x}(r))J\mathbf{q}_{\sigma r}\mathrm{d}r, \tag{54}$$

which we can rewrite as

$$\mathbf{q}_{\sigma t} = \mathcal{H}(\mathbf{x},\mathbf{q})_{\sigma t}, \qquad t\in(\sigma,T].$$

We claim that equation (54) admits a unique mild solution in the space $\mathcal{V}$. To prove this, it suffices to show that the term $(U_{\sigma t})_{0\le\sigma\le t\le T}$ defined before belongs to $\mathcal{V}$ and that $\mathcal{H}$ is a contraction in the space $\mathcal{V}$. Since, for any $\sigma < t$

$$\|e^{(t-\sigma)B}(I-B)Pg\|_{L_2(\Xi;X_\eta)} \le (t-\sigma)^{(\theta-1-\eta)}\|P\|_{L(H;X_\theta)}\|g\|_{L_2(\Xi;X_\eta)}$$

we have

$$\sup_{\sigma\in[0,T]}\sup_{t\in(\sigma,T]}(t-\sigma)^{p(1+\eta-\theta)}\|e^{(t-\sigma)B}(I-B)Pg\|^2_{L_2(\Xi;X_\eta)} < \infty.$$

This proves that $U \in \mathcal{V}$. Further, repeating almost identical passages as those leading to (50), we can prove that $\mathcal{H}(\mathbf{x},\cdot)$ is a contraction in the space $\mathcal{V}$ provided that β is chosen sufficiently large. This proves the claim.

Now we prove that $\mathbf{q}$ is the required version of $D\mathbf{x}$. Subtracting (54) from (52), we obtain, $\mathbb{P}$-a.s., for a.a. σ and t with $\sigma < t$

$$D_\sigma\mathbf{x}(t) - \mathbf{q}_{\sigma t} = \int_\sigma^t e^{(t-r)B}(I-B)P\nabla_u f(r, J\mathbf{x}(r))\,[J(D_\sigma\mathbf{x}(r) - \mathbf{q}_{\sigma r})]\,\mathrm{d}r.$$

Repeating the passages that led to (51), we obtain

$$|\!|\!|D_\sigma\mathbf{x} - \mathbf{q}|\!|\!|^2_\beta \le \alpha|\!|\!|D_\sigma\mathbf{x} - \mathbf{q}|\!|\!|^2_\beta$$

for some $\alpha < 1$, which clearly implies that $\mathbf{q}$ is a version of $D\mathbf{x}$.

To conclude the proof of 3 in Proposition 6 we need to prove that $\mathbf{x}(t) \in \mathbb{D}^{1,2}(X_\eta)$ for every $t \in [0,T]$. This assertion is clear for $t \in [0,s]$ since $\mathbf{x}(t) = x$ for $t < s$. For $t \in (s,T]$ we take a sequence $t_n \searrow t$ such that $\mathbf{x}(t_n) \in \mathbb{D}^{1,2}(X_\eta)$ and we note that by (46), the sequence

$$\mathbb{E}\int_0^T \|D_\sigma\mathbf{x}(t_n)\|^2_{L_2(\Xi;X_\eta)}\mathrm{d}\sigma$$

is bounded by a constant independent of n. Since $\mathbf{x}(t_n) \to \mathbf{x}(t)$ in $L^2(\Omega;X_\eta)$, it follows from the closedness of the operator D that $\mathbf{x} \in \mathbb{D}^{1,2}(X_\eta)$.

Now we prove point 4 in Proposition 6. To prove that $t \mapsto D\mathbf{x}(t)$ is continuous as a mapping from $[0,T]$ into $L^p(\Omega; L^2(0,T;X_\eta))$ we fix $t \in [0,T]$ and let $t_n^+ \searrow t$ and $t_n^- \nearrow t$. Then, since $D\mathbf{x}$ satisfies equation (54), we have

$$\begin{aligned}\mathbb{E}\int_0^T \|D_\sigma\mathbf{x}(t_n^+) - D_\sigma\mathbf{x}(t_n^-)\|^2_\eta\mathrm{d}\sigma &\le 2\int_0^T \left\|(e^{(t_n^+-\sigma)B} - e^{(t_n^--\sigma)B})(I-B)Pg\right\|^2_\eta\mathrm{d}\sigma \\ &+ 2\mathbb{E}\int_0^T\left\|\int_\sigma^{t_n^+} e^{(t_n^+-r)B}(I-B)P\nabla_u f(r,J\mathbf{x}(r))JD_\sigma\mathbf{x}(r)\mathrm{d}r\right. \\ &\left.- \int_\sigma^{t_n^-} e^{(t_n^--r)B}(I-B)P\nabla_u f(r,J\mathbf{x}(r))JD_\sigma\mathbf{x}(r)\mathrm{d}r\right\|^2_\eta \qquad (55)\end{aligned}$$

We estimate the two terms in the right member of the above inequality separately. Concerning the first one, which is deterministic we have

$$2\int_0^T \left\|(e^{(t_n^+-\sigma)B} - e^{(t_n^--\sigma)B})(I-B)Pg\right\|^2_\eta\mathrm{d}\sigma$$

$$= \int_{t_n^-}^{t_n^+} \left\| e^{(t_n^+-\sigma)B}(I-B)Pg \right\|_\eta^2 \mathrm{d}\sigma + \int_0^{t_n^-} \left\| (e^{(t_n^+-t_n^-)B} - I)e^{(t_n^- -\sigma)B}(I-B)Pg \right\|_\eta^2 \mathrm{d}\sigma$$

$$\leq C_{P,g} \int_{t_n^-}^{t_n^+} (t_n^+ - \sigma)^{2(\theta-\eta-1)} \mathrm{d}\sigma + C_{P,g} \|e^{(t_n^+-t_n^-)B} - I\|^2 \int_0^{t_n^-} (t_n^- - \sigma)^{2(\theta-1-\eta)} \mathrm{d}\sigma$$

We notice that both integrals in the last part of the above expression go to 0 as $n \to \infty$.

In a similar way, using the bound (46) on $D_\sigma \mathbf{x}$ for $p = \infty$, we conclude that also the second integral in (55) goes to 0 and the required continuity follows. □

Now we denote by $\{e_j\}_{j\in\mathbb{N}}$ a basis on the space Ξ and consider the standard real Wiener process $W^j(\tau) = \int_0^\tau \langle e_j, W(s)\rangle \mathrm{d}s$. We conclude the section by investigating the existence of the joint quadratic variation of $W^j, i \in \mathbb{N}$ with a process of the form $\{w(t, \mathbf{x}(t)) : \ t \in [0, T]\}$ for a given function $w : [0, T] \times X_\eta \to \mathbb{R}$, on an interval $[0, s] \subset [0, T)$. As usual this is defined as the limit in probability of

$$\sum_{i=1}^{n} (w(t_i, \mathbf{x}(t_i)) - w(t_{i-1}, \mathbf{x}(t_{i-1})))(W^j(t_i) - W^j(t_{i-1})),$$

where $\{t_i : 0 = t_0 < t_1 < \cdots < t_n, \ i = 0, \ldots, n\}$ is an arbitrary subdivision of $[0, t]$ whose mesh tends to 0. We do not require that convergence takes place uniformly in time. This definition is easily adapted to arbitrary interval of the form $[s, t] \subset [0, T)$. Existence of the joint quadratic variation is not trivial. Indeed, due to the occurrence of convolution type integrals in the definition of mild solution, it is not obvious that the process $\mathbf{x}$ is a semimartingale. Moreover, even in this case, the process $w(\cdot, \mathbf{x})$ might fail to be a semimartingale if w is not twice differentiable, since Itô formula does not apply. Nevertheless, the following result hold true. Its proof could be deduced from generalization of some result obtained in [18, pg. 193] to the infinite-dimensional case, but we prefer to give a simpler direct proof.

Proposition 7. *Suppose that $w \in C([0, T) \times X_\eta; \mathbb{R})$ is Gâteaux differentiable with respect to $\mathbf{x}$, and that for every $s < T$ there exist constants K and m (possibly depending on s) such that*

$$|w(t, x)| \leq K(1 + |x|)^m, \quad |\nabla w(t, x)| \leq K(1 + |x|)^m, \quad t \in [0, s], x \in X. \tag{56}$$

Let η and θ satisfy condition (13) in Theorem 3.6. Assume that for every $t \in [0, T)$, $x \in X_\eta$ the linear operator $k \mapsto \nabla w(t, x)(I - B)^{1-\theta}k$ (a priori defined for $k \in \mathcal{D}$) has an extension to a bounded linear operator $X \to \mathbb{R}$, that we denote by $[\nabla w(I - B)^{1-\theta}](t, x)$. Moreover, assume that the map $(t, x, k) \mapsto [\nabla w(I - B)^{1-\theta}](t, x)k$ is continuous from $[0, T) \times X_\eta \times X$ into $\mathbb{R}$. For $t \in [0, T), x \in X_\eta$, let $\{\mathbf{x}(t; s, x), t \in [s, T]\}$ be the solution of equation (29). Then the process $\{w(t, \mathbf{x}(t; s, x)), t \in [s, T]\}$ admits a joint quadratic variation process with W^j, for every $j \in \mathbb{N}$, on every interval $[s, t] \subset [s, T)$, given by

$$\int_s^t [\nabla w(I - B)^{1-\theta}](r, \mathbf{x}(r; s, x))(I - B)^\theta Pge_j \mathrm{d}r.$$

Proof. For simplicity we write the proof for the case $s = 0$ and we write $\mathbf{x}(t) = \mathbf{x}(t; s, x)$, $w(t) = w(t, \mathbf{x}(t))$. It follows from the assumptions that the mapping $(t, x, h) \mapsto \nabla w(t, x)h$ is continuous on $[0, T) \times X_\eta \times X$. By the chain rule for the Malliavin derivative operator (see Fuhrman and Tessitore [27]), it follows that for every $t < T$ we have $w(t) \in \mathbb{D}^{1,2}(X_\eta)$ and $Dw(t, \mathbf{x}(t)) = \nabla w(t, \mathbf{x}(t))D\mathbf{x}(t)$.

Let us now compute the joint quadratic variation of w and W^j on a fixed interval $[0,t] \subset [0,T)$. Let $0 = t_0 < t_1 \cdots < t_n = t$ be a subdivision of $[0,t] \subset [0,T]$ with mesh $\delta = \max_i(t_i - t_{i-1})$. By well-known rules of Malliavin calculus (see, e.g. [5]) we have

$$(w(t_i) - w(t_{i-1}))(W^j(t_i) - W^j(t_{i-1}) = (w(t_i) - w(t_{i-1}) \int_{t_{i-1}}^{t_i} \mathrm{d}W^j(t)$$
$$= \int_{t_{i-1}}^{t_i} D_\sigma(w(t_i) - w(t_{i-1}))\mathrm{d}\sigma + \int_{t_{i-1}}^{t_i} (w(t_i) - w(t_{i-1}))\hat{\mathrm{d}}W^j(\sigma),$$

where we use the symbol $\hat{\mathrm{d}}W^j$ to denote the Skorohod integral. We note that $D_\sigma w(t_{i-1}) = 0$ for $\sigma > t_{i-1}$. Therefore setting

$$U_\delta(\sigma) = \sum_{i=1}^n (w(t_i) - w(t_{i-1}))\mathbf{1}_{(t_{i-1},t_i]}(\sigma)$$

we obtain

$$\sum_{i=1}^n (w(t_i) - w(t_{i-1}))(W^j(t_i) - W^j(t_{i-1}))$$
$$= \int_0^t U_\delta(\sigma)\hat{\mathrm{d}}W^j(\sigma) + \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \nabla w(t_i, \mathbf{x}(t_i))D_\sigma \mathbf{x}(t_i)e_j \mathrm{d}\sigma.$$

Recalling the equation satisfied by $D\mathbf{x}$ (see (52)) we have

$$\begin{aligned}
&\sum_{i=1}^n (w(t_i) - w(t_{i-1}))(W^j(t_i) - W^j(t_{i-1})) \\
&\quad = \int_0^t U_\delta(\sigma)\hat{\mathrm{d}}W^j(\sigma) \\
&\quad + \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \nabla w(t_i, \mathbf{x}(t_i))e^{(t_i-\sigma)B}(I-B)Pge_j \mathrm{d}\sigma \\
&\quad + \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \nabla w(t_i, \mathbf{x}(t_i)) \int_\sigma^{t_i} e^{(r-\sigma)B}(I-B)P\nabla_u f(r, J\mathbf{x}(r))JD_\sigma \mathbf{x}(r)e_j \mathrm{d}r \; \mathrm{d}\sigma.
\end{aligned} \tag{57}$$

Now we let the mesh δ go to 0. We discuss the three terms in the right-member of the above inequality separately.

Using the continuity properties of the maps $t \mapsto \mathbf{x}(t)$ and $t \mapsto D\mathbf{x}(t)$ stated in Proposition 6 and taking into account the continuity properties of w and ∇w, the estimate (56) and the chain rule $D_\sigma w(t) = \nabla w(t, \mathbf{x}(t))D_\sigma \mathbf{x}(t)$, it is easy to see that the map $t \mapsto w(t) = w(t, \mathbf{x}(t))$ is continuous from $[0,T]$ to $\mathbb{D}^{1,2}(X_\eta)$. It follows that $U_\delta \to 0$ in $\mathbb{L}^{1,2}(X_\eta)$ and by the continuity of the Skorohod integral we conclude that the first term in the right-member of (57) goes to 0 as $\delta \to 0$.

According to the definition of $\nabla w(t, \mathbf{x}(t))$ the second term can be written as

$$\sum_{i=1}^n [\nabla w(I-B)^{1-\theta}](t_i, \mathbf{x}(t_i)) \int_{t_{i-1}}^{t_i} e^{(t_i-\sigma)B}(I-B)^\theta Pge_j \mathrm{d}\sigma$$
$$\leq \sum_{i=1}^n [\nabla w(I-B)^{1-\theta}](t_i, \mathbf{x}(t_i))(I-B)^\theta Pge_j(t_i - t_{i-1})$$

$$+ \sum_{i=1}^{n} [\nabla w(I - B)^{1-\theta}](t_i, \mathbf{x}(t_i)) \int_{t_{i-1}}^{t_i} (e^{(t_i-\sigma)B} - I)(I - B)^{\theta} P g e_j \mathrm{d}\sigma.$$

We notice that, for every $i = 0, \dots, n$,

$$\sup_{\sigma\in[t_{i-1},t_i]} \left|(e^{(t_i-\sigma)B} - I)(I - B)^{\theta} P g e_j\right| \le \sup_{\sigma\in[0,\delta]} \left|(e^{\sigma B} - I)(I - B)^{\theta} P g e_j\right| \to 0,$$

as $\delta \to 0$, by the strong continuity of the semigroup. From the properties of $[\nabla w(I-B)^{1-\theta}]$ and the continuity of the paths of $\mathbf{x}$ it follows that

$$\sum_{i=1}^{n} [\nabla w(I - B)^{1-\theta}](t_i, \mathbf{x}(t_i)) \int_{t_{i-1}}^{t_i} e^{(t_i-\sigma)B}(I - B)^{\theta} P g e_j \mathrm{d}\sigma$$
$$\to \int_0^t [\nabla w(I - B)^{1-\theta}](r, \mathbf{x}(r))(I - B)^{\theta} P g e_j \mathrm{d}r, \quad \mathbb{P} - a.s.$$

Now recalling that, by Proposition 6 - Point 2, $D_\sigma \mathbf{x} \in L^\infty(\Omega; L^2(0, T; X_\eta))$ and the boundedness of ∇f, we can estimate the third term in the right-member of (57) by

$$\left|\sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \nabla w(t_i, \mathbf{x}(t_i)) \int_{\sigma}^{t_i} e^{(r-\sigma)B}(I - B) P \nabla_u f(r, J\mathbf{x}(r)) J D_\sigma \mathbf{x}(r) e_j \mathrm{d}r \ \mathrm{d}\sigma\right|$$
$$\le C_{P,f} \sum_{i=1}^{n} \|\nabla w(t_i, \mathbf{x}(t_i))\|_{L(X_\eta)} \int_{t_{i-1}}^{t_i} \int_{\sigma}^{t_i} ((r - \sigma)^{\theta-1-\eta} + 1) \mathrm{d}r \ \mathrm{d}\sigma$$
$$\le C_{P,f,\theta,\eta} \sum_{i=1}^{n} \|\nabla w(t_i, \mathbf{x}(t_i))\|_{L(X_\eta)} [(t_i - t_{i-1})\delta^{\theta-\eta} + (t_i - t_{i-1})\delta]$$

where C is a positive constant depending only on P and f. But the last term goes to 0, $\mathbb{P}$-a.s., by the continuity properties of ∇w and the continuity of the paths of $\mathbf{x}$. □

7. **The backward stochastic differential equation.** In this section we consider the backward stochastic differential equation in the unknown (Y, Z):

$$\begin{cases} \mathrm{d}Y(\tau) = \psi(\tau, \mathbf{x}(\tau; t, x), Z(\tau))\mathrm{d}\tau + Z(\tau)\mathrm{d}W(\tau), & \tau \in [s, T], \\ Y(T) = \phi(\mathbf{x}(T; s, x)) \end{cases} \tag{58}$$

where $\mathbf{x}(\cdot; s, x)$ is the solution of the uncontrolled equation (29) (with the convention $\mathbf{x}(\tau; s, x) = x$ for $\tau \in [0, s)$) and ψ is the Hamiltonian function relative to the control problem described in Section 1. More precisely, for $t \in [0, T], x \in X_\eta, z \in \Xi^*$ we have

$$\psi(t, x, z) = \inf \{l(t, Jx, \gamma) + z \ r(t, Jx, \gamma) : \ \gamma \in \mathcal{U}\}. \tag{59}$$

For further use, we prove some additional properties of the function ψ:

Proposition 8. *Under Hypothesis 2.1 and 2.3 the following hold:*

1. *There exists a positive constant C such that for any $t \in [0, T], x_1, x_2 \in X_\eta$ and $z_1, z_2 \in \Xi^*$,*
$$|\psi(t, x_1, z_1) - \psi(t, x_2, z_2)| \le C|z_1 - z_2| + C(1 + \|x_1\|_\eta + \|x_2\|_\eta)\|x_1 - x_2\|_\eta;$$
2. *There exists a positive constant C such that*
$$\sup_{t\in[0,T]} |\psi(t, 0, 0)| \le C.$$

Proof. Point 2 follows from the fact that

$$|\psi(t,0,0)| = \left|\inf_{\gamma\in\mathcal{U}} l(t,0,\gamma)\right|.$$

Moreover, for all $\gamma \in \mathcal{U}$ we have

$$\begin{aligned}
&l(t,Jx,\gamma) + zr(t,Jx,\gamma)\\
&\le l(t,Jx',\gamma) + z'r(t,Jx',\gamma) + |l(t,Jx,\gamma) - l(t,Jx',\gamma)|\\
&\quad + |zr(t,Jx,\gamma) - z'r(t,Jx',\gamma)|\\
&\le l(t,Jx',\gamma) + z'r(t,Jx',\gamma) + |l(t,Jx,\gamma) - l(t,Jx',\gamma)|\\
&\quad + |(z-z')r(t,Jx,\gamma)| + |z'(r(t,Jx,\gamma) - r(t,Jx',\gamma)|\\
&\le l(t,Jx',\gamma) + z'r(t,Jx',\gamma) + C\|J\|\|x-x'\|_\eta(1+\|x\|_\eta+\|x'\|_\eta)\\
&\quad + C|z-z'| + C\|J\|\|x-x'\|_\eta|z'|,
\end{aligned}$$

and taking the infimum over γ,

$$\begin{aligned}
\psi(t,x,z) &\le \psi(t,x',z') + C\|J\|\|x-x'\|_\eta(1+\|x\|_\eta+\|x'\|_\eta) + C|z-z'|\\
&\quad + C\|J\|\|x-x'\|_\eta|z'|\\
&\le c|z-z'| + c\|x-x'\|_\eta(1+|z|+|z'|)(1+\|x\|_\eta+\|x'\|_\eta)
\end{aligned}$$

for some $c, C > 0$. Exchanging x, z with x', z' we get the conclusion. □

We make the following assumption.

Hypothesis 7.1. For almost every $t \in [0,T]$ the map $\psi(t,\cdot,\cdot)$ is Gâteaux differentiable on $X_\eta \times \Xi^*$ and the maps $(x,h,z) \mapsto \nabla_x\psi(t,x,z)[h]$ and $(x,z,\zeta) \mapsto \nabla_z\psi(t,x,z)[\zeta]$ are continuous on $X_\eta \times X \times \Xi^*$ and $X_\eta \times \Xi^* \times \Xi^*$ respectively.

Remark 7. Under the above assumption we immediately deduce the following estimates:

$$|\nabla_x\psi(t,x,z)[h]| \le C(1+\|x\|_\eta)\|h\| \quad \text{and} \quad |\nabla_z\psi(t,x,z)[\zeta]| \le C|\zeta|, \quad h \in X, \zeta \in \Xi^*.$$

We notice that Hypothesis 7.1 involves conditions on the function ψ and not on the functions l and r that determine ψ. However, Hypothesis 7.1 can be verified in concrete situations (for an example, see, for instance, [27, Ex. 2.7.1]).

The backward equation (58) is understood in the usual way: we look for a pair of processes (Y,Z), progressively measurable, which, for any $t \in [s,T]$

$$Y(t) + \int_t^T Z(\tau)\mathrm{d}W(\tau) = \phi(\mathbf{x}(T;s,x)) - \int_t^T \psi(\tau,\mathbf{x}(\tau;t,x),Z(\tau)\mathrm{d}\tau.$$

We can now state the following result on existence, uniqueness and smoothness of the solution to equation (58):

Proposition 9. *1. For all $x \in X_\eta$, $s \in [0,T]$ and $p \in [2,\infty)$ there exists a unique pair of processes (Y,Z) with $Y \in L^p(\Omega, C([s,T];\mathbb{R}))$, $Z \in L^p(\Omega, L^2(s,T;\Xi^*))$ solving the equation (58). In the following we will denote such a solution by $Y(\cdot;s,x)$ and $Z(\cdot;s,x)$.*

2. The map $(s,x) \mapsto (Y(\cdot;s,x), Z(\cdot;s,x))$ is continuous from $[0,T] \times X_\eta$ to $L^p(\Omega, C([s,T];\mathbb{R})) \times L^p(\Omega; L^2(s,T;\Xi^*))$.

3. *For all $s \in [0,T]$ the map $x \mapsto (Y(\cdot;s,x), Z(\cdot;s,x))$ is Gâteaux differentiable as a map from X_η to $L^p(\Omega, C([s,T];\mathbb{R})) \times L^p(\Omega; L^2(s,T;\Xi^*))$. Moreover, the map $(t,x,h) \mapsto (\nabla_x Y(\,\cdot\,;t,x)[h], \nabla_x Z(\,\cdot\,;t,x)[h])$ is continuous from $[0,T] \times X_\eta \times X$ to $L^p(\Omega, C([s,T];\mathbb{R})) \times L^p(\Omega; L^2(s,T;\Xi^*))$.*
4. *For any $x \in X_\eta, h \in X$, the pair of processes $(\nabla_x Y(\cdot;s,x)[h], \nabla_x Z(\cdot;s,x)[h])$ satisfies the equation*

$$\begin{cases} \mathrm{d}\nabla_x Y(\tau;s,x)[h] = \nabla_x \psi(\tau, \mathbf{x}(\tau;s,x), Z(\tau;s,x))\nabla_x \mathbf{x}(\tau;s,x)[h]\mathrm{d}\tau \\ \qquad +\nabla_z \psi(\tau, \mathbf{x}(\tau;s,x), Z(\tau;s,x))\nabla_x Z(\tau;s,x)[h]\mathrm{d}\tau + \nabla_x Z(\tau;s,x)[h]\mathrm{d}W(\tau), \\ \nabla_x Y(T;s,x)[h] = \nabla_x \phi(\mathbf{x}(T;s,x))\nabla_x \mathbf{x}(T;s,x)[h], \end{cases}$$

for $\tau \in [s,T]$, $s \in [0,T]$.

Proof. The claim follows directly from Proposition 4.8 in Fuhrman and Tessitore [27], from the differentiability stated in Proposition 3 and the chain rule (as stated in Lemma 2.1 in [27]). □

As in the previous section, starting from the Gâteaux derivatives of Y and Z, we introduce suitable auxiliary processes which will allow ourselves to express Z in terms of ∇Y and $(I-B)^{1-\theta}$ and then get the fundamental relation for the optimal control problem introduced in Section 1.

Proposition 10. *For every $p \geq 2$, $s \in [0,T]$, $x \in X_\eta, h \in X$ there exist two processes*

$$\{\Pi(t;s,x)[h] : \ t \in [0,T]\} \quad \textit{and} \quad \{Q(t;s,x)[h] : \ t \in [0,T]\}$$

with $\Pi(\cdot;s,x)[h] \in L^p(\Omega; C([0,T];\mathbb{R}))$ and $Q(\cdot;s,x)[h] \in L^p(\Omega; L^2([0,T];\Xi^*))$ such that if $s \in [0,T)$, $x \in X_\eta$ and $h \in \mathcal{D}$, where

$$\mathcal{D} := \left\{h \in X_{1-\theta} \subset X : \ (I-B)^{1-\theta}h \in X_\eta\right\},$$

then $\mathbb{P}$-a.s. the following identifications hold:

$$\Pi(t;s,x)[h] = \begin{cases} \nabla_x Y(t;s,x)(I-B)^{1-\theta}[h], & \textit{for all } t \in [s,T]; \\ \nabla_x Y(s;s,x)(I-B)^{1-\theta}[h], & \textit{for all } t \in [0,s); \end{cases} \tag{60}$$

$$Q(t;s,x)[h] = \begin{cases} \nabla_x Z(t;s,x)(I-B)^{1-\theta}[h], & \textit{for a.e.} t \in [s,T] \\ 0, & \textit{for all } t \in [0,s). \end{cases} \tag{61}$$

Moreover the map $(s,x,h) \mapsto \Pi(\cdot;s,x)[h]$ is continuous from $[0,T] \times X_\eta \times X$ to $L^p(\Omega; C([0,T];\mathbb{R}))$ and the map $(s,x,h) \mapsto Q(\cdot;s,x)[h]$ is continuous from $[0,T] \times X_\eta \times X$ into $L^p(\Omega; L^2([0,T];\Xi^*))$ and both maps are linear with respect to h. Finally, there exists a positive constant C such that

$$\mathbb{E} \sup_{t\in[s,T]} \|\Pi(t;s,x)[h]\|_\eta^p + \mathbb{E}\left(\int_s^T |Q(t;s,x)[h]|^2_{\Xi^*}\mathrm{d}t\right)^{p/2} \\ \leq C(T-s)^{(\theta-1-\eta)p}(1+\|x\|_\eta)^p\|h\|^p. \tag{62}$$

Proof. As in the proof of Proposition 5, we introduce a suitable stochastic differential equation which should give the pair $(\Pi(\cdot;s,x)[h], Q(\cdot;s,x)[h])$; more precisely, for fixed $s \in [0,T], x \in X_\eta$ and $h \in X$ we consider the following backward stochastic differential equation on the unknown $(\Pi(\cdot;s,x)[h],Q(\cdot;s,x)[h])$:

$$\begin{cases} \mathrm{d}\Pi(t;s,x)[h] = \nu(t;s,x)h\mathrm{d}t + \nabla_z\psi(t,\mathbf{x}(t;s,x),Z(t;s,x))Q(t;s,x)[h]\mathrm{d}s \\ \qquad\qquad + Q(t;s,x)[h]\mathrm{d}W(s) \qquad\qquad t\in[s,T] \\ \Pi(T;s,x)[h] = \eta(s,x)h, \end{cases} \tag{63}$$

where

$$\begin{aligned} &\nu(t;s,x)h \\ &\quad= \mathbf{1}_{[s,T]}(t)\nabla_x\psi(t,\mathbf{x}(t;s,x),Z(t;s,x))\left(\Theta(t;s,x)[h] + (I-B)^{1-\theta}e^{(t-s)B}h\right), \\ \eta(s,x)h &= \nabla_x\phi(\mathbf{x}(T;s,x))\left(\Theta(T;t,x)[h] + (I-B)^{1-\theta}e^{(T-s)B}h\right). \end{aligned}$$

We notice that the processes $\nu(\cdot;s,x)h$ and $\eta(\cdot;s,x)h$ are well-defined for every $h\in X$. Moreover, recalling the estimates on $\mathbf{x}(\cdot;s,x)$ obtained in Theorem 4.2, the properties of $\nabla_x\psi(t,x,z)$ (see Remark 7) and the bound (42) on $\Theta(\cdot;s,x)h$ proved in Proposition 5, we get

$$\begin{aligned} &\mathbb{E}\left(\int_0^T |\nu(t;s,x)h|^2\mathrm{d}t\right)^{p/2} \\ &\le C_1\ \mathbb{E}\left(\int_s^T (1+\|\mathbf{x}(t;s,x)\|_\eta)^2\left(\|\Theta(t;s,x)h\|_\eta + (t-s)^{\theta-1-\eta}\|h\|\right)^2\mathrm{d}t\right)^{p/2} \\ &\le C_2(1+\|x\|_\eta)^p\ \left[(T-s)^{p/2} + (T-s)^{(2(\theta-\eta)-1)p/2}\right]\|h\|^p \\ &\le C_3(1+\|x\|_\eta)^p\|h\|^p. \end{aligned}$$

where $C_i,\ i=1,2,3$ are suitable constants independent on t,x and h. In the same way, again by (18), (42) and by Hypothesis 2.3 we get

$$\begin{aligned} \mathbb{E}|\eta(s,x)h|^p &\le C\mathbb{E}\left[(1+\|\mathbf{x}(T;s,x)\|)(\|\Theta(T;s,x)h\|_\eta + (T-s)^{\theta-1-\eta})\|h\|\right]^p \\ &\le C(1+\|x\|_\eta)^p(T-s)^{(\theta-1-\eta)p}\|h\|^p. \end{aligned}$$

Hence, taking into account Proposition 4.3 in Fuhrman and Tessitore [27], we obtain that for every $s\in[0,T], x\in X_\eta, h\in X$ there exists a unique pair of processes $(\Pi(\cdot;s,x)[h],Q(\cdot;s,x)[h])$ solving the BSDE (63) such that $\Pi\in L^p(\Omega;C([0,T];\mathbb{R}))$ and $Q\in L^p(\Omega;L^2(0,T;\Xi^*))$. Moreover, inequality (62) holds.

Reasoning as in the proof of Proposition 5, when $h\in\mathcal{D}$ we get the representation given in (60) and (61). Clearly, the map $h\mapsto(\Pi(\cdot;s,x)[h],Q(\cdot;s,x)[h])$ is linear.

Now we prove its continuity. By estimate (62), it is sufficient to prove that it is continuous with respect to s and x for any fixed $h\in X$, and, again by Proposition 4.3 in Fuhrman and Tessitore [27], it is enough to prove that for all $h\in X, s_n\to s\in[0,T)$, $x_n\to x\in X_\eta$ we have

$$\begin{aligned} I_1^n + I_2^n + I_3^n := \mathbb{E}&\left(\int_0^T \|\nabla_z\psi(t,\mathbf{x}^n(t),Z^n(t)) - \nabla_z\psi(t,\mathbf{x}(t),Z(t))\|^2_{L(\Xi)}\,\mathrm{d}t\right)^{p/2} \\ &+\mathbb{E}\left(\int_0^T |\nu(t;s_n,x_n)h - \nu(t;s,x)h|_X^2\mathrm{d}t\right)^{p/2} \\ &+\mathbb{E}\ |\eta(s_n,x_n)h - \eta(s,x)h|^p \ \longrightarrow\ 0, \quad n\to\infty, \end{aligned}$$

where we set $\mathbf{x}^n(t) = \mathbf{x}(t; s_n, x_n)$ and $Z^n(t) = Z(t; s_n, x_n)$. We prove that the three integrals goes to 0 separately. To this end we notice that, by the continuous dependence of $\mathbf{x}$ and Z on the initial data, we have $\mathbf{x}^n \to \mathbf{x}$ and $Z^n \to Z$, respectively in $L^p(\Omega; C([0,T]; X_\eta))$ and $L^p(\Omega; L^2(0,T;\Xi^*))$. Thus from each subsequence in $\mathbb{N}$ we can extract a subsequence $\{n_i : i \in \mathbb{N}\}$ for which $\sum_i \|\mathbf{x}^{n_i} - \mathbf{x}\|_{L^p(\Omega;C([0,T];X_\eta)} < \infty$ and the series $\sum_i (\mathbf{x}^{n_i} - \mathbf{x})$ converges $\mathbb{P}$-a.s. and for all $t \in [0,T]$ to an element $\hat{\mathbf{x}}$ in $L^p(\Omega; C([0,T]; X_\eta))$, for which $\|\mathbf{x}^{n_i}(t)\|_\eta \le \|\mathbf{x}(t)\|_\eta + \|\hat{\mathbf{x}}(t)\|_\eta$, $\mathbb{P}$-a.s. for all $t \in [0,T]$. Similarly, there exists $\hat{Z}$ in $L^p(\Omega; L^2(0,T;\Xi^*))$ such that $\|Z^{n_i}(t)\|_{\Xi^*} \le \|Z(t)\|_{\Xi^*} + \|\hat{Z}(t)\|_{\Xi^*}$, $\mathbb{P}$-a.s. for all $t \in [0,T]$.

Now we use the above claim to prove that $I_i^n \to 0$, as $n \to \infty, i = 1,2,3$. Let us consider I_1^n. By the continuity assumptions stated in Hypothesis 7.1 and the convergence of $(\mathbf{x}^{n_i}, Z^{n_i})$ to $(\mathbf{x}, Z)$, $\mathbb{P}$-a.s. for all $t \in [0,T]$ we have that $\nabla_z \psi(t, \mathbf{x}^{n_i}(t), Z^{n_i}(t))$ converges to $\nabla_z \psi(t, \mathbf{x}(t), Z(t))$, $\mathbb{P}$-a.s for all $t \in [0,T]$. Further, by Remark 7, the function in the integral I_1 is bounded by a constant independent on $n \in \mathbb{N}$. Hence, by the dominated convergence theorem we have

$$\mathbb{E}\left(\int_0^T \|\nabla_z\psi(t,\mathbf{x}^{n_i}(t), Z^{n_i}(t)) - \nabla_z\psi(t,\mathbf{x}(t),Z(t))\|^2_{L(\Xi)}\,\mathrm{d}t\right)^{p/2} \longrightarrow 0.$$

To prove that $I_2^n \to 0$ as $n \to \infty$ we define

$$V^n(t) := \mathbf{1}_{[s_n,T]}(t)\left(\Theta(t;s_n,x_n)[h] + (I-B)^{1-\theta}e^{(t-s_n)B}h\right)$$

and

$$V(t) := \mathbf{1}_{[s,T]}(t)\left(\Theta(t;s,x)[h] + (I-B)^{1-\theta}e^{(t-s)B}h\right),$$

and we notice that

$$\nu(t;s_n,x_n) = \nabla_x\psi(t,\mathbf{x}^n(t),Z^n(t))V^n(t)$$

and

$$\nu(t;s,x) = \nabla_x\psi(t,\mathbf{x}(t),Z(t))V(t).$$

Then

$$\begin{aligned}
I_2^n \le\ & C_1\mathbb{E}\left(\int_0^T \left|\left(\nabla_x\psi(t;\mathbf{x}^n(t),Z^n(t)) - \nabla_x\psi(t;\mathbf{x}(t),Z(t))\right)V(t)\right|^2\mathrm{d}t\right)^{p/2}\\
&+ C_2\mathbb{E}\left(\int_0^T \left|\nabla_x\psi(\mathbf{x}^n(t),Z^n(t))(V^n(t)-V(t))\right|^2\mathrm{d}t\right)^{p/2}\\
=& I_{21}^n + I_{22}^n.
\end{aligned}$$

Taking into account the continuity assumption and the bound on $\nabla_x\psi(t,x,z)$ (see respectively Hypothesis 7.1 and Remark 7), we have

$$\begin{aligned}
&\mathbb{E}\sup_{t\in[0,T]}\left|\left(\nabla_x\psi(t;\mathbf{x}^n(t),Z^n(t)) - \nabla_x\psi(t;\mathbf{x}(t),Z(t))\right)\right|^2\\
&\quad\le\ C\ \mathbb{E}\sup_{t\in[0,T]}(1+\|\mathbf{x}^n(t)\|_\eta + \|\mathbf{x}(t)\|_\eta)^2\\
&\quad\le\ C(1+\|x_n\|_\eta + \|x\|_\eta)^2
\end{aligned}$$

Further $|V(t)|^2 \le c\mathbf{1}_{[s,T]}(t)(1+(t-s)^{2(\theta-\eta-1)})$ and $2(\theta-\eta-1) > -1$ (since $\theta - \eta > 1/2$. Hence, reasoning as it was done for I_1 we can apply the dominated convergence theorem and obtain that $I^n_{21} \to 0$, as $n \to \infty$. To show that $I^n_{22} \to 0$, as $n \to \infty$, we apply Hölder inequality to $\mathbb{E} \int_0^T |\nabla_x \psi(\mathbf{x}^n(t), Z^n(t))(V^n(t) - V(t))|^2 \,\mathrm{d}t$ to get

$$
\begin{aligned}
&\mathbb{E}\left(\int_0^T |\nabla_x \psi(\mathbf{x}^n(t), Z^n(t))(V^n(t) - V(t))|^2 \,\mathrm{d}t\right)^{p/2}\\
&\le \mathbb{E}\left(\left(\int_0^T \|\nabla_x \psi(t, \mathbf{x}^n(t), Z^n(t))\|^{2r} \mathrm{d}t\right)^{1/r} \left(\int_0^T \|V^n(t) - V(t)\|_\eta^{2q} \mathrm{d}t\right)^{1/q}\right)^{p/2}\\
&\le C\, \mathbb{E}\left(\left(\int_0^T (1 + \|\mathbf{x}^n(t)\|_\eta)^{2r} \,\mathrm{d}t\right)^{p/2r} \left(\int_0^T \|V^n(t) - V(t)\|_\eta^{2q} \mathrm{d}t\right)^{p/2q}\right)\\
&\le C\, \mathbb{E}\left(\left(1 + \sup_{t\in[0,T]} \|\mathbf{x}^n(t)\|_\eta^{2p}\right) \left(\int_0^T \|V^n(t) - V(t)\|_\eta^{2q} \mathrm{d}t\right)^{p/2q}\right)\\
&\le C\left(1 + \mathbb{E} \sup_{t\in[0,T]} \|\mathbf{x}^n(t)\|_\eta^{2p}\right)^{\frac{1}{2}} \left(\mathbb{E}\left(\int_0^T \|V^n(t) - V(t)\|_\eta^{2q} \mathrm{d}t\right)^{p/q}\right)^{\frac{1}{2}}
\end{aligned}
$$

for every r, q such that $1/r + 1/q = 1$. By the above inequality we see that $I^n_{22} \to 0,\ n \to \infty$ if $V^n \to V$ in $L^{2p}(\Omega; L^{2q}(0, T; X_\eta))$. To prove this limit we first note that

$$
\mathbb{E}\left(\int_0^T \left|\mathbf{1}_{[s_n,T]}(t)\Theta(t; s_n, x_n)[h] - \mathbf{1}_{[s,T]}(t)\Theta(t; s, x)[h]\right|^{2q}\right)^{p/q} \longrightarrow 0,
$$

since the map $(s, x) \mapsto \Theta(\cdot; s, x)$ is continuous with values in $L^p(\Omega; C([0, T]; X_\eta))$ for any $p \ge 2$. Hence it remains to show that

$$
\mathbf{1}_{[s_n,T]}(\cdot)(I - B)^{1-\theta} e^{(\cdot - s_n)B} h \longrightarrow \mathbf{1}_{[s,T]}(\cdot)(I - B)^{1-\theta} e^{(\cdot - s)B} h
$$

in $L^{2q}([0, T]; X_\eta)$. To this end we note that for all $s_n^+ \searrow s$ and $s_n^- \nearrow s$ we have

$$
\begin{aligned}
&\int_0^T \left\|(I - B)^{1-\theta}\left(\mathbf{1}_{[s_n^-,T]} e^{(t-s_n^-)B} - \mathbf{1}_{[s_n^+,T]} e^{(t-s_n^+)B}\right) h\right\|_\eta^{2q} \mathrm{d}t\\
&= \int_{s_n^-}^{s_n^+} \left\|(I - B)^{1-\theta} e^{(t-s_n^-)B} h\right\|_\eta^{2q} \mathrm{d}t + \int_{s_n^+}^T \left\|(I - B)^{1-\theta} e^{(t-s_n^+)B}\left(e^{(s_n^+-s_n^-)B} - I\right) h\right\|_\eta^{2q} \mathrm{d}t\\
&\le \|h\|^{2q} \int_{s_n^-}^{s_n^+} (t - s_n^-)^{2q(\theta-1-\eta)} \mathrm{d}t + \left\|\left(e^{(s_n^+-s_n^-)B} - I\right) h\right\|^{2q} \int_{s_n^+}^T (t - s_n^+)^{2q(\theta-1-\eta)} \mathrm{d}t \longrightarrow 0,
\end{aligned}
$$

provided $q > 1$, since $\theta - \eta > 1/2$.

In a similar way as it has been done for I^n_2, we can prove that $I^n_3 \to 0,\ n \to \infty$. □

We are now in the position to give a meaning to the expression $\nabla_x Y(t; s, x)(I - B)^{1-\theta}$ and, successively, to identify it with the process $Z(t; s, x)$.

Corollary 1. *Setting $v(s, x) = Y(s; s, x)$, we have $v \in C([0, T] \times X_\eta; \mathbb{R})$ and there exists a constant C such that $|v(s, x)| \le C(1 + \|x\|_\eta)^2$, $t \in [0, T]$, $x \in X_\eta$. Moreover v is Gâteaux differentiable with respect to x on $[0, T] \times X_\eta$ and the map $(s; x, h) \mapsto \nabla v(s, x)[h]$ is continuous.*

Moreover, for $s \in [0,T)$ and $x \in X_\eta$ the linear operator $h \mapsto \nabla v(s,x)(I-B)^{1-\theta}h$ - a priori defined for $h \in \mathcal{D}$ (with $\mathcal{D}$ as in 40) - has an extension to a bounded linear operator from X into $\mathbb{R}$, that we denote by $[\nabla v(I-B)^{1-\theta}](s,x)$.

Finally, the map $(s,x,h) \mapsto [\nabla v(I-B)^{1-\theta}](s,x)$ is continuous as a mapping from $[0,T) \times X_\eta \times X$ into $\mathbb{R}$ and there exists $C > 0$ such that

$$\|[\nabla v(I-B)^{1-\theta}(s,x)]h\| \le C(T-s)^{(\theta-\eta-1)}(1+\|x\|_\eta)\|h\|, \qquad t \in [0,T),\ x \in X_\eta,\ h \in X. \tag{64}$$

Proof. We recall that $Y(s;s,x)$ is deterministic.

Since the map $(s,x) \mapsto Y(\cdot;s,x)$ is continuous with values in $L^p(\Omega; C([0,T];\mathbb{R}))$, $p \ge 2$, then the map $(s,x) \mapsto Y(s;s,x)$ is continuous with values in $L^p(\Omega;\mathbb{R})$ and so the map $(s,x) \mapsto \mathbb{E}Y(s;s,x) = Y(s;s,x) = v(s,x)$ is continuous with values in $\mathbb{R}$.

Similarly $\nabla_x v(s,x) = \mathbb{E}\nabla_x Y(s;s,x)$ exists and has the required properties, by Proposition 9. Next we notice that $\Pi(s;s,x)h = \nabla_x Y(s;s,x)(I-B)^{1-\theta}h$. The existence of the required extensions and its continuity are direct consequence of Proposition 10 and estimate (64) follows directly from (62). □

Corollary 2. *For every $t \in [0,T]$, $x \in X_\eta$ we have, $\mathbb{P}$-a.s.*

$$Y(t;s,x) = v(t;\mathbf{x}(t;s,x)), \qquad \text{for all } t \in [s,T], \tag{65}$$

$$Z(t;s,x) = [\nabla v(I-B)^{1-\theta}](t;\mathbf{x}(t;s,x))(I-B)^{\theta}Pg, \qquad \text{for almost all } t \in [s,T]. \tag{66}$$

Proof. We start from the well-known equality: for $0 \le s \le r \le T$, $\mathbb{P}$-a.s.

$$\mathbf{x}(t;s,x) = \mathbf{x}(t;r,\mathbf{x}(r;s,x)), \qquad \text{for all } t \in [s,T].$$

It follows easily from the uniqueness of the backward equation (58) that $\mathbb{P}$-a.s.

$$Y(t;s,x) = Y(t;r,\mathbf{x}(r;s,x)), \qquad \text{for all } t \in [s,T].$$

Setting $r = t$ we arrive at (65).

To prove (66) we consider the joint quadratic variation of $Y(\cdot;s,x)$ and $W(\cdot)$ on an arbitrary interval $[s,t] \subset [s,T)$; from the backward equation (58) we deduce that this is equal to $\int_s^t Z(r;s,x)\mathrm{d}r$. On the other side, the same result can be obtained by considering the joint quadratic variation of $(v(t,\mathbf{x}(t;s,x)))_{t\in[s,T]}$ and W. Now by an application of Proposition 7 (whose assumptions hold true by Corollary 1) leads to the identity

$$\int_s^t Z(r;s,x)\mathrm{d}r = \int_s^t [\nabla v(I-B)^{1-\theta}](r,\mathbf{x}(t;s,x))(I-B)^{\theta}Pg\mathrm{d}r,$$

and (66) is proved. □

8. **The Hamilton-Jacobi-Bellman equation.** Let us consider again the solution $\mathbf{x}(t;s,x)$ of equation (17) and denote by $P_{s,t}$ its transition semigroup:

$$P_{s,t}[h](x) = \mathbb{E}h(\mathbf{x}(t;s,x)), \qquad x \in X_\eta, 0 \le s \le t \le T,$$

for any bounded measurable $h : X_\eta \to \mathbb{R}$. We notice that by the bound (18) this formula is meaningful for every h with polynomial growth. In the following $P_{s,t}$ will be considered as an operator acting on this class of functions.

Let us denote by $\mathcal{L}_t$ the generator of $P_{s,t}$:

$$\mathcal{L}_t[h](x) = \frac{1}{2}\text{Tr}[(I-B)Pg\nabla^2 h(x)g^*p^*(I-B)^*] \\ + \langle Bx + (I-B)Pf(t,Jx), \nabla h(x)\rangle,$$

where ∇h and $\nabla^2 h$ are first and second Gâteaux derivatives of h at the point $x \in X_\eta$ (here we are identified with elements of X and $L(X)$ respectively). This definition is formal, since it involves the terms $(I-B)Pg$ and $(I-B)Pf$ which - *a* priori - are not defined as elements of $L(X)$ and the domain of $\mathcal{L}_t$ is not specified.

The Hamilton-Jacobi-Bellman (HJB) equation for the optimal control is

$$\begin{cases} \frac{\partial}{\partial s}v(s,x) + \mathcal{L}_s[v(s,\cdot)](x) = \psi(s,x,\nabla v(s,x)(I-B)Pg), & s\in[0,T],\ x\in X_\eta, \\ v(T,x) = \phi(x). \end{cases} \tag{HJB}$$

This is a nonlinear parabolic equation for the unknown function $v : [0,T]\times X_\eta \to \mathbb{R}$. We define the notion of solution of the (HJB) by means of the variation of constant formula:

Definition 8.1. We say that a function $v : [0,T]\times X_\eta \to \mathbb{R}$ is a mild solution of the Hamilton - Jacobi - Bellman equation (HJB) if the following conditions hold:

1. $v \in C([0,T]\times X_\eta;\mathbb{R})$ and there exist constants $C, m \geq 0$ such that $|v(s,x)| \leq C(1+\|x\|_\eta)^m,\ s\in[0,T],\ x\in X_\eta$;
2. v is Gâteaux differentiable with respect to x on $[0,T)\times X_\eta$ and the map $(s,x,h)\mapsto \nabla v(s,x)[h]$ is continuous $[0,T)\times X_\eta\times X_\eta \to \mathbb{R}$;
3. For all $s\in[0,T)$ and $x\in X_\eta$ the linear operator $k\mapsto \nabla v(s,x)(I-B)^{1-\theta}k$ (a priori defined for $k\in\mathcal{D}$) has an extension to a bounded linear operator on $X\to\mathbb{R}$, that we denote by $[\nabla v(I-B)^{1-\theta}](s,x)$.

   Moreover the map $(s,x,k)\mapsto [\nabla v(I-B)^{1-\theta}](s,x)k$ is continuous $[0,T)\times X_\eta\times X\to\mathbb{R}$ and there exist constants $C,m\geq 0,\ \kappa\in[0,1)$ such that

$$\|[\nabla v(I-B)^{1-\theta}](s,x)\|_{L(X)} \leq C(T-s)^{-\kappa}(1+\|x\|_\eta)^m, \qquad s\in[0,T), x\in X_\eta.$$

4. The following equality holds for every $s\in[0,T]$, $x\in X_\eta$:

$$v(s,x) = P_{s,T}[\phi](x) - \int_s^T P_{s,r}\left[\psi(r,\cdot,[\nabla v(I-B)^{1-\theta}](r,\cdot)(I-B)^\theta Pg\right](x)\mathrm{d}r. \tag{67}$$

Remark 8. We notice that Proposition 8 implies that $|\psi(t,x,z)| \leq C(1+|z|+\|x\|_\eta^2)$, so that if v is a function satisfying the bound required in 3 of the above definition we have

$$\left|\psi(t,x,[\nabla v(I-B)^{1-\theta}](t,x)(I-B)^\theta Pg)\right| \leq C(T-t)^{-\kappa}(1+\|x\|_\eta)^{m+2}$$

and formula (67) is meaningful.

Now we are ready to prove that the solution of the equation (HJB) can be defined by means of the solution of the BSDE associated with the control problem (17).

Theorem 8.2. *Assume Hypothesis 2.1, 2.2 and 2.3, then there exists a unique mild solution of the equation (HJB). The solution is given by the formula*

$$v(s,x) = Y(s;s,x), \tag{68}$$

where $(\mathbf{x}, Y, Z)$ *is the solution of the forward-backward system* (17) *and* (58)*.*

Proof. We start by proving existence. By Corollary 1 the function v defined as in (68) has the regularity properties stated in Definition 8.1. In order to verify that equality (67) holds we first fix $s \in [0,T]$ and $x \in X_\eta$. We notice that

$$\psi(s, \cdot, [\nabla v(I-B)^{1-\theta}](s,\cdot)(I-B)^\theta Pg)(x) \\ = \psi(s, \cdot, [\nabla Y(I-B)^{1-\theta}](s,\cdot)(I-B)^\theta Pg)(x)$$

and we recall that

$$[\nabla v(I-B)^{1-\theta}](t; \mathbf{x}(t;s,x))(I-B)^\theta Pg = Z(t;s,x).$$

Hence

$$P_{s,t}\left[\psi(t, \cdot, [\nabla v(I-B)^{1-\theta}](t,\cdot)(I-B)^\theta Pg)\right](x) \\ = \mathbb{E}\left[\psi(t, \mathbf{x}(t;s,x), Z(t;s,x))\right]. \quad (69)$$

On the other hand, the backward equation gives

$$Y(s;s,x) + \int_s^T Z(r;s,x)\mathrm{d}W(r) = \phi(\mathbf{x}(T;s,x)) - \int_s^T \psi(r, \mathbf{x}(r;s,x), Z(r;s,x))\mathrm{d}r.$$

Taking the expectation we obtain

$$v(s,x) = P_{s,T}[\phi](x) - \mathbb{E}\int_s^T \psi(r, \mathbf{x}(r;s,x), Z(r;s,x))\mathrm{d}r$$

and substituting in the integral the expression obtained in (69) we get the required equality (67).

Now we consider uniqueness of the solution. Let v denote a mild solution. We look for a convenient expression for the process $v(t, \mathbf{x}(t;s,x))$, $t \in [s,T]$. By (67) and the Markov property of $\mathbf{x}$ we have

$$\begin{aligned}
& v(t, \mathbf{x}(t;s,x)) \\
= \; & P_{t,T}[\phi](\mathbf{x}(t;s,x)) \\
& - \int_t^T P_{t,r}\left[\psi(r, \cdot, [\nabla v(I-B)^{1-\theta}](r,\cdot)(I-B)^\theta Pg\right](\mathbf{x}(t;s,x))\mathrm{d}r \\
= \; & \mathbb{E}^{\mathcal{F}_t}\left[\phi(\mathbf{x}(T;s,x))\right] \\
& - \mathbb{E}^{\mathcal{F}_t}\left[\int_t^T \psi(r, \mathbf{x}(r;s,x), [\nabla v(I-B)^{1-\theta}](r, \mathbf{x}(r;s,x)))(I-B)^\theta Pg\mathrm{d}r\right] \\
= \; & \mathbb{E}^{\mathcal{F}_t}[\xi] \\
& + \int_s^t \psi(r, \mathbf{x}(r;s,x), [\nabla v(I-B)^{1-\theta}](r, \mathbf{x}(r;s,x)))(I-B)^\theta Pg\mathrm{d}r,
\end{aligned}$$

where we have defined

$$\xi := \phi(\mathbf{x}(T;s,x)) \\ - \int_s^T \psi(r, \mathbf{x}(r;s,x), [\nabla v(I-B)^{1-\theta}](r, \mathbf{x}(r;s,x))(I-B)^\theta Pg)\mathrm{d}r.$$

Now we notice that $\mathbb{E}^{\mathcal{F}_s}\xi = \mathbb{E}^{\mathcal{F}_s}[v(t, \mathbf{x}(t;s,x))] = v(s,x)$. Since $\xi \in L^2(\Omega;\mathbb{R})$ is $\mathcal{F}_T$-measurable, by a well-known representation theorem there exists $\tilde{Z} \in L^2_{\mathcal{F}}(\Omega \times [s,T]; L_2(\Xi;\mathbb{R}))$ such that $\mathbb{E}^{\mathcal{F}_t}[\xi] = \int_s^t \tilde{Z}(r)\mathrm{d}W(r) + v(s,x)$ We conclude that the process $v(t, \mathbf{x}(t; s, x))$, $t \in [s, T]$ is a real continuous semimartingale with canonical decomposition

$$v(t, \mathbf{x}(t; s, x)) = \int_s^t \tilde{Z}(r) \mathrm{d}W(r) + v(s, x) \\ + \int_s^t \psi(r, \mathbf{x}(r; s, x), [\nabla v(I - B)^{1-\theta}](r, \mathbf{x}(r; s, x)))(I - B)^{\theta} P g \mathrm{d}r, \quad (70)$$

into its continuous martingale part and continuous finite variation part.

Now we compute the joint quadratic variation process of both sides of the above equality with W on an arbitrary interval $[0, t] \subset [0, T)$. By the assumption made in Definition 8.1 - 3 we have that there exists a constant K_t such that $\|\nabla v(s, x)\|_{L(X)} \leq K_t(1 + \|x\|_\eta)^m$, for $s \in [0, t]$, $x \in X_\eta$; then we can apply Proposition 7 to conclude that the joint quadratic variation equals

$$\int_0^t [\nabla v(I - B)^{1-\theta}](r, \mathbf{x}(r; s, x))(I - B)^{\theta} P g \mathrm{d}r.$$

Computing the joint quadratic variation of the left-hand side of (70) with W yields the identity

$$\int_0^t [\nabla v(I - B)^{1-\theta}](r, \mathbf{x}(r; s, x))(I - B)^{\theta} P g \mathrm{d}r = \int_s^t \tilde{Z}(r) \mathrm{d}r.$$

Therefore, for a.a. $t \in [s, T]$, we have $\mathbb{P}$-a.s. $[\nabla v(I - B)^{1-\theta}](r, \mathbf{x}(r; s, x))(I - B)^{\theta} P g = \tilde{Z}(r)$, so substituting into (70) and taking into account that $v(T, \mathbf{x}(T; s, x)) = \phi(\mathbf{x}(T; s, x))$ we obtain, for $t \in [s, T]$,

$$v(t, \mathbf{x}(t; s, x)) + \int_t^T [\nabla v(I - B)^{1-\theta}](r, \mathbf{x}(r; s, x))(I - B)^{\theta} P g \mathrm{d}W(r) \\ = \phi(\mathbf{x}(T; s, x)) - \int_t^T \psi(r, \mathbf{x}(r; s, x), [\nabla v(I - B)^{1-\theta}](r, \mathbf{x}(r; s, x))(I - B)^{\theta} P g \mathrm{d}r.$$

Comparing with the backward equation (58) we notice that the pairs

$$(Y(t; s, x), Z(t; s, x)) \text{ and } (v(t, \mathbf{x}(t; s, x)), [\nabla v(I - B)^{1-\theta}](t, \mathbf{x}(t; s, x))(I - B)^{\theta} P g)$$

solve the same equation. By uniqueness, we have $Y(t; s, x) = v(t; \mathbf{x}(t; s, x)), t \in [s, T]$, and setting $t = s$ we obtain $Y(s; s, x) = v(s, x)$. □

9. **Synthesis of the optimal control.** In this section we proceed with the study of the optimal control problem associated to the stochastic Volterra equation (1).

We reformulate the optimal control problem in the weak sense, following the approach of [24]. For fixed $x \in X_\eta$ an admissible control system (*a.c.s.*) is given by $\mathbb{U} = (\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_t)_{t \geq 0}, \hat{\mathbb{P}}, (\hat{W}(t))_{t \geq 0}, \hat{\gamma})$ where

1. $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ is a complete probability space and $(\hat{\mathcal{F}}_t)_{t \geq 0}$ is a filtration on it satisfying the usual conditions;
2. $\hat{W}(t))_{t \geq 0}$ is a cylindrical $\hat{\mathbb{P}}$-Wiener process with values in Ξ and adapted to the filtration $(\hat{\mathcal{F}}_t)_{t \geq 0}$;
3. $\hat{\gamma}(t) \in \mathcal{U}$, $\hat{\mathbb{P}}$-a.s. for a.a. $t \in [0, T]$ where $\mathcal{U}$ is a fixed subset of U.

To each $(a.c.s.)$ $\mathbb{U}$ we associate the weak solution $u^{\mathbb{U}}$ of the Volterra equation

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{t} a(t-s)u^{\mathbb{U}}(s)\mathrm{d}s = Au^{\mathbb{U}}(t) + f(t,u(t)) \\ \qquad\qquad +g\,[\,r(t,u^{\mathbb{U}}(t),\hat{\gamma}(t)) + \dot{\hat{W}}(t)\,], \qquad t\in[0,T] \\ u^{\mathbb{U}}(t) = u_0(t), \qquad t\le 0. \end{cases}$$

We have showed in Section 3 that we can associate to such equation the controlled state equation

$$\begin{cases} \mathrm{d}\mathbf{x}^{\mathbb{U}}(t) = B\mathbf{x}^{\mathbb{U}}(t)\mathrm{d}t + (I-B)P\,f(t,J\mathbf{x}^{\mathbb{U}}(t)) \\ \qquad\qquad +(I-B)Pg(r(t,J\mathbf{x}^{\mathbb{U}}(t),\hat{\gamma}(t))\mathrm{d}t + \mathrm{d}\hat{W}(t) \\ \mathbf{x}^{\mathbb{U}}(0) = x. \end{cases}$$

Our purpose is to minimize a cost functional of the form

$$\mathbb{J}(u_0,\mathbb{U}) = \mathbb{E}\int_0^T l(r,u^{\mathbb{U}}(r),\hat{\gamma}(t))\mathrm{d}t + \mathbb{E}\phi(u^{\mathbb{U}}(T)).$$

To this end we will consider its translation in the state space setting, where the cost functional is of the form

$$\mathbb{J}(x,\mathbb{U}) = \mathbb{E}\int_0^T l(t,J\mathbf{x}^{\mathbb{U}}(t),\hat{\gamma}(t))\mathrm{d}t + \mathbb{E}\phi(J\mathbf{x}^{\mathbb{U}}(T)).$$

We will work under the assumptions given in Hypothesis 2.1, 2.2 and 2.3. We recall that the Hamiltonian corresponding to our control problem is given by

$$\psi(t,x,z) := \inf_{\gamma\in\mathcal{U}}\{l(t,Jx,\gamma) + zr(t,Jx,\gamma) :\ \gamma\in U\}$$

and we introduce the set of minimizers of (59):

$$\Gamma(t,x,z) := \{\gamma\in\mathcal{U} :\ l(t,Jx,\gamma) + zr(t,Jx,\gamma) = \psi(t,x,z)\}\,.$$

The existence of a minimizer for the Hamiltonian is not a direct consequence of our setting. Then we require it explicitly.

Hypothesis 9.1. $\Gamma(t,x,z)$ is not empty for all $t\in[0,T]$, $x\in X_\eta$ and $z\in\Xi^\star$ and there exists a measurable selection of Γ, i.e a Borel measurable function $\mu : [0,T]\times X_\eta\times\Xi\to U$ such that

$$\psi(t,x,z) = l(t,Jx,\mu(t,x,z)) + zr(t,Jx,\mu(t,x,z)) \qquad t\in[0,T],\ x\in X_\eta, z\in\Xi^\star. \tag{71}$$

Further, by v we will denote the solution of the Hamilton-Jacobi-Bellman equation relative to the above stated problem

$$\begin{cases} \frac{\partial}{\partial t}v(t,x) + \mathcal{L}_t[v(t,\cdot)](x) = \psi(t,x,\nabla v(t,x)(I-B)Pg), \quad t\in[0,T],\ x\in X_\eta, \\ v(T,x) = \phi(x). \end{cases} \tag{HJB}$$

As it was shown in the previous section, (HJB) admits a unique mild solution.

We wish to perform the standard synthesis of the optimal control problem, which consists in proving that the solution of the (HJB) equation is the value function of the control problem and allows to construct the optimal feedback law. Again we follow the approach of Fuhrman and Tessitore [27] with slight modifications.

The first step is to prove the so called *fundamental relation*, which gives a characterization of the value function of the control problem in terms of the solution of the Hamilton-Jacobi-Bellman equation.

Proposition 11. *Let v be the solution of (HJB). For every admissible control system $\mathbb{U} = (\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_t)_{t\geq 0}, \hat{\mathbb{P}}, (\hat{W}_t)_{t\geq 0}, \hat{\gamma})$ and for the corresponding trajectory $\mathbf{x}^{\mathbb{U}}$ starting at $x \in X_\eta$ we have*

$$\begin{aligned}\mathbb{J}(x, \hat{\gamma}) = v(0, x)& \\ &+ \mathbb{E}\int_0^T -\psi\left(\sigma, \mathbf{x}^{\mathbb{U}}(\sigma)), [\nabla v(I-B)^{1-\theta}](\sigma, \mathbf{x}^{\mathbb{U}}(\sigma))(I-B)^{\theta} Pg\right) \mathrm{d}\sigma \\ &+ \mathbb{E}\int_0^T [\nabla v(I-B)^{1-\theta}](\sigma, \mathbf{x}^{\mathbb{U}}(\sigma))(I-B)^{\theta} Pgr(\sigma, J\mathbf{x}^{\mathbb{U}}(r), \hat{\gamma}(\sigma))\mathrm{d}\sigma \\ &+ \mathbb{E}\int_0^T l(\sigma, J\mathbf{x}^{\mathbb{U}}(\sigma), \hat{\gamma}(\sigma))\mathrm{d}\sigma.\end{aligned}$$

Proof. The proof follows from the same arguments used in the proof of Theorem 7.2 in [27] and is, therefore, omitted. Just notice that in this case by Theorem 8.2 we have $Z(t; s, x) = [\nabla v(I-B)^{1-\theta}](t, \mathbf{x}(t; s, x))(I-B)^{\theta} Pg$ and the role of G in [27, Theorem 7.2] is here played by $(I-B)Pg$. □

A straightforward consequence of Proposition 11 is the so-called Verification Theorem.

Corollary 3. *For every admissible control system $\mathbb{U} = (\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_t)_{t\geq 0}, \hat{\mathbb{P}}, (\hat{W}_t)_{t\geq 0}, \hat{\gamma})$ and initial datum $x \in X_\eta$ we have $\mathbb{J}(x, \mathbb{U}) \geq v(0, x)$, and the equality holds if and only if the following feedback law*

$$\hat{\gamma}(t) = \mu(t, \mathbf{x}^{\mathbb{U}}(t), [\nabla v(I-B)^{1-\theta}](t, \mathbf{x}^{\mathbb{U}}(t))(I-B)^{\theta} Pg) \quad \hat{\mathbb{P}} - a.s. \text{ for a.a.} t \in [0, T],$$

is verified by the trajectory $\mathbf{x}^{\mathbb{U}}$ starting at x and corresponding to the control $\hat{\gamma}$. In this case the pair $(\hat{\gamma}(\cdot), \mathbf{x}^{\mathbb{U}}(\cdot))$ is optimal.

We are now in the position to state the existence and uniqueness of the so-called closed loop equation, which is given by

$$\begin{cases}\mathrm{d}\mathbf{x}^{\hat{\gamma}}(t) = B\mathbf{x}^{\hat{\gamma}}(t)\mathrm{d}t + (I-B)Pf(t, J\mathbf{x}^{\hat{\gamma}}(t)) \\ \qquad\qquad + (I-B)Pg(r(t, J\mathbf{x}^{\hat{\gamma}}(t), \hat{\gamma}(t, \mathbf{x}^{\hat{\gamma}}(t)))\mathrm{d}t + \mathrm{d}W(t)) \\ \mathbf{x}(0) = x.\end{cases} \tag{72}$$

where $\hat{\gamma}$ is given by

$$\hat{\gamma}(t, x) := \mu(t, x, [\nabla v(I-B)^{1-\theta}](t, x)(I-B)^{\theta} Pg). \tag{73}$$

The main result of this section thus reads as follows:

Proposition 12. *For every $t \in [0, T], x \in X_\eta$, the closed loop equation (72) admits a weak solution $(\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_t)_{t\geq 0}, \hat{\mathbb{P}}, (\hat{W}_t)_{t\geq 0}, \bar{\mathbf{x}}(t))_{t\geq 0})$ which is unique in law and setting*

$$\hat{\gamma}(t) = \mu(t, \bar{\mathbf{x}}(t), [\nabla v(I-B)^{1-\theta}](t, \bar{\mathbf{x}}(t))(I-B)^{\theta} Pg)$$

we obtain an optimal admissible control system $(\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_t), \hat{\mathbb{P}}, \hat{W}, \bar{\mathbf{x}}(t))_{t\geq 0}, \hat{\gamma})$.

Proof. Let us take an arbitrary set up $(\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_t)_{t\geq 0}, \mathbb{P}, (W_t)_{t\geq 0})$ and consider the solution $(\bar{\mathbf{x}}(t))_{t\geq 0}$ of the uncontrolled equation

$$\begin{cases}\mathrm{d}\bar{\mathbf{x}}(t) = B\bar{\mathbf{x}}(t)\mathrm{d}t + (I-B)Pf(t, J\bar{\mathbf{x}}(t)) + (I-B)Pg\mathrm{d}W(t) \\ \bar{\mathbf{x}}(0) = x,\end{cases} \tag{74}$$

which exists in virtue of Theorem (4.2). Now we define the control

$$\hat{\gamma}(t) = \mu(t, \bar{\mathbf{x}}(t), [\nabla v(I-B)^{1-\theta}](t, \bar{\mathbf{x}}(t))(I-B)^{\theta} P g)$$

and the process

$$\hat{W}(t) := W(t) - \int_s^{t \vee s} r(\sigma, \bar{\mathbf{x}}(\sigma), \hat{\gamma}(r)) \mathrm{d}r, \quad t \in [0, T].$$

Since the function r is bounded, by Girsanov theorem there exists a probability $\hat{\mathbb{P}}$ on $(\Omega, \mathcal{F})$ equivalent to $\mathbb{P}$, such that $\hat{W}$ is a $\hat{\mathbb{P}}$-Wiener process with respect to $(\mathcal{F}_t)$. Rewriting equation (74) in terms of $\hat{W}$ we conclude that $\bar{\mathbf{x}}$ is the required solution of (72). Now applying Proposition 11 and Corollary 3 to the a.c.s. $(\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_t)_{t\geq 0}, \hat{\mathbb{P}}, (\hat{W}_t)_{t\geq 0}, \hat{\gamma})$ with

$$\hat{\gamma}(t) = \mu(t, \bar{\mathbf{x}}(t), [\nabla v(I-B)^{1-\theta}](t, \bar{\mathbf{x}}(t))(I-B)^{\theta} P g)$$

we obtain the required conclusions. □

REFERENCES

[1] O. P. Agrawal, Solution for a fractional diffusion-wave equation defined in a bounded domain, *Nonlinear Dynamics*, **29** (2002), 145–155.

[2] A. Ambrosetti and G. Prodi, *A Primer of Nonlinear Analysis*, Cambridge Studies in Advanced Mathematics, 34, Cambridge University Press, Cambridge, 1995.

[3] J. P. Aubin and H. Frankowska, *Set-Valued Analysis, Systems and Control: Foundations and Applications 2*, Birkhäuser Boston Inc., Boston, MA, 1990.

[4] S. A. Belbas, A new method for optimal control of Volterra integral equations, *Appl. Math. Comput.*, **189** (2007), 1902–1915.

[5] S. Bonaccorsi, *Some Applications in Malliavin Calculus*, Ph.D thesis, Dept. Mathematics, Univ. Trento., 1998.

[6] S. Bonaccorsi and G. Desch, Volterra equations in Banach spaces with completely monotone kernels, *NoDEA Nonlinear Differential Equations Appl.*, **20** (2013), 557–594.

[7] S. Bonaccorsi and G. Desch, Controllability of a class of Volterra equations in Hilbert spaces with completely monotone kernel, *J. Evol. Equ.*, **12** (2012), 675–695.

[8] S. Bonaccorsi and E. Mastrogiacomo, An analytic approach to stochastic Volterra equations with completely monotone kernels, *J. Evol. Equ.*, **9** (2009), 315–339.

[9] S. Bonaccorsi, F. Confortola and E. Mastrogiacomo, Optimal control for stochastic Volterra equations with completely monotone kernels, *SIAM J. Control Optim.*, **50** (2012), 748–789.

[10] P. Briand and F. Confortola, SDEs with stochastic Lipschitz condition and quadratic PDEs in Hilbert spaces, *Stochastic Process. Appl.*, **118** (2008), 818–838.

[11] P. Briand and Y. Hu, BSDE with quadratic growth and unbounded terminal value, *Probab. Theory Related Fields*, **136** (2006), 604–618.

[12] P. Briand and Y. Hu, Quadratic BSDEs with convex generators and unbounded terminal conditions, *Probab. Theory Related Fields*, **141** (2008), 543–567.

[13] B. Coleman and M. Gurtin, Equipresence and constitutive equations for rigid heat conductors, *Z. Angew. Math. Phys.*, **18** (1967), 199–208.

[14] F. Confortola and E. Mastrogiacomo, Optimal control for stochastic heat equation with memory, *Evolution Equations and Control Theory*, **3** (2014), 35–58.

[15] G. Da Prato and J. Zabczyk, Evolution equations with white-noise boundary conditions, *Stoch. Stoch. Rep.*, **42** (1993), 167–182.

[16] G. Da Prato and J. Zabczyk, *Evolution Equations with White-Noise Boundary Conditions*, Cambridge Univ. Press, Cambridge, 1993.

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

[18] A. Debussche, M. Fuhrman and G. Tessitore, Optimal control of a stochastic heat equation with boundary-noise and boundary-control, *ESAIM Control Optim. Calc. Var.*, **13** (2007), 178–205 (electronic).

[19] W. Desch and R. K. Miller, Exponential stabilization of Volterra integral equations with singular kernels, *J. Integral Equations Appl.*, **1** (1988), 397–433.

[20] K. J. Engel and R. Nagel, *One-parameter Semigroups for Linear Evolution Equations*, Graduate Texts in Math., 194, Springer-Verlag, Berlin, 2000.

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

[22] N. El Karoui and S. Hamadène, BSDEs and risk-sensitive control, zero-sum and nonzerosum game problems of stochastic functional differential equations, *Stochastic Proc. Appl.*, **107** (2003), 145–169.

[23] G. Fabbri and B. Goldys, An LQ problem for the heat equation on the half line with Dirichlet boundary control and noise, *SIAM J. Control Optim.*, **48** (2009), 1473–1488.

[24] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag, New York, 1993.

[25] M. Fuhrman, A class of stochastic optimal control problems in Hilbert spaces: BSDEs and optimal control laws, state constraints, conditioned processes, *Stochastic Process. Appl.*, **108** (2003), 263–298.

[26] M. Fuhrman, Y. Hu and G. Tessitore, On a class of stochastic optimal control problems related to BSDEs with quadratic growth, *SIAM J. Control Optim.,* **45** (2006), 1279–1296 (electronic).

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

[28] M. Fuhrman and G. Tessitore, Existence of optimal stochastic controls and global solutions of forward-backward stochastic differential equations, *SIAM J. Control Optim.*, **43** (2004), 813–830.

[29] G. Gripenberg, S. O. Londen and O. Staffans, *Volterra Integral and Functional Equations*, Cambridge University Press, 1990.

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

[31] S. Hamadène and J. P. Lepeltier, Backward equations, stochastic control and zero-sum stochastic differential games, *Stochastics Stochastics Rep.*, **54** (1995), 221–231.

[32] K. W. Homan, *An Analytic Semigroup Approach to Convolution Volterra Equations.* Ph.D. Thesis, Delft University Press, 2003.

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

[34] N. V. Krylov and B. L. Rozovskii, Stochastic evolution equations, in *Current Problems in Mathematics, Vol. 14*, Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii, Moscow, 1979, 71–147, 256.

[35] G. E. Ladas and V. Lakshmikantham, *Differential Equations in Abstract Spaces,* Mathematics in Science and Engineering, 85, Academic Press, New York, 1972,

[36] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, 2013 reprint of the 1995 original, Modern Birkhäuser Classics, Basel, 1995.

[37] F. Masiero, A stochastic optimal control problem for the heat equation on the halfline with dirichlet boundary-noise and boundary-control, *Appl. Math. Optim.*, **62** (2010), 253–294.

[38] B. Maslowski, Stability of semilinear equations with boundary and pointwise noise, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.,* **22** (1995), 55–93.

[39] R. K. Miller, Linear Volterra integrodifferential equations as semigroups, *Funkcial. Ekvac.*, **17** (1974), 39–55.

[40] D. Nualart, *Malliavin Calculus and Related Topics,* Probability and its Applications (New York), Springer-Verlag, Berlin, 2006.

[41] J. Nunziato, On heat conduction in materials with memory, *Quart. Appl. Math.*, **29** (1971), 187–204.

[42] B. Øksendal and T. Zhang, The stochastic Volterra equation, in *Barcelona Seminar on Stochastic Analysis* (St. Feliu de Guíxols, 1991), Progr. Probab., 32, Birkhäuser, Basel, 1993, 168–202.

[43] B. Øksendal and T. Zhang, The general linear stochastic Volterra equations with anticipating coefficients, in *Stochastic Analysis and Applications* (Powys, 1995), World Sci. Publ., River Edge, NJ, 1996, 343–366.

[44] E. Pardoux, Nonlinear filtering, prediction and smoothing equations, *Stochastics*, **6** (1981/82), 193–231.

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

[46] J. Prüss, *Evolutionary Integral Equations and Applications*, Monographs in Mathematics, 87, Birkhäuser Verlag, Basel, 1993.

[47] W. R. Schneider and W. Wyss, Fractional diffusion and wave equations, *Journal of Mathematical Physics*, **30** (1989), 134–144.

[48] R. B. Sowers, Multidimensional reaction-diffusion equations with white noise boundary perturbations, *Ann. Probab.*, **22** (1994), 2071–2121.

[49] J. Yong and X. Y. Zhou, *Stochastic Controls. Hamiltonian Systems and HJB Equations*, Applications of Mathematics (New York), 43, Springer-Verlag, New York, 1999.

Received September 2013; revised August 2014.

E-mail address: fulvia.confortola@polimi.it
E-mail address: elisa.mastrogiacomo@unimib.it