The Annals of Applied Probability
2018, Vol. 28, No. 6, 3558–3599
https://doi.org/10.1214/18-AAP1397

VERIFICATION THEOREMS FOR STOCHASTIC OPTIMAL CONTROL PROBLEMS IN HILBERT SPACES BY MEANS OF A GENERALIZED DYNKIN FORMULA

BY SALVATORE FEDERICO AND FAUSTO GOZZI

Università di Siena and LUISS University

Verification theorems are key results to successfully employ the dynamic programming approach to optimal control problems. In this paper, we introduce a new method to prove verification theorems for infinite dimensional stochastic optimal control problems. The method applies in the case of additively controlled Ornstein–Uhlenbeck processes, when the associated Hamilton–Jacobi–Bellman (HJB) equation admits a *mild solution* (in the sense of [*J. Differential Equations* **262** (2017) 3343–3389]). The main methodological novelty of our result relies on the fact that it is not needed to prove, as in previous literature (see, e.g., [*Comm. Partial Differential Equations* **20** (1995) 775–826]), that the mild solution is a *strong solution*, that is, a suitable limit of classical solutions of approximating HJB equations. To achieve the goal, we prove a new type of Dynkin formula, which is the key tool for the proof of our main result.

1. Introduction.

In this paper, we introduce a new technique, based on a generalized Dynkin formula, to prove verification theorems for stochastic optimal control problems over infinite horizon in Hilbert spaces.

Verification theorems are key results to enable to solve in a closed way optimal control problems through the dynamic programming approach. Once a solution (in some sense to be precised) of the associated HJB equation is known to exists, the verification theorem provides a sufficient (sometimes also necessary) condition of optimality, which can be used to find optimal controls in feedback forms through the so-called closed loop equation. In the stochastic case, when the solution v is sufficiently smooth, the proof of such theorem is substantially based on an applying the Dynkin formula to the function v and to the state process. In our framework of discounted time-homogeneous infinite horizon problems, the dependence on time is known, so the HJB equation is elliptic and v only depends on the state variable. Hence, in the finite dimensional case, to employ the classical Dynkin formula, it is needed to know that $v \in C^2$. Fortunately, in the finite dimensional case, due to the presence of a powerful regularity theory (at least for nondegenerate second-order HJB equations) there is a wide class of problems for which actually v is

Received February 2017; revised December 2017.
MSC2010 subject classifications. 93E20, 70H20, 65H15, 49L20, 49N35.
Key words and phrases. Stochastic optimal control, infinite dimensional HJB equation, Dynkin's formula, transition semigroup, verification theorem, optimal feedback.

known to enjoy this regularity, hence the classical Dynkin formula applies and the verification theorem can be proved. On the other hand, if v is not known to be sufficiently smooth (e.g., when v is known to be only a viscosity solution), still in the finite dimensional case, other techniques have been developed to overcome the fact that the classical Dynkin formula is not applicable. We mention the following techniques:

– The technique developed in [33], dealing with *viscosity solutions*. In this case, the classical Dynkin formula is applied to test functions and only some weak results are obtained.
– The technique developed in [41]. Here, a solution $v \in C^1$ is obtained through the solution of a suitable backward SDE (BSDE). This technique applies to semilinear HJB equations and provides the verification theorem as a byproduct of the construction itself of the solution v. The latter feature is particularly meaningful, as it allows to completely bypass the problem of second-order regularity of v and the application of the classical Dynkin formula. On the other hand, the powerfulness of this approach is partly limited by the fact that it can be applied only when a *structural condition* is verified by the control operator.
– The technique developed in [32]: here, v is studied and treated as a *strong solution*, that is, as a suitable limit of classical solutions.

When the state space H is infinite dimensional the situation is much worse. First of all, the regularity needed to apply the classical Dynkin formula (see, e.g., [10], Section 4.4) is very demanding and does not allow to deal with many applied examples proposed and only partly studied in the literature. This is partly due to additional regularity assumptions on the coefficients needed in infinite dimension, partly due to the lack of a satisfactory regularity theory in infinite dimension. Hence, elaborating alternative methods is considerably more important than in the finite dimensional case. Clearly, the first attempt consists in trying to extend the techniques developed in the finite dimensional case to infinite dimensional one. On this side, so far the state of the art can be basically depicted as follows:

(a) There are no results concerning the case when v is a viscosity solution.

(b) Results with the BSDE approach have been elaborated in various papers (see, e.g., [21] in the infinite horizon case), but always under the structural condition. The latter requirement leaves out the treatments of important cases like boundary control of stochastic PDEs or delayed control of SDEs.

(c) Results dealing with strong solutions are available in [5, 26, 31].

The results we provide here are closer, in the conclusions, to the results mentioned in item (c) above. With respect to them, ours have a larger range of applicability and, not only in this sense, can be seen as a significant improvement of this technique, as we will comment more precisely afterwards.

We stress the fact that our method to prove the verification theorem is a novelty also in finite dimension: our results may be useful to treat also finite dimensional problems where only partial regularity properties of the value function are known. Here, we focus on the infinite dimensional case where the application is more meaningful.

We now illustrate the results and the novelties of our paper. We consider a class of stochastic optimal control problems in a real separable Hilbert space H, where the noise is additive and the control only appears in an additive form in the drift term. More precisely, the state equation is

$$dX(t) = \big[AX(t) + GL\big(u(t)\big)\big]\,dt + \sigma\,dW(t), \tag{1.1}$$

where $A : \mathcal{D}(A) \subseteq H \to H$, $G : K \to H$, $L : \Lambda \to K$, $\sigma : \Xi \to H$ are suitable operators, with K, Ξ being other real separable Hilbert spaces and Λ being a Polish space; W is a Ξ-valued cylindrical Browian motion; u is the control process taking values in Λ; X is the state process taking values in the Hilbert space H. The stochastic control problem consists in minimizing, over a set of admissible control processes, a cost functional in the form

$$\mathbb{E}\bigg[\int_0^\infty e^{-\lambda s} l\big(X(s), u(s)\big)\,ds\bigg],$$

where $\lambda > 0$ is a discount factor and l is a suitable real valued function. In this case, the associated HJB equation is an elliptic semilinear PDE in the space H:

$$\lambda v(x) - \frac{1}{2}\operatorname{Tr}\big[\sigma\sigma^* D^2 v(x)\big] - \langle Ax, Dv(x)\rangle_H - F_0\big(x, D^G v(x)\big) = 0,$$

where

$$F_0\big(x, D^G v(x)\big) = \inf_{u\in\Lambda}\big\{\langle L(u), D^G v(x)\rangle_K + l(x,u)\big\},$$

where $D^G v$ denotes the G-gradient of a function $v : H \to \mathbb{R}$ (see Section 2.2). Under reasonable assumptions, it is proved in [16] that such HJB equation admits a unique mild solution, that is, a solution of a suitable integral form of the above equation. Such solution admits G-gradient, that is, verifies the minimal differentiability requirement to give sense to the nonlinear Hamiltonian term F_0 in HJB above. Once one proves the existence of a mild solution v to the associated HJB equation, the approach of item (c) would require three nontrivial technical steps: first, proving that such a mild solution is indeed a strong solution (limit, in a suitable sense, of classical solutions of approximating HJB equations); second, applying Dynkin formula to the approximating classical solutions; third, passing to the limit the Dynkin formula. As one may expect, passing through all these steps requires additional hypotheses that may be nontrivial to check in practice (see, e.g., [31]). Our goal here is to bypass these steps through an alternative path. In fact, we show that the role of strong solutions is not essential. Indeed, relying on the theory of π-semigroups (see, e.g., [14], Appendix B and [43]), we prove a generalized (abstract) Dynkin formula—deserving interest in itself—which can be directly applied to mild solutions. The proof is quite involved and this is the reason why we consider here the case of stochastic control of equation of type (1.1), where the uncontrolled part of the state equation is of Ornstein–Uhlenbeck type.[1] Then, relying on this formula, we straightly prove a verification theorem. The new results on G-derivatives provided in [16] (see also [14], Chapter 4) enable us to apply our method to more general examples than the ones treated by the current literature; in particular, to cases where the structural condition required at item (b) above is not verified (see Section 6).

The main results of the paper are the abstract Dynkin formula (Theorem 4.8); the verification theorem (Theorem 5.6); the consequent Corollary 5.7 on sufficient conditions for the existence of optimal control processes in feedback form. Moreover, since the existence of optimal feedback controls might be easier to obtain when the optimal control problem is considered in the weak formulation, we also provide Corollary 5.8 in this direction. We underline that we do not provide general results on the existence of optimal control processes in feedback form, as such results strongly depend on the specific case at hand. To this regard, in Section 6—where we deal with two specific applications: optimal boundary control (of Neumann type) of the stochastic heat equation and optimal control of SDEs with delay in the control variable—we provide for the first example some results and comments on the existence of optimal feedback control processes.

The paper is organized as follows. After some preliminaries in Section 2 on spaces, notation and the notion of G-derivative recently extended in [16], we introduce our family of control problems in Section 3. Section 4 is devoted to prove our new Dynkin formula (Theorem 4.8), the methodological core of the paper. In Section 5, we prove our main results on the control problem: in Section 5.1, the verification theorem (Theorem 5.6); in Section 5.2, Corollary 5.7 on optimal feedbacks. Section 6 is devoted to illustrate the applications of our results to the aforementioned examples. Finally, the Appendix is devoted to prove few technical results needed to prove our Dynkin formula.

2. Preliminaries. In this section, we provide some preliminaries about spaces and notation used in the rest of the paper and recall from [16] the notion of G-derivative. We restrict the treatment of G-derivative to the case of real valued functions defined on Hilbert spaces and to constant operator maps G. This will be enough for the purposes of the present paper. For a more general theory and more details, we refer to the aforementioned paper [16].

[1]It is worth to stress that, even if in the case of Ornstein–Uhlenbeck dynamics the approach of strong solutions has already been successfully applied (see [31]), the method used here, other than being original, seems to be extendable to more general structures of state equations, where the strong solution approach would fail.

2.1. *Spaces and notation*. *Measurable bounded and continuous functions.* All the topological spaces are intended endowed with their Borel σ-algebra, denoted by $\mathcal{B}$. By measurable set (function), we always intend a *Borel* measurable set (function). If U is a topological space and V is a topological vector space, we denote by $B_b(U, V)$ the set of bounded measurable functions from U to V and by $C_b(U, V)$ the set of bounded continuous functions from U to V. If $V = \mathbb{R}$, we drop it in the latter notation. If V is complete, the spaces $B_b(U, V)$ and $C_b(U, V)$ are Banach spaces when endowed with the norm

$$|\varphi|_\infty = \sup_{x\in U} |\varphi(x)|_V. \tag{2.1}$$

Hilbert spaces. Let H be a Hilbert space. We denote its norm by $|\cdot|_H$ and its inner product by $\langle\cdot,\cdot\rangle_H$. We omit the subscript if the context is clear and if $H = \mathbb{R}$. If a sequence $(x_n)_{n\in\mathbb{N}} \subseteq H$, converges to $x \in U$ in the norm (strong) topology, we write $x_n \to x$.

We denote by H^* the topological dual of H, that is, the space of all continuous linear functionals defined on H. We always identify H^* with H through the standard Riesz identification.

Linear operators. Let H, K be real separable Hilbert spaces. We denote by $\mathcal{L}(H, K)$ the set of all bounded (continuous) linear operators $T : H \to K$ with norm $|T|_{\mathcal{L}(H,K)} := \sup_{x\in H, x\neq 0} \frac{|Tx|_K}{|x|_H}$, using for simplicity the notation $\mathcal{L}(H)$ when $H = K$. Moreover, we denote by $\mathcal{L}_u(H, K)$ the space of closed densely defined and possibly unbounded linear operators $T : \mathcal{D}(T) \subseteq H \to K$, where $\mathcal{D}(T)$ denotes the domain. We recall that $\mathcal{D}(T)$ is a Hilbert space when endowed with the graph norm $|x|_{\mathcal{D}(T)} = |x|_H + |Tx|_K$. The range of an operator $T \in \mathcal{L}_u(H, K)$ is denoted by $\mathcal{R}(T)$. Clearly, $\mathcal{L}(H, K) \subseteq \mathcal{L}_u(H, K)$. Given $T \in \mathcal{L}_u(H, K)$, we denote its adjoint operator by $T^* : \mathcal{D}(T^*) \subseteq K \to H$.

We denote by $\mathcal{L}_1(H)$ the set of trace class operators, that is, the operators $T \in \mathcal{L}(H)$ such that, given an orthonormal basis $\{e_k\}_{k\in\mathbb{N}}$ of H, the quantity

$$|T|_{\mathcal{L}_1(H)} := \sum_{k=1}^{\infty} \langle (T^*T)^{1/2} e_k, e_k\rangle_H$$

is finite (see [45], Section VI.6). The latter quantity is independent of the basis chosen and defines a norm making $\mathcal{L}_1(H)$ a separable Banach space. The trace of an operator $T \in \mathcal{L}_1(H)$ is denoted by $\mathrm{Tr}[T]$, that is, $\mathrm{Tr}[T] := \sum_{k=0}^{\infty} \langle Te_k, e_k\rangle_U$. The latter quantity is finite and, again, independent of the basis chosen. We denote by $\mathcal{L}_1^+(U)$ the subset of $\mathcal{L}_1(H)$ of self-adjoint nonnegative (trace class) operators on H. Note that, if $T \in \mathcal{L}_1^+(H)$, then $\mathrm{Tr}[T] = |T|_{\mathcal{L}_1(U)}$.

We denote by $\mathcal{L}_2(H, K)$ [subset of $\mathcal{L}(H, K)$] the space of Hilbert–Schmidt operators from H to K, that is, the spaces of operators such that, given an orthonormal basis $\{e_k\}_{k\in\mathbb{N}}$ of H, the quantity

$$|T|_{\mathcal{L}_2(H)} := \left(\sum_{k=0}^{\infty} |Te_k|_K^2\right)^{1/2}$$

is finite (see [45], Section VI.6). The latter quantity is independent of the basis chosen and defines a norm making $\mathcal{L}_2(H)$ a Banach space. It is actually a Hilbert space with the scalar product

$$\langle T, S\rangle_{\mathcal{L}_2(H,K)} := \sum_{k=0}^{\infty} \langle Te_k, Se_k\rangle_K,$$

where $\{e_k\}_{k\in\mathbb{N}}$ is any orthonormal basis of H.

Stochastic processes. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions. Given $p \in [1, +\infty)$, $T > 0$, and a Hilbert space U, we denote by $\mathcal{M}_{\mathcal{P}}^{p,T}(U)$ the set of all (equivalence classes of) progressively measurable processes $X\colon [0, T] \times \Omega \to U$ such that

$$|X|_{\mathcal{M}_{\mathcal{P}}^{p,T}(U)} := \left(\int_0^T \mathbb{E}\big[|X(s)|_U^p\big]\,ds\right)^{1/p} < \infty.$$

This is a Banach space with the norm $|\cdot|_{\mathcal{M}_{\mathcal{P}}^{p,T}(U)}$. Next, we denote by $\mathcal{M}_{\mathcal{P}}^{p,\mathrm{loc}}(U)$ the space of all (equivalence classes of) progressively measurable processes $X \in \mathcal{M}_{\mathcal{P}}^{p,T}(U)$ such that $X|_{[0,T]\times\Omega} \in \mathcal{M}_{\mathcal{P}}^{p,T}(U)$ for every $T > 0$. We denote by $\mathcal{K}_{\mathcal{P}}^{p,T}(U)$ the set of all (equivalence classes of) progressively measurable processes $X \in \mathcal{M}_{\mathcal{P}}^{p,T}(U)$ such that

$$[0, T] \to L^p(\Omega, U), \qquad t \mapsto X(t)$$

is continuous. This is a Banach space with the norm

$$|X|_{\mathcal{K}_{\mathcal{P}}^{p,T}(U)} := \sup_{s\in[0,T]} \big(\mathbb{E}|X(s)|_U^p\big)^{1/p}.$$

Next, we denote by $\mathcal{K}_{\mathcal{P}}^{p,\mathrm{loc}}(U)$ the space of all (equivalence classes of) progressively measurable processes $X\colon [0, +\infty) \times \Omega \to U$ such that $X|_{[0,T]\times\Omega} \in \mathcal{K}_{\mathcal{P}}^{p,T}(U)$ for every $T > 0$. We also say that elements of $\mathcal{K}_{\mathcal{P}}^{p,T}(U)$ and $\mathcal{K}_{\mathcal{P}}^{p,\mathrm{loc}}(U)$ are "p-mean continuous."

2.2. *G-derivative*. Here, we provide the notion of G-derivative for functions $f : H \to \mathbb{R}$, where H is a Hilbert space. The latter notion is considered in [16] when G is a map $G : U \to \mathcal{L}_u(Z, U)$, with U, Z Banach spaces. Here, we restrict to the case of constant G.

Recall that, if $f : H \to \mathbb{R}$, the Fréchet derivative of f at x (if it exists) is the (unique) linear functional $Df(x) \in H^* \cong H$ such that

$$\lim_{|h|_H\to 0} \frac{|f(x+h) - f(x) - \langle Df(x), h\rangle_H|}{|h|_H} = 0.$$

DEFINITION 2.1 (G-derivative). Let H, K be Hilbert spaces, let $f : H \to \mathbb{R}$ and $G \in \mathcal{L}_u(K, H)$. We say that f is continuously G-Fréchet differentiable at $x \in H$ (briefly, G-differentiable at $x \in H$) if there exists $D^G f(x) \in K^* \cong K$ (clearly, if it exists, then it is unique), called the G-derivative of f at x, such that

$$\lim_{k\in\mathcal{D}(G), |k|_K\to 0} \frac{|f(x+Gk) - f(x) - \langle D^G f(x), k\rangle_K|}{|k|_K} = 0. \tag{2.2}$$

We denote by $C_b^{1,G}(H)$ the space of all maps $f : H \to \mathbb{R}$ such that f is G-differentiable over H and $D^G f : H \to K$ belongs to $C_b(H, K)$. In the special case $K = H$ and $G = I$, we simply use the standard notation $C_b^1(H)$.

REMARK 2.2. Note that, in the definition of the G-derivative, one considers only the directions in H selected by the range of G. When $K = H$ and $G = I$ it reduces to the Fréchet derivative, that is, $Df = D^G f$. Clearly, if f is G-differentiable at x, then it is also G-Gateaux differentiable at x, in the sense that

$$\lim_{t\to 0} \frac{f(x+tGk) - f(x)}{t} = \langle D^G f(x), k\rangle_K \qquad \forall k \in \mathcal{D}(G); \tag{2.3}$$

moreover, the limit above is uniform in $k \in \mathcal{D}(G) \cap B_K(0, R)$, for every $R > 0$. Conversely, if there exists $k' \in K$ such that

$$\begin{gathered}\lim_{t\to 0} \frac{f(x+tGk) - f(x)}{t} = \langle k', k\rangle_K \\ \text{uniformly in } k \in \mathcal{D}(G) \cap B_K(0, R), \forall R > 0,\end{gathered} \tag{2.4}$$

then f is G-differentiable at $x \in H$ and $D^G f(x) = k'$.

The notion of G-derivative allows to deal with functions which are not Gateaux differentiable, as shown by the following example.

EXAMPLE 2.3. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x_1, x_2) := |x_1|x_2$. Clearly, f does not admit directional derivative in the direction $(1, 0)$ at the point $(x_1, x_2) = (0, 1)$. On the other hand, if we consider $G \in \mathcal{L}(\mathbb{R}^2) \cong \mathbb{R}^2$, defined by $G = (0, 1)$, then f admits G-Fréchet derivative at every $(x_1, x_2) \in \mathbb{R}^2$.

REMARK 2.4. Clearly, if f is Fréchet differentiable at some $x \in H$ and $G \in \mathcal{L}(K, H)$, it turns out that f is G-Fréchet differentiable at x and

$$D^G f(x) = G^* Df(x). \tag{2.5}$$

Also, if f is both Fréchet differentiable and G-differentiable at some $x \in H$, then $Df(x) \in \mathcal{D}(G^*)$ and (2.5) holds true. Indeed, we get by Fréchet differentiability

$$\lim_{s\to 0} \frac{f(x+sGk) - f(x)}{s} = \langle Df(x), Gk\rangle_H \qquad \forall k \in \mathcal{D}(G).$$

On the other hand, by G-Fréchet differentiability we also have

$$\lim_{s\to 0} \frac{f(x+sGk)-f(x)}{s} = \langle D^G f(x), k\rangle_K \qquad \forall k \in \mathcal{D}(G).$$

Hence

$$\left|\langle Df(x), Gk\rangle_H\right| = \left|\langle D^G f(x), k\rangle_K\right| \le \left|D^G f(x)\right|_K |k|_K \qquad \forall k \in \mathcal{D}(G).$$

It follows what claimed.

If G is unbounded, a function $f : H \to \mathbb{R}$ may be Fréchet-differentiable at some $x \in H$ and yet not G-Fréchet differentiable there, as shown by the following example.

EXAMPLE 2.5. Let H, K be Hilbert spaces, let $G : \mathcal{D}(G) \subsetneq K \to H$ be a closed densely defined unbounded linear operator on H, and let $G^* : \mathcal{D}(G^*) \subsetneq H \to K$ be its adjoint. Next, let $f : U \to \mathbb{R}$ be defined by $f(x) := \frac{1}{2}|x|_H^2$. Clearly, f is Fréchet differentiable at every $x \in H$ and $Df(x) = x$. On the other hand, if f was also G-differentiable at every $x \in H$, by Remark 2.4 it would follow $x \in \mathcal{D}(G^*)$ for every $x \in H$, that is, $\mathcal{D}(G^*) = H$, a contradiction.

3. Formulation of the stochastic optimal control problem. We are concerned with the optimal control of an Ornstein–Uhlenbeck process valued in a Hilbert space H. Precisely, let H, K, Ξ three real separable Hilbert spaces, let $(U, |\cdot|_U)$ be a real separable Banach space and let $\Lambda \subseteq U$ be measurable and endowed with the σ-algebra induced by $\mathcal{B}(U)$, the Borel σ-algebra of U. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge 0}, \mathbb{P})$ be a complete filtered probability space satisfying the usual conditions, let $W = (W_t)_{t\ge 0}$ be a Ξ-valued cylindrical Brownian motion (see [10], Chapter 4), and consider the controlled SDE

$$\begin{cases} dX(t) = \left[AX(t) + GL\big(u(t)\big)\right]dt + \sigma\, dW(t), & t \ge 0, \\ X(0) = x, \end{cases} \tag{3.1}$$

where the control process $u(\cdot)$, taking values in Λ, belongs to a suitable space of admissible controls and the coefficients A, G, L, σ satisfy the following assumptions, which will be standing and not repeated throughout the paper.

ASSUMPTION 3.1. (i) $A : \mathcal{D}(A) \subseteq H \to H$ is a closed densely defined linear operator generating a C_0-semigroup $\{e^{tA}\}_{t\ge 0}$ of operators of $\mathcal{L}(H)$.

(ii) $\sigma \in \mathcal{L}(\Xi, H)$, $e^{sA}\sigma\sigma^* e^{sA^*} \in \mathcal{L}_1(H)$ for all $s > 0$, and there exists $\gamma \in (0, 1/2)$ such that

$$\int_0^t s^{-2\gamma}\, \mathrm{Tr}\left[e^{sA}\sigma\sigma^* e^{sA^*}\right] ds < \infty \qquad \forall t \ge 0.$$

(iii) $G:\mathcal{D}(G)\subseteq K\to H$ is a closed densely defined[2] linear operator such that $e^{sA}G:\mathcal{D}(G)\to H$ can be extended for every $s>0$ to a continuous linear operator defined on K that we denote by $\overline{e^{sA}G}$. Moreover, there exists $C_G>0$, $a_G\in\mathbb{R}$ and $\beta\in[0,1)$ such that

$$|\overline{e^{sA}G}|_{\mathcal{L}(K,H)}\le C_G\big(s^{-\beta}\vee 1\big)e^{a_G s}\qquad \forall s>0. \tag{3.2}$$

(iv) $L:\Lambda\to K$ is measurable and $|L(u)|_K\le C_L(1+|u|_U)$ for some $C_L>0$.

REMARK 3.2. Since for every $t>0$ and $s\ge 0$ the operators $\overline{e^{(s+t)A}G}$ and $e^{sA}\overline{e^{tA}G}$ belong to $\mathcal{L}(K,H)$ and coincide on the dense subset $\mathcal{D}(G)\subseteq K$, we have

$$\overline{e^{(s+t)A}G}=e^{sA}\overline{e^{tA}G}\qquad \forall t>0,\forall s\ge 0. \tag{3.3}$$

This implies that the map $(0,+\infty)\to\mathcal{L}(K,H)$, $s\mapsto\overline{e^{sA}G}$ is strongly continuous, that is, $s\mapsto\overline{e^{sA}G}x$ is continuous for each $x\in H$.

We now take

$$p\in\left(\frac{1}{1-\beta},+\infty\right), \tag{3.4}$$

which will be fixed in the rest of the paper. We consider, as space of admissible controls, the space of processes

$$\begin{aligned}\mathcal{U}_p:=\Big\{&u:\Omega\times[0,+\infty)\to\Lambda \text{ prog. meas. and s.t.}\\ &\int_0^t\mathbb{E}\big[|u(s)|_U^p\big]\,ds<\infty\ \forall t\ge 0\Big\}.\end{aligned} \tag{3.5}$$

The reason for the choice of β in (3.2) and of p in (3.4) relies on the following result (cf. also [20], Proposition 8.8, and [23], Lemma 3.2), which will guarantee well-posedness of the controlled state equation (Proposition 3.4).

LEMMA 3.3. *Let E,V be real Banach spaces, let $\beta\in[0,1)$, $p>\frac{1}{1-\beta}$. Let $f\in L^p_{\mathrm{loc}}([0,+\infty);E)$ and let $g:(0,+\infty)\to\mathcal{L}(E,V)$ be strongly continuous[3] and such that $|g(s)|_{\mathcal{L}(E,V)}\le C_0(s^{-\beta}\vee 1)$ for some $C_0>0$ for every $s\in(0,+\infty)$. Then $F:\mathbb{R}^+\to V$ defined as Bochner integral by*

$$F(t):=\int_0^t g(t-s)f(s)\,ds,\qquad t\in\mathbb{R}^+,$$

is well-defined and continuous.

[2]The assumption that G is densely defined can be done without loss of generality, as one can always restrict K to $\overline{\mathcal{D}(G)}$.

[3]Meaning that $g(\cdot)e:(0,+\infty)\to V$ is continuous for each $e\in E$.

PROOF. Let $t > 0$. First of all, we note that the map

$$[0, t) \to V, \qquad s \mapsto g(t-s)f(s),$$

is measurable for each $t > 0$. Indeed, given $t > 0$ the above map can be seen as the composition $h_1 \circ h_2$ where

$$h_1 : (0, t] \times E \to V, \qquad h_1(s, e) = g(s)e;$$

$$h_2 : [0, t) \to (0, t] \times E, \qquad h_2(s) = \big(t - s, f(s)\big).$$

Now, h_2 is clearly measurable. Also h_1 is measurable, as it is continuous: indeed $g(\cdot)e$ is continuous for each $e \in E$ and $\{g(s)\}_{s\in[\varepsilon,t]} \subseteq \mathcal{L}(E, V)$ is a family of uniformly bounded operators for each $\varepsilon \in (0, t)$. Hence $h_1 \circ h_2$ is measurable.

Given the above, it makes sense to consider $\int_0^t g(t-s)f(s)\,ds$ in Bochner sense for each $t > 0$. By Hölder's inequality, setting $\kappa := -\frac{\beta p}{p-1} + 1 > 0$, we have for each $t > 0$

$$\begin{aligned}\int_0^t \big|g(t-s)f(s)\big|_V\,ds &\leq \int_0^t (t-s)^{-\beta}\big|f(s)\big|_V\,ds \\ &\leq \left(\int_0^t (t-s)^{-\beta\frac{p}{p-1}}\,ds\right)^{\frac{p-1}{p}} |f|_{L^p([0,T];\mathbb{R})} \\ &= \left(\frac{t^\kappa}{\kappa}\right)^{\frac{p-1}{p}} |f|_{L^p([0,T];\mathbb{R})}.\end{aligned}$$

This show, at once, that F is well-defined as Bochner integral in V and that $\lim_{t\to 0^+} F(t) = 0$, so F is continuous at 0.

Let us show now that F is continuous on each interval of the form $[t_0, T]$ with $t_0 \in (0, T)$. Set, for $\varepsilon \in (0, t_0)$,

$$F_\varepsilon(t) := \int_0^{t-\varepsilon} g(t-s)f(s)\,ds, \qquad t \in [t_0, T].$$

By dominated convergence we easily see that F_ε is continuous on $[t_0, T]$. Moreover, using again Hölder's inequality we have, for all $t \in [t_0, T]$,

$$\begin{aligned}\big|F(t) - F_\varepsilon(t)\big| &\leq \left(\int_{t-\varepsilon}^t (t-s)^{-\beta\frac{p}{p-1}}\,ds\right)^{\frac{p-1}{p}} |f|_{L^p([0,T];\mathbb{R})} \\ &= \left(\frac{\varepsilon^\kappa}{\kappa}\right)^{\frac{p-1}{p}} |f|_{L^p([0,T];\mathbb{R})}.\end{aligned}$$

This show $F_\varepsilon \to F$ uniformly in $[t_0, T]$, hence F is continuous in $[t_0, T]$, concluding the proof. □

PROPOSITION 3.4. *For each $u(\cdot) \in \mathcal{U}_p$, the process*

$$X\big(t; x, u(\cdot)\big) := e^{tA}x + \int_0^t e^{(t-s)A}\sigma\,dW(s) + \int_0^t \overline{e^{(t-s)A}G}L\big(u(s)\big)\,ds, \tag{3.6}$$

is well-defined and belongs to $\mathcal{K}_{\mathcal{P}}^{1,\mathrm{loc}}(H)$. *Moreover, it admits a version with continuous trajectories.*

PROOF. By Remark 3.2 and Assumption 3.1(iii)–(iv), we can apply Lemma 3.3 with

$$E = L^1(\Omega; K), \qquad V = L^1(\Omega; H), \qquad f(s) = L\big(u(s)\big)$$

and $g : (0, +\infty) \to \mathcal{L}(E, V)$ defined by

$$\big[g(s)Z\big](\omega) := \overline{e^{sA}G}Z(\omega), \qquad Z \in L^1(\Omega; K).$$

It follows that

$$t \longmapsto \int_0^t \overline{e^{(t-s)A}G}L\big(u(s)\big)\,ds \tag{3.7}$$

is well-defined as stochastic process and belongs to $\mathcal{K}_{\mathcal{P}}^{1,\mathrm{loc}}(H)$. We can repeat the argument employed above dealing now with trajectories. Fixing $\omega \in \Omega$ and applying Lemma 3.3 with

$$E = K, \qquad V = H, \qquad f(s) := L\big(u(s)(\omega)\big), \qquad g(s) = \overline{e^{sA}G},$$

it follows that the map

$$\mathbb{R}^+ \to H, \qquad t \mapsto \int_0^t \overline{e^{(t-s)A}G}L\big(u(s)(\omega)\big)\,ds$$

is continuous. The latter integral expression, for varying $\omega \in \Omega$, clearly provides a version of (3.7) with continuous trajectories.

On the other hand, in view of Assumption 3.1(ii), from [10], Theorem 5.2 and Theorem 5.11, we know that the *stochastic convolution*

$$W^A(t) := \int_0^t e^{(t-s)A}\sigma\,dW(s), \qquad t \geq 0,$$

is a (well-defined) stochastic process belonging to $\mathcal{K}_{\mathcal{P}}^{2,\mathrm{loc}}(H)$ and admitting a version with continuous trajectories, concluding the proof. □

We refer to the process (3.6) as the *controlled Ornstein–Uhlenbeck process* or *mild solution* of SDE (3.1). We always consider its version (unique, up to indistinguishability) with continuous trajectories.

Let $\lambda > 0$, $x \in H$, and let $l : H \times \Lambda \to \mathbb{R}$ be such that

$$l \text{ is measurable and bounded from below.}^{4} \tag{3.8}$$

[4] Cases where l is not bounded from below can be treated adding suitable growth conditions which depends on the specific problem at hand. We do not do it here for brevity. See also Remark 4.9 on this.

Consider the functional

$$J(x;u(\cdot)) = \mathbb{E}\left[\int_0^\infty e^{-\lambda s} l(X(s;x,u(\cdot)),u(s))\,ds\right], \qquad x \in H, u(\cdot) \in \mathcal{U}_p. \tag{3.9}$$

By (3.8), the functional above is well-defined (possibly with value $+\infty$) for all $x \in H$ and $u(\cdot) \in \mathcal{U}_p$. The stochastic optimal control problem consists in minimizing (3.9) over the set of admissible controls $\mathcal{U}_p$, that is, in solving the optimization problem

$$V(x) := \inf_{u(\cdot)\in\mathcal{U}_p} J(x;u(\cdot)), \qquad x \in H. \tag{3.10}$$

The function $V : H \to \mathbb{R} \cup \{+\infty\}$ is the so-called value function of the optimization problem. If $x \in H$ is such that $V(x) < \infty$ and $u^*(\cdot)$ is such that $V(x) = J(x; u^*(\cdot))$, then $u^*(\cdot)$ is called *optimal strategy* and the associated state trajectory is called *optimal state*; moreover, the couple $(u^*(\cdot), X(\cdot; x, u^*(\cdot)))$ is called an *optimal couple*.

4. Generalized Dynkin's formula. The aim of the present section is to prove an abstract Dynkin formula for the controlled Ornstein–Uhlenbeck process (3.6) composed with suitably smooth functions $\varphi : H \to \mathbb{R}$.

4.1. *Transition semigroups, generators and G-derivatives.* We consider the family of transition semigroups associated to (3.6) under constant controls. Precisely, we denote by $X^{(k)}(\cdot; x)$, where $k \in K$, the Ornstein–Uhlenbeck process starting at $x \in H$ with extra drift Gk; that is, the mild solution to

$$\begin{cases} dX(t) = [AX(t) + Gk]\,dt + \sigma\,dW(t), & t \geq 0, \\ X(0) = x. \end{cases} \tag{4.1}$$

Its explicit expression is

$$X^{(k)}(t;x) := e^{tA}x + \int_0^t e^{(t-s)A}\sigma\,dW(s) + \int_0^t \overline{e^{(t-s)A}G}k\,ds. \tag{4.2}$$

Correspondingly, we define the family of linear operators $\{P_t^{(k)}\}_{t\geq 0}$ in the space $C_b(H)$ as

$$P_t^{(k)}[\varphi](x) := \mathbb{E}[\varphi(X^{(k)}(t;x))], \qquad \varphi \in C_b(H), x \in H, t \geq 0. \tag{4.3}$$

In Proposition 4.3(i), we will show that the family $\{P_t^{(k)}\}_{t\geq 0}$ is a one-parameter semigroup of linear operators in the space $C_b(H)$. According to the related the literature, we call it the *transition semigroup* associated to the process $X^{(k)}$. Unfortunately, such semigroup is not in general a C_0-semigroup in $C_b(H)$, not even in the case $k = 0$. Indeed, in the framework of spaces of functions not vanishing at infinity, the C_0-property, that is, the fact that $\lim_{s\to 0^+} P_s^{(k)}\varphi = \varphi$ in the sup norm for every φ, fails even in basic cases. For instance, this property fails in the case of the Ornstein–Uhlenbeck semigroup in the space $C_b(\mathbb{R})$ (see, e.g., [4], Example 6.1, for a counterexample in $UC_b(\mathbb{R})$, or [8], Lemma 3.2, which implies this is a C_0-semigroup in $UC_b(\mathbb{R})$ if and only if the drift of the SDE vanishes). Even worse: given $\varphi \in C_b(H)$, the map $[0, +\infty) \to C_b(H)$, $t \mapsto P_t^{(k)}\varphi$ is not in general measurable, as shown in [16], Example 4.5. This prevents, for instance, to intend in Bochner sense, in the space $C_b(H)$ for each $g \in C_b(H)$, the integral defining the Laplace transform

$$\int_0^\infty e^{-\lambda s} P_s^{(k)}[g]\, ds. \tag{4.4}$$

Nevertheless, one can get, in a weaker sense, several statements of the classical theory of C_0-semigroups. This is performed, for example, by the theory of $\mathcal{K}$-semigroups (introduced in [4], see also [6], with the different terminology of *weakly continuous semigroups*) and π-semigroups (introduced in [43, 44]). Both theories (a survey of which can be found in Appendix B.5 of [14]) can be applied here getting substantially the same results. We employ the π-semigroups approach, as it seems more natural in our context. The definition of π-convergence can be found, for example, in [12], p. 111, where it is called *bp*-convergence (*bounded-pointwise* convergence) and in [43, 44]; the former in the space $C_b(H)$, the latter in the space $UC_b(H)$.

DEFINITION 4.1 (π-convergence). A sequence of functions $(f_n) \subseteq C_b(H)$ is said to be π-convergent to a function $f \in C_b(H)$ if

$$\sup_{n\in\mathbb{N}} |f_n|_{C_b(H)} < \infty \quad \text{and} \quad \lim_{n\to\infty} f_n(x) = f(x) \qquad \forall x \in H.$$

Such convergence is denoted by $f_n \xrightarrow{\pi} f$ or by $f = \pi\text{-}\lim_{n\to\infty} f_n$.

Now we recall the definition of π-semigroup as given in [43, 44]. Here, we state it in the space of continuous and bounded functions [the aforementioned references deal with the space of *uniformly* continuous and bounded functions, but also explain how to extend the definition to $C_b(H)$].

DEFINITION 4.2. A semigroup $\{P_t\}_{t\geq 0}$ of bounded linear operators on $C_b(H)$ is called a π-semigroup on $C_b(H)$ if it satisfies the following conditions:

(P1) There exist $M \geq 1$ and $\alpha \in \mathbb{R}$ such that $|P_t[f]|_\infty \leq Me^{\alpha t}|f|_\infty$ for every $t \in \mathbb{R}^+$, $f \in C_b(H)$.

(P2) For each $x \in H$ and $f \in C_b(H)$, the map $\mathbb{R}^+ \to \mathbb{R}$, $t \mapsto P_t[f](x)$ is continuous.

(P3) We have

$$\{f_n\}_{n\in\mathbb{N}} \subset C_b(H), \qquad f_n \xrightarrow{\pi} f \in C_b(H) \quad \Longrightarrow \quad P_t[f_n] \xrightarrow{\pi} P_t[f] \qquad \forall t \geq 0.$$

Define

$$\mathcal{D}(\mathcal{A}^{(k)}) := \left\{ \varphi \in C_b(H) : \exists \pi\text{-} \lim_{t \to 0^+} \frac{P_t^{(k)}[\varphi] - \varphi}{t} \right\} \tag{4.5}$$

and

$$\mathcal{A}^{(k)}[\varphi] := \pi\text{-} \lim_{t \to 0^+} \frac{P_t^{(k)}[\varphi] - \varphi}{t}, \qquad \varphi \in \mathcal{D}(\mathcal{A}^{(k)}). \tag{4.6}$$

It is proved (see [6], Lemma. 5.7, combined with the discussion of [43], Section 4.3) that, for φ sufficiently smooth,

$$\mathcal{A}^{(0)}[\varphi](x) = \frac{1}{2} \operatorname{Tr}\big[\sigma\sigma^* D^2 \varphi(x)\big] + \langle x, A^* D\varphi(x) \rangle. \tag{4.7}$$

We will use (4.7) to formally motivate the definition of mild solution (Definition 5.1) of the HJB equation associated to the control problem of Section 3.

PROPOSITION 4.3. *Let $k \in K$:*

(i) *The family of linear operators $\{P_t^{(k)}\}_{t \geq 0}$ defined in (4.3) is a π-semigroup on $C_b(H)$. We denote by $\mathcal{A}^{(k)}$ its infinitesimal generator.*

(ii) *The operator*

$$R_\lambda^{(k)}[g](x) := \int_0^\infty e^{-\lambda s} P_s^{(k)}[g](x)\, ds, \qquad g \in C_b(H), x \in H,$$

belongs to $\mathcal{L}(C_b(H))$ for every $\lambda > 0$ and is the resolvent of $\mathcal{A}^{(k)}$:

$$\big(\lambda - \mathcal{A}^{(k)}\big)^{-1} = R_\lambda^{(k)} \qquad \forall \lambda > 0.$$

(iii) *We have*[5]

$$\frac{d}{dt} P_t^{(k)}[\varphi](x) = P_t^{(k)}\big[\mathcal{A}^{(k)}[\varphi]\big](x) = \mathcal{A}^{(k)}\big[P_t^{(k)}[\varphi]\big](x)$$
$$\forall \varphi \in \mathcal{D}\big(\mathcal{A}^{(k)}\big), \forall x \in H, \forall t \geq 0.$$

PROOF. Claims (ii)–(iii) follow from [43], Proposition 3.2, Proposition 3.6, or [44], Proposition 6.2.7, Proposition 6.2.11,[6] once one proves claim (i), which we prove below.

[5] At $t = 0$, the derivative is intended as right derivative.

[6] These references deal mainly in the space of uniformly continuous and bounded functions—we warn that the author denotes by $C_b(H)$ the latter space. The extension to the space of continuous and bounded function—our space $C_b(H)$—is illustrated in [43], Section 5 and [44], Section 6.5.

Proof of (i). First of all, we prove that $\{P_t^{(k)}\}_{t\geq 0}$ is a semigroup of linear operators on $C_b(H)$. The fact that $P_0^{(k)} = I$ and that $P_t^{(k)} \in \mathcal{L}(C_b(H))$ for all $t \geq 0$ is immediate. The semigroup property of $\{e^{tA}\}_{t\geq 0}$ and (3.3) yield

$$X^{(k)}(t+s;x) = e^{sA}X^{(k)}(t;x) + \int_0^s e^{(s-r)A}\sigma\, dW(t+r) + \int_0^s \overline{e^{(s-r)A}G}k\, dr \qquad \forall t \geq 0, \forall s > 0.$$

The latter shows the strong Markov property of $X^{(k)}$ and then the fact that $\{P_t^{(k)}\}_{t\geq 0}$ satisfies the semigroup property follows as consequence (see, e.g., [10], Corollary 9.15).

Now we show the other properties of Definition 4.2. (P1) is obviously verified with $M = 1$ and $\alpha = 0$. (P2) of Definition 4.2 corresponds to

$$\mathbb{E}\big[f(X^{(k)}(t;x)\big] \overset{t\to t_0}{\longrightarrow} \mathbb{E}\big[f(X^{(k)}(t_0;x)\big] \qquad \forall f \in C_b(H), \forall x \in H, \forall t_0 \geq 0. \tag{4.8}$$

The latter follows from continuity of trajectories of $X^{(k)}(\cdot;x)$ and dominated convergence. Finally, (P3) of Definition 4.2 is verified by dominated convergence. $\square$

A key step toward the main goal of this section, that is, the proof of a generalized Dynkin formula for $\varphi(X(\cdot;x,u(\cdot))$ with a suitably regular φ, consists in showing the following decomposition of $\mathcal{A}^{(k)}$ when acting on the function φ

$$\varphi \in \mathcal{D}\big(\mathcal{A}^{(0)}\big) \cap C_b^{1,G}(H) \Rightarrow \begin{cases} \varphi \in \mathcal{D}\big(\mathcal{A}^{(k)}\big) & \forall k \in K, \\ \mathcal{A}^{(k)}[\varphi] = \mathcal{A}^{(0)}[\varphi] + \langle D^G\varphi(\cdot), k\rangle_K. \end{cases} \tag{4.9}$$

Looking at $\{P_t^{(k)}\}_{t\geq 0}$ as to a perturbation of $\{P_t^{(0)}\}_{t\geq 0}$, (4.9) is obtained in [25], Theorem 5.2, in the context of C_0-semigroups with respect to *mixed topology* of $C_b(H)$ and in [15], Theorem 4.6, in the context of *bi-continuous* semigroups. However, these references would require that $\varphi \in C_b^1(H)$, that A, σ are such that $C_b^1(H) \subseteq \mathcal{D}(\mathcal{A}^{(0)})$ and that $G \in \mathcal{L}(H)$. This would allow, in particular, to write the term $\langle D^G\varphi(\cdot), k\rangle_K$ in the formula above as $\langle D\varphi(\cdot), Gk\rangle_H$, simplifying a lot the framework. Here, we need to be sharper in this respect in order to cover other cases of interest in applications, for example, the case of unbounded G, occurring in boundary control problems. To this purpose, we introduce the class of functions

$$\mathcal{S}^{A,G}(H) := \Big\{\varphi \in C_b^{1,G}(H) : \lim_{t\to 0^+} \frac{\varphi(z(t) + \int_0^t \overline{e^{sA}G}k\, ds) - \varphi(z(t))}{t} = \langle D^G\varphi\big(z(0)\big), k\rangle_K \ \forall z \in C\big(\mathbb{R}^+; H\big)\Big\}. \tag{4.10}$$

Our generalized Dynkin formula will hold for functions in $\mathcal{D}(\mathcal{A}^{(0)}) \cap \mathcal{S}^{A,G}(H)$. In Appendix 6, we provide sufficient conditions on A, G, φ ensuring that $\varphi \in \mathcal{S}^{A,G}(H)$.

PROPOSITION 4.4. *Let $\varphi \in \mathcal{D}(\mathcal{A}^{(0)}) \cap \mathcal{S}^{A,G}(H)$. Then* (4.9) *holds.*

PROOF. Since $\varphi \in \mathcal{D}(\mathcal{A}^{(0)})$, we can write for every $x \in H$

$$\begin{aligned}
\mathcal{A}^{(k)}[\varphi](x) &= \lim_{t\to 0^+} \frac{P_t^{(k)}[\varphi](x) - \varphi(x)}{t} \\
&= \lim_{t\to 0^+} \frac{P_t^{(k)}[\varphi](x) - P_t^{(0)}[\varphi](x)}{t} + \lim_{t\to 0^+} \frac{P_t^{(0)}[\varphi](x) - \varphi(x)}{t} \\
&= \lim_{t\to 0^+} \frac{\mathbb{E}[\varphi(X^{(k)}(t;x)) - \varphi(X^{(0)}(t,x))]}{t} + \mathcal{A}^{(0)}[\varphi](x),
\end{aligned}$$

if the last limit exists. Observe that

$$X^{(k)}(t;x) = X^{(0)}(t;x) + \int_0^t \overline{e^{(t-s)A}G}k\,ds = X^{(0)}(t;x) + \int_0^t \overline{e^{sA}G}k\,ds.$$

Therefore, since $\varphi \in \mathcal{S}^{A,G}(H)$, continuity of $t \mapsto \int_0^t \overline{e^{sA}G}k\,ds$ and dominated convergence yield

$$\begin{aligned}
&\lim_{t\to 0^+} \frac{\mathbb{E}[\varphi(X^{(k)}(t;x)) - \varphi(X^{(0)}(t;x))]}{t} \\
&\quad = \lim_{t\to 0^+} \mathbb{E}\left[\frac{\varphi(X^{(0)}(t;x) + \int_0^t \overline{e^{sA}G}k\,ds) - \varphi(X^{(0)}(t;x))}{t}\right] \\
&\quad = \mathbb{E}\left[\lim_{t\to 0^+} \frac{\varphi(X^{(0)}(t;x) + \int_0^t \overline{e^{sA}G}k\,ds) - \varphi(X^{(0)}(t;x))}{t}\right] = \langle D^G\varphi(x), k\rangle_K.
\end{aligned}$$

The claim follows. □

4.2. *Proof of the generalized Dynkin's formula.* We introduce the linear space $\mathcal{K}^{s,p}$ of K-valued p-integrable càdlàg simple processes. An element $\kappa(\cdot) \in \mathcal{K}^{s,p}$ is of the form

$$\kappa(t) = \sum_{i=1}^n k_{i-1}\mathbf{1}_{[t_{i-1},t_i)}(t), \tag{4.11}$$

for some $n \in \mathbb{N}$, $0 = t_0 < t_1 < \cdots < t_n = +\infty$, and $\{k_i\}_{i=0,\dots,n-1}$ such that $k_i \in L^p(\Omega, \mathcal{F}_{t_i}, \mathbb{P}; K)$ for all $i = 0, \dots, n-1$. Processes in $\mathcal{K}^{s,p}$ are progressively measurable. By arguing as in the proof of Proposition 3.4, we get that, for any $\kappa(\cdot) \in \mathcal{K}^{s,p}$, the process

$$t \mapsto \int_0^t \overline{e^{(t-s)A}G}\kappa(s)\,ds$$

is well-defined, belongs to $\mathcal{M}^{1,\mathrm{loc}}_{\mathcal{P}}(H)$ and has a version with continuous trajectories. We will always refer to the version of this process (unique up to indistinguishability) having continuous trajectories. Given $\kappa(\cdot) \in \mathcal{K}^{s,p}$, we write

$$X^{\kappa(\cdot)}(t,x) := e^{tA}x + \int_0^t e^{(t-s)A}\sigma\, dW(s) + \int_0^t \overline{e^{(t-s)A}G}\kappa(s)\, ds.$$

Again arguing as in the proof of Proposition 3.4 we see that this process has a version with continuous trajectories. As above, we will always refer to this version (unique up to indistinguishability).

Recall that, if V_1, V_2 are two random variables with values, respectively, in two measurable spaces $(E_1, \mathcal{E}_1)$ and $(E_2, \mathcal{E}_2)$, a version of the conditional law of V_1 given V_2 is a family of probability measures $\{\mu(\cdot, v_2)\}_{v_2 \in E_2}$ on $(E_1, \mathcal{E}_1)$ such that, for every $f \in B_b(E_1 \times E_2; \mathbb{R})$, the map $v_2 \mapsto \int_{E_1} f(v_1, v_2)\mu(dv_1, v_2)$ is measurable and

$$\mathbb{E}\big[f(V_1, V_2)\big] = \int_{E_2} \nu(dv_2) \int_{E_1} f(v_1, v_2)\mu(dv_1, v_2),$$

where $\nu = \mathrm{Law}(V_2)$. This family, if it exists, is unique up to ν-null measure sets.

LEMMA 4.5. *Let $\kappa(\cdot) \in \mathcal{K}^{s,p}$ be in the form (4.11) and $t \in [t_{i-1}, t_i)$ for some $i = 1, \dots, n$. A version of the conditional law of $X^{\kappa(\cdot)}(t; x)$ given the couple $(X^{\kappa(\cdot)}(t_{i-1}; x), k_{i-1})$ is the family*

$$\mu_t(\cdot, x', k') := \mathrm{Law}\big(X^{(k')}(t - t_{i-1}; x')\big). \tag{4.12}$$

PROOF. The proof is standard (see [36], Chapter 2, Section 9, in finite dimension and in a much more general setting) and we omit it for brevity. □

LEMMA 4.6. *Let $\varphi \in \mathcal{D}(\mathcal{A}^{(0)}) \cap \mathcal{S}^{A,G}(H)$ and $\kappa(\cdot) \in \mathcal{K}^{s,p}$. Then*

$$\begin{aligned} &\frac{d}{dt}\mathbb{E}\big[\varphi\big(X^{\kappa(\cdot)}(t; x)\big)\big] \\ &\quad = \mathbb{E}\big[\mathcal{A}^{(0)}[\varphi]\big(X^{\kappa(\cdot)}(t; x)\big) + \big\langle \kappa(t), D^G\varphi\big(X^{\kappa(\cdot)}(t; x)\big)\big\rangle_K\big] \qquad \forall t \geq 0, \end{aligned} \tag{4.13}$$

where the derivative has to be intended as right derivative at the times $\{t_1, \dots, t_n\}$, where the simple process $\kappa(\cdot)$ jumps.

PROOF. Let $\kappa(\cdot) \in \mathcal{K}^{s,p}$ be as in (4.11), $t \in [t_{i-1}, t_i)$ for some $i = 1, \dots, n$, and $\varphi \in \mathcal{D}(\mathcal{A}^{(0)}) \cap C^{1,G}_b(H)$. Denote by ν the law of the couple $(X^{\kappa(\cdot)}(t_{i-1}; x), k_{i-1})$. By Lemma 4.5, we have

$$\begin{aligned} \mathbb{E}\big[\varphi\big(X^{\kappa(\cdot)}(t; x)\big)\big] &= \int_{H \times K} \nu(dx', dk') \int_H \varphi(\xi)\mu_t(d\xi; x', k') \\ &= \int_{H \times K} \nu(dx', dk')\mathbb{E}\big[\varphi\big(X^{(k')}(t - t_{i-1}; x')\big)\big] \\ &= \int_{H \times K} \nu(dx', dk') P^{(k')}_{t - t_{i-1}}[\varphi](x'). \end{aligned}$$

Now we differentiate under the integral sign using the fact that, by Proposition 4.4, $\varphi \in \mathcal{D}(\mathcal{A}^{(k')})$ and the fact that $(t, x') \mapsto P^{(k')}_{t-t_{i-1}}[\mathcal{A}^{(k')}[\varphi]](x')$ is bounded over $[t_{i-1}, t_i) \times H$. Then, using Propositions 4.3(i) and 4.4, we get

$$\begin{aligned}\frac{d}{dt}\mathbb{E}[\varphi(X^{\kappa(\cdot)}(t;x))] &= \int_{H\times K} \nu(dx', dk')\frac{d}{dt}P^{(k')}_{t-t_{i-1}}[\varphi](x') \\ &= \int_{H\times K} \nu(dx', dk')P^{(k')}_{t-t_{i-1}}[\mathcal{A}^{(k')}[\varphi]](x') \\ &= \int_{H\times K} \nu(dx', dk')\mathbb{E}[\mathcal{A}^{(k')}[\varphi](X^{(k')}(t-t_{i-1});x)] \\ &= \int_{H\times K} \nu(dx', dk')\int_H \mathcal{A}^{(k')}[\varphi](\xi)\mu_t(d\xi;x',k') \\ &= \mathbb{E}[\mathcal{A}^{(k_{i-1})}[\varphi](X^{\kappa(\cdot)}(t;x))] \\ &= \mathbb{E}[\mathcal{A}^{(0)}[\varphi](X^{\kappa(\cdot)}(t;x)) + \langle \kappa(t), D^G\varphi(X^{\kappa(\cdot)}(t;x))\rangle_K],\end{aligned}$$

the claim. $\square$

LEMMA 4.7. *For each $u(\cdot) \in \mathcal{U}_p$ and $T > 0$, there exists a sequence $\{\kappa_n\}_{n\in\mathbb{N}} \subset \mathcal{K}^{s,p}$ such that*

$$\kappa_n|_{[0,T]\times\Omega} \xrightarrow{\mathcal{M}^{p,T}_{\mathcal{P}}(H)} L(u(\cdot))|_{[0,T]\times\Omega},$$

$$X^{\kappa_n(\cdot)}(\cdot;x)|_{[0,T]\times\Omega} \xrightarrow{\mathcal{M}^{1,T}_{\mathcal{P}}(H)} X(\cdot;x,u(\cdot))|_{[0,T]\times\Omega}.$$

PROOF. Fix $T > 0$ and set $\kappa(\cdot) := L(u(\cdot))$. By standard arguments (see, e.g., [34], Chapter III, Lemma. 2.4, p.132),[7] we can construct a sequence $\{\kappa_n\}_{n\in\mathbb{N}} \subset \mathcal{K}^{s,p}$ such that

$$\kappa_n|_{[0,T]\times\Omega} \xrightarrow{\mathcal{M}^{p,T}_{\mathcal{P}}(H)} \kappa(\cdot)|_{[0,T]\times\Omega}.$$

Then, using the expression (3.6) for the state variable, the convergence

$$X^{\kappa_n(\cdot)}(\cdot;x)|_{[0,T]\times\Omega} \xrightarrow{\mathcal{M}^{1,T}_{\mathcal{P}}(H)} X(\cdot;x,u(\cdot))|_{[0,T]\times\Omega}$$

follows by simply applying dominated convergence. $\square$

[7] It is worth to point out some differences. First, we are dealing with càdlàg approximations (as it is more meaningful and natural to state Proposition 4.6) rather than with càdlàg (as in [34], Chapter III, Lemma. 2.4, p.132): this is not a problem as, from the point of view of integration, these classes coincide. Second, we are dealing with Hilbert-valued processes: therefore, more technical care is needed as the approximation is produced by Bochner integration.

THEOREM 4.8 (Dynkin's formula). *Let* $\varphi \in \mathcal{D}(\mathcal{A}^{(0)}) \cap \mathcal{S}^{A,G}(H)$. *Then, for every* $\lambda > 0$, $T > 0$, *and* $u(\cdot) \in \mathcal{U}_p$, *we have*

$$\mathbb{E}\big[e^{-\lambda T}\varphi\big(X\big(T;x,u(\cdot)\big)\big)\big]$$

$$= \varphi(x) + \mathbb{E}\bigg[\int_0^T e^{-\lambda t}\big[\big(\mathcal{A}^{(0)} - \lambda\big)[\varphi]\big(X\big(t;x,u(\cdot)\big)\big) + \big\langle L\big(u(t)\big), D^G\varphi\big(X\big(t;x,u(\cdot)\big)\big)\big\rangle_K\big]\,dt\bigg]. \tag{4.14}$$

PROOF. Let $u(\cdot) \in \mathcal{U}_p$ and take the approximating sequence $\{\kappa_n\}_{n\in\mathbb{N}}$ provided by Lemma 4.7. Then, applying, for each $n \in \mathbb{N}$, Lemma 4.6, we obtain from (4.13) (by taking the right derivatives at t_i), for all $t \geq 0$ and $\lambda > 0$,

$$\frac{d}{dt}e^{-\lambda t}\mathbb{E}\big[\varphi\big(X^{\kappa_n(\cdot)}(t;x)\big)\big]$$

$$= -\lambda e^{-\lambda t}\mathbb{E}\big[\varphi\big(X^{\kappa_n(\cdot)}(t;x)\big)\big] + e^{-\lambda t}\mathbb{E}\big[\mathcal{A}^{(0)}[\varphi]\big(X^{\kappa_n(\cdot)}(t;x)\big) + \big\langle\kappa_n(t), D^G\varphi\big(X^{\kappa_n(\cdot)}(t;x)\big)\big\rangle_K\big]. \tag{4.15}$$

Since the function $t \mapsto \mathbb{E}[e^{-\lambda t}\varphi(X^{\kappa_n(\cdot)}(t;x))]$ is everywhere continuous and stepwise differentiable, we can apply the fundamental theorem of calculus. So, integrating on $[0, T]$, we get

$$\mathbb{E}\big[e^{-\lambda T}\varphi\big(X^{\kappa_n(\cdot)}(T;x)\big)\big]$$

$$= \varphi(x) + \mathbb{E}\bigg[\int_0^T e^{-\lambda t}\big(\big(\mathcal{A}^{(0)} - \lambda\big)[\varphi]\big(X^{\kappa_n(\cdot)}(t;x)\big) + \big\langle\kappa_n(t), D^G\varphi\big(X^{\kappa_n(\cdot)}(t;x)\big)\big\rangle_K\big)\,dt\bigg].$$

Now, letting $n \to +\infty$, we get the claim by dominated convergence from Lemma 4.7, observing that φ, $D^G\varphi$, and $\mathcal{A}^{(0)}[\varphi]$ are bounded. $\square$

REMARK 4.9. The results of this section, in particular Theorem 4.8, can be extended, at the price of straightforward technical complications, to the case when the basic space of functions is (see [16])

$$C_m(H) := \bigg\{\phi : H \to \mathbb{R} \text{ continuous} : \sup_{x\in H}\frac{|\phi(x)|}{1 + |x|^m} < \infty\bigg\}. \tag{4.16}$$

Also the results of next Section 5 can be extended to this setting covering more general cases, in particular when the current cost of the control problem has polynomial growth in x. We do not do this here for brevity.

5. HJB equation, verification theorem and optimal feedbacks. By standard dynamic programming arguments, one formally associates to the control problem of Section 3 the following HJB equation for the value function (3.10):

$$\lambda v(x) - \frac{1}{2}\operatorname{Tr}\big[QD^2v(x)\big] - \langle Ax, Dv(x)\rangle_H - F\big(x, Dv(x)\big) = 0, \qquad x \in H, \tag{5.1}$$

where $Q = \sigma\sigma^*$ and the Hamiltonian F is defined by

$$F(x, p) := \inf_{u \in \Lambda} F_{\mathrm{CV}}(x, p; u), \qquad x \in H, p \in H, \tag{5.2}$$

where

$$F_{\mathrm{CV}}(x, p; u) := \langle GL(u), p \rangle_H + l(x, u), \qquad x \in H, u \in \Lambda, p \in H. \tag{5.3}$$

Note that this definition is only formal as $GL(u)$ may be not defined, since $L(u)$ may not belong to $\mathcal{D}(G)$. It is then convenient to introduce the modified Hamiltonian

$$F_0(x, q) := \inf_{u \in \Lambda} F_{0,\mathrm{CV}}(x, q; u), \qquad x \in H, q \in K, \tag{5.4}$$

where

$$F_{0,\mathrm{CV}}(x, q; u) := \langle L(u), q \rangle_K + l(x, u), \qquad x \in H, u \in \Lambda, q \in K. \tag{5.5}$$

Observing that

$$F(x, p) = F_0(x, G^* p) \qquad \forall p \in \mathcal{D}(G^*),$$

(5.1) can be formally rewritten as

$$\lambda v(x) - \frac{1}{2} \mathrm{Tr}[Q D^2 v(x)] - \langle Ax, Dv(x) \rangle_H - F_0(x, D^G v(x)) = 0, \qquad x \in H. \tag{5.6}$$

Note that, in principle, F_0 may take the value $-\infty$ somewhere. The concept of mild solution to (5.6) relies on Proposition 4.3(ii) and on (4.7), inspiring an integral form of (5.6) through the use of the semigroup $\{P_s^{(0)}\}_{s \geq 0}$.

DEFINITION 5.1. We say that a function $v : H \to \mathbb{R}$ is a mild solution to (5.6) if $v \in C_b^{1,G}(H)$, $F_0(\cdot, D^G v(\cdot)) \in C_b(H)$, and

$$v(x) = \int_0^\infty e^{-\lambda s} P_s^{(0)} [F_0(\cdot, D^G v(\cdot))](x)\, ds \qquad \forall x \in H. \tag{5.7}$$

REMARK 5.2. The problem of existence and uniqueness of mild solutions for equations in the form (5.6) is addressed in [16] and in [14], Chapter 4. In particular, existence and uniqueness of mild solutions is stated for sufficiently large $\lambda > 0$, under the following assumptions (see [16], Corollary 4.12, Theorem 3.8(ii) with $m = 0$):

(A1) $\overline{e^{tA} G}(K) \subseteq Q_t^{1/2}(H)$ for every $t > 0$, where $Q_t := \int_0^t e^{sA} \sigma\sigma^* e^{sA^*}\, ds$.
(A2) The operators[8]

$$\Gamma_G(t) : K \to H, \qquad \Gamma_G(t) := Q_t^{-1/2} \overline{e^{tA} G}, \qquad t \geq 0,$$

[8] Here, $Q_t^{-1/2}$ is the pseudo-inverse of $Q_t^{1/2}$.

which are well-defined by (A1) and bounded by the closed graph theorem, are such that the map $t \mapsto |\Gamma_G(t)|_{\mathcal{L}(K,H)}$ belongs to $L^1_{\mathrm{loc}}([0,+\infty),\mathbb{R})$ and is bounded in a neighborhood of $+\infty$.

(A3) The Hamiltonian F_0 satisfies, for suitable $C_{F_0} > 0$.

$$\begin{aligned} \bigl|F_0(x,q_1) - F_0(x,q_2)\bigr| &\le C_{F_0}|q_1 - q_2| \qquad \forall x \in H, \forall q_1, q_2 \in K, \\ \bigl|F_0(x,q)\bigr| &\le C_{F_0}\bigl(1 + |q|_K\bigr) \qquad \forall x \in H, \forall q \in K. \end{aligned}$$

Some results in the case of locally Lipschitz–Hamiltonian are available, up to now, only in special cases (see [10], Section 13.3.1 and [5]).

Due to Proposition 4.3(ii), a mild solution v of (5.1) enjoys the property of being a solution to the same equation also in a differential abstract way, that is, we have the following.

PROPOSITION 5.3. *Let v be a mild solution to* (5.6). *Then $v \in \mathcal{D}(\mathcal{A}^{(0)})$ and*

$$(\lambda - \mathcal{A}^{(0)})[v](x) = F_0\bigl(x, D^G v(x)\bigr) \qquad \forall x \in H. \tag{5.8}$$

PROOF. Using Proposition 4.3(ii), we rewrite (5.7) as

$$v(x) = \bigl(\lambda - \mathcal{A}^{(0)}\bigr)^{-1}\bigl[F_0\bigl(\cdot, D^G v(\cdot)\bigr)\bigr](x) \qquad \forall x \in H. \tag{5.9}$$

This entails $v \in \mathcal{D}(\mathcal{A}^{(0)})$ and, applying $\lambda - \mathcal{A}^{(0)}$ to both sides, we see that v solves (5.8). $\square$

REMARK 5.4. By Proposition 5.3, a mild solution v to (5.6) belongs to $\mathcal{D}(\mathcal{A}^{(0)})$. Hence, in order to apply Theorem 4.8 to it, we only need to assume that $v \in \mathcal{S}^{A,G}(H)$. This is what we indeed assume in all the next results of this section.

5.1. *Verification theorem.* The proof of the verification theorem relies in the so called *fundamental identity*.

PROPOSITION 5.5 (Fundamental identity). *Let* (3.8) *hold. Let v be a mild solution to* (5.6) *and assume that $v \in \mathcal{S}^{A,G}(H)$. Let $x \in H$ and let $u(\cdot) \in \mathcal{U}_p$ be such that*

$$J\bigl(x; u(\cdot)\bigr) := \mathbb{E}\biggl[\int_0^\infty e^{-\lambda t} l\bigl(X\bigl(t; x, u(\cdot)\bigr), u(t)\bigr)\,dt\biggr] < \infty. \tag{5.10}$$

Then, setting $X(\cdot) := X(\cdot; x, u(\cdot))$, we have

$$\begin{aligned} v(x) &= J\bigl(x; u(\cdot)\bigr) \\ &\quad + \mathbb{E}\biggl[\int_0^\infty e^{-\lambda t}\bigl(F_0\bigl(X(t), D^G v\bigl(X(t)\bigr)\bigr) - F_{0,\mathrm{CV}}\bigl(X(t), D^G v\bigl(X(t)\bigr); u(t)\bigr)\bigr)\,dt\biggr]. \end{aligned} \tag{5.11}$$

PROOF. Let $x \in H$, $T > 0$, and let $u(\cdot) \in \mathcal{U}_p$ be such that (5.10) holds. Set $X(\cdot) := X(\cdot; x, u(\cdot))$. Using Proposition 5.3 and applying the abstract Dynkin formula (Theorem 4.8), we get

$$
\begin{aligned}
(5.12)\quad & \mathbb{E}[e^{-\lambda T} v(X(T))] \\
&= v(x) + \mathbb{E}\left[\int_0^T e^{-\lambda t}[(\mathcal{A}^{(0)} - \lambda)[v](X(t)) + \langle L(u(t)), D^G v(X(t))\rangle_K]\, dt\right] \\
&= v(x) + \mathbb{E}\left[\int_0^T e^{-\lambda t}[-F_0(X(t), D^G v(X(t))) + \langle L(u(t)), D^G v(X(t))\rangle_K]\, dt\right].
\end{aligned}
$$

By (3.8) and (5.10), we have

$$
\mathbb{E}\left[\int_0^T e^{-\lambda t} l(X(t), u(t))\, dt\right] \in \mathbb{R} \qquad \forall T > 0.
$$

Then we can add and subtract the latter finite value in (5.12) and use (5.5) to get, rearranging the terms,

$$
\begin{aligned}
(5.13)\quad & \mathbb{E}[e^{-\lambda T} v(X(T))] - v(x) + \mathbb{E}\left[\int_0^T e^{-\lambda t} l(X(t), u(t))\, dt\right] \\
&= \mathbb{E}\left[\int_0^T e^{-\lambda t}[-F_0(X(t), D^G v(X(t))) + F_{0,\mathrm{CV}}(X(t), D^G v(X(t)); u(t))]\, dt\right].
\end{aligned}
$$

Now we let $T \to +\infty$. The right-hand side has a limit (possibly $+\infty$), as the integrand is positive. The left-hand side clearly converges to $J(x; u(\cdot)) - v(x)$. This implies that also the limit of the right-hand side is finite and

$$
\begin{aligned}
& J(x; u(\cdot)) - v(x) \\
&= \mathbb{E}\left[\int_0^\infty e^{-\lambda t}[-F_0(X(t), D^G v(X(t))) F_{0,\mathrm{CV}}(X(t), D^G v(X(t)); u(t))]\, dt\right].
\end{aligned}
$$

The claim follows rearranging the terms. □

THEOREM 5.6 (Verification theorem). *Let (3.8) hold. Let v be a mild solution to (5.6) and assume that $v \in \mathcal{S}^{A,G}(H)$. We have the following*:

(i) *$v \le V$ over H.*

(ii) *Let $x \in H$ and let $u^*(\cdot) \in \mathcal{U}_p$ be such that, setting $X^*(\cdot) := X(\cdot; x, u^*(\cdot))$, it is for a.e. $t \ge 0$ (almost surely)*

$$
F_0(X^*(t), D^G v(X^*(t))) = F_{0,\mathrm{CV}}(X^*(t), D^G v(X^*(t)); u^*(t)). \tag{5.14}
$$

Then $v(x) = V(x) = J(x; u^*(\cdot))$.

PROOF. (i) By (5.11), for all $u(\cdot) \in \mathcal{U}_p$ such that (5.10) holds, we have $v(x) \le J(x; u(\cdot))$, which yields this claim.

(ii) Let $u^*(\cdot)$ be such that (5.14) holds. If $J(x; u^*(\cdot)) < +\infty$, then, from (5.11), we immediately get $v(x) = J(x; u^*(\cdot))$, which, combined with item (i), yields the claim. We now prove that it cannot be $J(x; u^*(\cdot)) = +\infty$. Assume, by contradiction, that $J(x; u^*(\cdot)) = +\infty$. By (5.14), we have for a.e. $t \ge 0$ (almost sulrely)

$$l\big(X^*(t), u^*(t)\big) = F_0\big(X^*(t), D^G v\big(X^*(t)\big)\big) - \big\langle L\big(u^*(t)\big), D^G v\big(X^*(t)\big)\big\rangle_K. \tag{5.15}$$

By (5.8), $F_0(\cdot, D^G v(\cdot))$ is bounded. Hence, Assumption 3.1(iv), the fact that $u^*(\cdot) \in \mathcal{U}_p$ and (5.15) imply $\mathbb{E}[\int_0^T e^{-\lambda t} l(X^*(t), u^*(t))\, dt] < \infty$ for all $T > 0$. Then, we can argue as in the proof of Proposition 5.5 getting (5.13) with $u^*(\cdot)$ in this case and, using again (5.14),

$$\mathbb{E}\big[e^{-\lambda T} v\big(X^*(T)\big)\big] - v(x) + \mathbb{E}\biggl[\int_0^T e^{-\lambda t} l\big(X^*(t), u^*(t)\big)\, dt\biggr] = 0.$$

Letting $T \to +\infty$ we get $v(x) = J(x; u^*(\cdot)) = +\infty$, leading to a contradiction as v is finite. $\square$

5.2. *Optimal feedback controls.* As usual, the verification theorem is composed of two statements: the first one states that the solution to the HJB equation enjoys the property of being smaller than the value function; the second one is the most important from the point of view of the control problem, as it furnishes a sufficient condition of optimality [(5.14) in our case]. Then the problem becomes the so-called synthesis of an optimal control, that is, to produce a control $u^*(\cdot)$ verifying such condition. The answer relies in the study of the *closed loop equation*.

Let v be a mild solution to HJB equation (5.6). Assuming that the infimum of the map

$$\Lambda \to \mathbb{R}, \qquad u \mapsto F_{0,\mathrm{CV}}\big(x, D^G v(x); u\big)$$

is attained and defining the multivalued function (*feedback map*)

$$\begin{aligned} \Phi \colon H &\longrightarrow 2^{\Lambda}, \\ x &\longmapsto \operatorname*{arg\,min}_{u \in \Lambda} F_{0,\mathrm{CV}}\big(x, D^G v(x); u\big), \end{aligned} \tag{5.16}$$

the closed loop equation (CLE) associated with our problem and to v is indeed a stochastic differential inclusion:

$$dX(s) \in \big[AX(s) + GL\big(\Phi\big(X(s)\big)\big)\big]\, ds + \sigma\, dW(s). \tag{5.17}$$

We have the following result.

COROLLARY 5.7. *Let* (3.8) *hold. Let* v *be a mild solution to* (5.6) *and assume that* $v \in \mathcal{S}^{A,G}(H)$. *Let* $x \in H$ *and assume that the feedback map* Φ *defined in* (5.16) *admits a measurable selection* $\phi : H \to U$ *and consider the SDE*:

$$\begin{cases} dX(s) = [AX(s) + GL(\phi(X(s)))]\,ds + \sigma\,dW(s), \\ X(0) = x. \end{cases} \tag{5.18}$$

Assume that (5.18) *has a mild solution in* $\mathcal{M}^{1,\mathrm{loc}}_{\mathcal{P}}(U)$, *that is, there exists* $X_\phi(s;x) \in \mathcal{M}^{1,\mathrm{loc}}_{\mathcal{P}}(U)$ *such that*

$$\begin{aligned} X_\phi(t;x) := {} & e^{tA}x + \int_0^t e^{(t-s)A}\sigma\,dW(s) \\ & + \int_0^t \overline{e^{(t-s)A}G}L(\phi(X_\phi(s;x)))\,ds \qquad \forall t \ge 0. \end{aligned} \tag{5.19}$$

Define, for $s \ge 0$, $u_\phi(s) := \phi(X_\phi(s;x))$ *and assume that* $u_\phi(\cdot) \in \mathcal{U}_p$. *Then* $v(x) = V(x) = J(x; u_\phi(\cdot))$. *In particular, the couple* $(u_\phi(\cdot), X_\phi(\cdot;x))$ *is optimal at* x.

Moreover, if $\Phi(x)$ *is single-valued and the mild solution to* (5.18) *is unique, then the optimal control is unique.*

PROOF. Consider the couple $(u_\phi(\cdot), X_\phi(\cdot))$ and observe that $X_\phi(\cdot)$ is the unique mild solution (in the strong probabilistic sense) of the state equation associated to the control $u_\phi(\cdot)$, so that $X_\phi(\cdot;x) \equiv X(\cdot; x, u_\phi(\cdot))$. By construction such couple satisfies (5.14). Then, by Theorem 5.6(ii) we obtain that it is optimal.

Let us address now the uniqueness issue. We observe that, if $(\hat{u}(\cdot), X(\cdot; x, \hat{u}(\cdot)))$ is another optimal couple at x, we immediately have, by (5.11) and the fact that $v(x) = V(x)$,

$$\begin{aligned} \mathbb{E}\bigg[\int_0^\infty e^{-\lambda s}\big[& F_0(X(s;x,\hat{u}(\cdot)), D^G v(X(s;x,\hat{u}(\cdot)))) \\ & - F_{0,\mathrm{CV}}(X(s;x,\hat{u}(\cdot)), D^G v(X(s;x,\hat{u}(\cdot))); \hat{u}(s))\big]\,ds\bigg] = 0. \end{aligned}$$

As the integrand is always negative and as Φ is single-valued, this implies that $\mathbb{P} \times ds$-a.e. we have $\hat{u}(\cdot) = \Phi(X(\cdot; x, \hat{u}(\cdot)))$. This shows that $X(\cdot; x, \hat{u}(\cdot))$ solves (5.18). Then uniqueness of mild solutions to (5.18) gives the claim. □

We conclude the section commenting on the extension of our results to the case when the control problem is considered in the so-called *weak formulation*. So far, we have considered our family of stochastic optimal control problems in the *strong formulation*. It is possible to consider the problem also in the so-called weak formulation, that is, letting the filtered probability space and the cylindrical Brownian motion vary with the control strategy $u(\cdot)$ (see, e.g., [48], Chapter 2). More precisely, in the weak formulation, the control strategy is a 6-tuple

$(\overline{\Omega}, \overline{\mathcal{F}}, \{\overline{\mathcal{F}_t}\}_{t\geq 0}, \overline{\mathbb{P}}, \overline{W}, \overline{u}(\cdot))$. Calling $\overline{\mathcal{U}_p}$ the set such control strategies, the objective is to minimize the cost (3.9) over $\overline{\mathcal{U}_p}$. The resulting value function $\overline{V}$ is, in principle, smaller than V. The main advantage in choosing such formulation is that existence of optimal control strategies in feedback form is easier to obtain. The verification theorem above also holds when we consider the control problem in its weak formulation. Indeed, the proof of Theorem 5.6 works for every filtered probability space and any cylindrical Brownian motion on it. Hence, letting the filtered probability space and the cylindrical Brownian motion vary, one gets that $v \leq \overline{V}$ over H. Moreover, if (5.14) holds for a given control strategy[9] $\overline{u^*}(\cdot) \in \overline{\mathcal{U}_p}$, then we have $v(x) = \overline{V}(x) = J(x; \overline{u^*}(\cdot))$. One gets the following.

COROLLARY 5.8. *Let* (3.8) *hold. Let* v *be a mild solution to* (5.6) *and assume that* $v \in \mathcal{S}^{A,G}(H)$. *Let* $x \in H$ *and assume that the feedback map* Φ *defined in* (5.16) *admits a measurable selection* $\phi : H \to U$. *Assume that* (5.18) *has a* martingale solution[10] (*see* [10], *p.* 220, *or* [23], *Definition* 3.1, *p.* 75, *for the definition*) $\overline{X_\phi}(\cdot; x)$ *in some filtered probability space* $(\overline{\Omega}, \overline{\mathcal{F}}, \{\overline{\mathcal{F}_t}\}_{t\geq 0}, \overline{\mathbb{P}})$ *and for some* Ξ*-valued cylindrical Brownian motion* $\overline{W}$ *defined on it. Define, for* $s \geq 0$, $\overline{u_\phi}(s) = \phi(\overline{X_\phi}(s; x))$ *and assume* $\overline{u_\phi}(\cdot) \in \overline{\mathcal{U}_p}$.[11] *Then* $v(x) = \overline{V}(x) = J(x; \overline{u_\phi}(\cdot))$. *In particular* $(\overline{u_\phi}(\cdot), \overline{X_\phi}(\cdot; x))$ *is an optimal couple.*

6. Applications. In the present section, we provide two examples of application of our results.

The first example, fully developed, concerns the optimal control of the stochastic heat equation in a given space region $\mathcal{O} \subseteq \mathbb{R}^d$ when the control can be exercised only at the boundary $\partial\mathcal{O}$. Precisely, we consider the case when the control at the boundary enters through a Neumann-type boundary condition, corresponding to control the heat flow at the boundary. The existence and uniqueness of mild solutions to the associated elliptic HJB equation in this case is guaranteed (under suitable conditions) by the results of [16].

The second example concerns the optimal control of a stochastic differential equation with delay in the control process (see [29, 30] for the treatment of the same problem over finite horizon). In this case, the result we give needs to *assume* the existence of a mild solution to the associated elliptic HJB equation. The reason for that is that a theory of mild solutions for elliptic HJB equations associated to this kind problem has not been yet developed in the elliptic case. Indeed, unlike the first example, this kind of equations is not covered by the results of [16], due to the lack of G-smoothing. In this case, it is needed an *ad hoc* treatment of the equation,

[9]Elements of $\overline{\mathcal{U}_p}$ are, rigorously speaking, 6-tuples; however, for simplicity, we denote them simply by $\overline{u(\cdot)}$.
[10]Weak-mild solution in the terminology of [14].
[11]In the sense that the 6-tuple identified by u_ϕ belongs to $\overline{\mathcal{U}_p}$.

dealing with the specific case at hand, to show the existence of mild solutions (see, e.g., the aforementioned references [29, 30] in the parabolic case). Although a result of this kind for elliptic equation seems straightforward, a rigorous statement of this result has not been rigourously fixed yet. For this reason, we limit ourselves to provide a weaker result taking the existence of mild solutions to the associated HJB equation as an assumption and leaving the investigation of that for future work. Due to the lack of a rigourous background on which relying our results, we do not state in this case a theorem and just keep the arguments at the level of an informal exposition.

6.1. *Neumann boundary control of a stochastic heat equation with additive noise.* We consider the optimal control of a nonlinear stochastic heat equation in a given space region $\mathcal{O} \subseteq \mathbb{R}^d$ when the control can be exercised only at the boundary of $\mathcal{O}$.

6.1.1. *Problem setup.* Let $\mathcal{O}$ be an open, connected, bounded subset of $\mathbb{R}^d$ with regular (in the sense of [37], Section 6) boundary $\partial\mathcal{O}$.[12] We consider the controlled dynamical system driven by the following SPDE in the time interval $[0, +\infty)$:

$$\tag{6.1} \begin{cases} \dfrac{\partial y(t,\xi)}{\partial t} = \Delta y(t,\xi) + \sigma \dot{W}(t,\xi), & (t,\xi) \in [0,+\infty) \times \mathcal{O}, \\ y(0,\xi) = x(\xi), & \xi \in \overline{\mathcal{O}}, \\ \dfrac{\partial y(t,\xi)}{\partial n} = \gamma_0(t,\xi), & (t,\xi) \in [0,+\infty) \times \partial\mathcal{O}, \end{cases}$$

where:

- $y : [0,+\infty) \times \mathcal{O} \times \Omega \to \mathbb{R}$ is the stochastic process describing the evolution of the temperature distribution and is the *state variable* of the system;
- $\gamma_0 : [0,+\infty) \times \partial\mathcal{O} \times \Omega \to \mathbb{R}$ is the stochastic process representing the heat flow at the boundary; it is the *control variable* of the system and acts at the boundary of it: this is the reason of the terminology "boundary control";
- n is the outward unit normal vector at the boundary $\partial\mathcal{O}$;
- $x \in L^2(\mathcal{O})$ is the initial state (initial temperature distribution) in the region $\mathcal{O}$;
- W is a cylindrical Wiener process in $L^2(\mathcal{O})$;
- $\sigma \in \mathcal{L}(L^2(\mathcal{O}))$.

Assume that this equation is well posed (in some suitable sense, see below for the precise setting) for every given $\gamma_0(\cdot,\cdot)$ in a suitable set of admissible control processes and denote its unique solution by $y^{x,\gamma_0(\cdot,\cdot)}$ to underline the dependence

[12]We stress that such conditions may allow corners in the boundary: in particular, when $d = 2$ squares satisfy the required regularity.

of the state y on the control $\gamma_0(\cdot,\cdot)$ and on the initial datum x. The controller aims at minimizing, over the set of admissible controls, the objective functional

$$\mathbb{E}\left[\int_0^\infty e^{-\lambda t}\left(\int_{\mathcal{O}} \ell_1\big(y^{x,\gamma_0(\cdot,\cdot)}(t,\xi)\big)\,d\xi + \int_{\partial\mathcal{O}} \ell_2\big(\gamma_0(t,\xi)\big)\,d\xi\right)dt\right], \tag{6.2}$$

where $\ell_1, \ell_2 : \mathbb{R} \to \mathbb{R}$ are given measurable functions bounded from below and $\lambda > 0$ is a discount factor.

6.1.2. *Infinite dimensional setting*. We now rewrite the state equation (6.1) and the functional (6.2) in an infinite dimensional setting in the space $H := L^2(\mathcal{O})$. For more details, we refer to [16], Section 5, and references therein. Consider the realization of the Laplace operator with vanishing Neumann boundary conditions:[13]

$$\begin{cases} \mathcal{D}(A_N) := \left\{\phi \in W^{2,2}(\mathcal{O};\mathbb{R}) : \dfrac{\partial \phi}{\partial n} = 0 \text{ on } \partial\mathcal{O}\right\}, \\ A_N\phi := \Delta\phi \qquad \forall \phi \in \mathcal{D}(A_N). \end{cases} \tag{6.3}$$

It is well known (see, e.g., [39], Chapter 3) that A_N generates a strongly continuous analytic semigroup $\{e^{tA_N}\}_{t\geq 0}$ in H. Moreover, A_N is a self-adjoint and dissipative operator. In particular $(0,+\infty) \subset \varrho(A_N)$, where $\varrho(A_N)$ denotes the resolvent set of A_N. So, if $\delta > 0$, then $(\delta I - A_N)$ is invertible and $(\delta I - A_N)^{-1} \in \mathcal{L}(H)$. Moreover (see, e.g., [37], Appendix B) the operator $(\delta I - A_N)^{-1}$ is compact. Consequently, there exists an orthonormal complete sequence $\{e_k\}_{k\in\mathbb{N}}$ such that the operator A_N is diagonal with respect to it:

$$A_N e_k = -\mu_k e_k, \qquad k \in \mathbb{N}, \tag{6.4}$$

for a suitable sequence of eigenvalues $\{\mu_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}^+$ repeated according to their multiplicity (they are nonnegative due to dissipativity of A_N). We assume that such sequence is increasingly ordered. Then, $\mu_0 = 0$, as clearly the constant functions belong to Ker (A_N), and $\mu_k > 0$ for each $k \in \mathbb{N}_0 := \mathbb{N}\setminus\{0\}$, since, as an immediate consequence of the Gauss–Green formula, only the constant functions belong to Ker (A_N). Moreover, [46], Section 5.6.2, p. 395 (see also [37], Appendix B) provides also a growth rate for the sequence of eigenvalues; indeed

$$\mu_k \sim k^{2/d}. \tag{6.5}$$

We have (see, e.g., [37], Appendix B) the isomorphic identification

$$\mathcal{D}\big((\delta I - A_N)^\alpha\big) = H^{2\alpha}(\mathcal{O}) \qquad \forall \alpha \in \left(0, \frac{3}{4}\right), \forall \delta > 0, \tag{6.6}$$

[13] To be precise, $\mathcal{D}(A_N)$ is the closure in $H^2(\mathcal{O})$ of the set of functions $\phi \in C^2(\overline{\mathcal{O}})$ having vanishing normal derivative at the boundary $\partial\mathcal{O}$.

where $H^s(\mathcal{O})$ denotes the Sobolev space of exponent $s \in \mathbb{R}$. Next, consider the following problem with Neumann boundary condition:

$$\begin{cases} \Delta w(\xi) = \delta w(\xi), & \xi \in \mathcal{O}, \\ \dfrac{\partial w}{\partial n}(\xi) = \alpha(\xi), & \xi \in \partial\mathcal{O}. \end{cases} \tag{6.7}$$

Given any $\delta > 0$ and $\alpha \in L^2(\partial\mathcal{O})$, there exists a unique solution $N_\delta \alpha \in H^{3/2}(\mathcal{O})$ to (6.7). Moreover, the operator (*Neumann map*)

$$N_\delta : L^2(\partial\mathcal{O}) \to H^{3/2}(\mathcal{O}), \tag{6.8}$$

is continuous (see [38], Theorem 7.4). So, in view of (6.6), the map

$$N_\delta : L^2(\partial\mathcal{O}) \to \mathcal{D}\big((\delta I - A_N)^{\frac{3}{4}-\varepsilon}\big), \qquad \varepsilon \in (0, 3/4), \tag{6.9}$$

is continuous. In [16], Section 5, it is shown that the natural abstract reformulation of the original control problem in the space H is

$$\begin{cases} dX(t) = \big[A_N X(t) + G_N^{\delta,\varepsilon} L_N^{\delta,\varepsilon} \gamma(t)\big]\, dt + \sigma\, dW(t), \\ X(0) = x, \end{cases} \tag{6.10}$$

where $L_N^{\delta,\varepsilon} := (\delta I - A_N)^{\frac{3}{4}-\varepsilon} N_\delta \in \mathcal{L}(L^2(\partial\mathcal{O}); H)$, $G_N^{\delta,\varepsilon} := (\delta I - A_N)^{\frac{1}{4}+\varepsilon}$, and $u(t) := \gamma_0(t, \cdot) \in L^2(\partial\mathcal{O})$ for $t \geq 0$. We are now in the framework of (3.1), with $K = H$, $A = A_N$, $G = G_N^{\delta,\varepsilon}$, $L = L_N^{\delta,\varepsilon}$, and $U = L^2(\partial\mathcal{O})$. Let us consider, as set of admissible controls,

$$\mathcal{U}_p := \Big\{ u : [0, +\infty) \times \Omega \to \Lambda : \ u(\cdot) \text{ is } \{\mathscr{F}_t\}_{t \geq 0}\text{-prog. meas. and s.t.} \\ \int_0^t \mathbb{E}\big[|u(s)|^p_{L_2(\partial\mathcal{O})}\big]\, ds < \infty \ \forall t \geq 0 \Big\},$$

where $\Lambda \subseteq L^2(\partial\mathcal{O})$ and p will be specified later according to (3.4). Defining

$$l_1(x) := \int_{\mathcal{O}} \ell_1\big(x(\xi)\big)\, d\xi, \qquad l_2(u) := \int_{\partial\mathcal{O}} \ell_2\big(u(\xi)\big)\, d\xi,$$

and

$$l \colon H \times \Lambda \to \mathbb{R}, \qquad l(x, u) := l_1(x) + l_2(u),$$

the functional (6.2) can be rewritten in the Hilbert space framework as

$$J\big(x; u(\cdot)\big) := \mathbb{E}\Big[\int_0^\infty e^{-\lambda t} l\big(X\big(t; x, u(\cdot)\big), u(t)\big)\, dt\Big]. \tag{6.11}$$

6.1.3. *HJB equation and verification theorem.* Setting $Q := \sigma\sigma^*$, the HJB equation associated to the minimization of (6.11) is

$$\begin{aligned}\lambda v(x) - \frac{1}{2}\operatorname{Tr}\big[QD^2 v(x)\big] - \langle A_N x, Dv(x)\rangle_H - l_1(x)\\ - \inf_{u\in\Lambda}\big\{\langle L_N^{\delta,\varepsilon} u, D^{G_N^{\delta,\varepsilon}} v(x)\rangle_H + l_2(u)\big\} = 0.\end{aligned} \tag{6.12}$$

Since the semigroup $\{e^{tA_N}\}_{t\geq 0}$ is strongly continuous and analytic, then by [42], Theorem 6.13(c), $e^{tA_N} G_N^{\delta,\varepsilon}$ can be extended to $\overline{e^{tA_N} G_N^{\delta,\varepsilon}} = G_N^{\delta,\varepsilon} e^{tA_N} \in \mathcal{L}(H)$ for every $t > 0$ and

$$\big|\overline{e^{tA_N} G_N^{\delta,\varepsilon}}\big|_{\mathcal{L}(H)} \leq C t^{\frac{1}{4}+\varepsilon} \qquad \forall t > 0. \tag{6.13}$$

Hence, Assumption 3.1(i) and (iii) is satisfied with $A = A_N$, $G = G_N^{\delta,\varepsilon}$, and $\beta = \varepsilon + 1/4$. Consequently, recalling (3.4), we choose $p > \frac{1}{\frac{3}{4}-\varepsilon}$.

Now, assume the following:

(H1) σ satisfies Assumption 3.1(ii).

(H2) Conditions (A1) and (A2) of Remark 5.2 hold true with $G = G_N^{\delta,\varepsilon}$.

(H3) $\ell_1 \in C_b(\mathbb{R})$, so $l_1 \in C_b(H)$.[14] Moreover the map $q \mapsto F_1(q)$, defined by

$$F_1(q) := \inf_{u\in\Lambda}\big\{\langle L_N^{\delta,\varepsilon} u, q\rangle_H + l_2(u)\big\}, \qquad q \in H,$$

is Lipschitz continuous. These conditions imply that $F_0(x,q) = l_1(x) + F_1(q)$ satisfies condition (A3) of Remark 5.2.

Then, under such assumptions, by Remark 5.2, for sufficiently large $\lambda > 0$ there exists a unique mild solution v to (6.12). By definition of mild solution, we have $v \in C_b^{1,G}(H)$. Furthermore, Assumption A.2 is verified through Remark A.3 in this case. Hence Proposition A.4 applies yielding $v \in \mathcal{S}^{A,G}(H)$ and enabling the application of Theorem 5.6. We now discuss the validity of the above assumptions (H1)–(H3):

- *On the validity of* (H1). First of all, we note that in Assumption 3.1(ii), we can take γ as small as we want; indeed, if this assumption holds true for some $\bar{\gamma} \in (0, 1/2)$, then it holds true also for all $\gamma \in (0, \bar{\gamma})$. By (6.4), the operator e^{tA_N} is diagonal with respect to the orthonormal basis $\{e_k\}$ with eigenvalues $e^{-t\mu_k}$. Assumption 3.1(ii) rewrites as

$$\int_0^t \Big(s^{-2\gamma} \sum_{k\in\mathbb{N}} \langle e^{sA} Q e^{sA^*} e_k, e_k\rangle_H\Big)\, ds$$

[14] According to Remark 4.9, it is possible to deal with the case when ℓ_1, and so l_1, has polynomial growth.

$$= \int_0^t \left(s^{-2\gamma} \sum_{k \in \mathbb{N}} e^{-2\mu_k s} |\sigma e_k|_H^2 \right) ds < \infty \qquad \forall t \geq 0. \tag{6.14}$$

Applying Fubini–Tonelli's theorem and considering (6.5) we see that (6.14) holds if

$$\sum_{k \in \mathbb{N}_0} k^{\frac{2(2\gamma - 1)}{d}} |\sigma e_k|_H^2 < \infty. \tag{6.15}$$

Let $\theta \geq 0$ be such that

$$\limsup_{k \to \infty} \frac{|\sigma e_k|_H^2}{k^{-2\theta}} < \infty \tag{6.16}$$

[recall that $\sigma \in \mathcal{L}(H)$, so $\theta = 0$ always verifies (6.16)]. Considering that γ can be taken as small as we want and combining (6.15) and (6.16), we conclude that (H1) holds if we may take in (6.16)

$$\theta > \frac{1}{2} - \frac{1}{d}. \tag{6.17}$$

In particular, if $d = 1$, then (H1) holds true for all $\sigma \in \mathcal{L}(H)$.

- *On the validity of* (H2). By (6.5), we have, for $k \in \mathbb{N}$,

$$G_N^{\delta,\varepsilon} e_k = (\delta I - A_N)^{\frac{1}{4}+\varepsilon} e_k = g_k e_k \qquad \text{where } g_k := (\delta + \mu_k)^{\frac{1}{4}+\varepsilon}.$$

The operator $\overline{e^{tA_N} G_N^{\delta,\varepsilon}}$ is diagonal too with respect to $\{e_k\}_{k \in \mathbb{N}}$ and

$$\overline{e^{tA_N} G_N^{\delta,\varepsilon}} e_k = e^{-\mu_k t} g_k e_k = e^{-\mu_k t} (\delta + \mu_k)^{\frac{1}{4}+\varepsilon} e_k, \qquad k \in \mathbb{N}. \tag{6.18}$$

Assume now further that σ is diagonal with respect to $\{e_k\}_{k \in \mathbb{N}}$ and nondegenerate, that is, $\sigma e_k = \sigma_k e_k$ for every $k \in \mathbb{N}$, where $\sigma_k > 0$ for every $k \in \mathbb{N}$. Set $q_k := \sigma_k^2 > 0$ for $k \in \mathbb{N}$. Then Q_t is diagonal too. Moreover, and $Q_t e_0 = t q_0 e_0$ and

$$Q_t e_k = \frac{q_k}{2\mu_k} (1 - e^{-2\mu_k t}) e_k \qquad \text{if } k \in \mathbb{N}_0, \forall t \geq 0.$$

Hence, with the agreement $\frac{1 - e^{-2\mu_k t}}{2\mu_k} := t$ if $k = 0$, we have

$$\begin{aligned} \Gamma_G(t) e_k &:= Q_t^{-1/2} \overline{e^{tA_N} G_N^{\delta,\varepsilon}} e_k \\ &= \sqrt{\frac{2\mu_k}{(1 - e^{-2t\mu_k}) q_k}} e^{-\mu_k t} (\delta + \mu_k)^{\frac{1}{4}+\varepsilon} e_k \qquad \forall k \in \mathbb{N}. \end{aligned}$$

Since $|\Gamma_G(t)|_{\mathcal{L}(H)} = \sup_{k \in \mathbb{N}} |\Gamma_G(t) e_k|_H$, then, with the agreement that $\frac{2\mu_k}{e^{2t\mu_k} - 1} := t^{-1}$ if $k = 0$, conditions (A1) and (A2) of Remark 5.2 hold true if and only if

$$\exists \eta \in L^1_{\mathrm{loc}}([0,+\infty);\mathbb{R}) \text{ bounded in a neighborhood of } +\infty \text{ s.t.}$$

$$\sqrt{\frac{2\mu_k(\delta+\mu_k)^{\frac{1}{2}+2\varepsilon}}{(e^{2t\mu_k}-1)q_k}} \le \eta(t) \qquad \forall t>0, \forall k \in \mathbb{N}. \tag{6.19}$$

Assume that

$$\liminf_{k\to\infty} \frac{q_k}{k^{-2\theta}} > 0 \qquad \text{for some } \theta \ge 0, \tag{6.20}$$

and let $k_0 \in \mathbb{N}$ and $c_0 > 0$ be such that $q_k \ge c_0 k^{-2\theta}$ for some $c_0 > 0$ and every $k \ge k_0$. Considering (6.5), let $c_1, c_2 > 0$ and $k_0' \in \mathbb{N}$ be such that $c_1 k^{\frac{2}{d}} \le \mu_k \le c_2 k^{\frac{2}{d}}$ for every $k \ge k_0'$. Calling $\bar{k} := k_0 \vee k_0'$ it is clear that, for a suitable $C_0 > 0$,

$$\sup_{k<\bar{k}} \sqrt{\frac{2\mu_k(\delta+\mu_k)^{\frac{1}{2}+2\varepsilon}}{(e^{2t\mu_k}-1)q_k}} \le C_0 t^{-1/2}.$$

Hence, to prove (6.19) above, we take $k \ge \bar{k}$ and we rewrite (6.19) (up to a constant depending on c_0, c_1, c_2) as

$$\exists \eta \in L^1_{\mathrm{loc}}([0,+\infty);\mathbb{R}) \text{ bounded in a neighborhood of } +\infty \text{ s.t.}$$

$$\sqrt{\frac{k^{\frac{2}{d}}(\delta+k^{\frac{2}{d}})^{\frac{1}{2}+2\varepsilon}}{(e^{2tk^{\frac{2}{d}}}-1)k^{-2\theta}}} \le \eta(t) \qquad \forall t>0, \forall k \ge \bar{k}. \tag{6.21}$$

Noting that $C_1 := \sup_{s>0} \frac{s^{\frac{3}{2}+2\varepsilon+d\theta}}{e^s-1} < +\infty$, we can estimate

$$\frac{k^{\frac{2}{d}}(\delta+k^{\frac{2}{d}})^{\frac{1}{2}+2\varepsilon}}{(e^{2tk^{\frac{2}{d}}}-1)k^{-2\theta}} \le \frac{(1+\delta)^{\frac{1}{2}+2\varepsilon} k^{\frac{2}{d}(\frac{3}{2}+2\varepsilon)+2\theta}}{(e^{2tk^{\frac{2}{d}}}-1)} \le C_1 \frac{(1+\delta)^{\frac{1}{2}+2\varepsilon}}{(2t)^{\frac{3}{2}+2\varepsilon+d\theta}} \qquad \forall k \ge \bar{k}.$$

Therefore, (H2) is satisfied whenever (6.20) holds for some θ such that $\frac{3}{2}+2\varepsilon+d\theta < 2$. As $\varepsilon > 0$ can be taken arbitrarily small, we conclude that (H2) can be fulfilled if (6.20) holds for some θ such that

$$\frac{3}{2}+d\theta < 2 \quad \Longleftrightarrow \quad \theta < \frac{1}{2d}. \tag{6.22}$$

- *On the simultaneous validity of* (H1)–(H2). Looking at (6.17) and (6.22), we see that (H1)–(H2) can be simultaneously fulfilled by choosing a suitable $\varepsilon > 0$ if σ is diagonal with respect to $\{e_k\}_{k\in\mathbb{N}}$ and (6.20) is verified for some $\theta \ge 0$ such that

$$\frac{1}{2}-\frac{1}{d} < \theta < \frac{1}{2d}. \tag{6.23}$$

These requirements can be fulfilled only for dimension $d \le 2$.

- *On the validity of* (H3). This is guaranteed, for instance, if Λ is bounded, ℓ_1 is continuous and bounded, ℓ_2 is measurable.

6.1.4. *Optimal feedback controls*. In the framework of the previous subsection, we look now at the existence of optimal feedback controls.

THEOREM 6.1. *Let* (H1)–(H3) *of the previous subsection hold. Assume that the multi-valued map*

$$\Psi : H \to \Lambda, \qquad q \mapsto \arg\min_{u\in\Lambda}\{\langle L_N^{\delta,\varepsilon} u, q\rangle_H + l_2(u)\} \tag{6.24}$$

admits a Lipschitz continuous selection ψ and that $D^{G_N^{\delta,\varepsilon}} v$ is Lipschitz continuous. Set $\phi := \psi \circ D^{G_N^{\delta,\varepsilon}} v$. Then the SDE

$$\begin{cases} dX(t) = [A_N X(t) + G_N^{\delta,\varepsilon} L_N^{\delta,\varepsilon}(\phi(X(t)))]\,dt + \sigma\, dW(t), & t \geq 0, \\ X(0) = x, \end{cases} \tag{6.25}$$

admits a unique mild solution $X_\phi(\cdot; x) \in \mathcal{K}_{\mathcal{P}}^{1,\mathrm{loc}}(H)$ [*in the sense of* (5.19)] *admitting a version with continuous trajectories. As a consequence*, *Corollary* 5.7(i) *applies providing the optimality of the couple $(u_\phi(\cdot), X_\phi(\cdot; x))$, where $u_\phi(t) := \phi(X_\phi(t; x))$ for $t \geq 0$.*

PROOF. By the assumptions, the map ϕ is Lipschitz continuous, too. Then the proof follows the classical fixed point arguments as in standard results of existence and uniqueness of SDEs in infinite dimension; see, for example, [10], Theorem 7.5. Here we only need to take care of dealing with $\overline{e^{sA_N} G_N^{\delta,\varepsilon}}$ in place of e^{sA_N} in the convolution term and use (3.2) with $G = G_N^{\delta,\varepsilon}$. $\square$

The assumption that Ψ defined in (6.24) admits a Lipschitz continuous selection ψ is guaranteed, for example, if $\Lambda = U$, $l_2 : U \to \mathbb{R}$ is strictly convex,

$$\lim_{|u|_U \to +\infty} \frac{l_2(u)}{|u|_U} = +\infty,$$

l_2 is Fréchet differentiable, and Dl_2 has Lipschitz continuous inverse. Indeed, in this case the infimum in (6.24) is uniquely achieved (hence, Ψ is single-valued) at

$$u^*(q) = (Dl_2)^{-1}((L_N^{\delta,\varepsilon})^* q), \qquad q \in H.$$

Hence, if we are able to check that $D^{G_N^{\delta,\varepsilon}} v$ is Lipschitz continuous, we can then apply Corollary 5.7(i) in its strongest form to get uniqueness of the optimal control constructed.

On the other hand, checking that $D^{G_N^{\delta,\varepsilon}} v$ is Lipschitz continuous might be, in general, a very difficult task,[15] whereas mere continuity of $D^{G_N^{\delta,\varepsilon}} v$ is a condition

[15] This can be done assuming more regularity of ℓ_1—hence of l_1—and proving a suitable C^2 property of v; see, for example, the approach used in [31] or in [29].

already "contained" in the definition of mild solution to (6.12). Hence, it would be meaningful to provide a Peano-type result[16] of existence of mild solutions to CLE (6.25). This seems possible when a selection ψ of Ψ in (6.24) is known to be only continuous and bounded on bounded sets, as:

(i) the semigroup $\{e^{tA_N}\}_{t\geq 0}$ is compact;

(ii) as $D^{G_N^{\delta,\varepsilon}} v$ is continuous and bounded by construction, the map $\phi := \psi \circ D^{G_N^{\delta,\varepsilon}} v$ is continuous and bounded.

Indeed, in such a framework, it seems possible to use the methods of [7], Proposition 3 (see also [22]), passing through the use of the so-called Skorohod representation theorem, to construct martingale solutions to (6.25); hence, to construct optimal feedback controls in the weak formulation.

REMARK 6.2. In the specific case we are handling, where the diffusion term is just additive in the equation, a way to construct the solution in the original probability space Ω might consist in constructing a pathwise solution dealing with a parameterized family of deterministic problems with parameter $\omega \in \Omega$ (see [2], [9], Sections 14.2 and 15.2, [19, 40]). Once this is done, the problem is to prove that the family of solutions constructed ω by ω admits an adapted selection. The existence of a selection measurable with respect to $\mathcal{F}$ can be obtained using measurable selection theorems (see again [2]); proving that this selection is also adapted is a problematic task, which is still open. In the case, when one knows ex ante that the pathwise solution is unique for a.e. $\omega \in \Omega$, then F. Flandoli (personal communication) showed us how to accomplish this task. Unfortunately, in our case, the uniqueness of the solutions of the deterministic equations for a.e. $\omega \in \Omega$ only holds when the properties of the coefficients allow to find directly mild solutions to SDE (6.25).

6.2. *Stochastic optimal control with delay in the control variable.* Here, we consider an infinite horizon version of a control problem studied in [29, 30]. Consider the following linear controlled one dimensional SDE:

$$dy(t) = \left[a_0 y(t) + b_0 u(t) + \int_{-d}^{0} b_1(\xi) u(t+\xi)\, d\xi\right] dt + \sigma_0\, dW(t), \qquad t \geq 0, \tag{6.26}$$

under the initial conditions $y(0) = y_0$ and $u(\xi) = u_0(\xi)$ for $\xi \in [-d, 0)$, where y_0 is the inotial state u_0 is the history of the control at the initial date $t = 0$. In the equation above:

- $W = \{W(t)\}_{t\geq 0}$ is a standard one dimensional Brownian motion;
- $a_0, b_0, \sigma_0 \in \mathbb{R}$, $\sigma_0 > 0$;

[16] This is not straightforward: in infinite dimension Peano's theorem fails in general (see [24]).

- $d > 0$ represents the maximum delay the control takes to affect the system;
- $b_1(\cdot)$ is a (real-valued) function weighting the aftereffects of the control on the system; we consider here the case of distributed delay, that is, when $b_1 \in L^2([-d, 0], \mathbb{R})$.

The control u takes values in a closed subset $\Lambda \subseteq U := \mathbb{R}$ and belongs to $\mathcal{U}_2$ [defined by (3.5) with $p = 2$].

Such kind of equations (even in a deterministic framework) have been used to model the effect of advertising on the sales of a product [17, 27, 28], the effect of investments with time to build on growth [1, 13], to model optimal portfolio problems with execution delay [3], to model the interaction of drugs with tumor cells [35], p. 17.

Denoting by $y^{y_0,u_0,u(\cdot)}$ the unique solution to (6.26), the goal of the problem is to minimize, over all control strategies in $\mathcal{U}_2$, the following objective functional:

$$\mathbb{E}\left[\int_0^\infty e^{-\lambda t}\big(\ell_0\big(y^{y_0,u_0,u(\cdot)}(t)\big) + \ell_1\big(u(t)\big)\big)\,dt\right], \tag{6.27}$$

where $\ell_0 : \mathbb{R} \to \mathbb{R}$ and $\ell_1 : \Lambda \to \mathbb{R}$ are measurable and bounded from below. It is important to note that here ℓ_0 and ℓ_1 do not depend on the past of the state and/or control. This is a very common feature of many applied problems.

A standard way to approach these delayed control problems, introduced in [47] for the deterministic case and extended to the stochastic case in [27, 28], is to reformulate them as equivalent infinite dimensional control problems without delay.[17] The details are given in [29] for the finite horizon case, which is completely similar to the infinite horizon case, with the obvious changes (see also [17] for the infinite horizon case in a deterministic framework with a different embedding space). Consider the Hilbert space $H := \mathbb{R} \times L^2([-d, 0], \mathbb{R})$, set $b := (b_0, b_1(\cdot)) \in H$, and assume, without loss of generality, $|b|_H = 1$. The state equation (6.26) is rephrased in H as a linear SDE with state variable $X = (X_0, X_1(\cdot))$ as follows:

$$\begin{cases} dX(t) = \big[AX(t) + Gu(t)\big]\,dt + \sigma\, dW(t), & t \geq 0, \\ X(0) = x = \big(x_0, x_1(\cdot)\big), \end{cases} \tag{6.28}$$

where

$$\begin{aligned} \mathcal{D}(A) &= \big\{\big(x_0, x_1(\cdot)\big) \in \mathbb{R} \times W^{1,2}\big([-d, 0]; \mathbb{R}\big) : x_1(-d) = 0\big\}, \\ Ax &= \big(a_0 x_0 + x_1(0), -x_1'\big), \qquad x \in \mathcal{D}(A); \\ G &: \mathbb{R} \to H, \qquad G(u) = ub; \\ \sigma &: \mathbb{R} \to H, \qquad \sigma(z) = (\sigma_0 z, 0); \end{aligned}$$

[17] It must be noted that, under suitable restrictions on the data, one can treat (stochastic) optimal control problems with delay avoiding to look at them as infinite dimensional systems (see [18]). However, this is possible only in very special cases, leaving out many concrete applications.

$$
\begin{aligned}
x_0 &= y_0, \\
x_1(\xi) &= \int_{-d}^{\xi} b_1(\varsigma) u_0(\varsigma - \xi)\, d\varsigma, \qquad \xi \in [-d, 0].
\end{aligned}
$$

It is well known that A is the generator of a C_0-semigroup of linear bounded operators on H. Note that the infinite dimensional datum $x_1(\cdot)$ depends on the "initial past" $u_0(\cdot)$ of the control. It turns out that $X_0(t; x, u(\cdot)) = y^{y_0,u_0,u(\cdot)}$, so (6.27) is rewritten as

$$
J\big(x; u(\cdot)\big) := \mathbb{E}\bigg[\int_0^{\infty} e^{-\lambda t}\big(\ell_0\big(X_0\big(t; x, u(\cdot)\big) + \ell_1\big(u(t)\big)\big)\, dt\bigg]. \tag{6.29}
$$

Setting $Q := \sigma\sigma^*$, the HJB equation associated to the minimization of (6.29) is

$$
\begin{aligned}
\lambda v(x) &= \frac{1}{2}\operatorname{Tr}\big[Q D^2 v(x)\big] + \langle Ax, Dv(x)\rangle_H \\
&\quad + \inf_{u\in\Lambda}\big\{u D^G v(x) + \ell_1(u)\big\} + \ell_0(x_0), \qquad x \in H.
\end{aligned} \tag{6.30}
$$

Notice that $D^G = \frac{\partial}{\partial b}$, where the latter symbol denotes the directional derivative along the direction b. So, the nice feature of the equation above is that the nonlinearity on the gradient only involves the directional derivative D^G. Note also that here we do not have the so-called *structural condition* $G(\mathbb{R}) \subseteq \sigma(\mathbb{R})$; this prevents the use of techniques based on backward SDEs (see, e.g., [21]) to tackle the problem.

Now we discuss the assumptions of our main result Theorem 5.6 in this example. First of all, it is easy to check that Assumption 3.1 and (3.8). The third assumption, that is, the existence of a mild solution $v \in \mathcal{S}^{A,G}(H)$ to (6.30) needs to be discussed carefully.

In [29], the authors study a finite horizon optimal control problem with the same state equation (6.26) and a similar objective functional. Exploiting only partial smoothing properties of the transition semigroup associated to the state equation (6.28) with null control, the authors are able to provide, under suitable reasonable assumptions on the data, existence and uniqueness results for the parabolic HJB equation associated to the control problem.

We believe that the approach of [29] can be adapted to our infinite horizon case, getting a mild solution $v \in \mathcal{D}(\mathcal{A}^{(0)}) \cap C_b^{1,G}(H)$ to HJB (6.30). Then, to apply our theory one should prove that such function v is Lipschitz continuous on compact sets, which enables to apply Proposition A.6 (indeed Assumption A.5 is verified) to get $v \in \mathcal{S}^{A,G}(H)$. To get this goal, one can proceed as in [29] by assuming more regularity on the data of the problem. More precisely, assuming that $l_0 \in C_b^1(\mathbb{R})$ and that the Hamiltonian $p \mapsto \inf_{u\in\Lambda}\{up + \ell_1(u)\}$ is differentiable with Lipschitz continuous derivative, [29] proves that the mild solution $v \in C_b^1(H)$. This fact, in particular, implies the required Lipschitz continuity of v. In [30], the authors also provide a verification theorem for their finite horizon problem. They use an approximation procedure of the solution of the HJB equation, which our results allow to avoid here.

APPENDIX

Recall that, given $G \in \mathcal{L}_u(K, H)$, the pseudo-inverse $G^{-1} : \mathcal{R}(G) \to \mathcal{D}(G)$ is defined as the operator that associates to each $h \in \mathcal{R}(G)$ the element of $G^{-1}(\{h\})$ having minimum norm.[18] Note that $G^{-1}G : \mathcal{D}(G) \to \mathcal{D}(G)$ is bounded, so it can be extended to a bounded operator $\overline{G^{-1}G} \in \mathcal{L}(K)$.

LEMMA A.1. *We have*

$$\langle D^G f(x), \overline{G^{-1}G}k\rangle_K = \langle D^G f(x), k\rangle_K \qquad \forall k \in K, \forall x \in H. \tag{A.1}$$

PROOF. Assume first that $k \in \mathcal{D}(G)$. In this case $GG^{-1}Gk = Gk$. Then, using Remark 2.4, we write

$$\begin{aligned}\langle D^G f(x), \overline{G^{-1}G}k\rangle_K &= \lim_{s\to 0} \frac{f(x + sG\overline{G^{-1}G}k) - f(x)}{s}\\ &= \lim_{s\to 0} \frac{f(x + sGk) - f(x)}{s}\\ &= \langle D^G f(x), k\rangle_K \qquad \forall x \in H.\end{aligned}$$

If $k \notin \mathcal{D}(G)$, we can take a sequence $\{k_n\} \subseteq \mathcal{D}(G)$ converging to k. Considering (A.1) on k_n and passing to the limit the claim follows taking into account that $\overline{G^{-1}G}$ is bounded. □

ASSUMPTION A.2. The operator $G \in \mathcal{L}_u(K, H)$ is such that for every $k \in K$:

(i) there exists $\varepsilon > 0$ such that $\{\int_0^t \overline{e^{sA}G}k\, ds\}_{t\in(0,\varepsilon)} \subseteq \mathcal{R}(G)$;
(ii) $G^{-1}(\frac{1}{t}\int_0^t \overline{e^{sA}G}k\, ds) \to \overline{G^{-1}G}k$, as $t \to 0^+$.

REMARK A.3. Note that $\int_0^t e^{sA}h\, ds \in \mathcal{D}(A)$ for every $t > 0$ and $h \in H$. So, in view of the fact that $\overline{G^{-1}G}$ is bounded, Assumption A.2 is verified, in particular, if $K = H$, $\mathcal{D}(A) \subseteq \mathcal{D}(G)$ and, for sufficiently small $\varepsilon > 0$,

$$G\int_0^t e^{sA}h\, ds = \int_0^t \overline{e^{sA}G}h\, ds \qquad \forall t \in (0, \varepsilon), \forall h \in H.$$

This applies, for example, to the case when A is dissipative and generates an analytic semigroup, and $G = (\delta I - A)^\beta$ with $\delta > 0$ and $\beta \in (0, 1)$ (see the example of Section 6.1).

PROPOSITION A.4. *Let Assumption A.2 holds. Then $\mathcal{S}^{A,G}(H) = C_b^{1,G}(H)$.*

---

[18] Existence and uniqueness of such an element follows from the fact that G is a closed operator and applying the results of [11], Section II.4.29, p. 74.

PROOF. Fix $k \in K$, $z \in C(\mathbb{R}^+; H)$ and let $\varepsilon > 0$ be as in Assumption A.2(i). Noting that $GG^{-1}h = h$ for every $h \in \mathcal{R}(G)$, by Assumption A.2(i) we can write

$$\int_0^t \overline{e^{sA}G}k\,ds = Gk(t) \qquad \text{where } k(t) := G^{-1}\int_0^t \overline{e^{sA}G}k\,ds, \forall t \in (0,\varepsilon). \tag{A.2}$$

Moreover, by Assumption A.2(ii), we have

$$\frac{k(t)}{t} \xrightarrow{t\to 0^+} \overline{G^{-1}G}k. \tag{A.3}$$

Fix now $t \in (0,\varepsilon)$. Using (A.2), we write

$$\begin{aligned}
&\frac{\varphi(z(t) + \int_0^t \overline{e^{sA}G}k\,ds) - \varphi(z(t))}{t} \\
&\quad = \frac{\varphi(z(t) + Gk(t)) - \varphi(z(t)) - \langle D^G f(x(t)), k(t)\rangle_K}{t} \\
&\qquad + \left\langle D^G\varphi\big(z(t)\big), \frac{k(t)}{t}\right\rangle_K.
\end{aligned} \tag{A.4}$$

Mean value theorem applied to the function $[0,1] \to \mathbb{R}, \xi \mapsto f(x(t) + \xi Gk(t))$ yields (see also Remark 2.4)

$$\begin{aligned}
&\varphi\big(z(t) + Gk(t)\big) - \varphi\big(z(t)\big) \\
&\quad = \int_0^1 \frac{d}{d\xi}\varphi\big(z(t) + \xi Gk(t)\big)\,d\xi \\
&\quad = \int_0^1 \lim_{\eta\to 0}\frac{\varphi(z(t) + (\xi+\eta)Gk(t)) - \varphi(z(t) + \xi Gk(t))}{\eta}\,d\xi \\
&\quad = \int_0^1 \big\langle D^G\varphi\big(z(t) + \xi Gk(t)\big), k(t)\big\rangle_K\,d\xi.
\end{aligned}$$

Hence, (A.4) rewrites as

$$\begin{aligned}
&\frac{\varphi(z(t) + \int_0^t \overline{e^{sA}G}k\,ds) - \varphi(z(t))}{t} \\
&\quad = \int_0^1 \left\langle D^G\varphi\big(z(t) + \xi Gk(t)\big) - D^G\varphi\big(z(t)\big), \frac{k(t)}{t}\right\rangle_K d\xi \\
&\qquad + \left\langle D^G\varphi\big(z(t)\big), \frac{k(t)}{t}\right\rangle_K.
\end{aligned} \tag{A.5}$$

Moreover, we can estimate

$$\begin{aligned}
&\left|\left\langle D^G\varphi\big(z(t) + \xi Gk(t)\big) - D^G\varphi\big(z(t)\big), \frac{k(t)}{t}\right\rangle_K\right| \\
&\quad \le \big|D^G\varphi\big(z(t) + \xi Gk(t)\big) - D^G\varphi\big(z(t)\big)\big|_K \cdot \left|\frac{k(t)}{t}\right|_K \qquad \forall \xi \in [0,1].
\end{aligned} \tag{A.6}$$

Now we are going to take the limit for $t \to 0^+$ in (A.5). To this purpose, we observe that, as $D^G\varphi \in C_b(H, K)$ and $\{z(t)\}_{t\in(0,\varepsilon)}$ is compact in H, we have

$$\sup_{t\in(0,\varepsilon)} \big|D^G\varphi\big(z(t)+h\big) - D^G f\big(z(t)\big)\big|_K \overset{|h|\to 0^+}{\longrightarrow} 0. \tag{A.7}$$

By definition of $k(t)$ [see (A.2)], we have $|Gk(t)|_H \overset{t\to 0^+}{\longrightarrow} 0$. Hence, (A.7) provides

$$\sup_{\xi\in[0,1]} \big|D^G\varphi\big(z(t)+\xi Gk(t)\big) - D^G\varphi\big(z(t)\big)\big|_K \overset{t\to 0^+}{\longrightarrow} 0. \tag{A.8}$$

Hence, combining (A.3), (A.6) and (A.8), we get

$$\int_0^1 \left\langle D^G\varphi\big(z(t)+\xi Gk(t)\big) - D^G\varphi\big(z(t)\big), \frac{k(t)}{t}\right\rangle_K d\xi \overset{t\to 0^+}{\longrightarrow} 0. \tag{A.9}$$

Moreover, (A.3) and the continuity of the maps $t \mapsto z(t)$ and $x \mapsto D^G\varphi(x)$ entail

$$\left\langle D^G\varphi\big(z(t)\big), \frac{k(t)}{t}\right\rangle_K \overset{t\to 0^+}{\longrightarrow} \langle D^G\varphi\big(z(0)\big), \overline{G^{-1}Gk}\rangle_K. \tag{A.10}$$

Combining (A.5), (A.9), (A.10) and Lemma A.1, the claim follows. □

ASSUMPTION A.5. $G \in \mathcal{L}(K, H)$.

PROPOSITION A.6. *Let Assumption A.5 hold and let $\varphi \in C_b^{1,G}(H)$ be Lipschitz continuous on compact sets. Then $\varphi \in \mathcal{S}^{A,G}(H)$.*

PROOF. Let $k \in K$. Observe that, as $G \in \mathcal{L}(K, H)$, we have $k \in K = \mathcal{D}(G)$, $\overline{e^{sA}G}k = e^{sA}Gk$ for every $s > 0$, and

$$\lim_{t\to 0^+} \frac{1}{t}\int_0^t e^{sA}Gk\,ds \to Gk. \tag{A.11}$$

Let $t > 0$. We can split

$$\begin{aligned} &\frac{\varphi(z(t)+\int_0^t e^{sA}Gk\,ds) - \varphi(z(t))}{t} \\ &\quad= \frac{\varphi(z(t)+\int_0^t e^{sA}Gk\,ds) - \varphi(z(t)+tGk)}{t} \\ &\qquad+ \frac{\varphi(z(t)+tGk) - \varphi(z(t))}{t}. \end{aligned} \tag{A.12}$$

The set $\{z(t)+\int_0^t e^{sA}Gk\,ds\}_{t\in(0,1)} \cup \{z(t)+tGk\}_{t\in(0,1)} \subset K$ is precompact. Hence, by Lipschitz continuity of φ on compact sets, we have for some $C_0 > 0$ independent of $t \in (0, 1)$

$$\left|\frac{\varphi(z(t)+\int_0^t e^{sA}Gk\,ds) - \varphi(z(t)+tGk)}{t}\right| \le C_0\left|\frac{1}{t}\int_0^t e^{sA}Gk\,ds - Gk\right|. \tag{A.13}$$

We let now $t \to 0^+$ in (A.12). Combining with (A.13) and (A.11), we get

$$\begin{aligned}\lim_{t\to 0^+} & \frac{\varphi(z(t) + \int_0^t e^{sA} Gk\, ds) - \varphi(z(t))}{t} \\ &= \lim_{t\to 0^+} \frac{\varphi(z(t) + tGk) - \varphi(z(t))}{t},\end{aligned} \tag{A.14}$$

provided that the limit in the right-hand side above exists, as we are going to show. We write

$$\begin{aligned}\varphi\big(z(t) + tGk\big) - \varphi\big(z(t)\big) &= \int_0^1 \frac{d}{d\xi} \varphi\big(z(t) + \xi t Gk\big)\, d\xi \\ &= \int_0^1 \lim_{\eta\to 0} \frac{\varphi(z(t) + (\xi+\eta)tGk) - \varphi(z(t) + \xi t Gk)}{\eta}\, d\xi \\ &= \int_0^1 \big\langle D^G \varphi\big(z(t) + \xi t Gk\big), tk\big\rangle_K\, d\xi.\end{aligned}$$

By the equalities above and considering that $D^G\varphi \in C_b(H; K)$, we have

$$\begin{aligned}\lim_{t\to 0^+} \frac{\varphi(z(t) + tGk) - \varphi(z(t))}{t} &= \lim_{t\to 0^+} \int_0^1 \big\langle D^G \varphi\big(z(t) + \xi t Gk\big), k\big\rangle_K\, d\xi \\ &= \big\langle D^G \varphi\big(z(0)\big), k\big\rangle_K\end{aligned}$$

and the claim follows from (A.14). $\square$

Acknowledgments. The authors are sincerely grateful to Franco Flandoli, Ben Goldys and Mauro Rosestolato for fruitful discussions on Section 6.1.4 and Remark 6.2. The authors are also grateful to an anonymous referee for careful scrutiny and useful suggestions that led to an improved version of the paper.

REFERENCES

[1] AÏD, R., FEDERICO, S., PHAM, H. and VILLENEUVE, B. (2015). Explicit investment rules with time-to-build and uncertainty. *J. Econom. Dynam. Control* **51** 240–256. MR3294764

[2] BENSOUSSAN, A. and TEMAM, R. (1973). Équations stochastiques du type Navier–Stokes. *J. Funct. Anal.* **13** 195–222. MR0348841

[3] BRUDER, B. and PHAM, H. (2009). Impulse control problem on finite horizon with execution delay. *Stochastic Process. Appl.* **119** 1436–1469. MR2513141

[4] CERRAI, S. (1994). A Hille–Yosida theorem for weakly continuous semigroups. *Semigroup Forum* **49** 349–367. MR1293091

[5] CERRAI, S. (2001). Stationary Hamilton–Jacobi equations in Hilbert spaces and applications to a stochastic optimal control problem. *SIAM J. Control Optim.* **40** 824–852. MR1871456

[6] CERRAI, S. and GOZZI, F. (1995). Strong solutions of Cauchy problems associated to weakly continuous semigroups. *Differential Integral Equations* **8** 465–486. MR1306569

[7] CHOJNOWSKA-MICHALIK, A. and GOŁDYS, B. (1995). Existence, uniqueness and invariant measures for stochastic semilinear equations on Hilbert spaces. *Probab. Theory Related Fields* **102** 331–356. MR1339737

[8] DA PRATO, G. and LUNARDI, A. (1995). On the Ornstein–Uhlenbeck operator in spaces of continuous functions. *J. Funct. Anal.* **131** 94–114. MR1343161

[9] DA PRATO, G. and ZABCZYK, J. (1996). *Ergodicity for Infinite-Dimensional Systems. London Mathematical Society Lecture Note Series* **229**. Cambridge Univ. Press, Cambridge. MR1417491

[10] DA PRATO, G. and ZABCZYK, J. (2014). *Stochastic Equations in Infinite Dimensions*, 2nd ed. *Encyclopedia of Mathematics and Its Applications* **152**. Cambridge Univ. Press, Cambridge. MR3236753

[11] DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators. Part I*. Interscience, New York.

[12] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York. MR0838085

[13] FABBRI, G. and GOZZI, F. (2008). Solving optimal growth models with vintage capital: The dynamic programming approach. *J. Econom. Theory* **143** 331–373. MR2887332

[14] FABBRI, G., GOZZI, F. and ŚWIECH, A. (2017). *Stochastic Optimal Control in Infinite Dimension: Dynamic Programming and HJB Equations. Probability Theory and Stochastic Modelling* **82**. Springer, Cham. MR3674558

[15] FARKAS, B. (2004). Perturbations of bi-continuous semigroups with applications to transition semigroups on $C_b(H)$. *Semigroup Forum* **68** 87–107. MR2027610

[16] FEDERICO, S. and GOZZI, F. (2017). Mild solutions of semilinear elliptic equations in Hilbert spaces. *J. Differential Equations* **262** 3343–3389. MR3584895

[17] FEDERICO, S. and TACCONI, E. (2014). Dynamic programming for optimal control problems with delays in the control variable. *SIAM J. Control Optim.* **52** 1203–1236. MR3191586

[18] FEDERICO, S. and TANKOV, P. (2015). Finite-dimensional representations for controlled diffusions with delay. *Appl. Math. Optim.* **71** 165–194. MR3306054

[19] FLANDOLI, F. and SCHMALFUSS, B. (1999). Weak solutions and attractors for three-dimensional Navier–Stokes equations with nonregular force. *J. Dynam. Differential Equations* **11** 355–398. MR1695250

[20] FOLLAND, G. B. (1999). *Real Analysis: Modern Techniques and Their Applications*, 2nd ed. Wiley, New York. MR1681462

[21] FUHRMAN, M. and TESSITORE, G. (2004). Infinite horizon backward stochastic differential equations and elliptic equations in Hilbert spaces. *Ann. Probab.* **32** 607–660. MR2039938

[22] GATAREK, D. and GOLDYS, B. (1997). On invariant measures for diffusions on Banach spaces. *Potential Anal.* **7** 539–553. MR1467205

[23] GAWARECKI, L. and MANDREKAR, V. (2011). *Stochastic Differential Equations in Infinite Dimensions with Applications to Stochastic Partial Differential Equations*. Springer, Heidelberg. MR2560625

[24] GODUNOV, A. N. (1975). The Peano theorem in Banach spaces. *Funct. Anal. Appl.* **9** 53–55.

[25] GOLDYS, B. and KOCAN, M. (2001). Diffusion semigroups in spaces of continuous functions with mixed topology. *J. Differential Equations* **173** 17–39. MR1836243

[26] GOZZI, F. (1995). Regularity of solutions of a second order Hamilton–Jacobi equation and application to a control problem. *Comm. Partial Differential Equations* **20** 775–826. MR1326907

[27] GOZZI, F. and MARINELLI, C. (2006). Stochastic optimal control of delay equations arising in advertising models. In *Stochastic Partial Differential Equations and Applications—VII. Lect. Notes Pure Appl. Math.* **245** 133–148. Chapman & Hall/CRC, Boca Raton, FL. MR2227226

[28] GOZZI, F., MARINELLI, C. and SAVIN, S. (2009). On controlled linear diffusions with delay in a model of optimal advertising under uncertainty with memory effects. *J. Optim. Theory Appl.* **142** 291–321. MR2525792

[29] GOZZI, F. and MASIERO, F. (2017). Stochastic optimal control with delay in the control I: Solving the HJB equation through partial smoothing. *SIAM J. Control Optim.* **55** 2981–3012. MR3702860

[30] GOZZI, F. and MASIERO, F. (2017). Stochastic optimal control with delay in the control II: Verification theorem and optimal feedbacks. *SIAM J. Control Optim.* **55** 3013–3038. MR3702861

[31] GOZZI, F. and ROUY, E. (1996). Regular solutions of second-order stationary Hamilton–Jacobi equations. *J. Differential Equations* **130** 201–234. MR1409030

[32] GOZZI, F. and RUSSO, F. (2006). Verification theorems for stochastic optimal control problems via a time dependent Fukushima–Dirichlet decomposition. *Stochastic Process. Appl.* **116** 1530–1562. MR2269215

[33] GOZZI, F., ŚWIECH, A. and ZHOU, X. Y. (2005). A corrected proof of the stochastic verification theorem within the framework of viscosity solutions. *SIAM J. Control Optim.* **43** 2009–2019. MR2179474

[34] KARATZAS, I. and SHREVE, S. E. (1988). *Brownian Motion and Stochastic Calculus. Graduate Texts in Mathematics* **113**. Springer, New York. MR0917065

[35] KOLMANOVSKII, V. B. and SHAIKHET, L. E. (1996). *Control of Systems with Aftereffect. Translations of Mathematical Monographs* **157**. Amer. Math. Soc., Providence, RI. MR1415834

[36] KRYLOV, N. V. (1980). *Controlled Diffusion Processes. Applications of Mathematics* **14**. Springer, New York. MR0601776

[37] LASIECKA, I. (1980). Unified theory for abstract parabolic boundary problems—A semigroup approach. *Appl. Math. Optim.* **6** 287–333. MR0587501

[38] LIONS, J.-L. and MAGENES, E. (1972). *Non-homogeneous Boundary Value Problems and Applications. Vol. I.* Springer, Berlin.

[39] LUNARDI, A. (1995). *Analytic Semigroups and Optimal Regularity in Parabolic Problems. Progress in Nonlinear Differential Equations and Their Applications* **16**. Birkhäuser, Basel. MR1329547

[40] MIKULEVICIUS, R. and ROZOVSKII, B. L. (2005). Global L_2-solutions of stochastic Navier–Stokes equations. *Ann. Probab.* **33** 137–176. MR2118862

[41] PARDOUX, É. and PENG, S. (1992). Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic Partial Differential Equations and Their Applications* (*Charlotte*, *NC*, 1991). *Lect. Notes Control Inf. Sci.* **176** 200–217. Springer, Berlin. MR1176785

[42] PAZY, A. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations. Applied Mathematical Sciences* **44**. Springer, New York. MR0710486

[43] PRIOLA, E. (1999). On a class of Markov type semigroups in spaces of uniformly continuous and bounded functions. *Studia Math.* **136** 271–295. MR1724248

[44] PRIOLA, E. (1999). Partial differential equations with infinitely many variables. Ph.D. thesis, Tesi di Dottorato, Universita degli Studi di Milano.

[45] REED, M. and SIMON, B. (1980). *Methods of Modern Mathematical Physics. I: Functional Analysis*, 2nd ed. Academic Press, New York. MR0751959

[46] TRIEBEL, H. (1978). *Interpolation Theory, Function Spaces, Differential Operators. North-Holland Mathematical Library* **18**. North-Holland, Amsterdam. MR0503903

[47] VINTER, R. B. and KWONG, R. H. (1981). The infinite time quadratic control problem for linear systems with state and control delays: An evolution equation approach. *SIAM J. Control Optim.* **19** 139–153. MR0603086

[48] YONG, J. and ZHOU, X. Y. (1999). *Stochastic Controls: Hamiltonian Systems and HJB Equations. Applications of Mathematics* (*New York*) **43**. Springer, New York. MR1696772

DIPARTIMENTO DI ECONOMIA POLITICA E STATISTICA
UNIVERSITÀ DEGLI STUDI DI SIENA
P.ZZA S. FRANCESCO 7-8
53100 SIENA
ITALY
E-MAIL: salvatore.federico@unisi.it

DIPARTIMENTO DI ECONOMIA E FINANZA
LUISS UNIVERSITY
VIALE ROMANIA 32
00197 ROME
ITALY
E-MAIL: fgozzi@luiss.it