VARIATIONAL INEQUALITIES AND SMOOTH-FIT PRINCIPLE FOR SINGULAR STOCHASTIC CONTROL PROBLEMS IN HILBERT SPACES

SALVATORE FEDERICO, GIORGIO FERRARI, FRANK RIEDEL, AND MICHAEL RÖCKNER

Abstract. We consider a class of infinite-dimensional singular stochastic control problems. These can be thought of as spatial monotone follower problems and find applications in spatial models of production and climate transition. Let $(D, \mathcal{M}, \mu)$ be a finite measure space and consider the Hilbert space $H := L^2(D, \mathcal{M}, \mu; \mathbb{R})$. Let then X be an H-valued stochastic process on a suitable complete probability space, whose evolution is determined through an SPDE driven by a self-adjoint linear operator $\mathcal{A}$ and affected by a cylindrical Brownian motion. The evolution of X is controlled linearly via an H-valued control consisting of the direction and the intensity of action, a real-valued nondecreasing right-continuous stochastic process, adapted to the underlying filtration. The goal is to minimize a discounted convex cost-functional over an infinite time-horizon. By combining properties of semiconcave functions and techniques from viscosity theory, we first show that the value function of the problem V is a $C^{1,Lip}(H)$-viscosity solution to the corresponding dynamic programming equation, which here takes the form of a variational inequality with gradient constraint. Then, by allowing the decision maker to choose only the intensity of the control and requiring that the given control direction $\hat{n}$ is an eigenvector of the linear operator $\mathcal{A}$, we establish that the directional derivative $V_{\hat{n}}$ is of class $C^1(H)$, hence a second-order smooth-fit principle in the controlled direction holds for V. This result is obtained by exploiting a connection to optimal stopping and combining results and techniques from convex analysis and viscosity theory.

Keywords: infinite-dimensional singular stochastic control; viscosity solution; variational inequality; infinite-dimensional optimal stopping; smooth-fit principle.

MSC2020 subject classification: 93E20, 37L55, 35D40, 49J40, 60G40, 91B72.

1. Introduction

Singular control and optimal stopping problems arise frequently in Economics, Finance, Engineering, and related fields. Due to their inherent complexity, much analysis tends to focus on one-dimensional problems, where our understanding is relatively comprehensive. However, contemporary societal challenges present complex structures for singular control problems, that ask for a rigorous and sound mathematical basis. This paper introduces a framework for addressing singular control problems in the context of state processes governed by stochastic partial differential equations. By combining convex-analytic arguments and the theory of viscosity solutions, we show that the problem's value function V is a $C^{1,Lip}(H)$-viscosity solution to the corresponding dynamic programming equation. Furthermore, by exploiting a connection to a suitable family of simpler optimal stopping problems, we are able to further enhance regularity and prove that a second-order smooth-fit principle holds for V. Finally, we discuss potential applications in fields such as energy economics or climate modeling.

Let us describe the class of infinite-dimensional singular stochastic control problems and our contributions more precisely. Let $(D, \mathcal{M}, \mu)$ be a finite measure space, and consider the Hilbert space $H := L^2(D, \mathcal{M}, \mu; \mathbb{R})$. The state variable is described by a stochastic process X with values in H. Its evolution is determined by a stochastic partial differential equation (SPDE) driven by a self-adjoint linear operator $\mathcal{A}$ and a cylindrical Brownian motion. Next to technical requirements, we assume that the operator $\mathcal{A}$ generates a C_0-semigroup of positivity-preserving

Date: June 10, 2024.

contractions. As a benchmark case, one can consider the sum of the Laplacian operator and a multiplicative operator of the form $-\delta x$, for $\delta > 0$, representing a depreciation or dissipative term. The evolution of X is controlled linearly through an H-valued control, encompassing both direction and intensity of action, the latter being a real-valued nondecreasing right-continuous stochastic process adapted to the underlying filtration. The objective is to minimize a discounted convex cost-functional of the form

$$\mathcal{J}(x; I) := \mathsf{E}\bigg[\int_{0-}^{\infty} e^{-\rho t}\Big(G(X_t^{x,I})\mathrm{d}t + \langle q, \mathrm{d}I_t\rangle_H\Big)\bigg], \quad (x, I) \in H \times \mathcal{I}, \tag{1.1}$$

where $X^{x,I}$ is the state process starting at x and controlled via $I \in \mathcal{I}$ (cf. (2.12) below), G is a running cost function (see Assumption 2.8 below), $q \in H$ is a (stritly positive) proportional cost of action, and $\rho > 0$ is an intertemporal discount rate. The problem under study can thus be thought of as the infinite-dimensional version of the monotone follower problems addressed in [42, 43, 44], among others.

Our analysis begins by establishing preliminary regularity properties of the problem's value function, denoted as V. Specifically, assuming that the running cost function G is convex and semiconcave (with respect to the norm of H), we demonstrate that these properties are inherited by V. Thus, by adapting results from [14] to our infinite-dimensional setting, we find that $V \in C^{1,\mathrm{Lip}}(H)$.

We proceed by deriving the dynamic programming equation associated with (1.1) and demonstrate that V is a $C^{1,\mathrm{Lip}}(H)$-viscosity solution to it. The proof of this result relies on a Dynkin's formula for test functions of semimartingales and utilizes an equivalent representation of V, derived from a tailored application of the Radon-Nikodym theorem for vector-valued measures (see Lemma 2.11 and (2.19) below). It is noteworthy that the proof of the supersolution property of V employs a novel argument derived from an inequality stemming from Dynkin's formula and the dissipativity property of the operator $\mathcal{A}$. This approach is applicable in finite-dimensional settings as well, substantially reducing the technicalities usually associated with demonstrating the supersolution (or subsolution) property in minimization (or maximization) problems involving singular controls (see, e.g., [16], [34, Ch. VIII], [38], and [49], among others).

To further enhance the regularity of V, we introduce the assumption that the decision-maker can only control the intensity of action. The direction of action, denoted by $\hat{n} \in H$, is then taken to be an eigenvector of the operator $\mathcal{A}$. Under this requirement and further technical properties of G, we are able to show that the directional derivative of V in the direction of control, $V_{\hat{n}}$, is such that $V_{\hat{n}} \in C^1(H)$. This result can be read as a second-order smooth fit property of V, a regularity result of particular relevance in singular stochastic control problems (see the discussion in Section 5.1 below). The aforementioned smooth-fit property is obtained by identifying $V_{\hat{n}}$ as the value function of an optimal stopping problem (in the spirit of the finite-finite dimensional contribution [4]) and subsequently examining the regularity of its (sub)gradient. In particular, under a suitable nondegeneracy condition on the Brownian noise, assuming that the directional derivative $G_{\hat{n}}$ is semiconcave, and combining arguments from viscosity theory and convex analysis, we are able to show that $V_{\hat{n}}$ is Fréchet differentiable at any $x \in H$ and that the gradient $DV_{\hat{n}} \in C(H; \mathcal{D}(\mathcal{A}))$ (with the domain $\mathcal{D}(\mathcal{A})$ being endowed with the graph norm). For further details, please refer to Proposition 4.17.

In Section 6, we demonstrate the relevance of our framework in economic applications. For instance, we examine an irreversible investment problem in energy capacity and an energy balance climate model incorporating human impact. In the former, an energy producer seeks to maximize the net total expected surplus resulting from irreversible investments in energy production. In the latter, temperature is increased by human activities through carbon emissions and a social planner aims to minimize an intertemporal expected cost criterion, penalizing temperature deviations from an ideal level, such as pre-industrial temperatures.

Let us now discuss related literature and our contribution to it. The origin of singular stochastic control dates back to the early contributions by Bather and Chernoff [6] , and later by Beneš, Shepp, and Witsenhausen [7] and Karatzas [42, 43]. Those seminal papers deal with one-dimensional problems of so-called *monotone follower* type, in which a process with monotone paths (or, more generally, of bounded-variation) has to be chosen in order to track the evolution of a Brownian motion so that an expected cost criterion is minimized. Since then, the theory of singular stochastic control has attracted increasing attention, also boosted by its connection to optimal stopping (see [4], [9], and [44], just to cite a few) and its numerous applications in Economics and Finance. Among those, problems of optimal capacity expansion [4], optimal investment with transaction costs [58], optimal harvesting [2], and optimal dividends' distribution [41].

For stationary one-dimensional problems, or for two-dimensional degenerate problems with a suitable structure [31, 51], explicit solutions can be expected. Typically, these solutions are obtained through the guess-and-verify approach. This involves first determining a smooth solution to the problem's dynamic programming equation (in this case, a variational inequality with gradient constraints), and then verifying its optimality using a version of Itô's formula. Additionally, an optimal control is determined as a byproduct of this analysis. This is given in terms of the solution to a Skorokhod reflection problem at the so-called free boundary, i.e. the topological boundary of the region in which the gradient constraint is not active (the so-called no-action or continuation region).

For time-dependent problems or for stationary problems in dimension larger than one, the guess-and-verify approach is not feasible. This is because the dynamic programming equation now becomes a partial differential equation (PDE) with gradient constraints, for which explicit solutions are typically not available. As a consequence, direct probabilistic and analytical approaches are put in place in order to obtain regularity of the value function (typically under convexity requirements; see, e.g., [39, 40, 50, 57]) and, when possible, to characterize the optimal control as the minimal amount of effort needed to keep the underlying state process within the no-action region (see [23], [46], and references therein). As a matter of fact, differently to before, in multi-dimensional settings, the free boundary is not explicit and constructing a solution to the related Skorokhod reflection problem is far from trivial. We refer to the introduction of [23] for a discussion on this aspect. The aforementioned challenges explain why the number of contributions on singular stochastic control problems in multi-dimensional settings is still very limited.

The theory of regular stochastic control and of optimal stopping in infinite-dimensional (notably, Hilbert) spaces received a large attention in the last decades (see, e.g., the monography [27] for control problems, and [5], [18], [19], [29], [35], [36], [59] for optimal stopping). As previously discussed, we contribute to that bunch of literature by providing the viscosity property and C^1-regularity (smooth-fit) of the value function of a class of optimal stopping problems in Hilbert spaces. To the best of our knowledge, such a regularity result appears here for the first time, and it is therefore of independent interest.

On the other hand, the literature on singular stochastic control in infinite-dimensional spaces is very limited. The only three papers brought to our attention are [1] and [54], motivated by optimal harvesting, and ours [28]. In [54] the problem is posed for a quite general controlled SPDE, which also enjoys a space-mean dependence in [1]. The authors establish a necessary Maximum Principle, which is also sufficient assuming the concavity of the Hamiltonian function pertaining to the control problem under consideration. However, despite their significant contributions, there appears to be a foundational concern when dealing with (singularly controlled) SPDEs, particularly regarding the existence of a solution and the application of Itô's formula (refer to [47] for theory and results on SPDEs). Specifically, it is important to notice that in infinite-dimensional singular (stochastic) control problems, the precise interpretation of the integral with respect to the vector measure represented by the control process - and thus the exact interpretation of the controlled state equation - poses a nuanced issue that warrants careful consideration. Finally, our previous work [28] derives necessary and sufficient conditions for a class of singular stochastic control problems on an abstract partially ordered infinite-dimensional space. The main differences with respect to the present work are in the framework, the methodology, and the nature of the results. In [28], the controlled state process is fully degenerate and randomness comes into the problem only in a parametric form, thus making the underlying optimization problem not necessarily Markovian. Furthermore, the main result is obtained by the exclusive mean of convex analytic arguments, and no statement about the regularity of the value function is made. In this work, we deal with a singularly controlled SPDE and exploit the dynamic programming approach together with viscosity theory and convex analysis in order to achieve regularity results on the problem's value function.

Organization of the Paper. The rest of the paper is organized as follows. In Section 2 we provide the setting and introduce the problem. In Section 3 we then consider the variational inequality associated to the problem and prove preliminary regularity and viscosity property of its value function V. Under a suitable requirement on the direction of action, in Section 4 a connection to optimal stopping is derived and regularity of the optimal stopping problem's value function is proved. As a byproduct of that, in Section 5 a second-order smooth-fit property for V is then obtained. Finally, Section 6 proposes two applications in Economics, while Appendix A collects a result on semiconcave and semiconvex functions and Appendix B technical lemmata.

2. Setting and Problem Formulation

2.1. **Setting.** Let $(D, \mathcal{M}, \mu)$ be a finite standard Borel measure space and assume, without loss of generality for what follows, that $\mu(D) = 1$. Consider the separable Hilbert space

$$H := L^2(D, \mathcal{M}, \mu; \mathbb{R}).$$

The dual H^* is identified with H via the classical Riesz representation of H^*. The nonnegative cone of H is denoted by

$$H_+ := \{x \in H : \ x \geq 0\}.$$

We denote by $\mathcal{L}(H)$ the space of linear bounded operators on H and by $\mathcal{L}^+(H)$ the subspace of positivity-preserving operators of $\mathcal{L}(H)$; i.e., $P \in \mathcal{L}^+(H)$ if

$$x \in H_+ \implies Px \in H_+.$$

Throughout the paper, we consider a linear operator $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subseteq H \to H$ satisfying the following standing requirements.

Assumption 2.1. *$\mathcal{A}$ is self-adjoint, closed, densely defined, and such that, for some $\delta > 0$, we have*

$$\langle \mathcal{A}x, x\rangle_H \leq -\delta |x|_H^2, \quad \forall x \in H.$$

In particular (see, e.g., [26, Ch. II, Sec. 3] and [8, Ch. II-1, Sec. 2.10.1]), under Assumption 2.1, the operator $\mathcal{A}$ generates a C_0-semigroup of contractions $(e^{t\mathcal{A}})_{t\geq 0} \subseteq \mathcal{L}(H)$ and

$$|e^{t\mathcal{A}}|_{\mathcal{L}(H)} \leq e^{-\delta t}, \quad \forall t \geq 0.$$

Moreover, $0 \in \varrho(\mathcal{A})$ – with ϱ denoting the resolvent set – so that $\mathcal{A}$ is invertible and

$$\mathcal{A}^{-1} \in \mathcal{L}(H).$$

We also assume the following.

Assumption 2.2. *The C_0-semigroup of contractions $(e^{t\mathcal{A}})_{t\geq 0} \subseteq \mathcal{L}(H)$ is positivity-preserving; that is, $(e^{t\mathcal{A}})_{t\geq 0} \subseteq \mathcal{L}^+(H)$.*

Remark 2.3. *Sufficient conditions guaranteeing positivity of semigroups can be found, e.g., in* [3, Chap. C-II, Thm. 1.2, Thm. 1.8] *and* [20, Thm. 7.29 and Prop. 7.46]

Furthermore, we impose the next assumption.

Assumption 2.4. $\mathcal{D}(\mathcal{A}) \hookrightarrow L^\infty(D, \mathcal{M}, \mu; \mathbb{R})$.

Remark 2.5. *In the examples considered in Section 6, $\mathcal{D}(\mathcal{A})$ will be the Sobolev space $W^{2,2}(\mathcal{O})$ with appropriate boundary conditions, for $\mathcal{O}$ being an open, simply connected, and bounded set of $\mathbb{R}^n$ ($n < 4$) with smooth boundary. Assumption 2.4 is then verified in this setting by* [12, Cor. 9.15].

Let us now come to the probabilistic structure of our setup. We endow the time-interval $[0,\infty)$ with the Borel σ-algebra $\mathcal{B}([0,\infty))$. Also, let $(\Omega, \mathcal{F}, \mathbb{F}, \mathsf{P})$ be a filtered probability space, with filtration $\mathbb{F} := (\mathcal{F}_t)_{t\in[0,\infty)}$ satisfying the usual conditions, and let W be a cylindrical Wiener process on $(\Omega, \mathcal{F}, \mathbb{F}, \mathsf{P})$, taking values in another Hilbert space K. Finally, for future use, we denote by $\mathcal{T}$ the set of all $\mathbb{F}$-stopping times.

In the following, all the relationships involving $\omega \in \Omega$ as hidden random parameter are intended to hold P-almost surely. Also, in order to simplify the exposition, often we will not stress the explicit dependence of the involved random variables and processes with respect to $\omega \in \Omega$.

Let $\Delta \subseteq H_+$ be a convex cone of H_+ and set

$$\begin{aligned}\mathcal{M} \;&:=\; \big\{I : \Omega \times [0,\infty) \to H_+ : I.\ \text{is } \mathbb{F}-\text{adapted and such that } t \mapsto I_t \\ &\qquad \text{is càdlàg and with } I_t - I_{s-} \in \Delta \ \ \forall s,t \in [0,\infty) \text{ such that } t \geq s\big\}.\end{aligned} \tag{2.1}$$

Notice that, since any $I \in \mathcal{M}$ takes values in H_+, right-continuity is intended in the norm of H. In the following, we set $I_{0^-} := \mathbf{0} \in H_+$ for any $I \in \mathcal{M}$ (see Remark 2.6 below).

Any given $I \in \mathcal{M}$ can be seen as a (random) countably additive vector measure

$$I : \mathcal{B}([0,\infty)) \to H_+$$

of local finite variation, defined as

$$I([s,t]) := I_t - I_{s^-} \quad \forall s,t \in [0,\infty),\ t \geq s.$$

We denote by $|I|$ the variation of I; it is a nonnegative (optional random) measure on $([0,\infty), \mathcal{B}([0,\infty)))$ that, due to monotonicity of I, can be simply expressed as

$$|I|([s,t]) = |I_t - I_{s^-}|_H, \quad \forall s,t \in [0,\infty),\ s \leq t. \tag{2.2}$$

Remark 2.6. *By setting $I_{0^-} := 0$ for any $I \in \mathcal{M}$, we mean that we extend any $I \in \mathcal{M}$ by setting $I \equiv 0$ on $[-\varepsilon, 0)$, for a given and fixed $\varepsilon > 0$. In this way, the associated measures have a positive mass at initial time of size I_0. Notice that this is equivalent with identifying any control I with a countably additive measure $I : \mathcal{B}([0,\infty)) \to [0,\infty)$ of local finite variation defined as $I((s,t]) := I_t - I_s$, for every $s,t \in [0,\infty)$, $s < t$, plus a Dirac-delta at time 0 of amplitude I_0.*

Since H is a reflexive Banach space, by [24], Corollary 13 at p. 76 (see also Definition 3 at p. 61), there exists a Bochner measurable function $\hat{\vartheta} = \hat{\vartheta}(\omega) : [0,\infty) \to H_+$ such that

$$\int_{[0,T]} |\hat{\vartheta}_t|_H \mathrm{d}|I|_t < \infty \quad \forall T > 0 \quad \text{and} \quad \mathrm{d}I_t = \hat{\vartheta}_t\, \mathrm{d}|I|_t \quad \forall t \geq 0. \tag{2.3}$$

Notice that, seen as a stochastic process, $\hat{\vartheta} = (\hat{\vartheta}_t)_{t\geq 0}$ is $\mathbb{F}$-adapted, because so is I. Furthermore, given that the measures I and $|I|$ are equivalent by (2.2), one has

$$\hat{\vartheta}_t \neq \mathbf{0} \quad \text{for a.e. } t \geq 0. \tag{2.4}$$

The process $\hat{\vartheta}$ is clearly unique up to $\mathsf{P} \times |I|-$null measure sets.

Then, for a given H_+-valued $\mathbb{F}$-adapted process $f := (f_t)_{t\in[0,\infty)}$, recalling (2.3), for any $t \in [0,\infty)$ we define

$$\begin{aligned}\int_{0^-}^t \langle f_s, \mathrm{d}I_s\rangle_H &:= \int_{[0,t]} \langle f_s, \hat{\vartheta}_s\rangle_H \, \mathrm{d}|I|_s = \int_{[0,t]} \Big(\int_D f_s(\xi)\hat{\vartheta}_s(\xi)\mu(\mathrm{d}\xi)\Big)\mathrm{d}|I|_s \\ &= \int_D \Big(\int_{[0,t]} f_s(\xi)\hat{\vartheta}_s(\xi)\mathrm{d}|I|_s\Big)\mu(\mathrm{d}\xi),\end{aligned} \tag{2.5}$$

where the last step is possible due to Fubini-Tonelli's theorem. With regard to (2.3), we also set

$$\int_{0^-}^t e^{(t-s)\mathcal{A}}\mathrm{d}I_s := \int_{0^-}^t e^{(t-s)\mathcal{A}}\hat{\vartheta}_s\mathrm{d}|I|_s, \quad t \geq 0. \tag{2.6}$$

Thanks to (2.6), for any given $I \in \mathcal{M}$, we can then introduce the singularly continuous controlled dynamics

$$\mathrm{d}X_t^{x,I} = \mathcal{A}X_t^{x,I}\mathrm{d}t + \sigma\mathrm{d}W_t + \mathrm{d}I_t, \quad t \geq 0, \quad X_{0^-}^{x,I} = x \in H, \tag{2.7}$$

and define the unique mild solution to (2.7) as

$$X_t^{x,I} = e^{t\mathcal{A}}x + W_t^{\mathcal{A},\sigma} + \int_{0^-}^t e^{(t-s)\mathcal{A}}\mathrm{d}I_s, \quad t \geq 0. \tag{2.8}$$

Here,

$$W_t^{\mathcal{A},\sigma} := \int_0^t e^{(t-s)\mathcal{A}}\sigma\mathrm{d}W_s, \quad t \geq 0, \tag{2.9}$$

with σ satisfying the following standing condition.

Assumption 2.7. *$\sigma \in \mathcal{L}_2(K;H)$, where $\mathcal{L}_2(K;H)$ denotes the space of Hilbert-Schmidt operators from K to H.*

Denoting by $\mathcal{L}_1(H)$ the set of nonnegative trace-class operators on H and endowing $\mathcal{L}_1(H)$ with the usual norm

$$|Q|_{\mathcal{L}_1(H)} := \mathrm{Tr}[Q] = \sum_{k=0}^{\infty}\langle Qe_k, e_k\rangle_H,$$

where (e_k) is any orthonormal basis of H, we then have under Assumption 2.7 that

$$\sigma\sigma^* \in \mathcal{L}_1(H).$$

Notice that Assumptions 2.1 and 2.7 imply that the stochastic convolution (2.9) is well defined and continuous (see [21, Ch. 5]). For future use, we also note that, because of Assumptions 2.1 and 2.7, for all $m \in [1,\infty)$ one has for some $\bar{c}_m > 0$ (see [37])

$$\mathsf{E}\left[\sup_{t\geq 0} |W_t^{\mathcal{A},\sigma}|_H^m\right] \leq \bar{c}_m, \tag{2.10}$$

which, denoting the mild solution to (2.7) when I is the null control by $X_t^{x,0}$, implies

$$\mathsf{E}\left[\sup_{t\geq 0} |X_t^{x,0}|_H^m\right] \leq c_p(1 + |x|_H^m), \quad \forall x \in H, \tag{2.11}$$

for some other constant $c_m > 0$.

2.2. **Problem formulation.** We now move on by introducing the infinite-dimensional singular stochastic control problem which is the object of our study. Let

$$G : H \to \mathbb{R}$$

be a running cost function, satisfying the following requirements.

Assumption 2.8.
(i) G is convex; There exists $c_o, \kappa_1, \kappa_2 > 0$ and $p \geq 2$ such that

$$\kappa_1 |x|_H^p - \kappa_2 \leq G(x) \leq c_o(1 + |x|_H^p);$$

(ii) G is semiconcave with semiconcavity constant $c_1 > 0$; that is, there exists $c_1 > 0$ such that

$$\lambda G(x) + (1-\lambda)G(y) - G(\lambda x + (1-\lambda)y) \leq \frac{c_1}{2}\lambda(1-\lambda)|x-y|_H^2, \quad \forall x, y \in H, \ \lambda \in [0,1].$$

Notice that by Lemma A.1(ii) one has that $G \in C^{1,\mathrm{Lip}}(H)$.

Remark 2.9. *Benchmark examples satisfying Assumption 2.8 are the quadratic cost function*

$$G(x) = \frac{1}{2}|x - \overline{x}|_H^2, \qquad x \in H,$$

for some target level $\overline{x} \in H$, as well as

$$G(x) = \frac{1}{2}\langle x, h\rangle_H^2, \quad \text{or} \quad G(x) = \frac{1}{2}\langle Qx, x\rangle_H, \quad x \in H,$$

with $h \in H$, and with Q being positive semidefinite and symmetric.

Let now (cf. (2.1))

$$\begin{aligned} \mathcal{I} \ \ :=& \ \ \Big\{ I \in \mathcal{M} : \ \mathsf{E}\Big[\Big| \int_{0^-}^T |\hat{\vartheta}_s|_H \, \mathrm{d}|I|_s \Big|^p \Big] < \infty \text{ for } p \geq 2 \text{ as in} \\ & \qquad \text{Assumption 2.8 and } \forall T > 0 \Big\} \end{aligned} \tag{2.12}$$

be the class of admissible controls. For a discount rate $\rho > 0$ and for $q \in H_+$ such that $q \geq q_o\mathbf{1}$ for some $q_o > 0$ (being $\mathbf{1} \in H$ the constant unitary vector of H), recalling (2.5) we introduce the expected cost functional

$$\mathcal{J}(x; I) := \mathsf{E}\Big[\int_{0^-}^\infty e^{-\rho t}\Big(G(X_t^{x,I})\mathrm{d}t + \langle q, \mathrm{d}I_t\rangle_H \Big) \Big], \quad (x, I) \in H \times \mathcal{I}, \tag{2.13}$$

which is well-defined, although potentially infinite. The infinite-dimensional singular stochastic control problem under study is then

$$V(x) := \inf_{I \in \mathcal{I}} \mathcal{J}(x; I), \quad x \in H. \tag{2.14}$$

Remark 2.10. *Notice that the integrability condition in (2.12) is not required for the well posedness of (2.13), but it will be needed in the next section for the proof of the viscosity property of V.*

Given the structure of the cost functional (2.13), it is convenient to rewrite the decomposition (2.3) in a tailored way based on the instantaneous cost of control $\langle q, \mathrm{d}I_t\rangle_H$. To that end, recall that $\Delta \subseteq H_+$ is a convex cone of H_+ (cf. (2.12)), consider the convex set

$$\Theta := \big\{ \theta \in \Delta : \ \langle q, \theta\rangle_H = 1 \big\}, \tag{2.15}$$

and define

$$\begin{aligned} \mathcal{S} \ \ :=& \ \ \{ \nu : \Omega \times [0,\infty) \to [0,\infty) : \ \nu_\cdot \text{ is } \mathbb{F}-\text{adapted and such that } t \mapsto \nu_t \\ & \qquad \text{is càdlàg and nondecreasing} \}. \end{aligned} \tag{2.16}$$

In the sequel, we set $\nu_{0^-} := 0$ for any $\nu \in \mathcal{S}$ (see also Remark 2.6). Then, define

$$\begin{aligned}\mathcal{I}_0 \;:=\; & \Big\{(\vartheta,\nu) : \Omega\times[0,\infty) \to \Theta\times[0,\infty) : \; \vartheta_\cdot \text{ is } \mathbb{F}-\text{adapted}, \; \nu_\cdot \in \mathcal{S} \text{ and} \\ & \mathsf{E}\Big[\Big|\int_{0^-}^T |\vartheta_s|_H \,\mathrm{d}\nu_s\Big|^p\Big] \text{ for } p \geq 2 \text{ as in (2.12) and } \forall T>0\Big\}.\end{aligned} \tag{2.17}$$

Lemma 2.11. *For each $I \in \mathcal{I}$, there exists a couple $(\vartheta,\nu) \in \mathcal{I}_0$, with $\nu \sim |I|$, such that*

$$\mathrm{d}I_t = \vartheta_t \,\mathrm{d}\nu_t, \qquad \forall t \geq 0. \tag{2.18}$$

This couple is unique in the following sense: the optional random measure ν is unique and ϑ is unique up to $\mathsf{P}\otimes\nu-$null measure sets.

Proof. Due to (2.3), positivity of $\hat{\vartheta}$ and the fact that $q \geq q_o\mathbf{1}$, for some $q_o > 0$, we can write for any $t \geq 0$

$$\mathrm{d}I_t = \hat{\vartheta}_t \mathrm{d}|I|_t = \frac{\hat{\vartheta}_t}{\langle q, \hat{\vartheta}_t\rangle_H}\langle q, \hat{\vartheta}_t\rangle_H \mathrm{d}|I|_t = \vartheta_t\,\mathrm{d}\nu_t,$$

where

$$\vartheta_t := \frac{\hat{\vartheta}_t}{\langle q, \hat{\vartheta}_t\rangle_H}, \quad \text{and} \quad \mathrm{d}\nu_t := \langle q, \hat{\vartheta}_t\rangle_H \mathrm{d}|I|_t.$$

For $p \geq 2$ as in Assumption 2.8 (see also (2.12)), one clearly has that the integrability conditions required in (2.17) are met, because of the previous definitions of ϑ and $\mathrm{d}\nu$, and because $I \in \mathcal{I}$. This shows the first part of the claim.

Let us prove uniqueness. Assume that

$$\mathrm{d}I_t = \vartheta_t^{(1)}\mathrm{d}\nu_t^{(1)} = \vartheta_t^{(2)}\mathrm{d}\nu_t^{(2)}.$$

Then, for all $0 \leq a \leq b$,

$$\int_{[a,b]} \langle q, \vartheta_t^{(1)}\rangle_H \mathrm{d}\nu_t^{(1)} = \int_{[a,b]} \langle q, \vartheta_t^{(2)}\rangle_H \mathrm{d}\nu_t^{(2)},$$

implying, by definition of Θ,

$$\int_{[a,b]} \mathrm{d}\nu_t^{(1)} = \int_{[a,b]} \mathrm{d}\nu_t^{(2)}.$$

Hence, $\nu^{(1)} = \nu^{(2)} =: \nu$. Then, for all $0 \leq a \leq b$,

$$\int_{[a,b]} \mathrm{d}I_t = \int_{[a,b]} \vartheta_t^{(1)}\mathrm{d}\nu_t = \int_{[a,b]} \vartheta_t^{(2)}\mathrm{d}\nu_t,$$

implying $\vartheta^{(1)} = \vartheta^{(2)}$ up to $\mathsf{P}\otimes\nu$-null measure sets. □

Thanks to Lemma 2.11, we may identify $\mathcal{I}$ with $\mathcal{I}_0$ (cf. (2.12) and (2.17), respectively). Hence, hereafter, with a slight abuse of notation, we will often identify elements of the above sets. The cost functional (2.19) then rewrites as

$$\mathcal{J}(x;I) = \mathsf{E}\Big[\int_{0^-}^\infty e^{-\rho t}\Big(G(X_t^{x,I})\mathrm{d}t + \mathrm{d}\nu_t\Big)\Big], \qquad x \in H, \; I = (\vartheta,\nu) \in \mathcal{I}_0, \tag{2.19}$$

and the value function as

$$V(x) = \inf_{(\vartheta,\nu)\in\mathcal{I}_0} \mathcal{J}(x;I), \quad x \in H. \tag{2.20}$$

In the next Section 3, we will make use of both the equivalent representations (2.14) and (2.20). In particular, preliminary regularity properties of V (see Section 3.1 below) as well as Propositions 3.3 and 3.4 in Section 3.2 will be shown by using (2.14), while the viscosity property of V will be proved through (2.20) (see Proposition 3.5 and Theorem 3.7 in Section 3.2).

3. Regularity and Viscosity Property of V

In this section we first show via direct convex-analytic arguments that V is convex, has subquadratic growth and it is such that $V \in C^{1,\mathrm{Lip}}(H)$. Then, we prove that V is a viscosity solution to the associated Hamilton-Jacobi-Bellman equation, which in the present setting takes the form of a variational inequality with gradient constraint.

3.1. Preliminary properties of V.

Here we provide some a priori regularity properties of the value function $V : H \to \mathbb{R}$ as in (2.14).

Proposition 3.1.

(i) V is convex;

(ii) There exists $\hat{c}_o > 0$ such that, for $p \geq 2$ and $\kappa_2 > 0$ as in Assumption 2.8,

$$-\kappa_2 \leq V(x) \leq \hat{c}_o(1 + |x|_H^p);$$

(iii) V is locally Lipschitz;

(iv) V is semiconcave with semiconcavity constant $\hat{c}_1$; that is, there exists $\hat{c}_1 > 0$ such that

$$\lambda V(x) + (1-\lambda)V(y) - V(\lambda x + (1-\lambda)y) \leq \frac{\hat{c}_1}{2}\lambda(1-\lambda)|x-y|_H^2, \quad \forall x, y \in H,\ \lambda \in [0,1];$$

(v) $V \in C^{1,Lip}(H)$.

Proof. We prove each item separately.

Proof of (i). For $i = 1, 2$, let $x_i \in H$ and let $I^{(i)}$ be ε-optimal for the initial data x_i, for some $\varepsilon > 0$; that is,

$$V(x_i) + \varepsilon \geq \mathcal{J}(x_i; I^{(i)}).$$

For $\lambda \in [0,1]$, define

$$x_\lambda := \lambda x_1 + (1-\lambda)x_2, \qquad I^{(\lambda)} := \lambda I^{(1)} + (1-\lambda)I^{(2)}.$$

Given that the mapping $(x, I) \mapsto X^{x,I}$ is linear (see (2.9) and (2.8)), we have

$$X^{x_\lambda, I^{(\lambda)}} = \lambda X^{x_1, I^{(1)}} + (1-\lambda)X^{x_2, I^{(2)}}.$$

Then, using the convexity of G, we write from (2.14)

$$\begin{aligned}
&V(x_\lambda) \leq \mathcal{J}(x_\lambda; I^{(\lambda)}) = \mathsf{E}\Big[\int_{0^-}^\infty e^{-\rho t}\Big(G(X_t^{x_\lambda, I^{(\lambda)}})\mathrm{d}t + \langle q, \mathrm{d}I_t^{(\lambda)}\rangle_H\Big)\Big] \\
&\leq \lambda \mathsf{E}\Big[\int_{0^-}^\infty e^{-\rho t}\Big(G(X_t^{x_1, I^{(1)}})\mathrm{d}t + \langle q, \mathrm{d}I_t^{(1)}\rangle_H\Big)\Big] + (1-\lambda)\mathsf{E}\Big[\int_{0^-}^\infty e^{-\rho t}\Big(G(X_t^{x_2, I^{(2)}})\mathrm{d}t + \langle q, \mathrm{d}I_t^{(2)}\rangle_H\Big)\Big] \\
&= \lambda \mathcal{J}(x_1; I^{(1)}) + (1-\lambda)\mathcal{J}(x_2; I^{(2)}) \leq \lambda V(x_1) + (1-\lambda)V(x_2) + \varepsilon.
\end{aligned}$$

By arbitrariness of ε, the claim follows.

Proof of (ii). The bound from below is immediate given that $G \geq -\kappa_2$ on H, $q \in H_+$, and $I \in \mathcal{I}$. As for the bound from above, recall that $X^{x,0}$ denote the uncontrolled mild solution to

(2.7). Then, by Assumption 2.8(ii), (2.9) and the fact that $|e^{t\mathcal{A}}|_{\mathcal{L}(H)} \leq e^{-\delta t}|x|_H$, we can write

$$\begin{aligned} V(x) \leq \mathcal{J}(x;0) &= \mathsf{E}\left[\int_0^\infty e^{-\rho t} G(X_t^{x,0})\mathrm{d}t\right] \leq c_0 \int_0^\infty e^{-\rho t}\left(1 + \mathsf{E}[|X_t^{x,0}|_H^p]\mathrm{d}t\right) \\ &\leq c_0\left(\frac{1}{\rho} + 2^{p-1}\int_0^\infty e^{-\rho t}\left(|e^{t\mathcal{A}}x|_H^p + \mathsf{E}[|W_t^{\sigma,\mathcal{A}}|_H^p]\right)\mathrm{d}t\right) \\ &\leq c_0\left(\frac{1}{\rho} + 2^{p-1}\int_0^\infty e^{-(\rho+p\delta)t}|x|_H^p\mathrm{d}t + 2^{p-1}\frac{1}{\rho}\overline{c}_p\right) \\ &= c_0\left(\frac{1}{\rho}(1+2^{p-1}\overline{c}_p) + \frac{2^{p-1}}{\rho+p\delta}|x|_H^p\right), \end{aligned}$$

and the claim is proved.

Proof of (iii). It follows from the previous items and, e.g., [25, Cor. 2.4, p. 12].

Proof of (iv). Let $x, y \in H$ and $\lambda \in [0,1]$, and take an ε-optimal control $I^\varepsilon \in \mathcal{I}$ for $\lambda x + (1-\lambda)y$. Since the state equation is affine (cf. (2.8)), we have

$$X^{\lambda x+(1-\lambda)y, I^\varepsilon} = \lambda X^{x,I^\varepsilon} + (1-\lambda)X^{y,I^\varepsilon},$$

and the semiconcavity of G allows to write (cf. (2.13))

$$\begin{aligned} &\lambda V(x) + (1-\lambda)V(y) - V(\lambda x + (1-\lambda)y) \\ &\leq \lambda \mathcal{J}(x;I^\varepsilon) + (1-\lambda)\mathcal{J}(y;I^\varepsilon) - J(\lambda x + (1-\lambda)y; I^\varepsilon) + \varepsilon \\ &= \mathsf{E}\left[\int_0^\infty e^{-\rho t}\left(\lambda G(X_t^{x,I^\varepsilon}) + (1-\lambda)G(X_t^{y,I^\varepsilon}) - G(\lambda X_t^{x,I^\varepsilon} + (1-\lambda)X_t^{y,I^\varepsilon}\right)\mathrm{d}t\right] + \varepsilon \\ &= \mathsf{E}\left[\int_0^\infty e^{-\rho t}\frac{c_1}{2}\lambda(1-\lambda)|X_t^{x,I^\varepsilon} - X_t^{y,I^\varepsilon}|^2\mathrm{d}t\right] + \varepsilon = \frac{c_1}{2}\lambda(1-\lambda)\int_0^\infty e^{-\rho t}\left|e^{t\mathcal{A}}(x-y)\right|_H^2\mathrm{d}t + \varepsilon \\ &\leq \frac{c_1}{2}\lambda(1-\lambda)\int_0^\infty e^{-\rho t}e^{-2\delta t}|x-y|_H^2\mathrm{d}t + \varepsilon = \frac{c_1}{2(\rho+2\delta)}\lambda(1-\lambda)|x-y|_H^2 + \varepsilon. \end{aligned}$$

By arbitrariness of ε, the claim follows.

Proof of (v). This follows by Lemma A.1(ii). □

Note that, since V is finite on H and G is uniformly bounded from below on H, we can restrict the optimization (2.20) to the class of controls

$$\hat{\mathcal{I}}_0 := \left\{ I = (\vartheta, \nu) \in \mathcal{I}_0 : \ \mathsf{E}\left[\int_{0^-}^\infty e^{-\rho t}\mathrm{d}\nu_t\right] < \infty \right\}. \tag{3.1}$$

3.2. **Dynamic programming equation and viscosity solutions.** In this section we show that V is a viscosity solution to the dynamic programming equation associated to (2.20). To that end, we consider the variational inequality with gradient constraint

$$\max\left\{(\rho - \mathcal{G})v(x) - G(x),\ \sup_{\theta\in\Theta}\left\{-\langle Dv(x), \theta\rangle_H - 1\right\}\right\} = 0, \quad x \in H, \tag{3.2}$$

where Θ is as in (2.15) and where we have defined the second-order differential operator $\mathcal{G}$, formally associated to the process X^0 and acting on sufficiently smooth functions $v : H \to \mathbb{R}$, as

$$[\mathcal{G}v](x) := \langle \mathcal{A}x, Dv(x)\rangle_H + \frac{1}{2}\mathrm{Tr}\left[\sigma\sigma^* D^2 v(x)\right]. \tag{3.3}$$

Recall that $\mathcal{D}(\mathcal{A}) \subseteq H$ denotes the domain of the operator $\mathcal{A}$ and endow it with the graph norm $|\cdot|_{\mathcal{D}(\mathcal{A})}$; that is,

$$|x|_{\mathcal{D}(\mathcal{A})}^2 := |x|_H^2 + |\mathcal{A}x|_H^2, \qquad x \in \mathcal{D}(\mathcal{A}). \tag{3.4}$$

Hence, in order to provide the definition of viscosity solution to (3.2), we first introduce the suitable class of test functions, and we then have a Dynkin's formula for those functions (see Proposition 3.3 below). For a given but arbitrary $\widehat{C} > 0$, let

$$\mathcal{X} := \Big\{ \varphi \in C^2(H) : \ D\varphi \in C(H; \mathcal{D}(\mathcal{A})), \ \ \sigma\sigma^* D^2\varphi \in C(H; \mathcal{L}_1(H)), \ |\varphi(x)| \le \widehat{C}(1 + |x|_H^2),$$

$$|D\varphi(x)|_H + |\mathcal{A}D\varphi(x)|_H + \ |\sigma\sigma^* D^2\varphi(x)|_{\mathcal{L}_1(H)} \le \widehat{C} \ \ \forall x \in H \Big\}. \tag{3.5}$$

Remark 3.2. *In the context of classical stochastic control, the set $\mathcal{X}$, combined with the set of radial functions, is typically large enough to prove comparison results for viscosity solutions of the PDE (see* [27, Ch. 3]*), and therefore to characterize the value function of the control problem as the unique solution to the associated HJB equation.*

In this paper, we do not address the (relevant) topic of uniqueness (see also Remark 3.9 below). However, we will use test functionw from the set $\mathcal{X}$ – of quadratic form – in Section 4 in order to prove our regularity results (see, in particular, the proof of Proposition 4.15).

In the rest of this paper, for any measurable function $f : H \to \mathbb{R}$, we set, for any $\tau \in \mathcal{T}$ and any $I \in \mathcal{I}$,

$$e^{-\rho\tau} f(X_\tau^{x,I}) := \limsup_{t\to\infty} e^{-\rho t} f(X_t^{x,I}) \quad \text{on } \{\tau = +\infty\}. \tag{3.6}$$

We then have the next Dynkin's formula for functions in $\mathcal{X}$. In the subsequent analysis we shall use that any $I \in \mathcal{I}$ is identified with an element $(\vartheta, \nu) \in \mathcal{I}_0$ (see Lemma 2.11 and the discussion afterwards) and that the optimization in (2.20) can be actually performed over $\hat{\mathcal{I}}_0$ (cf. (3.1)).

Proposition 3.3. *Let $x \in H$, $\varphi \in \mathcal{X}$, $I \in \hat{\mathcal{I}}_0$, and $\tau \in \mathcal{T}$. Assume that*

$$\mathsf{E}\left[\int_{0^-}^{\tau} e^{-\rho t} \left| \langle D\varphi(X_t^{x,I}), \vartheta_t \rangle_H \right| \mathrm{d}\nu_t \right] < \infty. \tag{3.7}$$

Then, the following Dynkin's formula holds true:

$$\mathsf{E}\left[e^{-\rho\tau}\varphi(X_\tau^{x,I})\right] = \varphi(x) + \mathsf{E}\left[\int_0^\tau e^{-\rho t}[(\mathcal{G} - \rho)\varphi](X_t^{x,I})\mathrm{d}t + \int_{0^-}^\tau e^{-\rho t}\langle D\varphi(X_t^{x,I}), \vartheta_t\rangle_H \mathrm{d}\nu_t\right]. \tag{3.8}$$

Proof. Let $(\mathcal{A}_n)_{n\in\mathbb{N}}$ be the Yosida approximants of $\mathcal{A}$ (see, e.g., [26, Eq. (3.7), Ch. II]), and let $X^{n;x,I}$ denote the solution to

$$\mathrm{d}X_t^{n;x,I} = \mathcal{A}_n X_t^{n;x,I}\mathrm{d}t + \sigma \mathrm{d}W_t + \mathrm{d}I_t, \quad X_{0^-} = x \in H;$$

that is (cf. (2.8)),

$$X_t^{n;x,I} = e^{t\mathcal{A}_n}x + W_t^{\mathcal{A}_n,\sigma} + \int_{0^-}^t e^{(t-s)\mathcal{A}_n}\mathrm{d}I_s, \quad t \ge 0.$$

Also, set

$$[\mathcal{G}_n\varphi](x) := \langle \mathcal{A}_n x, D\varphi(x)\rangle + \frac{1}{2}\mathrm{Tr}\left[\sigma\sigma^* D^2\varphi(x)\right].$$

Thanks to [26, Lemma 3.4(ii), Ch. 2] and due to the fact that $\mathcal{A}$ is dissipative (cf. Assumption 2.1), we then have, as $n \uparrow \infty$,

$$\mathcal{A}_n y \to \mathcal{A}y \quad \text{and} \quad \sup_{n\in\mathbb{N}} |\mathcal{A}_n y|_H \le |\mathcal{A}y|_H \quad \forall y \in \mathcal{D}(\mathcal{A}), \tag{3.9}$$

From the second claim of (3.9), it follows that

$$\sup_{n\in\mathbb{N}} |\mathcal{A}_n y|_H \le |y|_{\mathcal{D}(A)}, \quad \forall y \in \mathcal{D}(\mathcal{A}). \tag{3.10}$$

Moreover, for $p \geq 2$ as in Assumption 2.8 (see also (2.12)), by Lemma B.1(ii) (see also [27, Prop. 1.132] for the regular control case and references therein) it holds

$$\lim_{n\to\infty} \mathsf{E}[\sup_{t\in[0,T]} |X_t^{n;x,I} - X_t^{x,I}|^p] = 0. \tag{3.11}$$

The latter implies – up to passing to a subsequence – that $X_t^{n;x,I} \to X_t^{x,I}$ and with probability one as $n \to \infty$. Also, by [52, Thm. 27.2], for any $\varphi \in \mathcal{X}$, $\tau \in \mathcal{T}$, and for $T > 0$, the following Dynkin's formula holds true

$$\begin{aligned} &\mathsf{E}[e^{-\rho(\tau\wedge T)}\varphi(X_{\tau\wedge T}^{n;x,I})] = \varphi(x) + \mathsf{E}\bigg[\int_0^{\tau\wedge T} e^{-\rho t}[(\mathcal{G}_n - \rho)\varphi](X_t^{n;x,I})\mathrm{d}t\bigg] \\ &+ \mathsf{E}\bigg[\int_{0-}^{\tau\wedge T} e^{-\rho t}\langle D\varphi(X_t^{n;x,I}), \vartheta_t\rangle_H \mathrm{d}\nu_t\bigg]. \end{aligned} \tag{3.12}$$

As a matter of fact, the growth condition $|\sigma\sigma^* D^2\varphi(x)|_{\mathcal{L}_1(H)} \leq \widehat{C}$, for some $\widehat{C} > 0$, as required for any test function $\varphi \in \mathcal{X}$ ensures that the local martingale term vanishes in expectation.

We now aim at taking $n \to \infty$ in (3.12). In this regard, the only convergences that require some attention are

$$\lim_{n\to\infty} \mathsf{E}\left[\int_0^{\tau\wedge T} e^{-\rho t}\langle X_t^{n;x,I}, \mathcal{A}_n D\varphi(X_t^{n;x,I})\rangle_H \mathrm{d}t\right] = \mathsf{E}\left[\int_0^{\tau\wedge T} e^{-\rho t}\langle X_t^{x,I}, \mathcal{A} D\varphi(X_t^{x,I})\rangle_H \mathrm{d}t\right] \tag{3.13}$$

and

$$\lim_{n\to\infty} \mathsf{E}\bigg[\int_{0-}^{\tau\wedge T} e^{-\rho t}\langle D\varphi(X_t^{n;x,I}), \vartheta_t\rangle_H \mathrm{d}\nu_t\bigg] = \mathsf{E}\bigg[\int_{0-}^{\tau\wedge T} e^{-\rho t}\langle D\varphi(X_t^{x,I}), \vartheta_t\rangle_H \mathrm{d}\nu_t\bigg]. \tag{3.14}$$

All the other convergences can be performed by using arguments as in the proof of [27, Prop. 1.164] (in the case of regular controls). Assuming then that limits as $n \uparrow \infty$ can be indeed interchanged with expectations and integrals in (3.12) (we will verify later that (3.13) and (3.14) do hold), we find that

$$\begin{aligned} &\mathsf{E}\left[e^{-\rho(\tau\wedge T)}\varphi(X_{\tau\wedge T}^{x,I})\right] = \varphi(x) + \mathsf{E}\bigg[\int_0^{\tau\wedge T} e^{-\rho t}[(\mathcal{G} - \rho)\varphi](X_t^{x,I})\mathrm{d}t\bigg] \\ &+ \mathsf{E}\bigg[\int_{0-}^{\tau\wedge T} e^{-\rho t}\langle D\varphi(X_t^{x,I}), \vartheta_t\rangle_H \mathrm{d}\nu_t\bigg], \end{aligned}$$

so that finally letting $T \uparrow \infty$ we obtain (3.8).

It thus remains to check the validity of (3.13) and (3.14). As for (3.14), recalling (3.5) one has

$$\Big|\mathbb{1}_{[0,\tau\wedge T]}(t)e^{-\rho t}\langle D\varphi(X_t^{n;x,I}), \vartheta_t\rangle_H\Big| \leq \widehat{C}|\vartheta_t|_H$$

and, because $(\vartheta, \nu) \in \mathcal{I}_0$ (cf. (2.17)),

$$\mathsf{E}\bigg[\int_{0-}^{T} \Big|\mathbb{1}_{[0,\tau\wedge T]}(t)e^{-\rho t}\langle D\varphi(X_t^{n;x,I}), \vartheta_t\rangle_H\Big| \,\mathrm{d}\nu_t\bigg] \leq \widehat{C}\mathsf{E}\bigg[\int_{0-}^{T} |\vartheta_t|_H \,\mathrm{d}\nu_t\bigg] < \infty.$$

Hence, the dominated convergence theorem implies (3.14), upon recalling that (up to a subsequence) $X_t^{n;x,I} \to X_t^{x,I}$ with probability one as $n \to \infty$.

We now move on by proving the validity of (3.13). Notice that, if $(x_n) \subseteq H$, $(y_n) \subset \mathcal{D}(\mathcal{A})$, and $x \in H$, $y \in \mathcal{D}(\mathcal{A})$, we have, using (3.10),

$$\begin{aligned}
|\langle x_n, \mathcal{A}_n y_n\rangle - \langle x, \mathcal{A}y\rangle_H| &\leq |\langle (x_n - x), \mathcal{A}_n y_n\rangle_H| + |\langle x, \mathcal{A}_n(y_n - y)\rangle_H| + |\langle x, (\mathcal{A}_n - \mathcal{A})y\rangle_H| \\
&\leq |x_n - x|_H |\mathcal{A}_n y_n|_H + |x|_H |\mathcal{A}_n(y_n - y)|_H + |x|_H |(\mathcal{A}_n - \mathcal{A})y|_H \\
&\leq |x_n - x|_H \, |y_n|_{\mathcal{D}(\mathcal{A})} + |x|_H \, |y_n - y|_{\mathcal{D}(\mathcal{A})} + |x|_H |(\mathcal{A}_n - \mathcal{A})y|_H. \qquad (3.15)
\end{aligned}$$

Therefore, by using Hölder's inequality, for a constant $C > 0$ that may vary from line to line, we find

$$\begin{aligned}
&\mathsf{E}\left[\int_0^{\tau\wedge T} e^{-\rho t}\Big|\langle X_t^{n;x,I}, \mathcal{A}_n D\varphi(X_t^{n;x,I})\rangle_H - \langle X_t^{x,I}, \mathcal{A} D\varphi(X_t^{x,I})\rangle_H\Big| \mathrm{d}t\right] \\
&\leq C\left(\mathsf{E}\left[\int_0^{\tau\wedge T} e^{-\rho t}|X_t^{n;x,I} - X_t^{x,I}|_H^2 \mathrm{d}t\right] \mathsf{E}\left[\int_0^{\tau\wedge T} e^{-\rho t}|D\varphi(X_t^{n;x,I})|^2_{\mathcal{D}(\mathcal{A})} \mathrm{d}t\right]\right)^{1/2} \\
&+ C\left(\mathsf{E}\left[\int_0^{\tau\wedge T} e^{-\rho t}|X_t^{x,I}|_H^2 \mathrm{d}t\right] \mathsf{E}\left[\int_0^{\tau\wedge T} e^{-\rho t}|D\varphi(X_t^{n;x,I}) - D\varphi(X_t^{x,I})|^2_{\mathcal{D}(\mathcal{A})} \mathrm{d}t\right]\right)^{1/2} \\
&+ \mathsf{E}\left[\int_0^{\tau\wedge T} e^{-\rho t}|X_t^{x,I}|_H \; |(\mathcal{A}_n - \mathcal{A})D\varphi(X_t^{x,I})|_H \mathrm{d}t\right].
\end{aligned}$$

We now verify that the dominated convergence theorem can be applied when taking limits as $n \uparrow \infty$ in order to show that the right-hand side of the last inequality converges to zero and thus (3.13) holds. We provide details only for the term

$$\mathsf{E}\left[\int_0^{\tau\wedge T} e^{-\rho t}|D\varphi(X_t^{n;x,I}) - D\varphi(X_t^{x,I})|^2_{\mathcal{D}(\mathcal{A})} \mathrm{d}t\right]$$

since all the others can be treated by similar arguments (thanks also to (3.11)). Recall (3.4) and (3.5). Then

$$\left|\mathbb{1}(t)_{[0,\tau\wedge T]} e^{-\rho t}|D\varphi(X_t^{n;x,I}) - D\varphi(X_t^{x,I})|^2_{\mathcal{D}(\mathcal{A})}\right| \leq \widehat{C}^2,$$

so that the dominated convergence theorem ensures that

$$\lim_{n\uparrow\infty} \mathsf{E}\left[\int_0^{\tau\wedge T} e^{-\rho t}|D\varphi(X_t^{n;x,I}) - D\varphi(X_t^{x,I})|^2_{\mathcal{D}(\mathcal{A})} \mathrm{d}t\right] = 0,$$

upon recalling again that (up to a subsequence) $X_t^{n;x,I} \to X_t^{x,I}$ with probability one as $n \to \infty$. □

The next proposition provides an inequality for the cube of the norm on H. This will be exploited in the proof of the viscosity property of V (see Theorem 3.7 below).

Proposition 3.4. *Let $x \in H$, $\varphi \in \mathcal{X}$, $I \in \hat{\mathcal{I}}_0$, and $\tau \in \mathcal{T}$. Assume that*

$$\mathsf{E}\left[\int_{0-}^{\tau} e^{-\rho t}|X_t^{x,I} - x|_H |\langle X_t^{x,I} - x, \vartheta_t\rangle_H| \mathrm{d}\nu_t\right] < \infty. \qquad (3.16)$$

Then, the following inequality form of Dynkin's formula holds true:

$$\begin{aligned}
&\mathsf{E}\left[e^{-\rho\tau}|X_\tau^{x,I} - x|_H^3\right] \\
&\leq \mathsf{E}\left[\int_0^{\tau} e^{-\rho t}\left(3|\sigma\sigma^*|_{\mathcal{L}_1(H)}|X_t^{x,I} - x|_H - \rho|X_t^{x,I} - x|_H^3\right)\mathrm{d}t + \int_{0-}^{\tau} e^{-\rho t} 3|X_t^{x,I} - x|_H \langle X_t^{x,I} - x, \vartheta_t\rangle_H \mathrm{d}\nu_t\right].
\end{aligned}$$

Proof. Notice that, setting

$$g(y) := |y - x|_H^3, \quad y, x \in H,$$

one has

$$Dg(y) = 3|y-x|_H(y-x), \quad D^2g(y) = 3|y-x|_H^{-1}(y-x)\otimes(y-x) + 3|y-x|_H \cdot \mathbf{Id}_H.$$

In particular

$$\mathrm{Tr}\left[\sigma\sigma^* D^2 g(y)\right] \leq |D^2 g(y)|_{\mathcal{L}(H)} \mathrm{Tr}\left[\sigma\sigma^*\right] \leq 6|y-x|_H |\sigma\sigma^*|_{\mathcal{L}_1(H)}.$$

The claim then follows by arguments as in the proof of Proposition 3.3, upon noticing that for the Yosida approximants $\langle \mathcal{A}_n X_t^{n;x,I}, X_t^{n;x,I}\rangle_H \leq 0$ (by dissipativity) and then taking the lim sup (in place of the lim) as $n \to \infty$ so to get rid of the unbounded term $\langle X_t^{x,I}, \mathcal{A} X_t^{x,I}\rangle_H$. See also [27, Prop. 1.166] for the regular control case. □

The next result is the dynamic programming principle for the singular stochastic control problem (2.20). We are not going to provide a proof here, as this would very much follow the arguments used in the finite-dimensional settings (see [22], [38], and [49], among others). As a matter of fact, the key steps in the proof are based on the flow property of the solution to the controlled dynamics, which is, in the present setting, guaranteed by the semigroup property.

Proposition 3.5. *[Dynamic Programming Principle for V] Recall (3.1). For each $\tau \in \mathcal{T}$, we have*

$$V(x) = \inf_{I:=(\vartheta,\nu)\in\hat{\mathcal{I}}_0} \mathsf{E}\left[\int_{0^-}^{\tau} e^{-\rho t}\left(G(X_t^{x,I})\mathrm{d}t + \mathrm{d}\nu_t\right) + e^{-\rho\tau}V(X_\tau^{x,I})\right], \quad x \in H. \tag{3.17}$$

We are then ready to provide the definition of viscosity solution to (3.2) and to prove the viscosity property of V as in (2.20).

Definition 3.6 (Viscosity solution)**.**

(i) We say that $v \in C(H)$ is a viscosity supersolution to (3.2) at $x \in H$ if, for every $\varphi \in \mathcal{X}$ such that $0 = v(x) - \varphi(x) = \min(v - \varphi)$, one has

$$\max\left\{(\rho - \mathcal{G})\varphi(x) - G(x),\ \sup_{\theta\in\Theta}\left\{-\langle D\varphi(x), \theta\rangle_H - 1\right\}\right\} \geq 0.$$

(ii) We say that $v \in C(H)$ is a viscosity subsolution to (3.2) at $x \in H$ if, for every $\varphi \in \mathcal{X}$ such that $0 = v(x) - \varphi(x) = \max(v - \varphi)$, one has

$$\max\left\{(\rho - \mathcal{G})\varphi(x) - G(x),\ \sup_{\theta\in\Theta}\left\{-\langle D\varphi(x), \theta\rangle_H - 1\right\}\right\} \leq 0.$$

(iii) We say that $v \in C(H)$ is a viscosity solution to (3.2) at $x \in H$ if it is both a viscosity super- and subsolution.

Theorem 3.7. *V is a viscosity solution to (3.2) at all $x \in H$.*

Proof. (*Subsolution property.*) Let $x \in H$, $\varphi \in \mathcal{X}$ be such that $0 = V(x) - \varphi(x) = \max(V - \varphi)$.

Step 1. For $\theta \in \Theta$ and $\zeta > 0$, consider the control

$$I_\cdot = (\vartheta_\cdot, \nu_\cdot) \equiv (\theta, \hat{\nu}) \in \hat{\mathcal{I}}_0,$$

with $\hat{\nu}_{0^-} = 0$ and $\hat{\nu}_t = \zeta$ for any $t \geq 0$. By Proposition 3.5, we have for $h > 0$

$$V(x) \leq \mathsf{E}\left[\int_0^h e^{-\rho t} G(X_t^{x,I})\mathrm{d}t + \zeta + e^{-\rho h}V(X_h^{x,I})\right]. \tag{3.18}$$

We now aim at taking limits as $h \to 0^+$ in (3.18). To that end, note that the limits in the right-hand side of (3.18) can be interchanged with the expectation: By the monotone convergence theorem for the integral term; because of the dominated convergence theorem for the third addend, since V has sub-polynomial growth (cf. Proposition 3.1), and since $X_h^{x,I} = X_h^{x,0} + \zeta\theta$

(cf. (2.7)) and $\mathsf{E}[\sup_{t\in[0,T]}|X_t^{x,0}|^p]<\infty$, for $p\geq 2$ and for any $T>0$ by (2.11). Hence, due to $X_h^{x,I}\to x+\zeta\theta$ as $h\to 0^+$, we find

$$V(x)\leq \zeta+V(x+\zeta\theta). \tag{3.19}$$

But then, (3.19) yields

$$\varphi(x)\leq \zeta+\varphi(x+\zeta\theta),$$

which, dividing by ζ and letting $\zeta\to 0^+$, in turn gives (cf. Proposition 3.1)

$$-\langle D\varphi(x),\theta\rangle_H\leq 1.$$

Since the latter holds for every $\theta\in\Theta$, we find

$$\sup_{\theta\in\Theta}\{-\langle D\varphi(x),\theta\rangle_H-1\}\leq 0.$$

Step 2. Let now $I=\mathbf{0}$ be the null control. Setting

$$\tau_R:=\inf\{t\geq 0:\ |X_t^{x,0}|_H\geq R\}$$

(with $\inf\emptyset=+\infty$) and letting $h>0$, by Proposition 3.5 we have

$$\varphi(x)\leq \mathsf{E}\left[\int_0^{\tau_R\wedge h}e^{-\rho t}G(X_t^{x,0})\mathrm{d}t+e^{-\rho(\tau_R\wedge h)}\varphi(X_{\tau_R\wedge h}^{x,0})\right].$$

Using now Proposition 3.3, we find

$$\mathsf{E}\bigg[\int_0^{\tau_R\wedge h}e^{-\rho t}\Big(\big[(\rho-\mathcal{G})\varphi\big](X_t^{x,0})-G(X_t^{x,0})\Big)\mathrm{d}t\bigg]\leq 0.$$

Dividing by h, recalling that $\varphi\in\mathcal{X}$ and using that $\mathsf{E}[\sup_{t\in[0,T]}|X_t^{x,0}|]<\infty$, for any $T>0$ by (2.11), we can invoke the integral mean-value and the dominated convergence theorems when letting $h\to 0^+$, and we obtain

$$(\rho-\mathcal{G})\varphi(x)-G(x)\leq 0.$$

Step 3. Combining the last two steps we obtain the desired subsolution property of V.

(*Supersolution property.*) Let now $x\in H$, $\varphi\in\mathcal{X}$ be such that $0=V(x)-\varphi(x)=\min(V-\varphi)$. Assume, by contradiction that there exists $\eta>0$ such that

$$\sup_{\theta\in\Theta}\{-\langle D\varphi(x),\theta\rangle_H-1\}\leq -2\eta \tag{3.20}$$

and

$$(\rho-\mathcal{G})\varphi(x)-G(x)\leq -2\eta. \tag{3.21}$$

By continuity, for a suitable $\varepsilon>0$,

$$\sup_{\theta\in\Theta}\{-\langle D\varphi(y),\theta\rangle_H-1\}\leq -\eta,\qquad \forall y\in B_{|\cdot|_H}(x,\varepsilon). \tag{3.22}$$

and

$$(\rho-\mathcal{G})\varphi(y)-G(y)\leq -\eta,\qquad \forall y\in B_{|\cdot|_H}(x,\varepsilon), \tag{3.23}$$

where

$$B_{|\cdot|_H}(x,\varepsilon):=\{y\in H:|y-x|_H\leq\varepsilon\}.$$

Let now $I=(\vartheta,\nu)\in\hat{\mathcal{I}}_0$ be arbitrary but fixed, and set

$$\tau_\varepsilon^I:=\inf\{t\geq 0:\ X_t^{x,I}\notin B_{|\cdot|_H}(x,\varepsilon)\},$$

with the convention $\inf\emptyset=\infty$, which still provides a sense to the formulae below. In the following, we are going simply to write τ_ε instead of τ_ε^I, unless it becomes important to stress the explicit dependence on I. By Assumption 2.4 and the fact that $\varphi \in \mathcal{X}$, we have for a constant $C > 0$

$$\sup_{B_{|\cdot|_H}(x,\varepsilon)} |D\varphi|_{L^\infty} \leq C \cdot \sup_{B_{|\cdot|_H}(x,\varepsilon)} |D\varphi|_{\mathcal{D}(\mathcal{A})} =: \widetilde{C} < \infty.$$

Then, since $q \geq q_o \mathbf{1}$, we have

$$\begin{aligned} &\mathsf{E}\bigg[\int_{0-}^{\tau_\varepsilon} e^{-\rho t} \Big|\langle D\varphi(X_t^{x,I}), \vartheta_t\rangle_H\Big| \mathrm{d}\nu_t\bigg] \leq \widetilde{C}\mathsf{E}\bigg[\int_{0-}^{\tau_\varepsilon} e^{-\rho t} \langle \mathbf{1}, \vartheta_t\rangle_H \mathrm{d}\nu_t\bigg] \\ &= \frac{\widetilde{C}}{q_o}\mathsf{E}\bigg[\int_{0-}^{\infty} e^{-\rho t} \langle q_o\mathbf{1}, \vartheta_t\rangle_H \mathrm{d}\nu_t\bigg] \leq \frac{\widetilde{C}}{q_o}\mathsf{E}\bigg[\int_{0-}^{\infty} e^{-\rho t} \langle q, \vartheta_t\rangle_H \mathrm{d}\nu_t\bigg] \\ &= \frac{\widetilde{C}}{q_o}\mathsf{E}\bigg[\int_{0-}^{\infty} e^{-\rho t} \mathrm{d}\nu_t\bigg] < \infty. \end{aligned} \tag{3.24}$$

Similarly, one gets

$$\mathsf{E}\left[\int_{0-}^{\tau_\varepsilon} e^{-\rho t} \Big|\langle X_t^{x,I} - x, \vartheta_t\rangle_H\Big| \mathrm{d}\nu_t\right] < \infty. \tag{3.25}$$

Hence, employing (3.22), (3.23), and Propositions 3.3 and 3.4, whose requirements are met due to (3.24) and (3.25), we find

$$\begin{aligned} &V(x) - \mathsf{E}[e^{-\rho\tau_\varepsilon} V(X_{\tau_\varepsilon}^{x,I})] \leq \varphi(x) - \mathsf{E}[e^{-\rho\tau_\varepsilon}\varphi(X_{\tau_\varepsilon}^{x,I})] \\ &= -\mathsf{E}\left[e^{-\rho\tau_\varepsilon}\left|X_{\tau_\varepsilon}^{x,I} - x\right|_H^3\right] + \mathsf{E}\left[e^{-\rho\tau_\varepsilon}\left|X_{\tau_\varepsilon}^{x,I} - x\right|_H^3\right] \\ &+ \mathsf{E}\left[\int_0^{\tau_\varepsilon} e^{-\rho t}(\rho - \mathcal{G})\varphi(X_t^{x,I})\mathrm{d}t - \int_{0-}^{\tau_\varepsilon} e^{-\rho t}\langle D\varphi(X_t^{x,I}), \vartheta_t\rangle_H \mathrm{d}\nu_t\right] \\ &\leq -\varepsilon^3\mathsf{E}\left[e^{-\rho\tau_\varepsilon}\mathbb{1}_{\{\tau_\varepsilon<\infty\}}\right] + \mathsf{E}\left[\int_0^{\tau_\varepsilon} e^{-\rho t} 3|\sigma\sigma^*|_{\mathcal{L}_1(H)}\varepsilon\mathrm{d}t + \int_{0-}^{\tau_\varepsilon} 3e^{-\rho t}\varepsilon^2\mathrm{d}\nu_t\right] \\ &+ \mathsf{E}\left[\int_{0-}^{\tau_\varepsilon} e^{-\rho t}(G(X_t^{x,I})\mathrm{d}t + \mathrm{d}\nu_t)\right] - \eta\mathsf{E}\left[\int_{0-}^{\tau_\varepsilon} e^{-\rho t}(\mathrm{d}t + \mathrm{d}\nu_t)\right]. \end{aligned}$$

The latter implies

$$\begin{aligned} &V(x) - \mathsf{E}\left[\int_{0-}^{\tau_\varepsilon} e^{-\rho t}(G(X_t^{x,I})\mathrm{d}t + \mathrm{d}\nu_t) + e^{-\rho\tau_\varepsilon}V(X_{\tau_\varepsilon}^{x,I})\right] \\ &\leq -\varepsilon^3\mathsf{E}\left[e^{-\rho\tau_\varepsilon}\right] + \mathsf{E}\left[\int_0^{\tau_\varepsilon} e^{-\rho t}3|\sigma\sigma^*|_{\mathcal{L}_1(H)}\varepsilon\mathrm{d}t + \int_{0-}^{\tau_\varepsilon} 3e^{-\rho t}\varepsilon^2\mathrm{d}\nu_t\right] - \eta\mathsf{E}\left[\int_{0-}^{\tau_\varepsilon} e^{-\rho t}(\mathrm{d}t + \mathrm{d}\nu_t)\right], \end{aligned} \tag{3.26}$$

where we have used that $e^{-\rho\tau_\varepsilon}\mathbb{1}_{\{\tau_\varepsilon=+\infty\}} = 0$ by (3.6).

Stressing now the dependency of τ_ε with respect to the given and fixed $I \in \hat{\mathcal{I}}_0$, taking the supremum on both terms of (3.26) with respect to $I \in \hat{\mathcal{I}}_0$, using (3.17), and considering the identity

$$e^{-\rho\tau_\varepsilon^I} + \rho\int_0^{\tau_\varepsilon^I} e^{-\rho t}\mathrm{d}t = 1,$$

we obtain

$$0 \leq \sup_{I\in\hat{\mathcal{I}}_0}\left\{-\varepsilon^3\mathsf{E}\left[\rho - \int_0^{\tau_\varepsilon^I} e^{-\rho t}\mathrm{d}t\right] + \mathsf{E}\left[\int_0^{\tau_\varepsilon^I} e^{-\rho t}\left(3|\sigma\sigma^*|_{\mathcal{L}_1(H)}\varepsilon - \eta\right)\mathrm{d}t + \int_{0-}^{\tau_\varepsilon^I} e^{-\rho t}(3\varepsilon^2 - \eta)\mathrm{d}\nu_t\right]\right\};$$

that is,

$$0 \leq -\rho\varepsilon^3 + \sup_{I \in \hat{\mathcal{I}}_0} \mathsf{E}\left[\int_0^{\tau_\varepsilon^I} e^{-\rho t}\left(3\varepsilon|\sigma\sigma^*|_{\mathcal{L}_1(H)} + \varepsilon^3 - \eta\right)\mathrm{d}t + \int_{0-}^{\tau_\varepsilon^I} e^{-\rho t}(3\varepsilon^2 - \eta)\mathrm{d}\nu_t\right],$$

Taking ε small enough with respect to η, which can be done without loss of generality, the integrands on the right-hand side of the last display formula are negative, so that the supremum is nonpositive. Hence, we reach a contradiction and the supersolution property of V is proved. □

Remark 3.8. *It is worth emphasizing that the argument of the proof of the supersolution property can be clearly applied also in the finite-dimensional setting. As a consequence, this novel argument is able to reduce substantially the technicalities that are typically needed in $\mathbb{R}^n$, $n \geq 1$, in order to show the supersolution (respectively, subsolution) property in minimization problems (respectively, maximization problems) involving singular controls (see, e.g.,* [16], [34, Ch. VIII], [38]*, and* [49]*, among others).*

Remark 3.9. *To the best of our knowledge, there are only two contributions in the literature in which the viscosity approach is developed for variational inequalities in infinite dimensional spaces: These are* [15] *and* [36]*. In both of them, the variational inequality is related to an obstacle problem and thus involves a constraint on the solution itself, and not on its gradient (as, instead, it is in our case).*

(i) In [15]*, a problem of optimal stopping for a stochastic process with memory is considered. The infinite-dimensional space in which the underlying state process takes values is that of continuous functions, and the evolution of the process is subject to the action of an unbounded operator (the first-order derivative). However, the Brownian noise is finite-dimensional in* [15].

(ii) In [36]*, the aim is to price American options defined on forward-rates models. The interest rate process takes values in an Hilbert space and it evolves according to nonlinear dynamics, which however do not involve an unbounded operator. In this framework, the authors are able to prove a comparison result and to apply it in order to (uniquely) characterize the value function of the underlying infinite-dimensional optimal stopping problem. Remark 4.10 below will articulate more on the relation between our paper and* [36].

4. Selecting a specific direction of action: A related optimal stopping problem

In order to achieve further regularity of V, we now specialize to the case in which the controller can only choose the intensity $\nu \in \mathcal{S}$ appearing in the decomposition $(\vartheta, \nu) \in \mathcal{I}_0$ of any admissible control $I \in \mathcal{I}$.

In particular, we select the convex cone

$$\Delta = \mathrm{Span}\{\hat{n}\},$$

where $\hat{n} \in H_+ \setminus \{\mathbf{0}\}$ and it is such that, without loss of generality, $\langle q, \hat{n}\rangle_H = 1$. Consequently (cf. (2.15)), $\Theta = \{\hat{n}\}$ and we make the following assumption, which will hold throughout the rest of the paper without further mention.

Assumption 4.1. *The vector $\hat{n} \in H_+ \setminus \{\mathbf{0}\}$ is an eigenvector of $\mathcal{A}$ with eigenvalue λ*[1].

Remark 4.2. (i) *It is worth stressing that Assumption 4.1 does not imply that there exists a singularly controlled component of the state variable X that is decoupled from the rest of the infinite-dimensional dynamics of X. To clarify this, suppose that $H = \mathbb{R}^2$ and consider*

$$\mathcal{A} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \hat{n} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

[1] Clearly, with $\lambda \leq -\delta < 0$, where δ is as in Assumption 2.1.

Notice $\hat{n}$ is an eigenvector of $\mathcal{A}$. In this case, the two components of the state process X are truly coupled through the operator $\mathcal{A}$ so that the action along the direction $\hat{n}$ affects indirectly also the first component of X.

(ii) *Sufficient conditions guaranteeing the validity of Assumption 4.1 can be found in* [3, 20]. *We refer to Section 4 in* [13] *for a detailed discussion.*

For future use, we recall here that the uncontrolled state-process uniquely solves in the mild sense (cf. (2.7))

$$\mathrm{d}X_t^{x,0} = \mathcal{A}X_t^{x,0}\mathrm{d}t + \sigma \mathrm{d}W_t, \quad X_0^{x,0} = x \in H;$$

that is (cf. (2.9)),

$$X_t^{x,0} = e^{t\mathcal{A}}x + \int_0^t e^{(t-s)\mathcal{A}}\sigma \mathrm{d}W_s = e^{t\mathcal{A}}x + W_t^{\mathcal{A},\sigma}, \quad t \geq 0. \tag{4.1}$$

Furthermore, $X^{x,0}$ has continuous sample paths.

For our subsequent analysis, we define

$$G_{\hat{n}}(x) := \langle DG(x), \hat{n}\rangle_H, \quad x \in H, \tag{4.2}$$

and introduce the optimal stopping problem

$$\sup_{\tau\in\mathcal{T}} \mathsf{E}\bigg[\int_0^\tau e^{-(\rho-\lambda)t}G_{\hat{n}}(X_t^{x,0})\mathrm{d}t - e^{-(\rho-\lambda)\tau}\bigg], \quad x \in H,$$

where $\mathcal{T}$ denotes the set of the $\mathbb{F}$-stopping times. The next result relates the previous optimal stopping problem to the directional derivative of V as in (2.20), in the direction $\hat{n}$. This can be thought of as an infinite-dimensional analogue of the result in [4], where, in a finite-dimensional setting, it is proven that a suitable optimal timing problem provides the marginal value of an irreversible investment problem of monotone follower type.

Theorem 4.3. *For any $x \in H$, one has*

$$V_{\hat{n}}(x) := \langle DV(x), \hat{n}\rangle_H = \sup_{\tau\in\mathcal{T}} \mathsf{E}\bigg[\int_0^\tau e^{-(\rho-\lambda)t}G_{\hat{n}}(X_t^{x,0})\mathrm{d}t - e^{-(\rho-\lambda)\tau}\bigg]. \tag{4.3}$$

Proof. The proof is organized in two steps.

Step 1. Let $x \in H$. Here we prove that there exists an optimal control $I^\star := (\hat{n}, \nu^\star) \in \mathcal{I}_0$ for $V(x)$ as in (2.20). Furthermore, if G is strictly convex, $I^\star$ is unique up to indistinguishability.

Let $(I^k)_{k\in\mathbb{N}} := (\hat{n}, \nu^k)_{k\in\mathbb{N}} \subseteq \mathcal{I}_0$ be a minimizing sequence for $V(x)$. Let us denote by $X^{x,k} := X^{x,I^k}$, $k \in \mathbb{N}$. Given that $V(x) \leq \hat{c}_o(1 + |x|_H^p)$, $x \in H$, (cf. Proposition 3.1) for $p \geq 2$ as in Assumption 2.8, we have for some $\varepsilon > 0$

$$\sup_{k\in\mathbb{N}} \mathsf{E}\bigg[\int_{0^-}^\infty e^{-\rho t}|X_t^{x,k}|_H^p\,\mathrm{d}t\bigg] \leq \frac{1}{\kappa_1}\Big(\hat{c}_o(1 + |x|_H^p) + \varepsilon + \frac{\kappa_2}{\rho}\Big). \tag{4.4}$$

Because

$$X_t^{x,k} = e^{t\mathcal{A}}x + W_t^{\mathcal{A},\sigma}{}_t + \widehat{n}\int_{0^-}^t e^{\lambda(t-s)}\mathrm{d}\nu_t^k,$$

simple estimates using $\lambda < 0$, (2.10), and $|e^{t\mathcal{A}}|_{\mathcal{L}(H)} \leq e^{-\delta t}$, for all $t \geq 0$, give, for a constant $C > 0$ (changing from line to line),

$$|\widehat{n}|_H^p e^{\lambda tp}\mathsf{E}\big[|\nu_t^k|^p\big] \leq C\Big(|x|_H^p + \overline{c}_p + \mathsf{E}\big[|X_t^{x,k}|^p\big]\Big). \tag{4.5}$$

This in turn yields by (4.4)

$$\sup_{k\in\mathbb{N}} \mathsf{E}\bigg[\int_0^\infty e^{-(\rho-\lambda p)t}|\nu_t^k|^p\mathrm{d}t\bigg] \leq C\Big(1 + |x|_H^p + \sup_{k\in\mathbb{N}} \mathsf{E}\bigg[\int_0^\infty e^{-\rho t}|X_t^{x,k}|^p\mathrm{d}t\bigg]\Big) < \infty. \tag{4.6}$$

Hence, by Banach-Saks theorem, there exist a subsequence of $(\nu^k)_{k\in\mathbb{N}}$ - still denoted by $(\nu^k)_{k\in\mathbb{N}}$ - and some $\nu^\star \in L^p(\Omega\times[0,\infty);\mathsf{P}\otimes e^{-(\rho-\lambda p)t}\mathrm{d}t)$ such that

$$\widehat{\nu}^j := \frac{1}{j+1}\sum_{k=0}^{j}\nu^k \to \nu^\star \quad \text{in } L^p(\Omega\times[0,\infty);\mathsf{P}\otimes e^{-(\rho-\lambda p)t}\mathrm{d}t). \tag{4.7}$$

Actually, up to passing to a further subsequence (again, relabeled in the following), such a convergence can be realized $\mathsf{P}\otimes e^{-(\rho-\lambda p)t}\mathrm{d}t$-a.e. Then, by arguing as in [44, Lemmata 4.5-4.7], the process $\nu^\star$ admits a modification - still denoted by $\nu^\star$ - that is right-continuous, nondecreasing and $\mathbb{F}$-adapted, and thus belongs to $\mathcal{S}$. Furthermore, given that $\nu^\star \in L^p(\Omega\times[0,\infty);\mathsf{P}\otimes e^{-(\rho-\lambda p)t}\mathrm{d}t)$ and that $t\mapsto \mathsf{E}[|\nu^\star_t|^p]$ is nondecreasing, it follows that the integrability condition required in (2.17) is also met (recall that, in this section, $\vartheta_t\equiv\hat{n}$). Set then $\widehat{I}^j := (\hat{n},\widehat{\nu}^j)\in\mathcal{I}_0$, $j\in\mathbb{N}$, and $I^\star := (\hat{n},\nu^\star)\in\mathcal{I}_0$.

Notice now that by making use of an integration by parts in the integrals with respect to $\mathrm{d}\widehat{\nu}^j$, one also has that $\mathsf{P}\otimes e^{-\rho t}\mathrm{d}t$-a.e. (cf. (2.9) and (2.8))

$$X^{x,\widehat{I}^j}_t = e^{t\mathcal{A}}x + W^{\mathcal{A},\sigma}_t + \int_{0^-}^t e^{(t-s)\mathcal{A}}\hat{n}\,\mathrm{d}\widehat{\nu}^j_s = e^{t\mathcal{A}}x + W^{\mathcal{A},\sigma}_t + \hat{n}\int_{0^-}^t e^{\lambda(t-s)}\mathrm{d}\widehat{\nu}^j_s \to X^{x,I^\star}_t,$$

and also P-a.s.

$$\int_{0^-}^\infty e^{-\rho t}\mathrm{d}\widehat{\nu}^j_t \to \int_{0^-}^\infty e^{-\rho t}\mathrm{d}\nu^\star_t.$$

By convexity of $(x,\nu)\mapsto\mathcal{J}(x;I)$ (cf. (2.19)) the sequence $(\widehat{I}^j)_{j\in\mathbb{N}} := (\hat{n},\widehat{\nu}^j)_{j\in\mathbb{N}}$ is minimizing as well, and we conclude by Fatou's lemma and the previous convergences that

$$V(x) = \liminf_{j\to\infty}\mathcal{J}(x;\widehat{I}^j) \ge \mathcal{J}(x;I^\star), \quad x\in H,$$

from which the optimality of $I^\star = (\hat{n},\nu^\star)\in\mathcal{I}_0$ for $V(x)$ follows.

Finally, the uniqueness claim follows from strict convexity of G, upon using arguments as those in the proof of Proposition 3.4 in [30].

Step 2. Here we prove (4.3). Since the proof very much follows the lines of those of Lemmata 3 and 4 in [4], we only sketch it.

Let $x\in H$, $I^\star = (\hat{n},\nu^\star)\in\mathcal{I}_0$ be optimal for $V(x)$ (cf. *Step 1* above), and for $\tau\in\mathcal{T}$ and $\varepsilon>0$, define

$$\xi_t := \begin{cases} I^\star_t, & 0\le t<\tau, \\ I^\star_t + \varepsilon e^{\lambda\tau}\hat{n}, & t\ge\tau. \end{cases} \tag{4.8}$$

The process $\xi\in\mathcal{I}_0$ and it has direction of action $\hat{n}$ and intensity of action ν^ξ such that $\nu^\xi_{0^-}=0$ and $\mathrm{d}\nu^\xi_t = \mathrm{d}\nu^\star_t + \varepsilon e^{\lambda t}\delta(t-\tau)$, for any $t\ge 0$. Furthermore, letting $X^{x-\varepsilon\hat{n},\xi}$ be the solution to (2.7) started at time 0^- from level $x-\varepsilon\hat{n}$ and controlled by ξ, we have that: $X^{x-\varepsilon\hat{n},\xi}\equiv X^{x-\varepsilon\hat{n},I^\star}$ on $[0,\tau)$, while $X^{x-\varepsilon\hat{n},\xi}\equiv X^{x,I^\star}$ on $[\tau,\infty)$.

Hence, by exploiting the convexity of G, Assumption 2.2, and the previous observations,

$$\begin{aligned} V(x) - V(x-\varepsilon\hat{n}) &\ge \mathcal{J}(x;I^\star) - \mathcal{J}(x-\varepsilon\hat{n};\xi) \\ &= \mathsf{E}\bigg[\int_0^\tau e^{-\rho t}\langle DG(X^{x-\varepsilon\hat{n},I^\star}_t), X^{x,I^\star}_t - X^{x-\varepsilon\hat{n},I^\star}_t\rangle_H\mathrm{d}t - \varepsilon e^{-(\rho-\lambda)\tau}\bigg] \\ &= \varepsilon\mathsf{E}\bigg[\int_0^\tau e^{-(\rho-\lambda)t}\langle DG(X^{x-\varepsilon\hat{n},I^\star}_t),\hat{n}\rangle_H\mathrm{d}t - e^{-(\rho-\lambda)\tau}\bigg] \\ &\ge \varepsilon\mathsf{E}\bigg[\int_0^\tau e^{-(\rho-\lambda)t}\langle DG(X^{x-\varepsilon\hat{n},0}_t),\hat{n}\rangle_H\mathrm{d}t - e^{-(\rho-\lambda)\tau}\bigg], \end{aligned} \tag{4.9}$$

from which

$$\liminf_{\varepsilon\to 0}\frac{1}{\varepsilon}\Big(V(x)-V(x-\varepsilon\hat{n})\Big)\geq \sup_{\tau\in\mathcal{T}}\mathsf{E}\bigg[\int_0^\tau e^{-(\rho-\lambda)t}G_{\hat{n}}(X_t^{x,0})\mathrm{d}t-e^{-(\rho-\lambda)\tau}\bigg]. \tag{4.10}$$

Under the usual convention $\inf\emptyset=+\infty$, let now $\tau^\star:=\inf\{t\geq 0:\,\nu_t^\star>0\}$ and $\tau^\varepsilon:=\inf\{t\geq 0:\,\nu_t^\star\geq\varepsilon e^{\lambda t}\}$, for $\varepsilon>0$. Clearly, $\tau^\varepsilon\downarrow\tau^\star$ as $\varepsilon\downarrow 0$. We then define

$$\eta_t:=\begin{cases}0, & 0\leq t<\tau^\varepsilon,\\ I_t^\star-\varepsilon e^{\lambda\tau^\varepsilon}\hat{n}, & t\geq\tau^\varepsilon,\end{cases} \tag{4.11}$$

and notice that denoting by $X^{x+\varepsilon\hat{n},\eta}$ the solution to (2.7) started at time 0^- from level $x+\varepsilon\hat{n}$ and controlled by η, we have that: $X^{x+\varepsilon\hat{n},\eta}\equiv X^{x+\varepsilon\hat{n},0}$ on $[0,\tau^\varepsilon)$, while $X^{x+\varepsilon\hat{n},\eta}\equiv X^{x,I^\star}$ on $[\tau^\varepsilon,\infty)$.

Convexity of G then yields

(4.12)

$$\begin{aligned}&V(x+\varepsilon\hat{n})-V(x)\leq\mathcal{J}(x+\varepsilon\hat{n};\eta)-\mathcal{J}(x;I^\star)\\&=\mathsf{E}\bigg[\int_0^{\tau^\varepsilon}e^{-\rho t}\langle DG(X_t^{x+\varepsilon\hat{n},0}),X_t^{x+\varepsilon\hat{n},0}-X_t^{x,I^\star}\rangle_H\mathrm{d}t-\int_{0^-}^{\tau^\varepsilon}e^{-\rho t}\mathrm{d}\nu_t^\star-\Big(\varepsilon e^{-\lambda\tau^\varepsilon}-\nu^\star_{\tau^\varepsilon}\Big)e^{-\rho\tau^\varepsilon}\bigg].\end{aligned}$$

Notice now that

$$X_t^{x+\varepsilon\hat{n},0}-X_t^{x,I^\star}=\hat{n}e^{\lambda t}\Big(\varepsilon-\int_{0^-}^t e^{-\lambda s}\mathrm{d}\nu_s^\star\Big).$$

Plugging the latter relation into (4.12), dividing by ε and adding and substracting terms, one arrives at

$$\begin{aligned}&\frac{1}{\varepsilon}\Big(V(x+\varepsilon\hat{n})-V(x)\Big)\leq\mathsf{E}\bigg[\int_0^{\tau^\star}e^{-(\rho-\lambda)t}\langle DG(X_t^{x,0}),\hat{n}\rangle_H\mathrm{d}t-e^{-(\rho-\lambda)\tau^\star}\bigg]\\&+\mathsf{E}\bigg[\int_0^{\tau^\star}e^{-(\rho-\lambda)t}\langle DG(X_t^{x+\varepsilon\hat{n},0})-DG(X_t^{x,0}),\hat{n}\rangle_H\mathrm{d}t\bigg]+\mathsf{E}\Big[e^{-(\rho-\lambda)\tau^\star}-e^{-(\rho-\lambda)\tau^\varepsilon}\Big]\\&+\mathsf{E}\bigg[\int_{\tau^\star}^{\tau^\varepsilon}e^{-(\rho-\lambda)t}\langle DG(X_t^{x+\varepsilon\hat{n},0}),\hat{n}\rangle_H\Big(1-\frac{1}{\varepsilon}\int_{0^-}^t e^{-\lambda s}\mathrm{d}\nu_s^\star\Big)\mathrm{d}t\bigg].\end{aligned} \tag{4.13}$$

Taking limits as $\varepsilon\downarrow 0$ it is not difficult to see that all the addends on the right-hand side of (4.13) but the first converge to zero. Hence,

$$\begin{aligned}&\limsup_{\varepsilon\to 0}\frac{1}{\varepsilon}\Big(V(x+\varepsilon\hat{n})-V(x)\Big)\leq\mathsf{E}\bigg[\int_0^{\tau^\star}e^{-(\rho-\lambda)t}G_{\hat{n}}(X_t^{x,0})\mathrm{d}t-e^{-(\rho-\lambda)\tau^\star}\bigg]\\&\leq\sup_{\tau\in\mathcal{T}}\mathsf{E}\bigg[\int_0^\tau e^{-(\rho-\lambda)t}G_{\hat{n}}(X_t^{x,0})\mathrm{d}t-e^{-(\rho-\lambda)\tau}\bigg].\end{aligned} \tag{4.14}$$

Combining (4.10) and (4.14) and using the fact that $V\in C^{1,\mathrm{Lip}}(H)$ (cf. Proposition 3.1(v)) we obtain (4.3) and thus complete the proof. □

We assume the next condition on the directional derivative $G_{\hat{n}}$.

Assumption 4.4. *$G_{\hat{n}}\in C^1(H)$ and $|DG_{\hat{n}}|_H\leq K_{G_{\hat{n}}}$.*

For our subsequent analysis, it is convenient to make an integration by parts and exploit the strong Markov property to write

$$V_{\hat{n}}(x)=-1-\Phi(x)+\sup_{\tau\in\mathcal{T}}\mathsf{E}\Big[e^{-(\rho-\lambda)\tau}\Phi(X_\tau^{x,0})\Big],\quad x\in H,$$

with

$$\Phi(x) := -\mathsf{E}\Big[\int_0^\infty e^{-(\rho-\lambda)t}\Big(G_{\hat{n}}(X_t^{x,0}) + \rho - \lambda\Big)\mathrm{d}t\Big], \quad x \in H. \tag{4.15}$$

Under Assumption 4.4, using (4.1) and that $|e^{t\mathcal{A}}|_{\mathcal{L}(H)} \le e^{-\delta t}$, one finds for any $x_1, x_2 \in H$ that (recall $\lambda \le -\delta < 0$)

$$\begin{aligned} |\Phi(x_2) - \Phi(x_1)|_H &\le \mathsf{E}\Big[\int_0^\infty e^{-(\rho-\lambda)t}\big|G_{\hat{n}}(X_t^{x_2,0}) - G_{\hat{n}}(X_t^{x_1,0})\big|_H \mathrm{d}t\Big] \\ &\le \mathsf{E}\Big[\int_0^\infty e^{-(\rho-\lambda)t}\big|e^{t\mathcal{A}}(x_2 - x_1)\big|_H \mathrm{d}t\Big] \le \frac{K_{G_{\hat{n}}}}{\rho - \lambda + \delta}|x_2 - x_1|_H. \end{aligned} \tag{4.16}$$

Actually, given that $G_{\hat{n}} \in C^1(H)$, it can be easily shown by direct calculations that $\Phi \in C^1(H)$.

With regard to those properties of Φ, in order to further investigate the C^1-regularity of $V_{\hat{n}}$ it then suffices to consider

$$U(x) := V_{\hat{n}}(x) + 1 + \Phi(x) = \sup_{\tau\in\mathcal{T}} \mathsf{E}\Big[e^{-(\rho-\lambda)\tau}\Phi(X_\tau^{x,0})\Big], \quad x \in H. \tag{4.17}$$

Proposition 4.5. *Let Assumption 4.4 hold. Then, one has*

$$|U(x_2) - U(x_1)|_H \le \frac{K_{G_{\hat{n}}}}{\rho - \lambda + \delta}|x_2 - x_1|_H, \quad x_1, x_2 \in H.$$

Proof. For $x_1, x_2 \in H$, recalling (4.1) and using (4.16) as well as $|e^{t\mathcal{A}}|_{\mathcal{L}(H)} \le e^{-\delta t}$ for any $t \ge 0$, one has

$$\begin{aligned} |U(x_2) - U(x_1)| &\le \sup_{\tau\in\mathcal{T}} \mathsf{E}\Big[e^{-(\rho-\lambda)\tau}\big|\Phi(X_\tau^{x_2,0}) - \Phi(X_\tau^{x_1,0})\big|\Big] \\ &\le \frac{K_{G_{\hat{n}}}}{\rho - \lambda + \delta}\sup_{\tau\in\mathcal{T}} \mathsf{E}\Big[e^{-(\rho-\lambda)\tau}\big|e^{\tau\mathcal{A}}(x_2 - x_1)\big|_H\Big] \le \frac{K_{G_{\hat{n}}}}{\rho - \lambda + \delta}|x_2 - x_1|_H. \end{aligned}$$

□

By continuity of U, the stopping region $\{x \in H : U(x) = \Phi(x)\}$ is closed. Hence, by standard theory of optimal stopping (see, e.g., [55, Ch. I, Sec. 2.2], whose results hold for an underlying process taking values in a measurable space), one has P-a.s. that

$$\tau^\star := \inf\{t \ge 0 : \ U(X_t^{x,0}) = \Phi(X_t^{x,0})\}, \quad x \in H,$$

is optimal.

Proposition 4.6 (Dynamic Programming Principle for U)**.** *For all stopping times $\theta \in \mathcal{T}$ we have*

$$U(x) = \sup_{\tau\in\mathcal{T}} \mathsf{E}\left[e^{-(\rho-\lambda)\tau}\mathbb{1}_{\tau<\theta}\Phi(X_\tau^x) + e^{-(\rho-\lambda)\theta}\mathbb{1}_{\tau\ge\theta}U(X_\theta^x)\right], \qquad \forall x \in H. \tag{4.18}$$

We refrain from providing a proof of the previous result. As a matter of fact, this would follow the same lines and steps as in the finite-dimensional settings (see, e.g., [56, Sec. 5.2] and references therein) and exploit the flow property of X, which is ensured here by the semigroup property.

Based on the dynamic programming principle, one has that the differential problem associated to U is the variational inequality

$$\min\big\{((\rho - \lambda) - \mathcal{G})u, \ u - \Phi\big\} = 0 \quad \text{on} \quad H, \tag{4.19}$$

with $\mathcal{G}$ as in (3.3), and for which, recalling the class (3.5), we now provide the definition of viscosity solution.

Definition 4.7.

(i) *We say that* $u \in C(H)$ *is a viscosity supersolution to* (4.19) *at* $x \in H$ *if, for every* $\varphi \in \mathcal{X}$ *such that* $0 = u(x) - \varphi(x) = \min(u - \varphi)$, *one has*

$$\min\left\{((\rho - \lambda) - \mathcal{G})\varphi,\ \varphi - \Phi\right\} \geq 0.$$

(ii) *We say that* $u \in C(H)$ *is a viscosity subsolution to* (4.19) *at* $x \in H$ *if, for every* $\varphi \in \mathcal{X}$ *such that* $0 = u(x) - \varphi(x) = \max(u - \varphi)$, *one has*

$$\min\left\{((\rho - \lambda) - \mathcal{G})\varphi,\ \varphi - \Phi\right\} \leq 0.$$

(iii) *We say that* $u \in C(H)$ *is a viscosity solution to* (4.19) *at* $x \in H$ *if it is both a viscosity super- and subsolution.*

Lemma 4.8. *Let* $x \in H$, $\varepsilon > 0$, *and let*

$$\theta := \inf\{t \geq 0 :\ X_t^{x,0} \notin \mathcal{B}_{|\cdot|_H}(x, \varepsilon)\} \wedge 1,$$

with $\mathcal{B}_{|\cdot|_H}(x,\varepsilon) := \{y \in H : |y - x|_H \leq \varepsilon\}$. *There exists* $m_o > 0$ *such that*

$$\mathsf{E}\left[\int_0^{\theta \wedge \tau} e^{-(\rho-\lambda)t}\mathrm{d}t + \frac{e^{-(\rho-\lambda)\tau}}{\rho - \lambda}\mathbb{1}_{\tau < \theta}\right] \geq m_o, \quad \forall \tau \in \mathcal{T}. \tag{4.20}$$

Proof. First of all, notice that

$$\begin{aligned}
&\mathsf{E}\left[\int_0^{\theta \wedge \tau} e^{-(\rho-\lambda)t}\mathrm{d}t + \frac{e^{-(\rho-\lambda)\tau}}{\rho - \lambda}\mathbb{1}_{\tau < \theta}\right] \\
&= \mathsf{E}\left[\left(\int_0^{\theta} e^{-(\rho-\lambda)t}\mathrm{d}t\right)\mathbb{1}_{\tau \geq \theta} + \left(\int_0^{\tau} e^{-(\rho-\lambda)t}\mathrm{d}t\right)\mathbb{1}_{\tau < \theta} + \frac{e^{-(\rho-\lambda)\tau}}{\rho - \lambda}\mathbb{1}_{\tau < \theta}\right] \\
&= \mathsf{E}\left[\frac{1 - e^{-(\rho-\lambda)\theta}}{\rho - \lambda}\mathbb{1}_{\tau \geq \theta} + \frac{1}{\rho - \lambda}\mathbb{1}_{\tau < \theta}\right].
\end{aligned}$$

Then assume, by aiming for a contradiction, that there exists a sequence $(\tau_n) \subseteq \mathcal{T}$ such that

$$\mathsf{E}\left[\frac{1 - e^{-(\rho-\lambda)\theta}}{\rho - \lambda}\mathbb{1}_{\tau_n \geq \theta} + \frac{1}{\rho - \lambda}\mathbb{1}_{\tau_n < \theta}\right] \to 0.$$

This means that

$$\mathsf{E}\left[\frac{1 - e^{-(\rho-\lambda)\theta}}{\rho - \lambda}\mathbb{1}_{\tau_n \geq \theta}\right] \to 0 \quad \text{and} \quad \mathsf{E}\left[\mathbb{1}_{\tau_n < \theta}\right] \to 0. \tag{4.21}$$

The second convergence above yields, passing to a subsequence still labeled in the same way,

$$\lim_n \mathbb{1}_{\tau_n \geq \theta} = 1 \text{ a.s.}.$$

But then, the dominated convergence theorem gives

$$\mathsf{E}\left[\frac{1 - e^{-(\rho-\lambda)\theta}}{\rho - \lambda}\mathbb{1}_{\tau_n \geq \theta}\right] \to \mathsf{E}\left[\frac{1 - e^{-(\rho-\lambda)\theta}}{\rho - \lambda}\right] > 0,$$

where the strict inequality in the formula above is clearly due to continuity of trajectories of the process $X^{x,0}$. Hence, we have found a contradiction with the first convergence in (4.21) and the proof is thus complete. □

Proposition 4.9. U *is a viscosity solution to* (4.19) *at all* $x \in H$.

Proof. We follow the ideas of the proof in the finite-dimensional setting proposed by [56, Thm. 5.2.1], but we simplify substantially the proof of the subsolution property thanks to Lemma 4.8.

Supersolution property. Recall (3.5) and let $\varphi \in \mathcal{X}$ be such that

$$0 = U(x) - \varphi(x) = \min(U - \varphi).$$

Clearly,

$$\varphi(x) = U(x) \geq \Phi(x).$$

Therefore, it remains to show that

$$((\rho - \lambda) - \mathcal{G})\varphi(x) \geq 0.$$

Setting

$$\tau_R := \inf\{t \geq 0 : \, |X_t^{x,0}|_H \geq R\}$$

(with $\inf \emptyset = +\infty$) and letting $h > 0$, by Proposition 3.5 we have

$$U(x) \geq \mathsf{E}[e^{-(\rho-\lambda)(\tau_R \wedge h)} U(X^{x,0}_{\tau_R \wedge h})],$$

so that

$$0 \geq \mathsf{E}[e^{-(\rho-\lambda)(\tau_R \wedge h)} U(X^{x,0}_{\tau_R \wedge h})] - U(x) \geq \mathsf{E}[e^{-(\rho-\lambda)(\tau_R \wedge h)} \varphi(X^{x,0}_{\tau_R \wedge h})] - \varphi(x).$$

On the other hand, Proposition 3.3, applied in the case of $I = \mathbf{0}$ (the null control), yields

$$\mathsf{E}[e^{-(\rho-\lambda)(\tau_R \wedge h)} \varphi(X^{x,0}_{\tau_R \wedge h})] = \varphi(x) + \mathsf{E}\left[\int_0^{\tau_R \wedge h} e^{-(\rho-\lambda)t} [(\mathcal{G} - (\rho - \lambda))\varphi](X_t^{x,0}) \mathrm{d}t\right].$$

Combining the last two display equations and dividing by h we obtain

$$\frac{1}{h}\mathsf{E}\left[\int_0^{\tau_R \wedge h} e^{-(\rho-\lambda)t} \Big[((\rho - \lambda) - \mathcal{G})\varphi\Big](X_t^{x,0}) \mathrm{d}t\right] \geq 0.$$

We conclude by taking $h \to 0^+$ and applying the integral mean-value theorem and the dominated convergence theorem, since $\varphi \in \mathcal{X}$.

Subsolution property. Let $\varphi \in \mathcal{X}$ be such that $0 = U(x) - \varphi(x) = \max(U - \varphi)$ and assume, by contradiction, that there exists some $\eta > 0$ such that

$$((\rho - \lambda) - \mathcal{G})\varphi(x) > 2\eta \;\; \text{and} \;\; U(x) - \Phi(x) > \frac{2\eta}{\rho - \lambda}.$$

By continuity,

$$((\rho - \lambda) - \mathcal{G})\varphi(y) > \eta \;\; \text{and} \;\; U(y) - \Phi(y) > \frac{\eta}{\rho - \lambda} \quad \forall \mathcal{B}_{|\cdot|_H}(x, \varepsilon),$$

where we recall that $\mathcal{B}_{|\cdot|_H}(x, \varepsilon) := \{y \in H : \, |y - x|_H \leq \varepsilon\}$.

Let us define the stopping time

$$\theta := \inf\{t \geq 0 : \; X_t^{x,0} \notin \mathcal{B}_{|\cdot|_H}(x, \varepsilon)\} \wedge 1.$$

Then, using Dynkin's formula of Proposition 3.3 (applied again in the case of I being the null control) we obtain for each $\tau \in \mathcal{T}$

$$\begin{aligned} &\mathsf{E}\left[e^{-(\rho-\lambda)(\theta \wedge \tau)} U(X^{x,0}_{\theta \wedge \tau})\right] - U(x) \leq \mathsf{E}\left[e^{-(\rho-\lambda)(\theta \wedge \tau)} \varphi(X^{x,0}_{\theta \wedge \tau})\right] - \varphi(x) \\ &= \mathsf{E}\left[\int_0^{\theta \wedge \tau} e^{-(\rho-\lambda)t} \Big[\big(\mathcal{G} - (\rho - \lambda)\big)\varphi\Big](X_t^{x,0}) \mathrm{d}t\right], \end{aligned}$$

which in turn, thanks to Lemma 4.8, implies

$$\begin{aligned}
U(x) &\geq \mathsf{E}\left[\eta \int_0^{\theta\wedge\tau} e^{-(\rho-\lambda)t}\mathrm{d}t + e^{-(\rho-\lambda)(\theta\wedge\tau)}U(X^{x,0}_{\theta\wedge\tau})\right]\\
&\geq \mathsf{E}\left[\eta \int_0^{\theta\wedge\tau} e^{-(\rho-\lambda)t}\mathrm{d}t + e^{-(\rho-\lambda)\tau}\mathbb{1}_{\tau<\theta}\left(\frac{\eta}{\rho-\lambda} + \Phi(X^{x,0}_{\tau})\right) + e^{-(\rho-\lambda)\theta}\mathbb{1}_{\tau\geq\theta}U(X^{x,0}_{\theta})\right]\\
&= \eta\mathsf{E}\left[\int_0^{\theta\wedge\tau} e^{-(\rho-\lambda)t}\mathrm{d}t + \frac{e^{-(\rho-\lambda)\tau}}{\rho-\lambda}\mathbb{1}_{\tau<\theta}\right] + \mathsf{E}\left[e^{-(\rho-\lambda)\tau}\mathbb{1}_{\tau<\theta}\Phi(X^{x,0}_{\tau}) + e^{-(\rho-\lambda)\theta}\mathbb{1}_{\tau\geq\theta}U(X^{x,0}_{\theta})\right]\\
&\geq \eta m_o + \mathsf{E}\left[e^{-(\rho-\lambda)\tau}\mathbb{1}_{\tau<\theta}\Phi(X^{x,0}_{\tau}) + e^{-(\rho-\lambda)\theta}\mathbb{1}_{\tau\geq\theta}U(X^{x,0}_{\theta})\right].
\end{aligned}$$

Taking the supremum over $\tau \in \mathcal{T}$ in the latter, we contradict (4.18), concluding the proof. □

Remark 4.10. *Comments as those collected in Remark 3.9 apply to (4.19). The analogy of our setting with [36] is at this point even tighter, as now (4.19) takes the form of an obstacle problem and thus involves a constraint on the solution itself (and not on its gradient as it was in the previous section).*

If one aims at a comparison theorem for (4.19), this might be proved by adapting the techniques of [36] in order to deal with the unbounded operator in the dynamics of the state process (not present in [36]). To that end, one should treat the unbounded term as in the regular control case by adding radial functions as test functions in the definition of viscosity solutions (see [27, Ch. 3]).

However, because our main aim is to provide regularity results for the optimal stopping problem under consideration, we refrain from studying the relevant question of uniqueness of the viscosity solution to (4.19), which is then left for future research.

Thanks to Assumption 2.1, we can introduce

$$|x|_{-1} := |\mathcal{A}^{-1}x|_H, \tag{4.22}$$

Then, in order to achieve the C^1-regularity of U, we strengthen the assumption on $G_{\hat{n}}$, by requiring the following, which will be standing in the rest of the section.

Assumption 4.11. *There exists $\kappa_o > 0$ such that $|G_{\hat{n}}(x)| \leq \kappa_o(1+|x|_{-1})$. Furthermore, $G_{\hat{n}}$ is semiconcave with respect to the norm $|\cdot|_{-1}$; that is, there exists $\kappa_1 > 0$ such that*

$$\lambda G_{\hat{n}}(x) + (1-\lambda)G_{\hat{n}}(y) - G_{\hat{n}}(\lambda x + (1-\lambda)y) \leq \frac{\kappa_1}{2}\lambda(1-\lambda)|x-y|^2_{-1}, \quad \forall x,y \in H,\ \lambda \in [0,1].$$

Remark 4.12. *Assumptions 2.8, 4.4, and 4.11 are satisfied, e.g., if*

$$G(x) = \frac{1}{2}(\langle x,h\rangle_H)^2, \quad \textit{or} \quad G(x) = \frac{1}{2}\langle Qx,x\rangle_H, \quad x \in H,$$

with $h \in \mathcal{D}(\mathcal{A})$, and with Q being positive semidefinite, symmetric and such that $Q\hat{n} \in \mathcal{D}(\mathcal{A})$.

Indeed, in these cases,

$$G_{\hat{n}}(x) = \langle x,h\rangle_H\langle h,\hat{n}\rangle_H, \quad \textit{respectively} \quad G_{\hat{n}}(x) = \langle x,Q\hat{n}\rangle_H, \quad x \in H,$$

so that $G_{\hat{n}}$ is clearly concave, belongs to $C^1(H)$ and it has sublinear growth with respect to $|\cdot|_H$. Moreover, in the first case, setting $k := \mathcal{A}h$, one has

$$\begin{aligned}
|G_{\hat{n}}(x)| &\leq |h|_H\big|\langle x,h\rangle_H\big| = |h|_H\big|\langle x,\mathcal{A}^{-1}k\rangle_H\big| = |h|_H\big|\langle x,\mathcal{A}^{-1}k\rangle_H\big|\\
&= |h|_H\big|\langle\mathcal{A}^{-1}x,k\rangle_H\big| \leq |h|_H|k|_H|x|_{-1}.
\end{aligned}$$

Similarly, in the second case, setting $k := \mathcal{A}Q\hat{n}$, one has

$$\begin{aligned}
|G_{\hat{n}}(x)| &\leq |\langle x,Q\hat{n}\rangle_H| = |\langle x,\mathcal{A}^{-1}k\rangle_H| = |\langle x,\mathcal{A}^{-1}k\rangle_H|\\
&= |\langle\mathcal{A}^{-1}x,k\rangle_H| \leq |k|_H|x|_{-1}.
\end{aligned}$$

Recall (2.9) and (4.22). By (4.1), it holds

$$|\mathcal{A}^{-1}X_s^{x,0}|_H \leq |e^{t\mathcal{A}}\mathcal{A}^{-1}x|_H + |\mathcal{A}^{-1}W_s^{\mathcal{A},\sigma}|_H.$$

Then, the contraction property of the semigroup $e^{t\mathcal{A}}$, estimate (2.10), and the fact that the norm $|\cdot|_{-1}$ is clearly dominated by the norm $|\cdot|_H$ give

$$\mathsf{E}\left[\sup_{s\geq 0}|X_s^{x,0}|_{-1}\right] \leq \bar{\kappa}(1+|x|_{-1}), \qquad \forall x \in H, \tag{4.23}$$

for some $\bar{\kappa} > 0$.

One then has the following preliminary result.

Proposition 4.13. *U is semiconvex with respect to the $|\cdot|_{-1}$ norm and there exists $\widehat{\kappa}_o > 0$ such that*

$$|U(x)| \leq \widehat{\kappa}_o(1+|x|_{-1}). \tag{4.24}$$

Furthermore, U is $|\cdot|_{-1}$-locally Lipschitz continuous.

Proof. Using (4.1) and that $|e^{t\mathcal{A}}|_{\mathcal{L}(H)} \leq e^{-\delta t}$, one easily finds from (4.15) that the semiconcavity with respect to the $|\cdot|_{-1}$ norm of $G_{\hat{n}}$ implies semiconvexity of Φ with respect to the $|\cdot|_{-1}$ norm as well. Hence, given that $x \mapsto X^{x,0}$ is linear, we find that U is semiconvex with respect to the $|\cdot|_{-1}$ norm being supremum of semiconvex functions (see [14, Prop. 2.1.5], whose proof does not suffer the dimensionality of the state space).

By (4.15), one also has that the growth condition on $G_{\hat{n}}$, together with (4.23), imply that $|\Phi(x)| \leq K_\Phi(1+|x|_{-1})$, for a suitable $K_\Phi > 0$. The latter, together with (4.24), in turn gives for $\widehat{\kappa}_o > 0$

$$|U(x)| \leq \sup_{\tau\in\mathcal{T}} \mathsf{E}\big[|\Phi(X_\tau^{x,0})|\big] \leq K_\Phi\big(1+\mathsf{E}\big[\sup_{s\geq 0}|X_s^{x,0}|_{-1}\big]\big) \leq \widehat{\kappa}_o(1+|x|_{-1}).$$

Finally, the last claim is due to [25, Cor. 2.4, p. 12] and to the fact that, being U semiconvex with respect to $|\cdot|_{-1}$ norm, by Lemma A.1(i) one can write

$$U(x) = U_0(x) - \frac{\widehat{\kappa}_1}{2}|x|_{-1}^2, \quad x \in H,$$

for some $\widehat{\kappa}_1 > 0$ and U_0 convex. □

In the rest of the paper, we are going to study the regularity properties of the (sub)gradient of U.

Proposition 4.14. *The following hold true:*

(i) $D^-U(x) \subseteq \mathcal{D}(\mathcal{A})$ at all $x \in H$;

(ii) The multivalued map $D^-U : H \to \mathcal{D}(\mathcal{A})$ is $|\cdot|_H$-to-$|\cdot|_{\mathcal{D}(\mathcal{A})}$ upper hemicontinuous, with $\mathcal{D}(\mathcal{A})$ being endowed with the graph norm.

Proof. Recall that by (i) in Lemma A.1 we can write

$$U(x) = U_0(x) - \frac{\widehat{\kappa}_1}{2}|x|_{-1}^2, \quad x \in H,$$

for some $\widehat{\kappa}_1 > 0$ and U_0 convex, so that, for any $x \in H$,

$$D^-U(x) = D^-U_0(x) - \widehat{\kappa}_1\mathcal{A}^{-1}x.$$

Hence, without loss of generality, up to replace U by U_0, we work in the rest of this proof under the assumption that U is convex. We now prove each item separately.

Proof of (i). Let $x \in H$, $(h_n) \subseteq \mathcal{D}(\mathcal{A})$ be such that $h_n \to 0$ with respect to $|\cdot|_H$, and set $\hat{h}_n := \mathcal{A}h_n$, so that also $|\hat{h}_n|_{-1} \to 0$. For $p \in D^-U(x)$, by convexity of U we have

$$U(x+\hat{h}_n) - U(x) \geq \langle p, \hat{h}_n\rangle_H,$$

which rewrites as

$$U(x+\hat{h}_n) - U(x) \geq \langle p, \mathcal{A}h_n\rangle_H. \tag{4.25}$$

Since U is $|\cdot|_{-1}$-continuous by Proposition 4.13, the linear functional

$$T_x : \mathcal{D}(\mathcal{A}) \to \mathbb{R}, \qquad T_x(h) = \langle p, \mathcal{A}h\rangle_H$$

is $|\cdot|_H$-upper semicontinuous by (4.25); that is, $\limsup_{n\to\infty} T_x(h_n) \leq 0$. Taking now the sequence $(-h_n) \subseteq \mathcal{D}(\mathcal{A})$ and exploiting the linearity of T_x, it also holds that T_x is $|\cdot|_H$-lower semicontinuous. Hence, it is $|\cdot|_H$-continuous, which means that $p \in \mathcal{D}(\mathcal{A})$.

Proof of (ii). Consider the function

$$u : \mathcal{D}(\mathcal{A}) \to \mathbb{R}, \qquad u(x) := U(\mathcal{A}x), \quad x \in H,$$

so that $U = u \circ \mathcal{A}^{-1}$. Then u is convex and by Proposition 4.13 one has

$$|u(x)| = |U(\mathcal{A}x)| \leq \widehat{\kappa}_o(1 + |\mathcal{A}x|_{-1}) = \widehat{\kappa}_o(1 + |\mathcal{A}^{-1}\mathcal{A}x|_H) = \widehat{\kappa}_o(1+|x|_H), \quad x \in H.$$

Hence, u is $|\cdot|_H$-locally Lipschitz continuous by [25, Cor. 2.4, p. 12] and thus it can be extended to a continuous function defined to the whole space H. Moreover, for all $x \in \mathcal{D}(\mathcal{A})$, we have

$$q \in D^- u(x) \iff p = \mathcal{A}q \in D^-U(\mathcal{A}x).$$

Hence, by using the definition of subgradient,

$$D^-u(x) = \mathcal{A}DU^-(\mathcal{A}x), \qquad \forall x \in \mathcal{D}(\mathcal{A}).$$

By convexity, the multivalued map

$$(\mathcal{D}(\mathcal{A}), |\cdot|_H) \to (H, |\cdot|_H), \quad x \mapsto D^-u(x)$$

is upper hemicontinuous. Because

$$D^-U(x) = \mathcal{A}^{-1}D^-u(\mathcal{A}^{-1}x),$$

it then follows that $D^-U : H \to \mathcal{D}(\mathcal{A})$ is $|\cdot|_H$-to-$|\cdot|_{\mathcal{D}(\mathcal{A})}$ upper hemicontinuous. □

Proposition 4.15. *Recall Assumption 2.7, let $x \in H$, $\hat{h} \in K$, and set $h := \sigma\hat{h}$. If $h \in \mathcal{D}(\mathcal{A})$, then U is differentiable at x along the direction h.*

Proof. We assume, without loss of generality, that $U(x) = 0$. Indeed, this can be always achieved by a translation. Again, by (i) in Lemma A.1 we can write

$$U(y) = U_0(y) - \frac{\widehat{\kappa}_1}{2}|y-x|^2_{-1}, \quad y \in H,$$

for some $\widehat{\kappa}_1 > 0$ and U_0 convex.

We now argue by contradiction and assume that U (hence U_0) is not differentiable along the direction $h := \sigma\hat{h}$, with $\hat{h} \in K$. By convexity of U_0 (cf. Proposition 4.13) and by Proposition 4.14, there exist $p, \bar{p} \in D^-U_0(x) \subseteq \mathcal{D}(\mathcal{A})$ such that

$$\langle p, h\rangle_H < \langle \bar{p}, h\rangle_H.$$

Again by convexity,

$$U_0(y) \geq \langle p, y-x\rangle_H \vee \langle \bar{p}, y-x\rangle_H, \quad \forall y \in H.$$

It is then possible to construct $(\widehat{\varphi}_n) \subseteq \mathcal{X}$ (cf. (3.5)) such that for any $n \in \mathbb{N}$

$$\begin{cases} \widehat{\varphi}_n(x) = U_0(x) = 0, \\ \widehat{\varphi}_n \leq U_0, \\ |D\widehat{\varphi}_n(x)|_H + |\mathcal{A}D\widehat{\varphi}_n(x)|_H \leq M < \infty, \\ \langle D^2\widehat{\varphi}_n(x)h, h\rangle_H \geq n, \\ \langle D^2\widehat{\varphi}_n(x)k, \bar{k}\rangle_H = 0 \quad \forall k, \bar{k} \in \operatorname{Span}\{h\}^\perp, \end{cases}$$

for some $M > 0$. Notice that it follows from the previous properties that

$$\frac{1}{2}\mathrm{Tr}\left[\sigma\sigma^* D^2\widehat{\varphi}_n(x)\right] \geq n. \tag{4.26}$$

As a matter of fact, the sequence

$$\widehat{\varphi}_n(y) = \frac{1}{2}\langle p+\bar{p}, y-x\rangle_H + \frac{n}{2}\langle y-x, h\rangle_H^2, \quad n \in \mathbb{N},\ y \in H,$$

realizes all the previous requirements, at least in a neighborhood of x (and can be easily modified in order to satisfy them globally on H).

Define now $\varphi_n(y) := \widehat{\varphi}_n(y) - \frac{\widehat{\kappa}_1}{2}|y-x|_{-1}^2$ for any $y \in H$. Then (cf. (4.26))

$$\begin{cases} \varphi_n(x) = U(x) = 0, \\ \varphi_n \leq U, \\ |D\varphi_n(x)|_H + |\mathcal{A}D\varphi_n(x)|_H \leq M < \infty, \\ \frac{1}{2}\mathrm{Tr}\left[\sigma\sigma^* D^2\varphi_n(x)\right] \geq n - \frac{1}{2}\kappa_1 \mathrm{Tr}\left[\sigma\sigma^*\mathcal{A}^{-1}\right]. \end{cases}$$

Because U is a viscosity supersolution at x of (4.19) (cf. Proposition 4.9), one finally has

$$\begin{aligned} 0 &\leq \left[\big((\rho-\lambda) - \mathcal{G}\big)\varphi_n\right](x) = -\frac{1}{2}\mathrm{Tr}\left[\sigma\sigma^* D^2\varphi_n(x)\right] - \langle x, \mathcal{A}D\varphi_n(x)\rangle_H \\ &\leq -n + \frac{1}{2}\kappa_1 \mathrm{Tr}\left[\sigma\sigma^*\mathcal{A}^{-1}\right] + M|x|_H \to -\infty \quad \text{as } n \to \infty, \end{aligned}$$

which gives the desired contradiction and thus completes the proof. □

Remark 4.16. *It is worth noticing that the proof of Proposition 4.15 only exploits the viscosity supersolution property of U, which is actually the easiest part to be shown in the proof of Proposition 4.9.*

The next proposition strenghts the regularity result of Proposition 4.15 in the case

$$\overline{\mathrm{R}(\sigma) \cap \mathcal{D}(\mathcal{A})} = H,$$

where $\mathrm{R}(\sigma)$ denotes the range of σ (with σ satisfying Assumption 2.7).

Proposition 4.17. *If*

$$\overline{R(\sigma) \cap \mathcal{D}(\mathcal{A})} = H,$$

then U is (Fréchet) differentiable at x, $DU(x) \in \mathcal{D}(\mathcal{A})$, and $DU \in C(H;\mathcal{D}(\mathcal{A}))$, with $\mathcal{D}(\mathcal{A})$ being endowed with the graph norm.

Proof. By following the arguments developed at the beginning of Proposition 4.14, we can here assume that U is convex. Then, due to Proposition 4.15 and convexity of U, we know that, for each $h \in R(\sigma) \cap \mathcal{D}(\mathcal{A})$, the set

$$\{\langle h, p\rangle_H : p \in D^- U(x)\}$$

is a singleton. Since $\overline{R(\sigma) \cap \mathcal{D}(\mathcal{A})} = H$, the set $D^-U(x)$ must be a singleton too. Again by convexity of U, it follows that U is differentiable at $x \in H$. Furthermore, using Proposition 4.14(i), we conclude that $DU(x) \in \mathcal{D}(\mathcal{A})$.

The last assertion finally follows from the previous conclusions and Proposition 4.14. □

5. Second-order smooth-fit for V

Thanks to Proposition 4.17 and Theorem 4.3 we finally achieve the second-order smooth-fit of V in the direction of $\hat{n}$. More precisely, we have the following corollary.

Corollary 5.1. *Under the assumptions of Proposition 4.17, $V \in C^{1,Lip}(H)$ and $V_{\hat{n}} \in C^1(H)$.*

The second-order regularity of the value function has been a fundamental aspect of singular stochastic control theory since its beginning. The underlying idea is that such smoothness of the value function lays the groundwork for characterizing the problem's free boundary and, consequently, for constructing an optimal control of reflection type.

In one-dimensional or two-dimensional fully degenerate stationary settings, where a guess-and-verify approach can be employed, imposing a suitable second-order regularity on the solution of the variational inequality enables the unique determination of certain otherwise free parameters. This, in turn, allows for the identification of the value function of the problem and the free boundary at which the state process should be optimally reflected (see [2, 33, 41, 51], among many others).

The validation of a second-order smooth-fit property in multiple dimensions still requires verification on a case-by-case basis. This is well described in the introduction of the seminal work by S.E. Shreve and H.M. Soner [57]: *"An important question is whether the principle of smooth fit can be expected to apply to multidimensional singular control problems, or is it strictly a one-dimensional phenomenon. Karatzas and Shreve [45] suggested that it might apply in higher dimensions. [...] Our discovery of a C^2-value function provides strong support for the belief in a widely applicable principle of smooth fit. Nevertheless, the argument of this paper depends heavily on the fact that only two dimensions are involved [...], and we have not found a way to obtain a similar result in higher dimensions."*

In suitable multiple-dimensional frameworks, second-order regularity of the derivative of the value function of the singular stochastic control problem along the direction of the control process has been obtained more recently through its relation to optimal stopping (see, e.g., [17, 23, 32]). Our result is thus situated within this body of literature and actually provides, for the first time, the validation of a second-order smooth-fit property in an infinite-dimensional setting.

6. Two applications in Economics

Here we discuss two economic models that can embedded into our setting.

6.1. An irreversible investment problem into energy capacity.

Let $\mathcal{O}$ be an open, simply connected, and bounded subset of $\mathbb{R}^n$, $n < 4$, with smooth boundary. We endow it with the Lebesgue measure μ on the Borel σ-algebra of $\mathcal{O}$, and consider the Hilbert space $H = L^2(\mathcal{O};\mu)$.

Within this mathematical framework, we consider a company which has to deliver energy to the locations of $\mathcal{O}$. We assume that, in absence of any investment by the company, the energy supply at time t and location ξ – denoted by $E^0(t,\xi)$ – evolves according to the parabolic PDE

$$
\begin{cases}
\dfrac{\partial E^0}{\partial t}(t,\xi) = \Delta E^0(t,\xi) - \delta E^0(t,\xi), \qquad (t,\xi) \in \mathbb{R}_+ \times \mathcal{O}, \\
E^0(0,\xi) = e(\xi), \quad \xi \in \mathcal{O}, \\
\dfrac{\partial E^0}{\partial \mathbf{n}}(t,\xi) = 0, \quad (t,\xi) \in \mathbb{R}_+ \times \partial\mathcal{O},
\end{cases}
\tag{6.1}
$$

where Δ is the Laplacian operator over $\mathcal{O}$, $\delta > 0$ is a depreciation factor, $e(\xi)$ is the initial value of the energy supply at location ξ, and $\mathbf{n}$ denotes the unitary outer normal vector at the boundary of $\mathcal{O}$; the Neumann boundary condition at $\partial\mathcal{O}$ models the fact that we assume there is no flux of energy at the boundary. Under suitable conditions on the domain, the above PDE is well posed as an abstract evolution equation in H. Precisely (see, e.g., [48, Sec.3.1]), defining the operator

$$\mathcal{L} : \mathcal{D}(\mathcal{L}) \subseteq H \to H,$$

where

$$\mathcal{D}(\mathcal{L}) = W_0^{2,2}(\mathcal{O}) := \left\{ f \in W^{2,2}(\mathcal{O}) : \ \frac{\partial f}{\partial \mathbf{n}} = 0 \text{ on } \mathcal{O} \right\}, \quad \mathcal{L} f := \Delta f - \delta f,$$

one has that $\mathcal{L}$ generates a strongly continuous semigroup of linear operators in H and (6.1) can be rewritten in abstract terms as

$$\mathrm{d}E_t^0 = \mathcal{L} E_t^0 \mathrm{d}t, \quad E_0 = e.$$

Assume now that the company can implement irreversible investment strategies $I_t(\xi)$ in order to adjust the production capacity. The energy supply over the region, i.e. the spatial process $E_t^I(\xi)$, then evolves according to the controlled abstract evolution equation

$$\mathrm{d}E_t^I = \mathcal{L} E_t^I \mathrm{d}t + \mathrm{d}I_t, \quad E_{0^-}^I = e.$$

Next, we model the total demand of energy as a spatial process $A_t(\xi)$ evolving according to the SPDE

$$\mathrm{d}A_t(\xi) = \mathcal{B} A_t(\xi) \mathrm{d}t - \sigma \mathrm{d}W_t(\xi), \quad A_0(\xi) = a(\xi),$$

for some initial maximal demand $a \in H_+$, some bounded nonnegative self-adjoint linear operator $\mathcal{B} \in \mathcal{L}(H)$ such that $\mathcal{L} - \mathcal{B}$ satisfies Assumption 2.2, and for σ and W satisfying the requirements of Section 2. The set of admissible investment strategies is therefore naturally modeled by the class $\mathcal{I}$ (cf. (2.12)). Moreover, we model the inverse demand function for energy assuming that it depends on the location $\xi \in \mathcal{O}$ and is linear in the quantity $E_t^I(\xi)$ that is being delivered; that is,

$$p_t(\xi, E_t^I(\xi)) = A_t(\xi) - B(\xi) E_t^I(\xi),$$

for some function $B(\xi) > 0$. Here, $A_t(\xi)/B(\xi)$ is the maximal possible demand at location ξ at time t. We set $B \equiv 1$ in the following (just for simplicity). Then, the total surplus at a given time t and location ξ is given by

$$U_t(\xi) = \int_0^{E_t(\xi)} p_t(\xi, z) \mathrm{d}z = A_t(\xi) E_t(\xi) - \frac{1}{2} \left(E_t(\xi) \right)^2,$$

so that the overall total surplus at time t is

$$S_t = \int_D U_t(\xi) \mu(\mathrm{d}\xi) = \langle A_t, E_t^I \rangle_H - \frac{1}{2} \langle E_t^I, E_t^I \rangle_H = -\frac{1}{2} \langle E_t^I - A_t, E_t - A_t \rangle_H + \frac{1}{2} \langle A_t, A_t \rangle_H.$$

As the last term is independent of the control I, it is irrelevant in the optimization process we are going to define and we ignore it in the following.

Set now $X_t^I := E_t^I - A_t$ and $\mathcal{A} := \mathcal{L} - \mathcal{B}$. Then we have

$$\mathrm{d}X_t^I = \mathcal{A} X_t^I \mathrm{d}t + \sigma \mathrm{d}W_t + \mathrm{d}I_t, \quad X_{0^-}^I = e - a =: x, \tag{6.2}$$

The latter controlled SPDE satisfies our Assumptions 2.1, 2.2, and 2.4.

For an intertemporal discount factor $\rho > 0$, the energy producer aims at maximizing the total expected surplus, net of the investment costs. Assuming that the cost of investment may vary with location ξ and that it is modeled by a local price $q \in H_+$, bounded away from zero – that is, there exists $q_o > 0$ such that $q(\xi) \geq q_o$ for every $\xi \in \mathcal{O}$ – the company's optimization problem is

$$\sup_{I \in \mathcal{I}} \mathsf{E} \left[\int_0^\infty e^{-\rho t} \Big(-\frac{1}{2} |X_t|_H^2 - \langle q, \mathrm{d}I_t \rangle_H \Big) \right].$$

The latter is clearly equivalent to minimizing the cost functional (2.13) with $G(x) = |x|_H^2$.

6.2. **An energy balance climate model with human impact.** Let us consider a one-dimensional Energy Balance Climate Model with Human Impact (see, e.g., [10]; the basic climate model dates back to Gerald North, see [53]).

We first describe quickly the basics of an energy balance climate model. The earth's temperature is taken to be the result of incoming radiation from the sun and outgoing radiation through reflection. We consider temperature on the hemisphere, modeled by the half-circle that we identify with the interval $D := [-1, 1]$, where $\xi \in [-1, 1]$ is the sine of latitude. $\mathcal{M}$ is the Borel σ-algebra on D, and μ is taken to be the Lebesgue measure on $(D, \mathcal{M})$.

We now model the temperature evolution $T_t(\xi)$ over time t at location $\xi \in D$. The incoming radiation at ξ is denoted by $R(\xi) = QS(\xi)\alpha(\xi)$. Here, $S(\xi)$ is the solar energy arriving at latitude ξ, Q is the solar constant divided by 4, and $\alpha(\xi)$ describes the amount of heat absorbed at location ξ (co-albedo); in general, it depends on temperature and location, but here as in [53] we assume that it is just a function of $\xi \in D$. The outgoing infrared radiation at location ξ is linear in temperature, say $\gamma + \eta T_t(\xi)$, for two constants γ and $\eta > 0$. On the surface, we have a typical heat transport that is modeled via the second derivative $\frac{\mathrm{d}^2}{\mathrm{d}\xi^2}$, or more generally, by the operator

$$(\mathcal{B}f)(\xi) := \frac{\mathrm{d}}{\mathrm{d}\xi}\left(D(\xi)\frac{\mathrm{d}}{\mathrm{d}\xi}f(\xi)\right),$$

for some diffusion coefficient D driving the heat transport. The overall resulting energy-balance operator

$$\big(\mathcal{Q}f\big)(\xi) := QS(\xi)\alpha(\xi) - \gamma - \eta f(\xi) + \big(\mathcal{B}f\big)(\xi) \tag{6.3}$$

describes the energy balance of the earth without human impact. The equilibrium temperature distribution $T^\star(\xi)$ is given by the solution to the partial differential equation $\mathcal{Q}T^\star = 0$, subject to appropriate boundary conditions (for instance, both zero Neumann or periodic boundary conditions can be chosen).

We now add human impact due to carbon emissions. Let global cumulative human carbon emissions be described by the (real-valued) process $\nu \in \mathcal{S}$ (cf. (2.16)).

The temperature evolution at time t is then described by

$$\mathrm{d}T_t = \mathcal{Q}T_t\mathrm{d}t + \sigma\mathrm{d}W_t + \mathbf{1}\mathrm{d}\nu_t, \quad T_{0^-} = x \in H, \tag{6.4}$$

with $\mathbf{1}$ being the unitary vecor in H, and with σ and W as specified in Section 2. In particular, the Brownian motion W takes care of noise and unmodeled influences.

The dynamics (6.4) does not fit exactly into our setting because of the constant (in time) drift term

$$b(\xi) := QS(\xi)\alpha(\xi) - \gamma,$$

but this problem can be easily fixed. Define

$$X_t := T_t - T^\star, \qquad \mathcal{A}f := -\eta f + \mathcal{B}f,$$

so that

$$\mathcal{Q}f = \mathcal{A}f + b.$$

Recalling that $T^\star$ is an equilibrium distribution for the temperature, we may write

$$\begin{aligned}
\mathrm{d}X_t = \mathrm{d}(T_t - T^\star) &= \mathcal{Q}T_t\mathrm{d}t - \mathcal{Q}T^\star\mathrm{d}t + \sigma\mathrm{d}W_t + \mathbf{1}\mathrm{d}\nu_t \\
&= \mathcal{A}T_t\mathrm{d}t - \mathcal{A}T^\star\mathrm{d}t + \sigma\mathrm{d}W_t + \mathbf{1}\mathrm{d}\nu_t \\
&= \mathcal{A}X_t\mathrm{d}t + \sigma\mathrm{d}W_t + \mathbf{1}\mathrm{d}\nu_t
\end{aligned}$$

and we are back in our setting. In particular, the operator $\mathcal{A}$ with the aforementioned boundary conditions satisfies Assumptions 2.1, 2.2, and 2.4 (for the null Neumann boundary conditions one, see, e.g., [48, Sec.3.1]; for the periodic ones, see Section 5.2 in [28]).

Assume now that a decision maker (for instance, in this context, the United Nations) has an ideal profile of temperature $\widehat{T}$ in mind; this could be the pre-industrial equilibrium temperature $T^\star$ or a temperature distribution in its vicinity. The planner thus aims at minimizing the average square distance $|T-\widehat{T}|_H^2$ to that ideal temperature and measures the cost of investment into capacity by some price $q > 0$. Then, the resulting minimization problem in terms of the controlled process X^ν is (cf. (2.12))

$$\inf_{\nu \in \mathcal{S}} \mathsf{E}\bigg[\int_0^\infty e^{-\rho t}\Big(\big|X_t^\nu + T^\star - \widehat{T}\big|_H^2 \mathrm{d}t + q\mathrm{d}\nu_t\Big)\bigg],$$

for some intertemporal discount rate $\rho > 0$. This falls into our setting for $G(x) = |x+T^\star-\widehat{T}\big|_H^2$. Such a specification greatly simplifies the full economic model which would be beyond the scope of the current paper. Compare [11] for an attempt in that direction.

Acknowledgements. Funded by the *Deutsche Forschungsgemeinschaft* (DFG, German Research Foundation) - Project-ID 317210226 - SFB 1283. Salvatore Federico is grateful to Andrzej Swiech for the valuable discussions on viscosity solutions to variational inequalities in Hilbert spaces that led to the formulation of Remarks 3.9 and 4.10.

Appendix A. A result on semiconcave and semiconvex functions

We state here some properties of semiconcave and semiconvex functions on H.

Lemma A.1. (i) *Let H be endowed with an arbitrary norm $|\cdot|$ and take $F : H \to \mathbb{R}$ be semiconvex (with respect to the norm $|\cdot|$); that is, there exists $c_F > 0$ such that*

$$\lambda F(x) + (1-\lambda)F(y) - F(\lambda x + (1-\lambda)y) \geq -\frac{c_F}{2}\lambda(1-\lambda)|x-y|^2, \quad \forall x, y \in S, \ \lambda \in [0,1].$$

Then

$$x \mapsto F(x) + \frac{c_F}{2}|x|^2$$

is convex.

(ii) *Let $F : H \to \mathbb{R}$ be both semiconcave and semiconvex (with respect to the norm $|\cdot|_H$); that is, there exists $c_F > 0$ such that*

$$|\lambda F(x) + (1-\lambda)F(y) - F(\lambda x + (1-\lambda)y)| \leq \frac{c_F}{2}\lambda(1-\lambda)|x-y|_H^2,$$

for every $x, y \in S$ and $\lambda \in [0,1]$. Then $F \in C^{1,Lip}(H)$.

Proof. Proof of (i). This follows from [14, Prop. 1.1.3] (equivalences (a)-(c)), whose proof is not affected by the dimensionality of the space under consideration.

Proof of (ii). This follows by adapting the convex-analytic arguments of the finite-dimensional setting in [14, Ch.3]. We report these here for completeness. Let $x \in S$. Since F is semiconvex and semiconcave, we have that both the supergradient $D^+F(x)$ and the subgradient $D^-F(x)$ are not empty. We now show that F is differentiable at $x \in H$. Suppose, by the aim of contradiction, that $p_1 \in D^-F(x)$ and $p_2 \in D^+F(x)$ and that $p_1 \neq p_2$. Then, for $h \in H$, by definition of subgradient and by Proposition 3.3.1 in [14, Ch.3] (whose proof does not suffer the dimensionality of the considered space) in the case of linear modulus $\omega(r) = \frac{c_F}{2}r$, $r \geq 0$, we find

$$\langle p_1, h\rangle_H - \frac{c_F}{2}|h|_H^2 \leq F(x+h) - F(x) \leq \langle p_2, h\rangle_H + \frac{c_F}{2}|h|_H^2. \tag{A-1}$$

Therefore, we obtain

$$0 \geq \langle p_1 - p_2, h\rangle_H - c_F|h|_H^2.$$

Given the arbitrariness of h, we can now take $h = \alpha(p_1 - p_2)$, for $\alpha > c_F$ so to achieve a contradiction. Hence, $p_1 = p_2$, which, using (A-1), leads to

$$\frac{|F(x+h) - F(x) - \langle p_1, h\rangle_H|}{|h|_H} \leq \frac{c_F}{2}|h|_H, \tag{A-2}$$

and therefore, in particular, to the differentiability of F at $x \in H$ with $DF(x) = p_1$.

We now show that $DF : H \to H$ is Lipschitz continuous. To that end, we borrow arguments from the proof of Theorem 3.3.7 in [14, Ch.3], which, once more, does not suffer the dimensionality of the considered space. By (A-2) with $p_1 = DF(x)$, for any $x, y \in H$ we have

$$|F(y) - F(x) - \langle DF(x), y - x\rangle_H| \leq \frac{c_F}{2}|y - x|_H^2. \tag{A-3}$$

Then, for each $x, v, w \in H$, we have

$$\langle DF(x+w), v\rangle_H \leq F(x+w+v) - F(x+w) + \frac{c_F}{2}|v|_H^2,$$

and

$$\langle DF(x), v\rangle_H \geq F(x+v) - F(x) - \frac{c_F}{2}|v|_H^2.$$

Hence, from the last two display equations, we obtain

$$\begin{aligned}
&\langle DF(x+w) - DF(x), v\rangle_H \leq F(x+w+v) - F(x+w) - F(x+v) + F(x) + c_F|v|_H^2 \\
&= F(x+w+v) - \frac{1}{2}F(x+2v) - \frac{1}{2}F(x+2w) \\
&- F(x+w) + \frac{1}{2}F(x+2w) + \frac{1}{2}F(x) \\
&- F(x+v) + \frac{1}{2}F(x+2v) + \frac{1}{2}F(x) + c_F|v|_H^2 \\
&\leq 2c_F\left(|w|_H^2 + |v|_H^2\right),
\end{aligned}$$

where in the last estimate we used the semiconcavity and the semiconvexity of F. The latter estimate now implies the claimed Lipschitz property, since

$$|DF(x+w) - DF(x)|_H = \frac{1}{|w|_H}\max_{|v|_H = |w|_H}\langle DF(x+w) - DF(x), v\rangle_H \leq 2c_F|w|_H.$$

□

Appendix B. Some technical results

Lemma B.1. *Let $x \in H$, $I \in \mathcal{I}$, $(\mathcal{A}_n)_{n\in\mathbb{N}}$ be the Yosida approximants of $\mathcal{A}$, and denote by $X^{n;x,I}$ the unique mild solution to*

$$\mathrm{d}X_t^{n;x,I} = \mathcal{A}_n X_t^{n;x,I}\mathrm{d}t + \sigma\mathrm{d}W_t + \mathrm{d}I_t, \quad X_{0^-} = x \in H;$$

that is,

$$X_t^{n;x,I} = e^{t\mathcal{A}_n}x + W_t^{\mathcal{A}_n,\sigma} + \int_{0^-}^t e^{(t-s)\mathcal{A}_n}\mathrm{d}I_s, \quad t \geq 0. \tag{B-1}$$

For $p \geq 2$ as in (2.12), for any $T > 0$ and for some $M > 0$ it holds:

(i) $\mathsf{E}\big[\sup_{t\in[0,T]}\big|X_t^{n;x,I}\big|_H^p\big] \leq M\Big(1 + |x|_H^p + \mathsf{E}\Big[\Big|\int_0^T\big|\hat{\vartheta}_s\big|_H\mathrm{d}|I|_s\Big|^p\Big]\Big)$;

(ii) $\lim_{n\uparrow\infty}\mathsf{E}\big[\sup_{t\in[0,T]}\big|X_t^{n;x,I} - X_t^{x,I}\big|_H^p\big] = 0$.

Proof. We start by proving (i). By (B-1), the Burkholder-Davis-Gundy inequality (see Theorem 1.111 in [27]), the fact that $|e^{t\mathcal{A}_n}|_{\mathcal{L}(H)} \leq e^{-\frac{n\delta}{n+\delta}t}$ (cf. Equation (B-14) in [27]), and an estimate analogous to (2.10), we have, for some constant $M > 0$ (possibly depending on p and changing from line to line),

$$\begin{aligned}
\mathsf{E}\Big[\sup_{t\in[0,T]}\big|X_t^{n;x,I}\big|_H^p\Big] &\leq M\Big(|x|_H^p + \mathsf{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^t e^{(t-s)\mathcal{A}_n}\sigma \mathrm{d}W_s\Big|_H^p\Big] \\
&\quad + \mathsf{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^t e^{(t-s)\mathcal{A}_n}\hat{\vartheta}_s \mathrm{d}|I|_s\Big|_H^p\Big]\Big) \\
&\leq M\Big(|x|_H^p + \mathsf{E}\Big[\int_0^T |\sigma\sigma^*|_{\mathcal{L}_1(H)}\mathrm{d}s\Big]^{\frac{p}{2}} + \mathsf{E}\Big[\Big|\int_0^T \big|\hat{\vartheta}_s\big|_H \mathrm{d}|I|_s\Big|^p\Big]\Big) \\
&\leq M\Big(1 + |x|_H^p + \mathsf{E}\Big[\Big|\int_0^T \big|\hat{\vartheta}_s\big|_H \mathrm{d}|I|_s\Big|^p\Big]\Big),
\end{aligned}$$

where the last expectation is finite due to the fact that $I \in \mathcal{I}$ (cf. (2.12)). This proves the first claim.

As for (ii) notice that, for a constant $M > 0$ changing from line to line,

$$\begin{aligned}
\mathsf{E}\Big[\sup_{t\in[0,T]}\big|X_t^{n;x,I} - X_t^{x,I}\big|_H^p\Big] &\leq M\Big(\sup_{t\in[0,T]}\big|\big(e^{t\mathcal{A}_n} - e^{t\mathcal{A}}\big)x\big|_H^p \\
&\quad + \mathsf{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^t \Big(e^{(t-s)\mathcal{A}_n} - e^{(t-s)\mathcal{A}}\Big)\sigma \mathrm{d}W_s\Big|_H^p\Big] \\
&\quad + \mathsf{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^t \Big(e^{(t-s)\mathcal{A}_n} - e^{(t-s)\mathcal{A}}\Big)\hat{\vartheta}_s \mathrm{d}|I|_s\Big|_H^p\Big]\Big).
\end{aligned}$$

The first two addends on the right-hand side of the latter display equation converge to zero as $n \uparrow \infty$ as in the proof of [27, Thm. 1.131]. In order to deal with the third addend, define

$$\psi_n(s) := \sup_{t\in[s,T]}\Big|\Big(e^{(t-s)\mathcal{A}_n} - e^{(t-s)\mathcal{A}}\Big)\hat{\vartheta}_s\Big|_H, \quad s \in [t,T],$$

which is such that $\psi_n(s) \to 0$ as $n \uparrow \infty$ P-a.s. by [27, Prop. B.34]. Since now $|\psi_n(s)| \leq 2|\hat{\vartheta}_s|_H$, and, for any $T > 0$, $\int_0^T |\hat{\vartheta}_s|_H \mathrm{d}|I|_s < \infty$ P-a.s. by (2.3) and $\mathsf{E}[|\int_0^T |\hat{\vartheta}_s|_H \mathrm{d}|I|_s|^p] < \infty$ because $I \in \mathcal{I}$, the dominated convergence theorem gives

$$\lim_{n\uparrow\infty}\mathsf{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^t \Big(e^{(t-s)\mathcal{A}_n} - e^{(t-s)\mathcal{A}}\Big)\hat{\vartheta}_s \mathrm{d}|I|_s\Big|_H^p\Big] \leq \lim_{n\uparrow\infty}\mathsf{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^t \psi_n(s)\mathrm{d}|I|_s\Big|_H^p\Big] = 0$$

□

By arguing as in the proof of Lemma B.1(i) one can also prove the following.

Lemma B.2. *Let $x \in H$, $I \in \mathcal{I}$ and let $X^{x,I}$ denote the unique mild solution to (2.7); that is,*

$$X_t^{x,I} = e^{t\mathcal{A}}x + W_t^{\mathcal{A},\sigma} + \int_{0^-}^t e^{(t-s)\mathcal{A}}\mathrm{d}I_s, \quad t \geq 0.$$

Let $p \geq 2$ as in (2.12). Then, for any $T > 0$ and for some $M > 0$ it holds:

$$\mathsf{E}\Big[\sup_{t\in[0,T]}\big|X_t^{x,I}\big|_H^p\Big] \leq M\Big(1 + |x|_H^p + \mathsf{E}\Big[\Big|\int_0^T \big|\hat{\vartheta}_s\big|_H \mathrm{d}|I|_s\Big|^p\Big]\Big).$$

References

[1] Agram, N., Hilbert, A., Øksendal, B. (2020). Singular Control of SPDEs with Space-Mean Dynamics. *Math. Control Relat. Fields*, **10(2)** 425-441.

[2] Alvarez, L.H.R. (2000). Singular Stochastic Control in the Presence of a State-dependent Yield Structure. *Stoch. Process. Appl.* **86** 323-343.

[3] Arendt, W., Grabosch, A., Greiner, G., Groh, U., Lotz, H.P., Moustakas, U., Nagel, N., Neubrander, F., Schlotterbeck, U. (1986). *One-Parameter Semigroups of Positive Operators.*Lecture Notes in Math. 1184, Springer-Verlag, Berlin.

[4] Baldursson, F., Karatzas, I. (1997). Irreversible Investment and Industry Equilibrium. *Fin. Stoch.* **1**, 69–89.

[5] Barbu, V., Marinelli, C. (2008). Variational Inequalities in Hilbert Spaces with Measures and Optimal Stopping Problems. *Appl. Math. Optim.* **58(2)** 237-262.

[6] Bather, J.A., Chernoff, H. (1966). Sequential decisions in the control of a spaceship. *Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability* **3** 181–207.

[7] Beneš, V.E., Shepp, L.A., Witsenhausen, H.S. (1980). Some Solvable Stochastic Control Problems. *Stochastics* **4(1)** 39-83.

[8] Bensoussan, A., Da Prato G., Delfour M.C., Mitter S.K. (2006). *Representation and Control of Infinite Dimensional Systems* (Second Edition). Birkhäuser.

[9] Boetius, F., Kohlmann, M. (1998). Connections between Optimal Stopping and Singular Stochastic Control. *Stochastic Process. Appl.* **77** 253-281.

[10] Brock, W.A., Engström, G., Grass, D., Xepapadeas, A. (2013). Energy Balance Climate Models and General Equilibrium Optimal Mitigation Policies. *J. Econ. Dyn. Con.* **37(12)** 2371–2396.

[11] Brock, W.A., Engström, G., Xepapadeas, A. (2013). Spatial Climate-Economic Models in the Design of Optimal Climate Policies across Locations. *Eur. Econ. Rev.* **69** 78–103.

[12] Brezis, H. (2010). *Functional Analysis, Sobolev Spaces and Partial Differential Equation.* Springer.

[13] Calvia, A., Federico, S., Gozzi, F. (2021). State Constrained Control Problems in Banach Lattices and Applications. *SIAM J. Control Optim.* **59** 4481-4510.

[14] Cannarsa, P., Sinestrari C. (2004). *Semiconcave Functions, Hamilton–Jacobi Equations, and Optimal Control.* Birkhauser.

[15] Chang, M.H., Peng, T., Pemy, M. (2012). Viscosity Solution of Optimal Stopping Problem for Stochastic Systems with Bounded Memory. *Stoch. Anal. Appl.* **30** 1102–1135.

[16] Chiarolla, M.B. (1997). Singular Stochastic Control of a Singular Diffusion Process. *Stoch. Stoch. Rep.* **62(1-2)** 31-63.

[17] Chiarolla, M.B., Haussmann, U. (2000). Controlling Inflation: The Infinite Horizon Case. *Appl Math Optim.* **41** 25-50.

[18] Chiarolla, M.B., De Angelis, T. (2016). Optimal stopping of a Hilbert space valued diffusion: an infinite dimensional variational inequality. *Appl. Math. Optim.* **73(2)** 271-312.

[19] Chow, P.L., Menaldi, J.L. (2006). Variational Inequalities for the Control of Stochastic Partial Differential Equations. In: *Stochastic Partial Differential Equations and Applications II*: Proceedings of a Conference held in Trento, Italy February 1–6, 1988 (pp. 42-52). Springer Berlin Heidelberg.

[20] Clément, P., Heijmans, H.J.A.M., Angenent, S., Van Duijn, C.J., De Pagter, B. (1987). *One Parameter Semigroups.* CWI Monographs 5, North-Holland Publishing Co., Amsterdam.

[21] Da Prato, G., Zacbzyk, J. (2014). *Stochastic Equations in Infinite Dimensions* (Second Edition). Encyclopedia of Mathematics and its Applications **152**. Cambridge University Press.

[22] De Angelis, T., Milazzo, A. (2023). Dynamic Programming Principle for Classical and Singular Stochastic Control with Discretionary Stopping. *Appl. Math. Optim.* **88(7)**.

[23] Dianetti, J., Ferrari, G. (2023). Multidimensional Singular Control and Related Skorokhod Problem: Sufficient Conditions for the Characterization of Optimal Controls. *Stoch. Process. Appl.* **162** 547–592.

[24] Diestel, J., Uhl, J.J. Jr. (1977). *Vector Measures.* Mathematical Surveys-Number 15. American Mathematical Society.

[25] Ekeland, I., Temam, R. (1999). *Convex Analysis and Variational Problems.* SIAM Press.

[26] Engel K.J., Nagel R. (1995). *One-parameter Semigroups for Linear Evolution Equations.* Springer, Graduate Texts in Mathematics **194**.

[27] Fabbri, G., Gozzi, F., Swiech, A. (2017). *Stochastic Optimal Control in Infinite Dimension: Dynamic Programming and HJB Equations.* Probability Theory and Stochastic Modelling **82**. Springer.

[28] Federico, S., Ferrari, G., Riedel, F., Röckner, M. (2021). On a Class of Infinite-Dimensional Singular Stochastic Control Problems. *SIAM J. Control Optim.* **59(2)** 1680–1704.

[29] Federico, S., Øksendal, B. (2011). Optimal Stopping of Stochastic Differential Equations with Delay Driven by a Lévy Noise. *Potential Anal.* **34(2)** 181–198.
[30] Federico, S., Pham, H. (2014). Characterization of the Optimal Boundaries in Reversible Investment Problems. *SIAM J. Control Optim.* **52(4)** 2180–2213.
[31] Ferrari, G. (2015). On an Integral Equation for the Free-Boundary of Stochastic, Irreversible Investment Problems. *Ann. Appl. Probab.* **25(1)** 150–176.
[32] Ferrari, G. (2018). On the Optimal Management of Public Debt: A Singular Stochastic Control Problem. *SIAM J. Control Optim.* **56(3)** 2036-2073.
[33] Ferrari, G., Koch, T. (2021). An Optimal Extraction Problem with Price Impact. *Appl. Mathe. Optim.* **83** 1951-1990.
[34] Fleming, W.H., Soner, H.M. (2005). *Controlled Markov Processes and Viscosity Solutions*. 2nd Edition. Springer.
[35] Fuhrman, M., Masiero, F., Tessitore, G. (2017). Reflected BSDEs, Optimal Control and Stopping for Infinite-Dimensional Systems. *ESAIM Control Optim. Calc. Var.* **23(4)** 1419–1445.
[36] Gatarek, D., Swiech, A. (1999). Optimal Stopping in Hilbert Spaces and Pricing of American Options. *Math. Meth. Oper. Res.* **50** 135-147.
[37] Hausenblas, E., Seidler, J. (2001). A Note on Maximal Inequality for Stochastic Convolution. *Czechoslov. Math. J.* **51(126)** 785-790.
[38] Haussmann, U.G., Suo, W. (1995). Singular Optimal Stochastic Controls II: Dynamic Programming. *SIAM J. Control Optim.* **33(3)** 937-959.
[39] Hynd, R., Mawi, H. (2016). On Hamilton-Jacobi-Bellman Equations with Convex Gradient Constraints. *Interface Free Bound.* **18** 291-315.
[40] Hynd, R. (2013). Analysis of Hamilton-Jacobi-Bellman Equations Arising in Stochastic Singular Control. *ESAIM Control Optim. Calc. Var.* **19** 112-128.
[41] Jeanblanc-Picqué, M., Shiryaev, A.N. (1995). Optimization of the Flow of Dividends. *Russian Math. Surveys*, **50(2)**, 257–277.
[42] Karatzas, I. (1981). The Monotone Follower Problem in Stochastic Decision Theory. *Appl. Math. Optim.* **7** 175–189.
[43] Karatzas, I. (1983). A Class of Singular Stochastic Control Problems. *Adv. Appl. Prob.* **15** 225-254.
[44] Karatzas, I., Shreve, S.E. (1984). Connections Between Optimal Stopping and Singular Stochastic Control I. Monotone Follower Problems. *SIAM J. Control Optim.* **22** 856–877.
[45] Karatzas, I., Shreve, S.E. (1986). Equivalent Models for Finite-fuel Stochastic Control. *Stochastics* **17** 245-276.
[46] Kruk, L. (2000). Optimal Policies for n-Dimensional Singular Stochastic Control Problems Part I: The Skorokhod Problem. *SIAM J. Control. Optim.* **38(5)** 1603-1622.
[47] Liu, W., Röckner, M. (2015). *Stochastic Partial Differential Equations: An Introduction*. Universitext, Springer.
[48] Lunardi A. (1995). *Analytic Semigroups and Optimal Regularity in Parabolic Problems*. Birkhäuser.
[49] Ma. J., Yong, J. (1993). Regular-Singular Stochastic Controls for Higher Dimensional Diffusions - Dynamic Programming Approach. IMA Preprint Series 1170.
[50] Menaldi, J.L., Taksar, M.I. (1989). Optimal Correction Problem of a Multidimensional Stochastic System. *Automatica* **25(2)**, 223-232.
[51] Merhi, A., Zervos, M. (2007). A Model for Reversible Investment Capacity Expansion. *SIAM J. Control Optim.* **46(3)** 839-876.
[52] Metivier, M. (1982). *Semimartingales: A Course on Stochastic Processes*. De Gruyter.
[53] North, G.R. (1975). Analytical Solution to a Simple Climate Model with Diffusive Heat Transport. *J. Atmos. Sci.* **32(7)** 1301–1307.
[54] Øksendal, B., Sulem, A., Zhang, T. (2014). Singular Control and Optimal Stopping of SPDEs and Backward SPDEs with Reflection. *Math. Oper. Res.* **39(2)** 464–486.
[55] Peskir, G., Shiryaev, A.N. (2006). *Optimal Stopping and Free-Boundary Problems*. Springer.
[56] Pham, H. (2009). *Continuous-time Stochastic Control and Optimization with Financial Applications*. Springer.
[57] Shreve, S.E., Soner, H.M. (1989). Regularity of the Value Function for a Two-Dimensional Singular Stochastic Control Problem. *SIAM J. Control Optim.* **27(4)** 876 – 907.
[58] Shreve, S.E., Soner, H.M. (1994). Optimal Investment and Consumption with Transaction Costs. *Ann. Appl. Probab.* **4(3)** 609 – 692.
[59] Zabczyk, J. (1997). Stopping Problems on Polish Spaces. *Ann. Univ. Mariae Curie-Skłodowska Sect. A* **51(1)** 181–199.

S. Federico: Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato 5, 40126, Bologna, Italy
Email address: s..federico@unibo.it

G. Ferrari: Center for Mathematical Economics (IMW), Bielefeld University, Universitätsstrasse 25, 33615, Bielefeld, Germany
Email address: giorgio.ferrari@uni-bielefeld.de

F. Riedel: Center for Mathematical Economics (IMW), Bielefeld University, Universitätsstrasse 25, 33615, Bielefeld, Germany
Email address: frank.riedel@uni-bielefeld.de

M. Röckner: Faculty of Mathematics, Bielefeld University, Universitätsstrasse 25, 33615, Bielefeld, Germany
Email address: roeckner@math.uni-bielefeld.de