ESAIM: COCV 23 (2017) 39–70
DOI: 10.1051/cocv/2015039

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

SENSITIVITY RESULTS IN STOCHASTIC OPTIMAL CONTROL: A LAGRANGIAN PERSPECTIVE *

J. Backhoff[1] and F.J. Silva[2]

Abstract. In this work we provide a first order sensitivity analysis of some parameterized stochastic optimal control problems. The parameters and their perturbations can be given by random processes and affect the state dynamics. We begin by proving a one-to-one correspondence between the adjoint states appearing in a weak form of the stochastic Pontryagin principle and the Lagrange multipliers associated to the state equation when the stochastic optimal control problem is seen as an abstract optimization problem on a suitable Hilbert space. In a first place, we use this result and classical arguments in convex analysis, to study the differentiability of the value function for convex problems submitted to linear perturbations of the dynamics. Then, for the linear quadratic and the mean variance problems, our analysis provides the stability of the optimizers and the C^1-differentiability of the value function, as well as explicit expressions for the derivatives, even when the data perturbation is not convex in the sense of [R.T. Rockafellar, Conjugate duality and optimization. Society for Industrial and Applied Mathematics, Philadelphia, Pa. (1974).

Mathematics Subject Classification. 93E20, 49Q12, 47J30, 49N10, 91G10.

Received September 4, 2014. Accepted July 16, 2015.

1. Introduction

One of the most important results in stochastic optimal control theory is Pontryagin's Principle, introduced and refined by [2, 5, 15, 20, 24] among others (see [31], Chap. 3, Sect. 7 for a historical account). In its simplest form, it says that almost surely the optimal control minimizes an associated *Hamiltonian.* This Hamiltonian depends on the optimal state and an *adjoint pair*, which solves an associated Backward Stochastic Differential Equation (BSDE for short). Roughly speaking, the mentioned necessary condition appears as one perturbs the optimal control and analyzes up to first order (or second-order, if the volatility term is controlled and the set of admissible controls is non-convex) the impact of such perturbation on the cost function. A natural

Keywords and phrases. Stochastic control, Pontryagin principle, Lagrange multipliers, sensitivity analysis, LQ problems, mean variance portfolio selection problem.

* *This work was initiated during a one month visit to the Hausdorff Research Institute for Mathematics at the University of Bonn in the framework of the Trimester Program Stochastic Dynamics in Economics and Finance.*

[1] Fakultät für Mathematik, Universität Wien, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. julio.backhoff@univie.ac.at

[2] Institut de recherche XLIM-DMI, UMR-CNRS 7252 Faculté des sciences et techniques Université de Limoges, 87060 Limoges, France. francisco.silva@unilim.fr

Article published by EDP Sciences

question that arises is whether by regarding the stochastic optimal control problem as an infinite dimensional optimization problem in an appropriate functional setting, the usual machinery of optimization theory yields an interpretation of the aforementioned adjoint states. From this perspective it is conceivable that fundamental tools such as convex-duality, Lagrange multipliers and non-smooth analysis (to name a few) may shed new lights and provide new interpretations into the field of stochastic optimal control.

The idea of dealing with stochastic optimal control problems from the point of view of abstract optimization theory is not new. In a remarkable article [3], the author extends to the stochastic case the results of [28] obtained in the deterministic framework. For convex problems, he proves essentially that the solutions of the original optimization problem and its dual, in the sense of convex analysis, must fulfil the conditions appearing in Pontryagin's Principle. In the non-convex case, a very interesting analysis is performed in [21] where the author uses non-smooth analysis techniques to tackle the case of a non-linear controlled Stochastic Differential Equation (SDE for short)[3].

In this article we develop a rigorous functional framework under which the Lagrangian approach to stochastic optimal control becomes fruitful. As a matter of fact, we relate the adjoint states appearing in the Pontryagin principle with the Lagrange multipliers of the associated optimization problem, thus extending the results of [3] in the convex case, by using a different method. In several interesting cases, this result allows us to perform a *first order sensitivity analysis of the value function*, under *random functional perturbations* of the dynamics. To the best of our knowledge, this type of sensitivity result had been obtained for finite dimensional perturbations of the initial condition (see the works [21, 32, 33]) only.

We restrict ourselves to a finite-horizon, brownian setting, yet consider the case of non-linear controlled SDEs with random coefficients and the control being present both in the drift and diffusion parts, pointwise convex constraints on the controls, and finite dimensional constraints of expectation-type on the final state. In mathematical language, we deal with problems of the form:

$$\left.\begin{array}{ll} \inf\limits_{(x,u)} \mathbb{E}\left[\displaystyle\int_0^T \ell(\omega,t,x(t),u(t))\mathrm{d}t + \varPhi(\omega,x(T))\right] & \\ \text{s.t.} \quad x(t) = x_0 + \int_0^t f(s,x(s),u(s))\mathrm{d}s + \int_0^t \sigma(s,x(s),u(s))\mathrm{d}W(s), \ \forall\, t\in[0,T[, & \\ \qquad \mathbb{E}\left(\varPhi_E(x(T))\right) = 0, \quad \mathbb{E}\left(\varPhi_I(x(T))\right)\le 0, \quad u(\omega,t)\in U \ \text{ a.s.} & \end{array}\right\} \qquad (CP)$$

where ℓ, $\varPhi$, f, σ, x_0, $\varPhi_E$, $\varPhi_I$ are the data of the problem, which can be random, satisfying some natural requirements detailed in Section 3, and $U \subseteq \mathbb{R}^m$ is a convex set. Under standard assumptions, we have that for every square integrable and progressively measurable control u, there exists a unique solution $x[u]$ of the SDE in (CP). In this sense, problem (CP) can be reformulated in terms of u only and the SDE constraint can be eliminated. However, we have chosen to work with the pair (x,u) and keep the SDE constraint in order to associate to it a Lagrange multiplier, in view of the important consequences of this approach in the sensitivity analysis of the optimal cost of (CP) (see Sect. 4).

By defining a Hilbert space topology on the space of Itô processes, we naturally deduce that whenever the Lagrange multipliers associated to the SDE constraint in (CP) exist they must be Itô processes themselves. With this methodology we can prove a one-to-one simple relationship between the aforementioned Lagrange multipliers and the adjoint states appearing in a weak form of Pontryagin's principle. More concretely, we say that (p,q) is a *weak-Pontryagin multiplier* at a solution (x,u) if the same conditions appearing in the usual Pontryagin principle hold true (see [24], Thm. 3), except for the condition of minimization of the Hamiltonian which is replaced by the weaker statement corresponding to its first order optimality condition (see Sect. 3.1 for a detailed exposition). Thus, it is easily seen that every adjoint pair appearing in the usual Pontryagin principle is a weak-Pontryagin multiplier. In Theorem 3.12 we prove that given a weak-Pontryagin multiplier (p,q), the

[3] More recently, in *e.g.* [13, 19], the Lagrange multiplier technique has been applied formally in order to derive optimality conditions. However, no connexions with Pontryagin's principle are analyzed.

process

$$\lambda(\cdot) := p(0) + \int_0^{\cdot} p(s)\mathrm{d}s + \int_0^{\cdot} q(s)\mathrm{d}W(s), \tag{1.1}$$

is a Lagrange multiplier associated to the SDE constraint in (CP). Conversely, every Lagrange multiplier $\lambda(\cdot) = \lambda_0 + \int_0^{\cdot} \lambda_1(s)\mathrm{d}s + \int_0^{\cdot} \lambda_2(s)\mathrm{d}W(s)$, associated to this constraint, satisfies that $\lambda_0 = \lambda_1(0)$ and (λ_1, λ_2) is a weak-Pontryagin multiplier. Let us stress that the main difficulty of this result lies in first having identified the proper Hilbertian topology useful for our problem and then making a link between certain adjoint operators on Itô processes and linear BSDEs (see also ([31], Chap. 7, Sect. 2) for a related analysis).

One advantage of identifying the Lagrange multipliers of an optimization problem is that, under some precise conditions, these multipliers allow to perform a first-order sensitivity analysis of the value function as a function of the problem parameters. In a nutshell, if the optimization problem at hand is convex (this is the case of convex costs and linear equality constraints) or smooth and stable with respect to parameter perturbations (*e.g.* if the optimizers converge as we vary the parameters, and the functions involved are at least continuously differentiable) then the sensitivity of the value function in terms of the perturbation is related to the derivative of the Lagrangian with respect to the parameters taken in the perturbation direction (see *e.g.* [8], Sect. 4.3).

Using our identification of Lagrange and weak-Pontryagin multipliers we establish in Section 4 our main results. In a first part we rely on classical duality theory for convex problems (see *e.g.* [29]) and we prove in Theorem 4.2 that for stochastic optimal control problems with convex costs and linear dynamics, an additive (random, time-dependent) perturbation $(\Delta f, \Delta\sigma)$ to the drift and diffusion parts of the controlled SDE changes the value function (up to first order) by exactly

$$\mathbb{E}\left(\int_0^T p(t)^\top \Delta f(t)\mathrm{d}t\right) + \mathbb{E}\left(\int_0^T \mathrm{tr}\left[q^\top(t)\Delta\sigma(t)\right]\mathrm{d}t\right),$$

where (p, q) is (in this case) the unique adjoint state appearing in the Pontryagin's principle. Let us notice that this type of result could also be obtained by invoking the duality theory developed by Bismut in [3]. A simple corollary of this is that if one perturbs a deterministic optimal control problem by a small (brownian) noise term, the value function remains unaltered up to first-order, as was observed in [21] by other methods. Then in Theorem 4.4 we provide a version of the previous result when in addition final constraints are considered. We remark that in this case, due to the possible non-uniqueness of the Lagrange multipliers, the directional derivative is not necessarily a linear function of the perturbations. Despite that, at the present point, we cannot extend the previous sensitivity analysis to general non-convex problems, we do tackle in a second part some cases of non-additive parameter perturbations of convex stochastic optimal control problems. This is an important improvement from what was outlined in the previous lines for additive perturbations, as in practice parameter error/inaccuracy can propagate in very complicated fashions if for instance this error is amplified by the decision (control) variable. This is the setting we face in two benchmark examples we deal with in this article; the stochastic Linear-Quadratic (LQ) control problem and the Mean-Variance portfolio selection problem, which is an LQ problem with a constraint on the expected value of the final state. In these problems, it is natural to consider perturbations of the matrices appearing in the dynamics that *multiply* either the state or the control. We should underline that for these types of perturbations, classical arguments based on convex analysis as in [29] and the results in [3] are not applicable for the sensitivity analysis and more recent results on optimization theory have to be invoked (see [7,8]). The main tool here is again the identification in Theorem 3.12 and the stability result in Proposition 4.1 regarding a weak continuity property for the solutions of linear SDE and BSDE in terms of the parameters, which in both mentioned examples allows us to prove the convergence of the solutions of the perturbed problems.

As suggested by their name, in a stochastic LQ problem one seeks to minimize a quadratic functional of the state and control variables, which are related through a linear SDE. Such problems are to be found everywhere in engineering and economics and we refer the reader to [4,12,30,31] and the references therein for an exposition of the theory. Our main results here are a *strong stability property* for the solutions of parameterized unconstrained convex LQ problems (see Prop. 4.8) and Theorem 4.9, where we provide a complete sensitivity analysis for the value function in terms of the parameters. More precisely, we prove that the optimal cost depends in a *continuously differentiable* manner on the various parameters and we give explicit expressions for the associated derivatives. From the practical point of view, this result can have interesting applications. As matter of fact, recall that the resolution of deterministic LQ problems can be done through the resolution of an associate deterministic backward Riccati differential equation. The analogous result holds true in the stochastic framework [30], but in that case the Riccati equation is a highly nonlinear BSDE. Therefore, for small random perturbations of the matrices of a deterministic LQ problem, it seems reasonable to approximate the value function of the perturbed problem as the value of the deterministic one plus a first order term, which can be calculated in terms of the solution of the *deterministic* Riccati equation (see Rem. 4.10(i)).

In the classical Mean-Variance portfolio selection problem, one seeks to find the portfolio rendering the least variance of the terminal wealth with a guaranteed fixed expected return. This is a very central topic in finance and economics, and we refer the reader to [34] (random coefficients), [23] (case with jumps), among others. As for the general LQ case, our major contributions here are Proposition 4.14, dealing with an *stability* analysis for the optimal solutions in terms of the perturbation parameters (the initial capital, deterministic interest/saving rates, the desired return, the drift and the diffusion coefficients) and Theorem 4.16, where we prove that the optimal cost is C^1 with respect to those perturbations on a suitable open set.

To the best of our knowledge, the aforementioned results for the LQ and mean variance problems, regarding the strong stability of the minimizers, the C^1-differentiability of the value functions and the computation of the derivatives for general random perturbations of the dynamics, are novel in the literature.

The article is structured as follows. In Section 2 we introduce relevant notation and present the mentioned Hilbert space topology in the space of Itô processes. Next, in Section 3 we define the optimal control problem and we establish the one-to-one relationship between Lagrange multipliers and weak-Pontryagin multipliers. Then in Section 4 we take advantage of the Lagrange point of view and analyze the differentiability properties of the value function with respect to its parameters in the case of linear perturbations (Sect. 4.1) of convex problems, the case of stochastic Linear-Quadratic problems (Sect. 4.2) and Mean-Variance portfolio optimization problem (Sect. 4.3). In the Appendix, we present some useful relations between Backward Differential Equations (BSDEs) and some adjoint operators (see Sect. A.1), as well as (see Sect. A.2) some technical proofs regarding the differentiability of the data in the functional spaces introduced in Section 3. The general reader may want to first visit the Appendix, Section A.1, before undertaking Section 3, as it completely justifies the necessity of our Hilbertian setting on the space of Itô processes. The expert reader may however simply use the Appendix, Section A.2, as a reference for some technical results.

2. Preliminaries and functional framework

Let $T > 0$ and consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, on which a d-dimensional ($d \in \mathbb{N}^*$) Brownian motion $W(\cdot)$ is defined. We suppose that $\mathbb{F} = \{\mathcal{F}_t\}_{0\leq t\leq T}$ is the natural filtration, augmented by all $\mathbb{P}$-null sets in $\mathcal{F}$, associated to $W(\cdot)$. We recall that $\mathbb{F}$ is right-continuous. Given $\beta, p \in [1, \infty]$ and $n \in \mathbb{N}$ let us consider the Banach spaces

$$(L_{\mathbb{F}}^{\beta,p})^n := \left\{ v \in L^\beta\left(\Omega; L^p\left([0,T]; \mathbb{R}^n\right)\right);\ (t,\omega) \to v(t,\omega) := v(\omega)(t) \text{ is } \mathbb{F}\text{-progressively measurable} \right\}.$$

We write $\|\cdot\|_{\beta,p}$ for the natural norms, whenever $\beta < \infty$, given by:

$$\|v\|_{\beta,p} := \left[\mathbb{E}\left(\|v(\omega)\|^\beta_{L^p([0,T])}\right)\right]^{\frac{1}{\beta}},$$

and we set for the extreme cases the following:

$$\|v\|_{\beta,\infty} := \left[\mathbb{E}\left(\operatorname*{ess\,sup}_{t\in[0,T]} |v(t,\omega)|^\beta\right)\right]^{\frac{1}{\beta}} \quad \text{and} \quad \|v\|_{\infty,p} := \operatorname*{ess\,sup}_{\omega\in\Omega} \|v(\omega)\|_{L^p([0,T])}.$$

The case $\beta = p = 2$ is of particular interest since $(L^{2,2}_{\mathbb{F}})^n$ is a Hilbert space endowed with the scalar product

$$\langle v_1, v_2\rangle_{L^2} := \mathbb{E}\left(\int_0^T v_1(t)^\top v_2(t)\mathrm{d}t\right).$$

We set $(\mathcal{M}^2_c)^n$ for the set consisting of $\mathbb{F}$-adapted, $\mathbb{R}^n$-valued square integrable martingales $x(\cdot)$ satisfying that $x(0) = 0$. Recall that in the brownian filtration $\mathbb{F}$, every martingale admits a version having $\mathbb{P}$-almost surely (a.s.) continuous trajectories (see [26], Thm. 3.5, Chap. V). In particular, the elements in $(\mathcal{M}^2_c)^n$ can be identified with $\mathbb{F}$-progressively measurable processes. Let us also recall that for every $x \in (\mathcal{M}^2_c)^n$, the martingale representation theorem (see *e.g.* [16], Chap. 2, Thm. 6.6) provides the existence of a unique $x_2 \in (L^{2,2}_{\mathbb{F}})^{n\times d}$ such that

$$x(t) = \int_0^t x_2(s)\mathrm{d}W(s) \qquad \forall\, t \in [0,T], \tag{2.1}$$

where, denoting $x_2^{ij} := (x_2^j)^i$,

$$\left(\int_0^\cdot x_2(s)\mathrm{d}W(s)\right)^i := \sum_{j=1}^d \int_0^\cdot x_2^{ij}(s)\mathrm{d}W^j(s) \qquad \text{for all } i = 1,\dots,n.$$

Note that relation (2.1), Doob's inequality and the Itô isometry for the stochastic integral imply that, endowed with the scalar product

$$\langle x, y\rangle_{\mathcal{M}^2_c} := \mathbb{E}\left(x(T)^\top y(T)\right),$$

$(\mathcal{M}^2_c)^n$ is a Hilbert space which is a closed subspace of $(L^{2,\infty}_{\mathbb{F}})^n$. We now consider a larger Hilbert space, called Itô space, which is fundamental in the rest of the article. In order to provide a rigorous definition let us consider the application $I : \mathbb{R}^n \times (L^{2,2}_{\mathbb{F}})^n \times (L^{2,2}_{\mathbb{F}})^{n\times d} \to (L^{2,\infty}_{\mathbb{F}})^n$ defined as

$$I(x_0, x_1, x_2)(\cdot) := x_0 + \int_0^\cdot x_1(s)\mathrm{d}s + \int_0^\cdot x_2(s)\mathrm{d}W(s). \tag{2.2}$$

The proof of the following result is omitted, since it is a straightforward consequence of Doob's inequality and the fact that the only continuous martingales with finite-variation are the constants.

Lemma 2.1. *The application I is well defined, injective and $\exists\, c > 0$ such that*

$$\|I(x_0, x_1, x_2)\|_{2,\infty} \le c\left(|x_0| + \|x_1\|_{2,2} + \sum_{j=1}^d \|x_2^j\|_{2,2}\right). \tag{2.3}$$

We consider the space $\mathcal{I}^n$ of $\mathbb{R}^n$-valued Itô processes defined by

$$\mathcal{I}^n := I\left(\mathbb{R}^n \times (L^{2,2}_{\mathbb{F}})^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}\right).$$

By Lemma 2.1 have that $\mathcal{I}^n$ is a linear space which can be identified with $\mathbb{R}^n \times (L^{2,2}_{\mathbb{F}})^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$. For $x \in \mathcal{I}^n$ we set $(x_0, x_1, x_2) = I^{-1}(x)$ and we define the scalar product

$$\begin{aligned}\langle x, y\rangle_{\mathcal{I}} &:= x_0^\top y_0 + \langle x_1, y_1\rangle_{L^2} + \sum_{j=1}^d \langle x_2^j, y_2^j\rangle_{L^2} \qquad \forall\, x, y \in \mathcal{I}^n,\\ &= x_0^\top y_0 + \langle x_1, y_1\rangle_{L^2} + \sum_{j=1}^d \left\langle \int_0^\cdot x_2^j(s)\mathrm{d}W^j(s), \int_0^\cdot y_2^j(s)\mathrm{d}W^j(s)\right\rangle_{\mathcal{M}^2_c}\\ &= x_0^\top y_0 + \mathbb{E}\left(\int_0^T x_1(t)^\top y_1(t)\mathrm{d}t\right) + \mathbb{E}\left(\int_0^T \mathrm{tr}\left[x_2(t)^\top y_2(t)\right]\mathrm{d}t\right).\end{aligned} \tag{2.4}$$

and we define the norm $\|x\|_{\mathcal{I}} := \sqrt{\langle x, x\rangle_{\mathcal{I}}}$.

Lemma 2.2. *The space* $(\mathcal{I}^n, \|\cdot\|_{\mathcal{I}})$ *is a Hilbert space which is continuously embedded in* $(L^{2,\infty}_{\mathbb{F}})^n$.

Proof. The result is a direct consequence of the fact that $\mathbb{R}^n \times (L^{2,2}_{\mathbb{F}})^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$ is a Hilbert space and Lemma 2.1. □

Remark 2.3.

(i) We can thus identify $\mathcal{I}^n$ with $\mathbb{R}^n \times (L^{2,2}_{\mathbb{F}})^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$ and by (2.1) with $\mathbb{R}^n \times (L^{2,2}_{\mathbb{F}})^n \times (\mathcal{M}^2_c)^n$.
(ii) We will identify the topological dual $\mathcal{I}^*$ with the space $\mathcal{I}$ itself.
(iii) In the next section we give conditions under which the solution of a SDE (our controlled state equation) belongs to $\mathcal{I}^n$, suggesting that indeed this space is the right one for our purposes.

3. Optimal control problem and Lagrange multipliers

Let us introduce some notation and assumptions. For a differentiable function $(a,b) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \psi(a,b) \in \mathbb{R}^{n_3}$ we denote by $\psi_a(a,b) \in \mathbb{R}^{n_3\times n_1}$ and $\psi_b(a,b) \in \mathbb{R}^{n_3\times n_2}$ the corresponding Jacobian matrices. Let $f : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $\sigma : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{n\times d}$. In what follows we use the notations $f = (f^i)_{(1\leq i\leq n)}$ and $\sigma = (\sigma^{ij})_{1\leq i\leq n,\, 1\leq j\leq d}$, where each f^i and σ^{ij} is real valued. The columns of σ are written σ^j for $j = 1,\dots,d$. We suppose that:

(H1) The maps $\psi = f^j, \sigma^{ij}$ satisfy:
(i) ψ is $\mathcal{F}_T \otimes \mathcal{B}([0,T] \times \mathbb{R}^n \times \mathbb{R}^m)$-measurable.
(ii) For a.a. $(\omega,t) \in \Omega \times [0,T]$ the mapping $(x,u) \to \psi(\omega,t,x,u)$ is C^1, the application $(\omega,t) \in \Omega \times [0,T] \to \psi(\omega,t,\cdot,\cdot) \in C^1(\mathbb{R}^n \times \mathbb{R}^m)$ is progressively measurable and there exists $c_1 > 0$ such that almost surely in (ω,t)

$$\begin{cases} |\psi(\omega,t,x,u)| \leq c_1\,(1 + |x| + |u|)\,, \\ |\psi_x(\omega,t,x,u)| + |\psi_u(\omega,t,x,u)| \leq c_1, \end{cases} \tag{3.1}$$

Remark 3.1. Note that under **(H1)** for every $(x,u) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m$ we have that $(\omega,t) \to \psi(\omega,t,x(\omega,t),u(\omega,t))$ is progressively measurable, and so $\int_0^{\cdot} f(\omega,t,x(\omega,t),u(\omega,t))\mathrm{d}t$ and $\int_0^{\cdot} \sigma(\omega,t,x(\omega,t),u(\omega,t))\mathrm{d}W(t)$ are two a.s. continuous progressively measurable processes. The latter is also a square integrable continuous martingale.

Let us consider the application $G : \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m \to \mathcal{I}^n$ defined by

$$G(x,u) := \int_0^{\cdot} f(s,x(s),u(s))\mathrm{d}s + \int_0^{\cdot} \sigma(s,x(s),u(s))\mathrm{d}W(s) - x(\cdot). \tag{3.2}$$

Lemma 3.2. *Under* **(H1)** *the mapping* G *is Lipschitz continuous and Gâteaux differentiable. Its Gâteaux derivative* $DG(x,u) : \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m \mapsto \mathcal{I}^n$ *is given by*

$$\begin{aligned} DG(x,u)(z,v)(\cdot) = &\textstyle\int_0^{\cdot} \left[f_x(t,x(t),u(t))z(t) + f_u(t,x(t),u(t))v(t)\right]\mathrm{d}t \\ &+ \textstyle\int_0^{\cdot} \left[\sigma_x(t,x(t),u(t))z(t) + \sigma_u(t,x(t),u(t))v(t)\right]\mathrm{d}W(t) - z(\cdot), \end{aligned} \tag{3.3}$$

for all $(z,v) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m$. *Moreover, for every* $u,\, v \in (L^{2,2}_{\mathbb{F}})^m$ *and every* $x \in \mathcal{I}^n$ *we have that* $DG(x,u)(\cdot,v) : \mathcal{I}^n \mapsto \mathcal{I}^n$ *is bijective.*

Proof. See the Appendix, Section A.2. □

Remark 3.3. Note that under our assumptions G is Lipschitz. Therefore, by classical results (see *e.g.* [8], Prop. 2.49) we have that G is Hadamard differentiable, *i.e.*

$$\lim_{\tau\to 0,(z',v')\to(z,v)} \frac{G(x+\tau z', u+\tau v')(\cdot) - G(x,u)(\cdot)}{\tau} = DG(x,u)(z,v)(\cdot) \quad \text{in } \mathcal{I}^n.$$

In general, it is not clear that G is C^1. However, if f and σ are affine functions of the pair (x,u), it can be easily checked that $(x,u) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m \to DG(x,u) \in L(\mathcal{I}^n,\mathcal{I}^n)$ is continuous ($L(\mathcal{I}^n,\mathcal{I}^n)$ is the space of bounded linear applications from $\mathcal{I}^n$ to $\mathcal{I}^n$), which implies that G is continuously differentiable.

Now, let

$$\ell : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}, \quad \Phi : \Omega \times \mathbb{R}^n \to \mathbb{R} \quad \Phi_E : \Omega \times \mathbb{R}^n \to \mathbb{R}^{n_E}, \quad \Phi_I : \Omega \times \mathbb{R}^n \to \mathbb{R}^{n_I}.$$

(H2) We suppose that:
(i) The maps ℓ and $\psi = \Phi$, Φ_E^i, Φ_I^j ($1 \le i \le n_E$ and $1 \le j \le n_I$) are respectively $\mathcal{F}_T \otimes \mathcal{B}([0,T] \times \mathbb{R}^n \times \mathbb{R}^m)$ and $\mathcal{F}_T \otimes \mathcal{B}(\mathbb{R}^n)$ measurable.
(ii) For a.a. (ω, t) the maps $(x,u) \to \ell(\omega,t,x,u)$ and $x \to \psi(\omega,x)$ are C^1. The application $(\omega,t) \in \Omega \times [0,T] \to \ell(\omega,t,\cdot,\cdot) \in C^1(\mathbb{R}^n \times \mathbb{R}^m)$ is progressively measurable. In addition, there exists $c_2 > 0$ such that almost surely in (ω,t) we have that

$$\begin{cases} |\ell(\omega,t,x,u)| \le c_2 \left(1 + |x| + |u|\right)^2, \\ |\ell_x(\omega,t,x,u)| + |\ell_u(\omega,t,x,u)| \le c_2 \left(1 + |x| + |u|\right), \\ |\psi(\omega,x)| \le c_2 \left(1+|x|\right)^2, \ |\psi_x(\omega,x)| \le c_2 \left(1+|x|\right). \end{cases} \tag{3.4}$$

We define $F : \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m \to \mathbb{R}$, $G_E : \mathcal{I}^n \to \mathbb{R}^{n_E}$ and $G_I : \mathcal{I}^n \to \mathbb{R}^{n_I}$ as

$$\begin{aligned} F(x,u) :=& \mathbb{E}\left(\int_0^T \ell(t,x(t),u(t))\mathrm{d}t + \Phi(x(T)) \right), \\ G_E^i(x) :=& \mathbb{E}\left(\Phi_E^i(x(T)) \right) \qquad \forall\, i = 1,\dots,n_E, \\ G_I^j(x) :=& \mathbb{E}\left(\Phi_I^j(x(T)) \right) \qquad \forall\, j = 1,\dots,n_I. \end{aligned} \tag{3.5}$$

Lemma 3.4. *The functions F, G_E and G_I are continuously differentiable (in the Fréchet sense) and* $\forall\,(x,u)$, $(z,v) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m$ *we have that*

$$\begin{aligned} DF(x,u)(z,v) =& \mathbb{E}\left(\int_0^T \left[\ell_x(t,x(t),u(t))z(t) + \ell_u(t,x(t),u(t))v(t)\right]\mathrm{d}t + \Phi_x(x(T))z(T) \right), \\ DG_E^i(x,u)(z,v) =& \mathbb{E}\left((\Phi_E^i)_x(x(T))z(T) \right) \qquad \forall\, 1 \le i \le n_E, \\ DG_I^j(x,u)(z,v) =& \mathbb{E}\left((\Phi_I^j)_x(x(T))z(T) \right) \qquad \forall\, 1 \le j \le n_I. \end{aligned} \tag{3.6}$$

Proof. See the Appendix, Section A.2. □

Let $U \subseteq \mathbb{R}^m$ be a non-empty, closed and convex set and define

$$\mathcal{U} := \left\{ u \in (L^{2,2}_{\mathbb{F}})^m \ ; \ u(\omega,t) \in U \ \text{ for a.a. } (\omega,t) \in \Omega \times [0,T] \right\}. \tag{3.7}$$

We consider the optimal control problem

$$\text{Min}_{x \in \mathcal{I}^n, u \in (L^{2,2}_{\mathbb{F}})^m} \ \ F(x,u) \ \text{ s.t. } \ G(x,u) + x_0 = 0, \ \ G_E(x) = 0 \ \text{ and } \ G_I(x) \le 0, \ \ u \in \mathcal{U}. \tag{$\mathcal{SP}$}$$

Remark 3.5. Usually the optimal control problem above is stated only in terms of u and optimality theory is studied in this framework (see *e.g.* [1,9]). Indeed, under our assumptions, for every $u \in (L^{2,2}_{\mathbb{F}})^m$ there exists a unique $x[u] \in \mathcal{I}^n$ such that $G(x[u],u) + x_0 = 0$. Therefore, problem ($\mathcal{SP}$) can be equivalently written as

$$\text{Min}_{\ u} \ \ F(x[u],u) \ \text{ s.t. } \ G_E(x[u]) = 0 \ \text{ and } \ G_I(x[u]) \le 0, \ \ u \in \mathcal{U}. \tag{$\mathcal{SP}'$}$$

We have preferred to consider the minimization problem in terms of the pair (x,u) and thus to maintain explicitly the constraint $G(x,u) + x_0 = 0$ in order to associate a Lagrange multiplier to it.

Definition 3.6.
(i) We say that $(x,u) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m$ is feasible for ($\mathcal{SP}$) if $G(x,u) + x_0 = 0$, $G_E(x) = 0$, $G_I(x) \le 0$ and $u \in \mathcal{U}$. The set of feasible pairs for problem ($\mathcal{SP}$) is denoted by $F(\mathcal{SP})$.
(ii) We say that $(\bar{x},\bar{u}) \in F(\mathcal{SP})$ is a local solution of ($\mathcal{SP}$) iff $\exists\ \varepsilon > 0$ such that $F(\bar{x},\bar{u}) \le F(x,u)$ for all $(x,u) \in F(\mathcal{SP})$ satisfying that $\|x - \bar{x}\|_{\mathcal{I}} + \|u - \bar{u}\|_{2,2} \le \varepsilon$.

3.1. Weak-Pontryagin multipliers and Lagrange multipliers

Given $\alpha \geq 0$ the *Hamiltonian* $H[\alpha] : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \to \mathbb{R}$ is defined as

$$H[\alpha](\omega,t,x,u,p,q) := \alpha\ell(\omega,t,x,u) + p^\top f(\omega,t,x,u) + \sum_{j=1}^{d}(q^j)^\top \sigma^j(\omega,t,x,u). \tag{3.8}$$

Definition 3.7 (Weak-Pontryagin multiplier)**.** We say that $0 \neq (\bar\alpha,\bar p,\bar q,\bar\lambda_E,\bar\lambda_I) \in \mathbb{R}\times\mathcal{I}^n\times(L^{2,2}_{\mathbb{F}})^{n\times d}\times\mathbb{R}^{n_E}\times\mathbb{R}^{n_I}$ is a generalized weak-Pontryagin multiplier at $(\bar x,\bar u)$ if

$$\begin{aligned}
\mathrm{d}\bar p(t) &= -\, H_x[\bar\alpha](t,\bar x(t),\bar u(t),\bar p(t),\bar q(t))^\top \mathrm{d}t + \bar q(t)\mathrm{d}W(t),\\
\bar p(T) &= \bar\alpha\Phi_x(\bar x(T))^\top + (\Phi_E)_x(\bar x(T))^\top\bar\lambda_E + (\Phi_I)_x(\bar x(T))^\top\bar\lambda_I,\\
0 &\leq H_u[\bar\alpha](\omega,t,\bar x(t),\bar u(t),\bar p(t),\bar q(t))(v-\bar u(\omega,t)) \quad \forall v\in U \text{ a.a. } (\omega,t)\in\Omega\times[0,T],\\
0 &< |\bar\alpha| + |\bar\lambda_I| + |\bar\lambda_E|,\\
0 &= \bar\lambda_I^j G_I^j(\bar x(T)) \quad \forall\, j=1,\dots,n_I,\\
0 &\leq \bar\lambda_I^j \quad \forall\, j=1,\dots,n_I \text{ and } 0\leq\bar\alpha.
\end{aligned} \tag{3.9}$$

If $\bar\alpha > 0$ (and therefore can be normalized to $\bar\alpha = 1$), we say that $0 \neq (\bar p,\bar q,\bar\lambda_E,\bar\lambda_I)$ is a weak-Pontryagin multiplier at $(\bar x,\bar u)$. The set of weak-Pontryagin multipliers is denoted by $\Lambda_{wP}(\bar x,\bar u)$.

It is well known that the following stochastic *weak-Pontryagin minimum principle* holds (see *e.g.* [24], Sect. 6 for the case where control constraints and finitely many equality constraints are considered and [31], Thm. 6.1, Chap. 3 for our more general case).

Theorem 3.8 (Weak-Pontryagin minimum principle)**.** *Assume that* **(H1)**$-$**(H2)** *hold and let* $(\bar x,\bar u) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m$ *be a local solution of* (SP)*. Then, there exists at least one generalized weak-Pontryagin multiplier at* $(\bar x,\bar u)$*.*

Remark 3.9.

(i) In view of Remark 3.5, Pontryagin principles are usually stated for a local solution $\bar u$ of $(\mathcal{SP}')$. However, we easily check that $\bar u$ is a local solution of $(\mathcal{SP}')$ if and only if $(\bar x,\bar u)$ is a local solution of $(\mathcal{SP})$.

(ii) We called the result of Theorem 3.8 a weak-Pontryagin minimum principle, since in general more information can be obtained. In fact, even when U is not convex, under a Lipschitz type assumption on the second derivatives of the data, a second pair of adjoint processes can be introduced in such a manner that the optimal $\bar u$ minimizes an associated Hamiltonian in U. In the particular case when U is convex, (3.9) is an easy consequence of this result (see *e.g.* [24,31], Chap. 3).

The *Lagrangian* $\mathcal{L} : \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m \times \mathbb{R} \times \mathcal{I}^n \times \mathbb{R}^{n_E} \times \mathbb{R}^{n_I} \to \mathbb{R}$ associated to problem $(\mathcal{SP})$ is defined by

$$\mathcal{L}(x,u,\alpha,\lambda_{\mathcal{I}},\lambda_E,\lambda_I) := \alpha F(x,u) + \langle\lambda_{\mathcal{I}}, G(x,u)+x_0\rangle_{\mathcal{I}} + \lambda_E^\top G_E(x) + \lambda_I^\top G_I(x), \tag{3.10}$$

where $G : \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m \mapsto \mathcal{I}^n$ is defined in (3.2) and F, G_E and G_I are defined in (3.5).

Definition 3.10. We say that $0 \neq (\bar\alpha,\bar\lambda_{\mathcal{I}},\bar\lambda_E,\bar\lambda_I)$ is a generalized Lagrange multiplier at $(\bar x,\bar u)$ if

$$\begin{aligned}
0 &= D_x\mathcal{L}(\bar x,\bar u,\bar\alpha,\bar\lambda_{\mathcal{I}},\bar\lambda_E,\bar\lambda_I),\\
0 &\leq D_u\mathcal{L}(\bar x,\bar u,\bar\alpha,\bar\lambda_{\mathcal{I}},\bar\lambda_E,\bar\lambda_I)(v-\bar u) \quad \forall v\in\mathcal{U},\\
0 &< |\bar\alpha| + |\bar\lambda_I| + |\bar\lambda_E|,\\
0 &= \bar\lambda_I^j G_I^j(\bar x(T)) \quad \forall\, j=1,\dots,n_I,\\
0 &\leq \bar\lambda_I^j \quad \forall\, j=1,\dots,n_I \text{ and } 0\leq\bar\alpha.
\end{aligned} \tag{3.11}$$

If $\bar\alpha > 0$ (and therefore can be normalized to $\bar\alpha = 1$) we will say that $0 \neq (\bar\lambda_{\mathcal{I}},\bar\lambda_E,\bar\lambda_I)$ is a Lagrange multiplier at $(\bar x,\bar u)$ and we will eliminate the $\bar\alpha$ from the arguments of $\mathcal{L}$. The set of Lagrange multipliers is denoted by $\Lambda_L(\bar x,\bar u)$.

Remark 3.11. If no final constraints are present, we will eliminate (λ_E, λ_I) from the arguments of $\mathcal{L}$.

Using the relations between adjoint operators in $\mathcal{I}^n$ and BSDEs (see the Appendix, Sect. A.1), we can prove the following

Theorem 3.12. *Let $(\bar{x}, \bar{u}) \in F(\mathcal{SP})$. If $(\bar{\alpha}, \bar{p}, \bar{q}, \bar{\lambda}_E, \bar{\lambda}_I)$ is a generalized weak-Pontryagin multiplier at $(\bar{x}, \bar{u})$ then $(\bar{\alpha}, \bar{\lambda}_{\mathcal{I}}, \bar{\lambda}_E, \bar{\lambda}_I)$ is a generalized Lagrange multiplier at $(\bar{x}, \bar{u})$, where*

$$\bar{\lambda}_{\mathcal{I}}(\cdot) := \bar{p}(0) + \int_0^{\cdot} \bar{p}(s)\mathrm{d}s + \int_0^{\cdot} \bar{q}(s)\mathrm{d}W(s). \tag{3.12}$$

Conversely, if $(\bar{\alpha}, \bar{\lambda}_{\mathcal{I}}, \bar{\lambda}_E, \bar{\lambda}_I)$ is a generalized Lagrange multiplier at $(\bar{x}, \bar{u})$ then $(\bar{\alpha}, (\bar{\lambda}_{\mathcal{I}})_1, (\bar{\lambda}_{\mathcal{I}})_2, \bar{\lambda}_E, \bar{\lambda}_I)$ is a generalized weak-Pontryagin multiplier at $(\bar{x}, \bar{u})$ and $(\bar{\lambda}_{\mathcal{I}})_0 = (\bar{\lambda}_{\mathcal{I}})_1(0)$.

Remark 3.13. If $\bar{\alpha} = 1$ we can replace in the statement of the theorem "generalized weak-Pontryagin multiplier" by "weak-Pontryagin multiplier" and "generalized Lagrange multiplier" by "Lagrange multiplier".

Proof. For notational convenience we set $\ell_x(t) := \ell_x(t, \bar{x}(t), \bar{u}(t))$, $\sigma_x(t) := \sigma_x(t, \bar{x}(t), \bar{u}(t))$ with analogous definitions for $f_u(t)$ and $\sigma_u(t)$. Let $(\bar{\alpha}, \bar{p}, \bar{q}, \bar{\lambda}_E, \bar{\lambda}_I)$ be a generalized weak-Pontryagin multiplier at $(\bar{x}, \bar{u})$. In order to prove that $(\bar{\alpha}, \bar{\lambda}_{\mathcal{I}}, \bar{\lambda}_E, \bar{\lambda}_I)$, with $\bar{\lambda}_{\mathcal{I}}$ given by (3.12), is a generalized Lagrange multiplier at $(\bar{x}, \bar{u})$ it suffices to show that the first two relations in (3.11) hold true. For the first one, for every $z \in \mathcal{I}^n$, Lemma 3.2 and Lemma 3.4 imply that

$$\begin{aligned}
\bar{\alpha} D_x F(x, u) z &= \mathbb{E}\left(\int_0^T \bar{\alpha}\ell_x(t) z(t)\mathrm{d}t + \bar{\alpha}\Phi_x(\bar{x}(T)) z(T)\right), \\
\langle \bar{\lambda}_E, D_x G_E(\bar{x}) z\rangle &= \mathbb{E}\left(\bar{\lambda}_E^\top (\Phi_E)_x(\bar{x}(T)) z(T)\right), \\
\langle \bar{\lambda}_I, D_x G_I(\bar{x}) z\rangle &= \mathbb{E}\left(\bar{\lambda}_I^\top (\Phi_I)_x(\bar{x}(T)) z(T)\right), \\
\langle \bar{\lambda}_{\mathcal{I}}, D_x G(\bar{x}) z\rangle_{\mathcal{I}} &= \mathbb{E}\left(\int_0^T \left[(\bar{\lambda}_{\mathcal{I}})_1(t)^\top f_x(t) + \sum_{j=1}^d (\bar{\lambda}_{\mathcal{I}})_2^j(t)^\top \sigma_x^j(t)\right] z(t)\mathrm{d}t\right) - \langle \bar{\lambda}_{\mathcal{I}}, z\rangle_{\mathcal{I}}.
\end{aligned} \tag{3.13}$$

Using Proposition A.3 in the Appendix, with $a = \bar{\alpha}\ell_x(t)$ and $g^\top = \bar{\alpha}\Phi_x(\bar{x}(T)) + \bar{\lambda}_E^\top (\Phi_E)_x(\bar{x}(T)) + \bar{\lambda}_I^\top (\Phi_I)_x(\bar{x}(T))$, we get with (A.4)

$$\begin{aligned}
D_x\mathcal{L}(x, u, \alpha, \lambda_{\mathcal{I}}, \lambda_E, \lambda_I) z &= \left\langle \hat{p}(0) + \int_0^{\cdot} \hat{p}(t)\mathrm{d}t + \int_0^{\cdot} \hat{q}(t)\mathrm{d}W(t) - \bar{\lambda}_{\mathcal{I}}, z\right\rangle_{\mathcal{I}} + \langle \lambda_{\mathcal{I}}, (A_{f_x} + B_{\sigma_x}) z\rangle_{\mathcal{I}}, \\
&= \left\langle \hat{p}(0) + \int_0^{\cdot} \hat{p}(t)\mathrm{d}t + \int_0^{\cdot} \hat{q}(t)\mathrm{d}W(t) + (A_{f_x} + B_{\sigma_x})^* \bar{\lambda}_{\mathcal{I}} - \bar{\lambda}_{\mathcal{I}}, z\right\rangle_{\mathcal{I}},
\end{aligned}$$

where $(\hat{p}, \hat{q}) \in \mathcal{I}^n \times (L_{\mathbb{F}}^{2,2})^{n\times d}$ is the unique solution of

$$\begin{aligned}
\mathrm{d}\hat{p}(t) &= -\bar{\alpha}\ell_x(t)^\top \mathrm{d}t + \hat{q}(t)\mathrm{d}W(t), \\
\hat{p}(T) &= \bar{\alpha}\Phi_x(\bar{x}(T))^\top + (\Phi_E)_x(\bar{x}(T))^\top \bar{\lambda}_E + (\Phi_I)_x(\bar{x}(T))^\top \bar{\lambda}_I.
\end{aligned}$$

By Corollary A.4 in the Appendix we get that

$$D_x\mathcal{L}(\bar{x}, \bar{u}, \alpha, \lambda_{\mathcal{I}}, \lambda_E, \lambda_I) z = \left\langle p(0) + \int_0^{\cdot} p(t)\mathrm{d}t + \int_0^{\cdot} q(t)\mathrm{d}W(t) - \bar{\lambda}_{\mathcal{I}}, z\right\rangle_{\mathcal{I}}, \tag{3.14}$$

where $(p, q) \in \mathcal{I}^n \times (L_{\mathbb{F}}^{2,2})^{n\times d}$ is the unique solution of

$$\begin{aligned}
\mathrm{d}p(t) &= -\left[\bar{\alpha}\ell_x(t)^\top + f_x(t)^\top (\bar{\lambda}_{\mathcal{I}})_1(t) + \sigma_x(t)^\top (\bar{\lambda}_{\mathcal{I}})_2(t)\right]\mathrm{d}t + q(t)\mathrm{d}W(t), \\
p(T) &= \bar{\alpha}\Phi_x(\bar{x}(T))^\top + (\Phi_E)_x(\bar{x}(T))^\top \bar{\lambda}_E + (\Phi_I)_x(\bar{x}(T))^\top \bar{\lambda}_I.
\end{aligned} \tag{3.15}$$

Since $((\bar{\lambda}_{\mathcal{I}})_1, (\bar{\lambda}_{\mathcal{I}})_2) = (\bar{p}, \bar{q})$, by (3.9) we get that $p(T) - \bar{p}(T) = 0$ and $\mathrm{d}[p - \bar{p}](t) = [q(t) - \bar{q}(t)]\mathrm{d}W(t)$ which yields to $p = \bar{p}$, $q = \bar{q}$ and in particular $p(0) = \bar{p}(0)$, hence the first relation in (3.11) follows from (3.14). In order to prove the second relation in (3.11) it suffices to note that for all $v \in \mathcal{U}$

$$D_u\mathcal{L}(\bar{x}, \bar{u}, \bar{\alpha}, \bar{\lambda}_{\mathcal{I}}, \bar{\lambda}_E, \bar{\lambda}_I)(v - \bar{u}) = \mathbb{E}\left(\int_0^T H_u[\bar{\alpha}](\omega, t, \bar{x}(t), \bar{u}, \bar{p}(t), \bar{q}(t))(v(t) - \bar{u}(t))\mathrm{d}t\right) \geq 0. \tag{3.16}$$

Now, let $(\bar{\alpha}, \bar{\lambda}_{\mathcal{I}}, \bar{\lambda}_E, \bar{\lambda}_I)$ be a generalized Lagrange multiplier at $(\bar{x}, \bar{u})$. By the first relation in (3.11) and (3.14) we obtain that

$$\bar{\lambda}_{\mathcal{I}} = p(0) + \int_0^{\cdot} p(t)\mathrm{d}t + \int_0^{\cdot} q(t)\mathrm{d}W(t),$$

where $(p, q) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$ solves (3.15). Therefore, we get $(\bar{\lambda}_{\mathcal{I}})_1 = p$ and $(\bar{\lambda}_{\mathcal{I}})_2 = q$ and $(\bar{\lambda}_{\mathcal{I}})_0 = p(0) = (\bar{\lambda}_{\mathcal{I}})_1(0)$. Thus (3.15) implies that $((\bar{\lambda}_{\mathcal{I}})_1, (\bar{\lambda}_{\mathcal{I}})_2)$ satisfies the first and second relations in (3.9). Finally, by the second relation in (3.11) and expression (3.16), we obtain the third relation in (3.9) following the same argument as in the proof of ([10], Thm. 1.5). □

As a consequence of the above result we obtain the following sufficient condition, under convexity assumptions. The proof is standard, but since it is very short we provide it for the reader's convenience.

Corollary 3.14 (Sufficient condition for convex problems)**.** *Suppose that F and G_I are convex and that G and G_E are affine.*

(i) *Let $(\bar{x}, \bar{u}) \in F(\mathcal{SP})$ and suppose that $(\bar{p}, \bar{q}, \bar{\lambda}_E, \bar{\lambda}_I) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^{n\times d} \times \mathbb{R}^{n_E} \times \mathbb{R}^{n_I}$ is a weak-Pontryagin multiplier at $(\bar{x}, \bar{u})$. Then, the pair $(\bar{x}, \bar{u})$ solves $(\mathcal{SP})$.*

(ii) *The set of weak-Pontryagin multipliers is independent of the solutions of $(\mathcal{SP})$. More precisely, let $(\bar{x}^1, \bar{u}^1)$, $(\bar{x}^2, \bar{u}^2) \in F(\mathcal{SP})$ be two solutions of $(\mathcal{SP})$. Then, $(\bar{p}, \bar{q}, \bar{\lambda}_E, \bar{\lambda}_I)$ is a weak-Pontryagin multiplier at $(\bar{x}^1, \bar{u}^1)$ if and only if it is a weak-Pontryagin multiplier at $(\bar{x}^2, \bar{u}^2)$.*

Proof. By Theorem 3.12, $\bar{\lambda}_{\mathcal{I}} \in \mathcal{I}^n$ defined by (3.12) is such that $(\bar{\lambda}_{\mathcal{I}}, \bar{\lambda}_E, \bar{\lambda}_I)$ is a Lagrange multiplier. Now, let (x, u) be feasible for $(\mathcal{SP})$, then by the convexity of $\mathcal{L}(\cdot, \cdot, 1, \bar{\lambda}_{\mathcal{I}}, \bar{\lambda}_E, \bar{\lambda}_I)$,

$$\begin{aligned} F(x, u) \geq \mathcal{L}(x, u, 1, \bar{\lambda}_{\mathcal{I}}, \bar{\lambda}_E, \bar{\lambda}_I) &\geq \mathcal{L}(\bar{x}, \bar{u}, 1, \bar{\lambda}_{\mathcal{I}}, \bar{\lambda}_E, \bar{\lambda}_I) + D_x\mathcal{L}(\bar{x}, \bar{u}, 1, \bar{\lambda}_{\mathcal{I}}, \bar{\lambda}_E, \bar{\lambda}_I)(x - \bar{x}) \\ &\quad + D_u\mathcal{L}(\bar{x}, \bar{u}, 1, \bar{\lambda}_{\mathcal{I}}, \bar{\lambda}_E, \bar{\lambda}_I)(u - \bar{u}). \end{aligned}$$

Since $\mathcal{L}(\bar{x}, \bar{u}, 1, \bar{\lambda}_{\mathcal{I}}, \bar{\lambda}_E, \bar{\lambda}_I) = F(\bar{x}, \bar{u})$ assertion (i) follows from (3.11). Assertion (ii) is a direct consequence of Theorem 3.12 and the fact that for convex problems the set of Lagrange multipliers $\Lambda_L(\bar{x}, \bar{u})$ does not depend on $(\bar{x}, \bar{u})$ (see *e.g.* [8], Thm. 3.6). □

4. Sensitivity analysis

In this section we take advantage of the Lagrange multiplier interpretation of the adjoint state (p, q) in order to obtain some sensitivity results for the optimal cost when the problem dynamics and final constraints are perturbed. We will first consider general convex problems and linear perturbations of the dynamics. Next, we study in detail the case of Linear Quadratic (LQ) stochastic problems and the mean variance portfolio selection problem, where the perturbations are performed also in the matrices multiplying the state and control variables. We shall study these last two problems separately, since although they belong to a same family, their specific structures mean that we need to employ slightly different arguments and assume different hypotheses. In any case a stability result for the solutions of the parameterized problems is needed and will be a consequence of the following result:

Proposition 4.1. *The following assertions hold:*

(i) *Let $x^k \in \mathcal{I}^n$ be a sequence converging weakly to $x \in \mathcal{I}^n$. Then x^k converges weakly to x in $(L^{2,2}_{\mathbb{F}})^n$ and for all $t \in [0, T]$ we have that $x^k(t)$ converges weakly to $x(t)$ in $(L^2_{\mathcal{F}_t})^n$.*

(ii) *Let* $x_0^k \in \mathbb{R}^n$*,* $A^k \in (L_{\mathbb{F}}^{\infty,\infty})^{n\times n}$*,* $(C^j)^k \in (L_{\mathbb{F}}^{\infty,\infty})^{n\times n}$*,* $\xi_1^k \in (L_{\mathbb{F}}^{2,2})^n$*,* $(\xi_2^j)^k \in (L_{\mathbb{F}}^{2,2})^n$ $(j=1,\dots,d)$*. Suppose that* $(x_0^k, A^k, (C^j)^k)$ *converges strongly to* (x_0, A, C^j) *and that* $(\xi_1^k, (\xi_2^j)^k)$ *converges weakly to* (ξ_1, ξ_2^j)*. Then, the solutions* x^k *of*

$$\begin{aligned} \mathrm{d}x^k(t) &= \left[A^k(t)x^k(t) + \xi_1^k(t)\right]\mathrm{d}t + \sum_{j=1}^d \left[(C^j)^k(t)x^k(t) + (\xi_2^j)^k(t)\right]\mathrm{d}W^j(t), \\ x^k(0) &= x_0^k, \end{aligned}$$

converge weakly in $\mathcal{I}^n$ *to the solution* x *of*

$$\begin{aligned} \mathrm{d}x(t) &= [A(t)x(t) + \xi_1(t)]\,\mathrm{d}t + \sum_{j=1}^d \left[C^j(t)x(t) + \xi_2^j(t)\right]\mathrm{d}W^j(t), \\ x(0) &= x_0. \end{aligned} \tag{4.1}$$

(iii) *Let* $D^k \in (L_{\mathbb{F}}^{\infty,\infty})^{n\times n}$*,* $(E^j)^k \in (L_{\mathbb{F}}^{\infty,\infty})^{n\times n}$ $(j=1,\dots,d)$*,* $\xi_3^k \in (L_{\mathbb{F}}^{2,2})^n$ *and* $\xi_4^k \in (L_{\mathcal{F}_T}^2)^n$*. Suppose that* $(D^k, (E^j)^k)$ *converges strongly to* (D, E^j) *and* (ξ_3^k, ξ_4^k) *converge weakly to* (ξ_3, ξ_4)*. Then, the solution* (p^k, q^k) *of*

$$\begin{aligned} \mathrm{d}p^k(t) &= \left[D^k(t)p^k(t) + \sum_{j=1}^d (E^j)^k(t)(q^j)^k(t) + \xi_3^k(t)\right]\mathrm{d}t + q^k(t)\mathrm{d}W(t), \\ p^k(T) &= \xi_4^k. \end{aligned} \tag{4.2}$$

converges weakly in $\mathcal{I}^n \times (L_{\mathbb{F}}^{2,2})^{n\times d}$ *to the solution* (p,q) *of*

$$\begin{aligned} \mathrm{d}p(t) &= \left[D(t)p(t) + \sum_{j=1}^d E^j(t)q^j(t) + \xi_3(t)\right]\mathrm{d}t + q(t)\mathrm{d}W(t), \\ p(T) &= \xi_4. \end{aligned} \tag{4.3}$$

Proof of (i). It follows directly from Lemma 2.2 and the fact that $(L_{\mathbb{F}}^{2,\infty})^n$ is continuously embedded in $(L_{\mathbb{F}}^{2,2})^n$ and $(L_{\mathcal{F}_t}^2)^n$, for all $t \in [0,T]$. □

Proof of (ii). Since $|x_0^k|$, $\|A^k\|_{\infty,\infty}$, $\|(C^j)^k\|_{\infty,\infty}$, $\|(D^j)^k\|_{\infty,\infty}$, $\|\xi_1^k\|_{2,2}$ and $\|(\xi_2^j)^k\|_{2,2}$ are bounded, by the classical proof for the stability of linear SDEs (see *e.g.* ([31], Chap. 6, Sect. 4), we have that $\|x^k\|_{2,\infty}$ is uniformly bounded in k. Therefore for any subsequence there exists $\hat{x} \in (L_{\mathbb{F}}^{2,2})^n$ such that for a further subsequence x^k converges weakly in $(L_{\mathbb{F}}^{2,2})^n$ to $\hat{x}$. Using that $A^k x^k$, $(C^j)^k x^k$ converge weakly in $(L_{\mathbb{F}}^{2,2})^n$ to $A\hat{x}$, $C^j\hat{x}$, respectively, we see that x^k converges weakly in $\mathcal{I}^n$ to

$$\tilde{x}(\cdot) := x_0 + \int_0^{\cdot} [A(t)\hat{x}(t) + \xi_1(t)]\,\mathrm{d}t + \sum_{j=1}^d \int_0^{\cdot} \left[C^j(t)\hat{x}(t) + \xi_2^j(t)\right]\mathrm{d}W(t).$$

By (i) we have that $\tilde{x} = \hat{x}$ and since (4.1) has a unique solution (and so independent of the given subsequence) the result follows. □

Proof of (iii). We argue in a similar manner. Note that since (ξ_3^k, ξ_4^k) is bounded in $(L_{\mathbb{F}}^{2,2})^n \times (L_{\mathcal{F}_T}^2)^n$ and $\|D^k\|_{\infty,\infty}$, $\|(E^j)^k\|_{\infty,\infty}$ are bounded, following the lines of the proof ([31], Chap. 7, Thm. 2.2) we obtain that $\|p^k\|_{2,\infty} + \sum_{j=1}^d \|(q^j)^k\|_{2,2}$ is uniformly bounded in k. So for any subsequence there exists

$(\hat{p},\hat{q}) \in (L^{2,2}_{\mathbb{F}})^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$ such that, except for some further subsequence, (p^k, q^k) converge to $(\hat{p},\hat{q})$ weakly in $(L^{2,2}_{\mathbb{F}})^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$. Since $D^k p^k$ and $(E^j)^k (q^j)^k$ converge weakly in $(L^{2,2}_{\mathbb{F}})^n$ respectively to $D\hat{p}$, $E^j \hat{q}^j$, we easily obtain that p^k converges weakly in $\mathcal{I}^n$ to

$$\tilde{p}(\cdot) := \tilde{p}(0) + \int_0^{\cdot} \left[D(t)\hat{p}(t) + \sum_{j=1}^{d} E^j(t)\hat{q}^j(t) + \xi_3(t) \right] \mathrm{d}t + \int_0^{\cdot} \hat{q}(t)\mathrm{d}W(t), \tag{4.4}$$

where

$$\tilde{p}(0) := \mathbb{E}\left(\xi_4 - \int_0^T \left[D(t)\hat{p}(t) + \sum_{j=1}^{d} E^j(t)\hat{q}^j(t) + \xi_3(t) \right] \mathrm{d}t \right).$$

By (i) we obtain that $\tilde{p} = \hat{p}$, and $\hat{p}(T) = \xi_4$ using that $p^k(T) = \xi_4^k$ converges weakly in $(L^2_{\mathcal{F}_T})^n$ to ξ_4. From this fact and (4.4), we have that $(\hat{p},\hat{q})$ solves (4.3). Finally, since the solution of (4.3) is unique, the result follows. □

4.1. Convex problems and linear perturbations of the dynamics

Let us define the perturbation space $\mathcal{P}_1 := \mathbb{R}^n \times (L^{2,2}_{\mathbb{F}})^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$ and let $P := (x_0, \hat{f}, \hat{\sigma}) \in \mathcal{P}_1$. We consider the problem

$$\begin{aligned} &\inf_{(x,u)\in\mathcal{I}^n\times(L^{2,2}_{\mathbb{F}})^m} \mathbb{E}\left(\int_0^T \ell(t,\omega,x(t),u(t))\mathrm{d}t + \Phi(\omega,x(T)) \right) \\ &\text{s.t.} \quad \begin{cases} \mathrm{d}x(t) = [f(t,\omega,x(t),u(t)) + \hat{f}(t,\omega)]\mathrm{d}t + [\sigma(t,\omega,x(t),u(t)) + \hat{\sigma}(t,\omega)]\mathrm{d}W(t), \\ x(0) = x_0, \\ u \in \mathcal{U}. \end{cases} \end{aligned} \tag{$P_{1,P}$}$$

We suppose that (ℓ, Φ, f, σ) satisfies assumptions **(H1)**−**(H2)** in Section 3 and $\mathcal{U}$ is given by (3.7). In addition, we will need the following convexity assumption:

(H3) For almost all $(t,\omega) \in [0,T]\times\Omega$ (respectively $\omega \in \Omega$), the function $\ell(t,\omega,\cdot,\cdot)$ (respectively $\Phi(\omega,\cdot)$) is convex. Moreover, we assume that a.s. in $[0,T]\times\Omega$ the functions $f(t,\omega,\cdot,\cdot)$ and $\sigma(t,\omega,\cdot,\cdot)$ are affine.

We define *the value function* $v : \mathcal{P}_1 \to \mathbb{R}\cup\{-\infty\}$ as the function that associates to P the optimal cost for problem $(P_{1,P})$. Note that under **(H1)**−**(H2)** the feasible set for $(P_{1,P})$ is not empty, and therefore v is well defined.

Let us recall the notion of Hadamard directional differentiability. Given two Banach spaces $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Z}, \|\cdot\|_{\mathcal{Z}})$ a map $f : \mathcal{X} \to \mathcal{Z}$ is directionally differentiable at x if for all $h \in \mathcal{X}$ the limit in $\mathcal{Z}$

$$Df(x,h) := \lim_{\tau\downarrow 0} \frac{f(x+\tau h) - f(x)}{\tau}$$

exists. If in addition, for all $h \in \mathcal{X}$ the equality in $\mathcal{Z}$

$$Df(x,h) = \lim_{\tau\downarrow 0,\ h'\to h} \frac{f(x+\tau h') - f(x)}{\tau}$$

holds true, then we say that f is directionally differentiable at x in the Hadamard sense.

In the next theorem we prove that given a nominal parameter P, where problem $(P_{1,P})$ admits at least one solution, and a *random functional perturbation* ΔP acting linearly on the dynamics, the value function v admits directional derivatives that can be expressed in terms of a *unique adjoint state*, which, in view of the convexity of the problem, is independent of the solution of $(P_{1,P})$. The existence of such adjoint state has been proved for the first time in [3] and, as pointed out in Remark 3.9(i), it also follows from the more general results in [24]. Due to its simplicity, we provide here a short and direct proof of the existence and uniqueness of such an adjoint state using classical results in abstract optimization theory (see *e.g.* [22, 27, 35]) and the identification of Lagrange multipliers and adjoint states in Theorem 3.12.

Theorem 4.2. *Assume* **(H1)**–**(H3)** *and that for* $P \in \mathcal{P}_1$ *problem* $(P_{1,P})$ *admits at least one solution. Then, there exists* $(\bar{p}, \bar{q}) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$ *such that for every solution* $(\bar{x}, \bar{u})$ *of* $(P_{1,P})$*, the pair* $(\bar{p}, \bar{q})$ *is the unique weak-Pontryagin multiplier associated to* $(\bar{x}, \bar{u})$*. Moreover, the value function* v *is continuous at* P*, Hadamard and Gâteaux directionally differentiable at* P *and its directional derivative* $Dv(P; \cdot) : \mathcal{P}_1 \to \mathbb{R}$ *is given by*

$$Dv(P; \Delta P) = \bar{p}(0)^\top \Delta x_0 + \mathbb{E}\left(\int_0^T \bar{p}(t)^\top \Delta f(t)\mathrm{d}t\right) + \mathbb{E}\left(\int_0^T \mathrm{tr}\left[\bar{q}(t)^\top \Delta\sigma(t)\right]\mathrm{d}t\right), \tag{4.5}$$

for all $\Delta P = (\Delta x_0, \Delta f, \Delta\sigma) \in \mathcal{P}_1$.

Proof. Recall the definition of the map I in (2.2). Let us write the problem $(P_{1,P})$ as

$$\inf_{(x,u)\in\mathcal{I}^n\times(L^{2,2}_{\mathbb{F}})^m} F(x, u) + \chi_{\mathcal{U}}(x, u) \ \text{ subject to } G(x, u) + I(P) = 0,$$

where $\chi_{\mathcal{U}} : \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m \to \mathbb{R}\cup\{+\infty\}$ is the convex, proper, l.s.c. function defined as $\chi_{\mathcal{U}}(x, u) = 0$ if $u \in \mathcal{U}$ and $+\infty$ otherwise and

$$G(x, u)(\cdot) := \int_0^{\cdot} f(t, \omega, x(t), u(t))\mathrm{d}t + \int_0^{\cdot} \sigma(t, \omega, x(t), u(t))\mathrm{d}W(t) - x(\cdot).$$

For every $(x, u) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m$ and $v \in (L^{2,2}_{\mathbb{F}})^m$, Lemma 3.2 implies that $DG(x, u)(\cdot, v)$ is surjective. Therefore, the following regularity condition is trivially satisfied (see *e.g.* [7], Sect. 3.2)

$$0 \in \mathrm{int}\left\{G(x, u) + P + DG(x, u)\left(\mathcal{I}^n \times \mathcal{U}\right)\right\}. \tag{4.6}$$

Thus, by classical results in convex optimization (see *e.g.* [6], Sect. 4.3.2, Example 4.51 or [8], Sect. 2.5) (x, u) is a solution of $(P_{1,P})$ iff there exists $\lambda \in \mathcal{I}^n$ such that

$$(0, 0) \in \partial_{(x,u)}(F(x, u) + \chi_{\mathcal{U}}(x, u)) + DG(x, u)^*\lambda. \tag{4.7}$$

Since F is differentiable in $\mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m$, in particular it is continuous in $\mathcal{I}^n \times \mathcal{U}$, and so (see *e.g.* [8], Rem. 2.170)

$$\partial_{(x,u)}(F(x, u) + \chi_{\mathcal{U}}(x, u)) = \partial_{(x,u)}F(x, u) + \partial_{(x,u)}\chi_{\mathcal{U}}(x, u) \subseteq (D_xF(x, u), D_uF(x, u)) + \{0\} \times N_{\mathcal{U}}(u),$$

where $N_{\mathcal{U}}(u) := \{v^* \in (L^{2,2}_{\mathbb{F}})^m \ ; \ \langle v^*, v - u\rangle_{L^2} \leq 0, \quad \forall\ v \in \mathcal{U}\}$ is the normal cone to $\mathcal{U}$ at u. Using that $DG(x, u)^*\lambda = (D_xG(x, u)^*\lambda, D_uG(x, u)^*\lambda)$, we obtain with (4.7)

$$(0, 0) \subseteq (D_xF(x, u), D_uF(x, u)) + \{0\} \times N_{\mathcal{U}}(u) + (D_xG(x, u)^*\lambda, D_uG(x, u)^*\lambda),$$

which is equivalent to

$$D_x\mathcal{L}(x, u, \lambda) = 0 \quad \text{and} \quad D_u\mathcal{L}(x, u, \lambda)(v - u) \geq 0 \quad \forall v \in \mathcal{U}. \tag{4.8}$$

Therefore, $\lambda \in \Lambda_L(x, u)$ and by Theorem 3.12 and the convexity of the associated Hamiltonian we have that $(\bar{p}, \bar{q}) := (\lambda_1, \lambda_2)$ is weak-Pontryagin multiplier. Now, let $\lambda^1_{\mathcal{I}}$, $\lambda^2_{\mathcal{I}} \in \Lambda_L(x, u)$. By the first equation in (4.8), we get that

$$\left\langle (D_xG(x, u))^* (\lambda^1_{\mathcal{I}} - \lambda^2_{\mathcal{I}}), z\right\rangle_{\mathcal{I}} = 0 \ \ \forall\ z \in \mathcal{I}^n, \quad \text{or} \quad (D_xG(x, u))^* (\lambda^1_{\mathcal{I}} - \lambda^2_{\mathcal{I}}) = 0.$$

Since, by Lemma 3.2, $D_xG(x, u) : \mathcal{I}^n \mapsto \mathcal{I}^n$ is surjective we get that $D_xG(x, u)^*$ is injective, which implies that $\lambda^1_{\mathcal{I}} = \lambda^2_{\mathcal{I}}$ and by Theorem 3.12 the weak-Pontryagin multiplier is unique. The independence of the set $\Lambda_L(\cdot)$ over the set of solutions of $(P_{1,P})$ is a consequence of Corollary 3.14(ii). Finally, the continuity, the Gâteaux and Hadamard differentiability of v and expression (4.5) for $Dv(P; \Delta P)$ are a direct translation of ([29], Thm. 17) using the uniqueness of the Lagrange multiplier. ☐

In the following remark we underline some simple consequences of Theorem 4.2:

Remark 4.3.

(i) The gradient of v at P, *i.e.* the Riesz representative of the bounded linear application $Dv(P;\cdot)$, is given by

$$\bar{p}(0) + \int_0^{\cdot} \bar{p}(t)\mathrm{d}t + \int_0^{\cdot} \bar{q}(t)\mathrm{d}W(t).$$

(ii) It is well known (see *e.g.* [8], Sect. 2.2 and the references therein) that for real-valued functions defined on finite dimensional spaces, Gâteaux differentiability together with Hadamard differentiability imply Fréchet differentiability. Therefore, if the perturbations for problem $(P_{1,P})$ are finite dimensional, then v is Fréchet differentiable at P. This is the case, for example, if the initial condition is perturbed and/or the perturbations of the dynamics have the form $\Delta f(t,\omega) = \xi_0(t,\omega)A_0$, $(\Delta\sigma(t,\omega))^j = \xi_j(t,\omega)A_j$ with ξ_0, $\xi_j \in (L_{\mathbb{F}}^{\infty,\infty})^{n\times n}$ $(j = 1,\dots,d)$ being fixed, and A_0, $A_j \in \mathbb{R}^n$ being the perturbation parameters. In fact, defining the new states

$$\mathrm{d}y_0 = 0, \ \text{ for } t \in [0,T], \ \ y_0(0) = A_0, \ \ \mathrm{d}y_j = 0, \ \text{ for } t \in [0,T], \ \ y_j(0) = A_j \ \text{ for } j = 1,\dots,d,$$

the new dynamical system is affine w.r.t. $(x,(y_0,y_j))$ and the perturbations are performed over the initial condition. Let us point out that the Fréchet differentiability of the value function under finite-dimensional perturbations in our convex framework can also be deduced using ([31], Chap. 5, Cor. 4.5).

(iii) Suppose that the nominal problem is deterministic (and thus $\bar{q} = 0$) and only the $\mathrm{d}W(t)$ part of the dynamics is perturbed, *i.e.* $\Delta x_0 = 0$, $\Delta f \equiv 0$. Then, by (4.5) we directly obtain that $Dv(P;\Delta P) = 0$. This fact was already observed by Loewen [21] for finite dimensional perturbations.

(iv) A close look at the proof Theorem 4.2 shows that even if $\ell(\omega,t,\cdot,\cdot)$ and $\Phi(\omega,\cdot)$ are not convex, we can apply the abstract optimization results (see *e.g.* [8], Sect. 3.1) in order to derive existence and uniqueness of a Lagrange multiplier at a local solution $\bar{u}$. More precisely, using (4.6) it is possible to show (see [8], Lem. 3.7) that if (x,u) is a solution of problem $(P_{1,P})$ then $(z,v) = (0,0)$ is a solution of

$$\inf_{(z,v)\in\mathcal{I}^n\times(L_{\mathbb{F}}^{2,2})^m} DF(x,u)(z,v) \ \text{ such that } \ DG(x,u)(z,v) = 0, \ \ v \in T_{\mathcal{U}}(u), \tag{LP}$$

(where $T_{\mathcal{U}}(u)$, defined as the closure in $(L_{\mathbb{F}}^{2,2})^m$ of $\bigcup_{\tau>0}\tau^{-1}(\mathcal{U}-u)$, is the tangent cone to $\mathcal{U}$ at u, see [8], Prop. 2.55). Problem (LP) is a convex one and we can proceed exactly as in the proof of Theorem 4.2 in order to show the existence and uniqueness of a Lagrange multiplier λ at $(0,0)$. It is easy to see that λ is a Lagrange multiplier at $(0,0)$ for problem (LP) iff λ is a Lagrange multiplier at (x,u) for problem $(P_{1,P})$. Therefore, by Theorem 3.12 this argument provides a simple proof of the existence of weak-Pontryagin multipliers for stochastic problems with non-convex cost and linear dynamics. Let us point out that it is not clear that the general result of [24] for the case of nonlinear dynamics, even in the form of a weak-Pontryagin principle, can be derived with the Lagrange multipliers method. In fact, the main issue is the apparent lack of C^1 differentiability of $G(x,u)$ in the non-affine case (see Rem. 3.3) and the non-convexity of $\mathcal{U}$ in [24].

We consider now the case where final state constraints are also included in the system. We set as parameter set the space $\mathcal{P}_2 := \mathbb{R}^n \times (L_{\mathbb{F}}^{2,2})^n \times (L_{\mathbb{F}}^{2,2})^{n\times d} \times \mathbb{R}^{n_E} \times \mathbb{R}^{n_I}$. Let $P := (x_0, \hat{f}, \hat{\sigma}, \delta_E, \delta_I) \in \mathcal{P}_2$ and consider the problem

$$\begin{aligned} &\inf_{(x,u)\in\mathcal{I}^n\times(L_{\mathbb{F}}^{2,2})^m} \mathbb{E}\left(\int_0^T \ell(t,\omega,x(t),u(t))\mathrm{d}t + \Phi(\omega,x(T))\right) \\ &\text{s.t.} \quad \begin{cases} \mathrm{d}x(t) = [f(t,\omega,x(t),u(t)) + \hat{f}(t,\omega)]\mathrm{d}t \\ \qquad\qquad + [\sigma(t,\omega,x(t),u(t)) + \hat{\sigma}(t,\omega)]\mathrm{d}W(t), \\ x(0) = x_0, \\ \mathbb{E}\left(\Phi_E^i(\omega,x(T))\right) = -\delta_E^i \quad \text{for all } i = 1,\dots,n_E, \\ \mathbb{E}\left(\Phi_I^j(\omega,x(T))\right) \le -\delta_I^j \quad \text{for all } j = 1,\dots,n_I, \\ u \in \mathcal{U}. \end{cases} \end{aligned} \tag{$P_{2,P}$}$$

We will assume that:

(H4) For almost all $\omega \in \Omega$ the functions $\Phi_E^i(\omega,\cdot)$ $(i=1,\dots,n_E)$ are affine and $\Phi_I^j(\omega,\cdot)$ $(j=1,\dots,n_I)$ are convex.

Recall that G is defined in (3.2) and G_E, G_I are defined in (3.5). Define $\hat{G}:\mathcal{I}^n\times(L_{\mathbb{F}}^{2,2})^m\to\mathcal{I}^n\times\mathbb{R}^{n_E}\times\mathbb{R}^{n_I}$ as

$$\hat{G}(x,u):=(G(x,u),G_E(x),G_I(x)).$$

Thus, problem $(P_{2,P})$ can be written as

$$\inf_{(x,u)\in\mathcal{I}^n\times(L_{\mathbb{F}}^{2,2})^m} F(x,u)+\chi_{\mathcal{U}}(x,u)\ \text{ subject to } \hat{G}(x,u)+\hat{P}\in K,$$

where $\hat{P}:=(I(x_0,\hat{f},\hat{\sigma}),\delta_E,\delta_I)$ and K is the closed and convex set defined as

$$K:=\{(z,a,b)\in\mathcal{I}^n\times\mathbb{R}^{n_E}\times\mathbb{R}^{n_I}\ ;\ z=0,\quad a=0,\quad b^j\le 0\qquad \text{for all } j=1,\dots,n_I\}.$$

The *Robinson's constraint qualification condition* for problem $(P_{2,P})$ at a feasible point (x,u) is (see *e.g.* [7], Sect. 3.2)

$$\exists\,\varepsilon>0\quad\text{such that}\quad B(0,\varepsilon)\subseteq\hat{G}(x,u)+\hat{P}+D\hat{G}(x,u)(\mathcal{I}^n\times(\mathcal{U}-u))-K, \tag{4.9}$$

where $B(0,\varepsilon)$ denotes the open ball in $\mathcal{I}^n\times\mathbb{R}^{n_E}\times\mathbb{R}^{n_I}$. Compared to (4.6), which automatically holds in the absence of state constraints, we will need to impose (4.9):

Theorem 4.4. *Assume* **(H1)**$-$**(H4)** *and that for $P\in\mathcal{P}_2$ problem $(P_{2,P})$ admits at least one solution $(\bar{x},\bar{u})$. Suppose in addition that (4.9) is satisfied at $(\bar{x},\bar{u})$. Then, the set of weak-Pontryagin multipliers $\Lambda_{wP}(\bar{x},\bar{u})\subset\mathcal{I}^n\times(L_{\mathbb{F}}^{2,2})^{n\times d}\times\mathbb{R}^{n_E+n_I}$ at any solution $(\bar{x},\bar{u})$ is a non-empty, weakly compact set, which is independent of the solution $(\bar{x},\bar{u})$. Moreover, the value function v is continuous at P, Hadamard directionally differentiable at P and its directional derivative $Dv(P;\cdot):\mathcal{P}_2\to\mathbb{R}$ is given by*

$$\begin{aligned}Dv(P;\Delta P)=\max_{(p,q,\lambda_E,\lambda_I)\in\Lambda_{wP}(\bar{x},\bar{u})}&\Big\{p(0)^\top\Delta x_0+\mathbb{E}\left(\textstyle\int_0^T p(t)^\top\Delta f(t)\mathrm{d}t\right)+\mathbb{E}\left(\textstyle\int_0^T\mathrm{tr}\left[q(t)^\top\Delta\sigma(t)\right]\mathrm{d}t\right)\\&+\lambda_E^\top\Delta\delta_E+\lambda_I^\top\Delta\delta_I\Big\},\end{aligned}\tag{4.10}$$

for all $\Delta P=(\Delta x_0,\Delta f,\Delta\sigma,\Delta\delta_E,\Delta\delta_I)\in\mathcal{P}_2$.

Proof. Since the problem is convex, using assumption (4.9) and reasoning exactly as in the proof of Theorem 4.2 (or applying directly [8], Thm. 2.165) we get that the set of Lagrange multipliers (or equivalently the set of solutions of the dual problem) is nonempty, bounded, and weakly compact. The identification of Lagrange multipliers with $\Lambda_{wP}(\bar{x},\bar{u})$ in Theorem 3.12 and Corollary 3.14(ii) prove the assertions for $\Lambda_{wP}(\bar{x},\bar{u})$. The fact that v is directionally Hadamard differentiable and that formula (4.10) holds follow directly from ([8], Thm. 2.151). □

Remark 4.5. In the absence of control constraints we have that (4.9) is equivalent to the following *Mangasarian–Fromovitz* condition

$$\left.\begin{array}{l}\text{(a)}\ \ (DG(\bar{x},\bar{u}),DG_E(\bar{x})):\mathcal{I}^n\times(L_{\mathbb{F}}^{2,2})^m\mapsto\mathcal{I}^n\times\mathbb{R}^{n_E}\ \text{ is surjective and}\\ \text{(b)}\ \ \exists\,(\hat{z},\hat{v})\in\mathrm{Ker}DG(\bar{x},\bar{u})\ \text{ such that } \hat{z}\in\mathrm{Ker}DG_E(\bar{x}),\ \ DG_I^j(\bar{x})\hat{z}<0\ \forall\,j=1,\dots,,n_I\end{array}\right\}\tag{MF}$$

and, since (G,G_E) is affine and G_I^j $(j=1,\dots,n_I)$ are convex, (4.9) is also equivalent to the *Slater condition*

$$\left.\begin{array}{l}\text{(i)}\ \ (DG(\bar{x},\bar{u}),DG_E(\bar{x})):\mathcal{I}^n\times(L_{\mathbb{F}}^{2,2})^m\mapsto\mathcal{I}^n\times\mathbb{R}^{n_E}\ \text{ is surjective and}\\ \text{(ii)}\ \ \exists\,(\hat{z},\hat{v})\in\mathcal{I}^n\times(L_{\mathbb{F}}^{2,2})^m\ ;\ G(\hat{z},\hat{v})=0,\ \ G_E(\hat{z})=0,\ \ G_I^j(\hat{z})<0\ \forall\,j=1,\dots,,n_I.\end{array}\right\}\tag{S}$$

Note that if no inequality constraints are present (which can be written as $n_I=0$), the qualification condition for $(P_{2,P})$ is given by (MF)(a). In this case, as in Theorem 4.2, we get the uniqueness of the multiplier and thus, using (4.10), v is also Gâteaux differentiable at P.

4.2. Multiplicative perturbations in the Linear Quadratic framework

In this part we adopt the framework of *unconstrained* Linear Quadratic (LQ) stochastic control problems with random coefficients (see *e.g.* [4, 12, 30, 31] and the references therein). More precisely, let us consider the problem

$$\begin{array}{ll}\inf_{(x,u)\in\mathcal{I}^n\times(L^{2,2}_{\mathbb{F}})^m} F(x,u) := \frac{1}{2}\mathbb{E}\left(\int_0^T \left[x(t)^\top Q(t)x(t) + u(t)^\top N(t)u(t)\right]\mathrm{d}t + x(T)^\top M x(T)\right) \\ \text{s.t. } \begin{cases} \mathrm{d}x(t) = [A(t)x(t)+B(t)u(t)+e(t)]\mathrm{d}t + \sum_{j=1}^d [C^j(t)x(t)+D^j(t)u(t)+f^j(t)]\mathrm{d}W^j(t), \\ x(0) \ = x_0. \end{cases}\end{array} \tag{$P_{3,P}$}$$

We shall view $P = (x_0, A, B, C^j, D^j, e, f^j)$ $(j = 1,\dots,d)$ as parameters for the problem $(P_{3,P})$. Thus, we consider as parameter space

$$\mathcal{P}_3 = \mathbb{R}^n \times (L^{\infty,\infty}_{\mathbb{F}})^{n\times n} \times (L^{\infty,\infty}_{\mathbb{F}})^{n\times m} \times (L^{\infty,\infty}_{\mathbb{F}})^{(n\times n)\times d} \times (L^{\infty,\infty}_{\mathbb{F}})^{(n\times m)\times d} \times (L^{2,2}_{\mathbb{F}})^{n} \times (L^{2,2}_{\mathbb{F}})^{n\times d}.$$

It is well known (see [4], Thm. 2.1) that given $P \in \mathcal{P}_3$ and $u \in (L^{2,2}_{\mathbb{F}})^m$ the linear SDE in $(P_{3,P})$ admits a unique solution in $\mathcal{I}^n$. We will also need the following result:

Lemma 4.6. *The constraint function $G : \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m \times \mathcal{P}_3 \mapsto \mathcal{I}^n$ defined by:*

$$G(x,u,P) := -x(\cdot) + x_0 + \int_0^{\cdot} [A(t)x(t)+B(t)u(t)+e(t)]\mathrm{d}t + \int_0^{\cdot} \sum_{j=1}^d [C^j(t)x(t)+D^j(t)u(t)+f^j(t)]\mathrm{d}W^j(t),$$

is continuously Fréchet differentiable. Furthermore, $D_{(x,u)}G(x,u,P)$ is onto.

Proof. That G is well defined is a simple application of Lemma 2.2. Following the lines of the proof of Lemma 3.2 (see Sect. A.2 in the Appendix) we have that G is Gâteaux differentiable at any $(x,u,P) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m \times \mathcal{P}_3$ and for every $(x',u') \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m$ and $P' = (x_0', A', B', \{(C^j)'\}, \{(D^j)'\}, e', \{(f^j)'\}) \in \mathcal{P}_3$ we have that

$$\begin{aligned} DG(x,u,P)(x',u',P') = &\int_0^{\cdot}[Ax' + Bu' + e']\mathrm{d}t + \int_0^{\cdot} \sum_{j=1}^d [C^j x'(t) + D^j u' + (f')^j]\mathrm{d}W^j(t) \\ &+ \int_0^{\cdot}[A'x + B'u]\mathrm{d}t + \int_0^{\cdot} \sum_{j=1}^d [(C^j)'X + (D^j)'u]\mathrm{d}W^j(t) + x_0' - x'(\cdot). \end{aligned}$$

Thus, for every (x_1,u_1), $(x_2,u_2) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m$ and P_1, $P_2 \in \mathcal{P}_3$ we have that

$$\|DG(x_1,u_1,P_1)(x',u',P') - DG(x_2,u_2,P_2)(x',u',P')\|^2_{\mathcal{I}}$$

is given by

$$\begin{array}{l}\mathbb{E}\left(\int_0^T \left|(A_1-A_2)x' + (B_1-B_2)u' + A'(x_1-x_2) + B'(u_1-u_2)\right|^2 \mathrm{d}t\right) + \\ \mathbb{E}\left(\sum_{j=1}^d \int_0^T \left|(C_1^j - C_2^j)x' + (D_1^j - D_2^j)u' + (C^j)'(x_1-x_2) + (D^j)'(u_1-u_2)\right|^2 \mathrm{d}t\right).\end{array}$$

Therefore, if $\|P'\| = 1$, we find that $\|DG(x_1,u_1,P_1)P' - DG(x_2,u_2,P_2)P'\|^2_{\mathcal{I}}$ is bounded by

$$c\left(\|x_1-x_2\|^2_{\mathcal{I}} + \|u_1-u_2\|_2^2 + \|A_1-A_2\|^2_\infty + \|B_1-B_2\|^2_\infty + \sum_{j=1}^d\left[\|C_1^j - C_2^j\|^2_\infty + \|D_1^j - D_2^j\|^2_\infty\right]\right),$$

for some $c > 0$, where we used Lemma 2.2 to make $\|\cdot\|_{\mathcal{I}}$ appear. Thus, G is Gâteaux differentiable with a continuous directional derivative, and so G is indeed Fréchet continuously differentiable. The surjectivity of $D_{(x,u)}G(x,u,P)$ follows from Lemma 3.2. □

We make the following convexity assumption:
(H5) The matrix processes $Q : [0,T] \times \Omega \mapsto \mathbb{R}^{n\times n}$, $N : [0,T] \times \Omega \mapsto \mathbb{R}^{m\times m}$ are essentially bounded and progressively measurable, whereas the matrix $M : \Omega \mapsto \mathbb{R}^{n\times n}$ is essentially bounded and $\mathcal{F}_T$-measurable. In addition Q, N and M are a.s. non-negative symmetric matrices and further there exists $\delta > 0$ such that $N \succeq \delta I$.
By ([4], Thm. 3.1) we have that under **(H5)** problem $(P_{3,P})$ admits a unique solution $(x[P], u[P])$.
Moreover, by ([4], Thm. 3.2) (or Thm. 4.2) we obtain the existence of a unique weak-Pontryagin multiplier $(p[P], q[P]) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$ such that

$$\begin{aligned}
\mathrm{d}x(t) =& [A(t)x(t) + B(t)u(t) + e(t)]\mathrm{d}t + \sum_{j=1}^{d}[C^j(t)x(t) + D^j(t)u(t) + f^j(t)]\mathrm{d}W^j(t),\\
u(t) =& -N(t)^{-1}\left[B(t)^\top p(t) + \sum_{j=1}^{d} D^j(t)^\top q^j(t)\right],\\
\mathrm{d}p(t) =& -[A(t)^\top p(t) + \sum_{j=1}^{d} C^j(t)^\top q^j(t) + Q(t)x(t)]\mathrm{d}t + \sum_{j=1}^{d} q^j(t)\mathrm{d}W^j(t),\\
x(0) =& x_0,\ p(T) = Mx(T),
\end{aligned} \tag{4.11}$$

where we have omitted the dependence on P in order to simplify the notation. We want to obtain now an energy estimate for $(x[P], u[P], p[P], q[P])$ in terms of P, in the spirit of ([30], Thm. 2.2). Because we need to keep track of the constants that will appear (since they depend on model parameters, which we shall later vary) we prove the following technical Lemma; the expert reader may skip it.

Lemma 4.7. *Under* **(H5)** *there exists a continuous function $\beta : \mathcal{P}_3 \to \mathbb{R}$ such that*

$$\|x[P]\|^2_{\mathcal{I}} + \|u[P]\|^2_{2,2} + \|p[P]\|^2_{\mathcal{I}} + \sum_{j=1}^{d} \|q^j[P]\|^2_{2,2} \le \beta(P).$$

Proof. For notational convenience we will omit the dependence on P of $(x[P], u[P], p[P], q[P])$. A close look at the classical proof for the stability of solutions to linear SDEs (see *e.g.* [31], Chap. 6, Sect. 4) and of linear BSDEs (see *e.g.* [31], Chap. 7, Thm. 2.2) gives that

$$\begin{aligned}
\|x\|^2_{2,\infty} \le& \kappa_0(P)\left(\|u\|^2_{2,2} + |x_0|^2 + \|e\|^2_{2,2} + \sum_{j=1}^{d}\|f^j\|^2_{2,2}\right),\\
\|p\|^2_{2,\infty} + \sum_{j=1}^{d}\|q^j\|^2_{2,2} \le& \kappa_1(P)\mathbb{E}\left(|M(T)x(T)|^2 + \int_0^T |Q(t)x(t)|^2\mathrm{d}t\right),
\end{aligned} \tag{4.12}$$

where

$$\kappa_0 = \kappa_0(\|A\|_{\infty,\infty}, \|B\|_{\infty,\infty},, \sum_{j=1}^{d}\|C^j\|_{\infty,\infty},, \sum_{j=1}^{d}\|D^j\|_{\infty,\infty}),\quad \text{and}\quad \kappa_1 = \kappa_1\left(\|A\|_{\infty,\infty}, \sum_{j=1}^{d}\|C^j\|_{\infty,\infty}\right),$$

are continuous functions. Recall that for a symmetric non-negative matrix $L \in \mathbb{R}^{n\times n}$ one has that $k_L L \succeq L^2$ for k_L equals the largest eigenvalue of L. It is easy to check that $k_L \le n \max_{i,j\in\{1,\dots,n\}} |L^{ij}|$. Applying this we see that

$$\begin{aligned}
\int_0^T |Q(t)x(t)|^2\mathrm{d}t \le& c\int_0^T x(t)^\top Q(t)x(t)\mathrm{d}t,\\
|M(T)x(T)|^2 \le& cx(T)^\top M(T)x(T),
\end{aligned} \tag{4.13}$$

where $c = n\max\{\|Q\|_{\infty,\infty}, \|M\|_\infty\}$. Now, combining Lemma A.2 and (4.11), we get

$$\mathbb{E}\left(x(T)^\top M(T)x(T) + \int_0^T \left[x^\top Qx + u^\top Nu\right]\mathrm{d}t\right) = p(0)^\top x_0 + \mathbb{E}\left(\int_0^T \left[p^\top e + \sum_{j=1}^d (q^j)^\top f^j\right]\mathrm{d}t\right). \tag{4.14}$$

Therefore, by the second inequality in (4.12), (4.13) and (4.14) we have that

$$\|p\|_{2,\infty}^2 + \sum_{j=1}^d \|q^j\|_{2,2}^2 \le c\kappa_1\left\{|p(0)||x_0| + \mathbb{E}\left(\int_0^T \left|p^\top e + \sum_{j=1}^d (q^j)^\top f^j\right|\mathrm{d}t\right)\right\}.$$

Using now the inequality $2ab \le a^2 + b^2$ for all a, $b \in \mathbb{R}$, we get that

$$\|p\|_{2,\infty}^2 + \sum_{j=1}^d \|q^j\|_{2,2}^2 \le \kappa_2 \mathbb{E}\left(\left[\int_0^T \left(\frac{|x_0|}{T} + |e|\right)\mathrm{d}t\right]^2 + \int_0^T \sum_{j=1}^d |f^j|^2 \mathrm{d}t\right), \tag{4.15}$$

where κ_2 depends continuously on c and κ_1 only, and so the r.h.s. is clearly a continuous function of the model parameters. On the other hand, by (4.13) we have that

$$\delta\|u\|_{2,2}^2 \le p(0)^\top x_0 + \mathbb{E}\left(\int_0^T \left|p^\top e + \sum_{j=1}^d (q^j)^\top f^j\right|\mathrm{d}t\right).$$

Using (4.15) we obtain that $\|u\|_{2,2}^2$ is bounded by a continuous function of P. Therefore, from the first equation in (4.12) we get that $\|x\|_{2,\infty}^2$ is bounded by a continuous function of P. Thus, noting that

$$\begin{aligned}
&x_1[P] = Ax[P] + Bu[P] + e \quad \text{and} \quad x_2^j[P] = C^j(t)x[P](t) + D^j(t)u[P](t) + f^j(t),\\
&p_1[P] = -[A(t)^\top p[P](t) + \textstyle\sum_{j=1}^d C^j(t)^\top q[P]^j(t) + Q(t)x[P](t)] \quad \text{and} \quad p_2^j[P] = q[P]^j,\\
&p_0[P] = \mathbb{E}\left(Mx[P](T) - \textstyle\int_0^T p_1[P](t)\mathrm{d}t\right),
\end{aligned}$$

we obtain that $\|x[P]\|_{\mathcal{I}}^2 + \|p[P]\|_{\mathcal{I}}^2$ is bounded by a continuous function of P. The result follows. □

We prove now a crucial stability result for the solutions of $(P_{3,P})$ in terms of P. More precisely, let P^k and $P \in \mathcal{P}_3$ be such that $P^k \to P$ as $k \to \infty$. We have the following stability result for $(x^k, u^k, p^k, q^k) := (x[P^k], u[P^k], p[P^k], q[P^k])$.

Proposition 4.8. *Call* $(\bar{x}, \bar{u}, \bar{p}, \bar{q}) := (x[P], u[P], p[P], q[P])$ *and suppose that* **(H5)** *holds true. Then:*

(i) *Convergence of the value functions holds:* $v(P^k) \to v(P)$.

(ii) $(x^k, u^k, p^k, q^k) \to (\bar{x}, \bar{u}, \bar{p}, \bar{q})$ *strongly in* $\mathcal{I}^n \times (L_{\mathbb{F}}^{2,2})^m \times \mathcal{I}^n \times (L_{\mathbb{F}}^{2,2})^{n\times d}$.

Proof of (i).
Define $\hat{x}^k$ as the solution of the following SDE:

$$\begin{aligned}
&\mathrm{d}\hat{x}^k(t) = [A^k(t)\hat{x}^k(t) + B^k(t)\bar{u}(t) + \mathrm{e}^k(t)]\mathrm{d}t + \sum_{j=1}^d [(C^j)^k(t)\hat{x}^k(t) + (D^j)^k(t)\bar{u}(t) + (f^j)^k(t)]\mathrm{d}W^j(t),\\
&\hat{x}^k(0) = x_0^k.
\end{aligned}$$

By definition, $(\hat{x}^k, \bar{u}) \in F(P_{3,P^k})$ and by the first estimate in (4.12) we have $\hat{x}^k$ is bounded in $(L_{\mathbb{F}}^{2,\infty})^n$, uniformly in k. Now, $\hat{z}^k := \hat{x}^k - \bar{x} \in \mathcal{I}^n$ satisfies

$$\begin{aligned}
\mathrm{d}\hat{z}^k(t) &= [A(t)\hat{z}^k(t) + \delta^k A\hat{x}^k + \delta^k B(t)\bar{u}(t) + \delta^k e(t)]\mathrm{d}t\\
&\quad + \sum_{j=1}^d [C^j(t)\hat{z}^k(t) + \delta^k C^j(t)\hat{x}^k + \delta^k D^j(t)\bar{u}(t) + \delta^k f^j(t)]\mathrm{d}W^j(t)\\
\hat{z}^k(0) &= \delta^k x_0,
\end{aligned}$$

where $\delta^k A := A^k - A$, $\delta^k B := B^k - B$ and $\delta^k e := \mathrm{e}^k - e$ with an analogous definition for $\delta^k x_0, \delta^k C^j$, $\delta^k D^j$, $\delta^k f^j$. By the convergence $P^k \to P$, the boundedness of $\hat{x}^k$ in $(L^{2,\infty}_{\mathbb{F}})^n$ and classical bounds for linear SDEs (see *e.g.* [31], Chap. 6, Sect. 4), we get that $\hat{z}^k \to 0$ in $(L^{2,\infty}_{\mathbb{F}})^n$. This, implies that $|F(\hat{x}^k, \bar{u}) - F(\bar{x}, \bar{u})|$ tends to zero as $k \uparrow \infty$. Therefore, we get

$$v(P^k) \leq F(\hat{x}^k, \bar{u}) = F(\bar{x}, \bar{u}) + o(1) = v(P) + o(1),$$

which implies that $\limsup_{k\uparrow\infty}[v(P^k) - v(P)] \leq 0$. Analogously, if $\tilde{x}^k$ is the solution of

$$\begin{aligned} \mathrm{d}\tilde{x}^k(t) &= [A(t)\tilde{x}^k(t) + B(t)u^k(t) + e(t)]\mathrm{d}t + \sum_{j=1}^d [C^j(t)\tilde{x}^k(t) + D^j(t)u^k(t) + f^j(t)]\mathrm{d}W^j(t), \\ \tilde{x}^k(0) &= x_0, \end{aligned}$$

we have that $(\tilde{x}^k, u^k) \in F(P_{3,P})$. In addition, $\tilde{z}^k := x^k - \tilde{x}^k$ satisfies

$$\begin{aligned} \mathrm{d}\tilde{z}^k(t) = {} & [A^k(t)\tilde{z}^k(t) + \delta^k A\tilde{x}^k + \delta^k B(t)u^k(t) + \delta^k e(t)]\mathrm{d}t \\ & + \sum_{j=1}^d [(C^j)^k(t)\tilde{z}^k(t) + \delta^k C^j(t)\tilde{x}^k + \delta^k D^j(t)u^k(t) + \delta^k f^j(t)]\mathrm{d}W^j(t), \\ \tilde{z}^k(0) = {} & \delta^k x_0. \end{aligned}$$

By Lemma 4.7 we see that u^k is bounded in $(L^{2,2}_{\mathbb{F}})^m$. So as before we get that $\tilde{x}^k$ is bounded in $(L^{2,\infty}_{\mathbb{F}})^n$, and similarly obtain that $\tilde{z}^k \to 0$ in $(L^{2,\infty}_{\mathbb{F}})^n$ and so $|F(\tilde{x}^k, u^k) - F(x^k, u^k)| \to 0$. Thus, we obtain

$$v(P) \leq F(\tilde{x}^k, u^k) = F(x^k, u^k) + o(1) = v(P^k) + o(1),$$

which implies that $\liminf_{k\uparrow\infty}[v(P^k) - v(P)] \geq 0$, proving the convergence of the value functions. □

Proof of (ii). Since P^k converges to P, Lemma 4.7 implies the existence of $(\hat{x}, \hat{u}, \hat{p}, \hat{q})$ such that, up to some subsequence, (x^k, u^k, p^k, q^k) converges weakly in $\mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m \times \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$ to $(\hat{x}, \hat{u}, \hat{p}, \hat{q})$. By Proposition 4.1, we easily get that $(\hat{x}, \hat{u}, \hat{p}, \hat{q})$ satisfies (4.11). By Corollary 3.14, we have that $(\hat{x}, \hat{u})$ is a solution of $(P_{3,P})$, which by uniqueness implies that $(\hat{x}, \hat{u}) = (\bar{x}, \bar{u})$ and so $(\hat{p}, \hat{q}) = (\bar{p}, \bar{q})$. On the other hand, using the elementary fact that for every sequences a_k, b_k of real numbers such that $a^k + b^k \to a + b$ and $a \leq \liminf a^k$, $b \leq \liminf b^k$ we have that $a^k \to a$ and $b^k \to b$, we get, by the lower semicontinuity of the three terms appearing in F, that $\mathbb{E}\left[\int_0^T u^k(t)^\top N u^k(t)\mathrm{d}t\right] \to \mathbb{E}\left[\int_0^T u(t)^\top N u(t)\mathrm{d}t\right]$ and so by expanding $\mathbb{E}\left[\int_0^T (u^k(t) - u(t))^\top N(u^k(t) - u(t))\mathrm{d}t\right]$ and **(H5)** we conclude that $u^k \to \bar{u}$ strongly in $(L^{2,2}_{\mathbb{F}})^m$. Setting $z^k := x^k - \bar{x}$ and $v^k = u^k - \bar{u}$, we have

$$\begin{aligned} \mathrm{d}z^k(t) = {} & [A(t)z^k(t) + \delta^k A x^k + B(t)v^k + \delta^k B(t)u^k(t) + \delta^k e(t)]\mathrm{d}t \\ & + \sum_{j=1}^d [C^j(t)z^k(t) + \delta^k C^j(t)x^k + D^j(t)v^k + \delta^k D^j(t)u^k(t) + \delta^k f^j(t)]\mathrm{d}W^j(t), \\ z^k(0) = {} & \delta^k x_0. \end{aligned}$$

Since $v^k \to 0$ in $(L^{2,2}_{\mathbb{F}})^m$, using the first estimate of (4.12) and the fact that (x^k, u^k) is bounded in $\mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m$, we obtain that $x^k \to x$ strongly in $(L^{2,\infty}_{\mathbb{F}})^n$ and consequently, passing to the $(L^{2,2}_{\mathbb{F}})^n$ limit in x^k_1 and x^k_2, also in $\mathcal{I}^n$. Finally, setting $\hat{p}^k := p^k - \bar{p}$ and $\hat{q}^k := q^k - \bar{q}$, we have that

$$\begin{aligned} \mathrm{d}\hat{p}^k(t) = {} & -[A(t)^\top \hat{p}^k(t) + \delta^k A(t)p^k(t) + \textstyle\sum_{j=1}^d [C^j(t)^\top (\hat{q})^j(t) + \delta^k C^j(t)^\top (q^j)^k(t)] + Q(t)z^k(t)]\mathrm{d}t \\ & + \textstyle\sum_{j=1}^d (\hat{q}^j)^k(t)\mathrm{d}W^j(t), \\ \hat{p}^k(T) = {} & M z^k(T). \end{aligned}$$

Then, applying the classical estimates for linear BSDEs (see *e.g.* [31], Chap. 7, Thm. 2.2) and using that $z^k(T) \to 0$ strongly in $(L^2_{\mathcal{F}_T})^n$, and that (p^k, q^k) remain bounded in $\mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$, we get that $(\hat{p}^k, \hat{q}^k) \to (0, 0)$ strongly in $(L^{2,2}_{\mathbb{F}})^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$. By passing to the limit in $\hat{p}^k_1$ and $\hat{p}^k_2$ we obtain the desired result. □

Define now the value function $v : \mathcal{P}_3 \mapsto \mathbb{R}$ of problem $(P_{3,P})$ as a function of the parameters. Note that, under **(H5)**, v is well defined. With the previous proposition, we can prove the following sensitivity result:

Theorem 4.9. *Suppose that* **(H5)** *holds. Then, v is of class C^1.*
Moreover, at any $P = (x_0, A, B, \{C^j\}, \{D^j\}, e, \{f^j\}) \in \mathcal{P}_3$ the directional derivative is given by

$$\begin{aligned} Dv(P; \Delta P) = &\bar{p}(0)\Delta x_0 + \mathbb{E}\left(\int_0^T \bar{p}(t)^\top \left[\Delta A(t)\bar{x}(t) + \Delta B(t)\bar{u}(t) + \Delta e(t)\right] \mathrm{d}t\right) \\ &+ \mathbb{E}\left(\int_0^T \sum_{j=1}^d \bar{q}^j(t)^\top \Big[\Delta C^j(t)\bar{x}(t) + \Delta D^j(t)\bar{u}(t) + \Delta f^j(t)\Big] \mathrm{d}t\right), \end{aligned} \tag{4.16}$$

where $\Delta P := (\Delta x_0, \Delta A, \Delta B, \{\Delta C^j\}, \{\Delta D^j\}, \Delta e, \{\Delta f^j\})$ and $(\bar{x}, \bar{u}, \bar{p}, \bar{q}) = (x[P], u[P], p[P], q[P])$.

Proof. The Hadamard differentiability property for v and expression (4.16) follow from the surjectivity result in Lemma 4.6, the strong stability of the solutions proved in Proposition 4.8, the identification of the Lagrange multipliers with the weak-Pontryagin multipliers proved in Theorem 3.12 and ([8], Thm. 4.24), dealing with sensitivity results for the optimal value in optimization problems in Banach spaces. Moreover, using again Proposition 4.8 and expression (4.16) we easily check that $Dv(\cdot) : \mathcal{P}_3 \to L(\mathcal{P}_3, \mathbb{R})$ is continuous, which implies the C^1 property. □

Remark 4.10.
(i) Note that if the nominal problem is deterministic then

$$Dv(P; \Delta P) = \bar{p}(0)\Delta x_0 + \int_0^T \bar{p}(t)^\top \left[\mathbb{E}\left(\Delta A(t)\right)\bar{x}(t) + \mathbb{E}\left(\Delta B(t)\right)\bar{u}(t) + \mathbb{E}[\Delta e(t)]\right] \mathrm{d}t$$

Therefore, the first order term of $v(P+\Delta P) - v(P)$ can be computed with the help of a deterministic differential Riccati equation. This could be useful in practice, since it provides a first order approximation for the value $v(P+\Delta P)$ of the stochastic LQ problem, whose solution is typically characterized in terms of Riccati backward stochastic differential equations, which are more difficult to solve than their deterministic counterparts.

(ii) It could be interesting to study the extension of the above result for the case of indefinite control weight costs, *i.e.* when N is not necessarily definite positive (see [11,31], Chap. 6 and references therein).

4.3. Mean-Variance portfolio selection

Suppose that a market consists of $d+1$ assets $S^0, S^1, \ldots, S^d$ whose prices are defined by

$$\begin{aligned} &\mathrm{d}S^0(t) = rS^0(t), \quad \text{for } t \in [0,T], \quad S^0(0) = 1, \\ &dS(t) = \mathrm{diag}(S(t))\mu(t)\mathrm{d}t + \mathrm{diag}(S(t))\sigma(t)\mathrm{d}W(t) \;\; \text{for} \;\; t \in [0,T], \; S(0) = S_0 \in \mathbb{R}^d, \end{aligned} \tag{4.17}$$

where $S := (S^1, \ldots, S^d)$ and for $a \in \mathbb{R}^d$ the matrix $\mathrm{diag}(a) \in \mathbb{R}^{d\times d}$ is defined as $\mathrm{diag}(a)_{ij} = \delta_{ij}a_i$ for all $i, j \in \{1, \ldots, d\}$ (δ_{ij} is the Kronecker symbol). The precise properties on the processes $r \in L^\infty([0,T]; \mathbb{R})$, $\mu \in (L^{\infty,\infty}_{\mathbb{F}})^d$ and $\sigma \in (L^{\infty,\infty}_{\mathbb{F}})^{d\times d}$ shall be given shortly and will imply that the financial market is arbitrage-free and complete (see *e.g.* [18], Chap. 1,Thms. 4.2 and 6.6).

Given an initial wealth $x \in \mathbb{R}$ and a *self-financing portfolio* $\pi \in (L^{2,2}_{\mathbb{F}})^d$ measured in units of wealth, the associated *wealth process* X is defined through the SDE:

$$\begin{aligned} \mathrm{d}X(t) = &\{r(t)X(t) + \pi(t)^\top(\mu(t) - r(t)\mathbf{1})\}\mathrm{d}t + \pi(t)^\top\sigma(t)\mathrm{d}W(t) \;\; \text{for all} \;\; t \in [0,T], \\ X(0) = &x. \end{aligned} \tag{4.18}$$

where $\mathbf{1}$ denotes the vector of ones in $\mathbb{R}^d$. For $A \in \mathbb{R}$ we consider the problem (see *e.g.* [14,23,34]):

$$\inf_{(X,\pi)\in\mathcal{I}^1\times(L^{2,2}_{\mathbb{F}})^d} \mathbb{E}\left([X - A]^2\right), \;\; \text{such that (4.18) is verified and} \;\; \mathbb{E}\left(X(T)\right) = A. \tag{MVP}$$

We then see that the aim is to minimize the risk (variance) subject to a guaranteed mean-return at the final time T.

We intend to compute the sensitivities of this problem with respect to its parameters. Let us define as *parameter space* $\mathcal{P}_4 := \mathbb{R} \times L^\infty([0,T]) \times \mathbb{R} \times (L^{\infty,\infty}_{\mathbb{F}})^d \times (L^{\infty,\infty}_{\mathbb{F}})^{d\times d}$. We will further say that $P = (x, r, A, \mu, \sigma)$ belongs to $\hat{\mathcal{P}}_4$ if $P \in \mathcal{P}_4$, $\sigma\sigma^\top \succeq \delta I_{d\times d}$ for some $\delta > 0$, and

$$\sum_{i=1}^{d} \left| \mathbb{E}\left(\int_0^T [\mu_i(t) - r(t)]\mathrm{d}t \right) \right| > 0. \tag{4.19}$$

Note that $\hat{\mathcal{P}}_4$ is an open subset of $\mathcal{P}_4$. Let us call $v(P) :=$ *value of* (MVP), the corresponding optimal value function (as a function of the model parameters). On a first step we prove some estimates relating the norms of the portfolio and wealth. As in the LQ-case, we compute the constants rather explicitly to show that they will not explode when we vary the model parameters; the expert reader may want to also skip it.

Lemma 4.11. *If $P = (x, r, A, \mu, \sigma) \in \hat{\mathcal{P}}_4$ and X satisfies (4.18), then*

$$\|\pi\|_{2,2}^2 \le \frac{2}{\delta}\mathbb{E}\left[X(T)^2\right]\left(1 + 2T\left(\|r\|_\infty + \|\sigma^{-1}\{\mu - r\mathbf{1}\}\|^2_{\infty,\infty}\right)\mathrm{e}^{2\left(\|r\|_\infty + \|\sigma^{-1}\{\mu - r\mathbf{1}\}\|^2_{\infty,\infty}\right)T}\right), \tag{4.20}$$

$$\|X\|_{2,2}^2 \le T\mathbb{E}\left(|X(T)|^2\right)\mathrm{e}^{2\left(\|r\|_\infty + \|\sigma^{-1}\{\mu - r\mathbf{1}\}\|^2_{\infty,\infty}\right)T}. \tag{4.21}$$

Proof. By classical results on SDEs (*e.g.* [4], Thm. 2.1) we have that $X \in L^{2,\infty}_{\mathbb{F}}$. Let us set $Z = \sigma^\top\pi$. We have that

$$X(t) = x + \int_0^t \left[r(s)X(s) + Z^\top\sigma^{-1}(s)\{\mu(s) - r(s)\mathbf{1}\}\right]\mathrm{d}s + \int_0^t Z^\top\mathrm{d}W(s).$$

By Itô's formula we have that

$$|X(t)|^2 = |X(T)|^2 - 2\int_t^T X(s)\mathrm{d}X(s) - \int_t^T |Z(s)|^2\mathrm{d}s.$$

Using Lemma A.1 we see that $\int_0^\cdot X\pi^\top\sigma\mathrm{d}W$ is a martingale, and so taking the expectation in the above expression and omitting the time arguments, we get:

$$\begin{aligned}
\mathbb{E}\left(|X(t)|^2 + \textstyle\int_t^T |Z|^2\mathrm{d}s\right) &= \mathbb{E}\left(|X(T)|^2 - 2\textstyle\int_t^T r|X|^2\mathrm{d}s - 2\int_t^T XZ^\top\sigma^{-1}\{\mu - r\mathbf{1}\}\mathrm{d}s\right),\\
&\le \mathbb{E}\left(|X(T)|^2 + 2\|r\|_\infty\textstyle\int_t^T |X|^2\mathrm{d}s + 2\|\sigma^{-1}\{\mu - r\mathbf{1}\}\|_{\infty,\infty}\int_t^T |X||Z|\mathrm{d}s\right),\\
&\le \mathbb{E}\left(|X(T)|^2 + 2\left(\|r\|_\infty + \|\sigma^{-1}\{\mu - r\mathbf{1}\}\|^2_{\infty,\infty}\right)\textstyle\int_t^T |X|^2\mathrm{d}s + \frac{1}{2}\int_t^T |Z|^2\mathrm{d}s\right),
\end{aligned}$$

from which

$$\mathbb{E}\left(|X(t)|^2 + \tfrac{1}{2}\int_t^T |Z|^2\mathrm{d}s\right) \le \mathbb{E}\left(|X(T)|^2 + 2\left(\|r\|_\infty + \|\sigma^{-1}\{\mu - r\mathbf{1}\}\|^2_{\infty,\infty}\right)\int_t^T |X|^2\mathrm{d}s\right). \tag{4.22}$$

Since the above inequality implies that

$$\mathbb{E}\left(|X(t)|^2\right) \le \mathbb{E}\left(|X(T)|^2 + 2\left(\|r\|_\infty + \|\sigma^{-1}\{\mu - r\mathbf{1}\}\|^2_{\infty,\infty}\right)\int_t^T |X|^2\mathrm{d}s\right).$$

by Gronwall's Lemma we obtain that

$$\mathbb{E}\left(|X(t)|^2\right) \le \mathbb{E}\left(|X(T)|^2\right)\mathrm{e}^{2\left(\|r\|_\infty + \|\sigma^{-1}\{\mu - r\mathbf{1}\}\|^2_{\infty,\infty}\right)T} \tag{4.23}$$

and (4.20) follows from estimate (4.22), Fubini's theorem, the definition of Z and the fact that $\sigma\sigma^\top \ge \delta I_{d\times d}$. Finally, estimate (4.21) is a consequence of (4.23) and Fubini's theorem. □

For $P \in \mathcal{P}_4$ let us write the dynamic constraint (4.18) as $G(X, \pi, P) = 0$ with

$$G(X,\pi,P) = x + \int_0^{\cdot} \left[r(t)X(t) + \pi(t)^\top \{\mu(t) - r(t)\mathbf{1}\} \right] \mathrm{d}t + \int_0^{\cdot} \pi(t)^\top \sigma(t) \mathrm{d}W(t) - X(\cdot)$$

and further consider $\hat{G}(X, \pi, P) = (G(X, \pi, P), \mathbb{E}[X(T)] - A)$. Let us prove first:

Lemma 4.12. *The function $\hat{G} : \mathcal{I}^1 \times (L^{2,2}_{\mathbb{F}})^d \times \mathcal{P}_4 \mapsto \mathcal{I}^1 \times \mathbb{R}$ is continuously Fréchet differentiable. Furthermore, if $P \in \hat{\mathcal{P}}_4$, then $D_{(x,\pi)}\hat{G}(X, \pi, P) : \mathcal{I}^1 \times (L^{2,2}_{\mathbb{F}})^d \mapsto \mathcal{I}^1 \times \mathbb{R}$ is onto.*

Proof. The Fréchet differentiability of $\hat{G}$ can be proved following exactly the same lines of the proof in Lemma 4.6 and using that the second component of $\hat{G}$ is a continuous linear functional. For the surjectivity claim, suppose that $P \in \hat{\mathcal{P}}_4$ and that we are given $Y \in \mathcal{I}^1$ and $\xi \in \mathbb{R}$. Then we need to find $(Z, \nu) \in \mathcal{I}^1 \times (L^{2,2}_{\mathbb{F}})^d$ such that:

$$\begin{aligned} -Z(\cdot) + \textstyle\int_0^{\cdot} \left[rZ + \nu^\top(\mu - r\mathbf{1}) \right] \mathrm{d}t + \int_0^{\cdot} \nu^\top \sigma \mathrm{d}W(t) &= Y_0 + \textstyle\int_0^{\cdot} Y_1 \mathrm{d}t + \int_0^{\cdot} Y_2 \mathrm{d}W(t), \\ \mathbb{E}[Z(T)] &= \xi. \end{aligned} \tag{4.24}$$

Let $i \in \{1, \ldots, d\}$ be such that $\kappa := \mathbb{E}\left(\int_0^T [\mu_i(t) - r(t)]\right) \mathrm{d}t \neq 0$. Then, consider the portfolio ν with $\nu^j = 0$ for $j \neq i$ and

$$\nu^i(t) := \left(\frac{\xi + \mathrm{e}^{\int_0^T r(t)\mathrm{d}t} \left[Y_0 + \mathbb{E}\left(\int_0^T \mathrm{e}^{-\int_0^t r(s)\mathrm{d}s} Y_1(t) \mathrm{d}t \right) \right]}{\mathrm{e}^{\int_0^T r(t)\mathrm{d}t} \kappa} \right) \mathrm{e}^{\int_0^t r(s)\mathrm{d}s}.$$

Then, defining $Z \in \mathcal{I}^1$ as the solution of

$$\begin{aligned} \mathrm{d}Z(t) &= \left[r(t)Z(t) + \nu^\top(\mu - r\mathbf{1}) - Y^1(t) \right] \mathrm{d}t + \left[\nu^\top \sigma - Y^2(t) \right] \mathrm{d}W(t), \quad \text{for all } t \in [0,T], \\ Z(0) &= -Y_0, \end{aligned}$$

we easily check that (Z, ν) satisfies (4.24). □

We now show that problem (MVP) is attained. From here onwards $P := (x, r, A, \mu, \sigma) \in \hat{\mathcal{P}}_4$ will denote a tuple of (reference, nominal) parameters. We denote by $v(P)$ the value of (MVP) under parameters P.

Lemma 4.13. *We have that $v(P) < \infty$, and further this value is attained at a unique feasible pair $(X[P], \pi[P])$. Moreover, there exists a unique weak-Pontryagin multiplier*

$$(p[P], q[P], \lambda_E[P]) \in \mathcal{I} \times (L^{2,2}_{\mathbb{F}})^{1\times d} \times \mathbb{R}$$

satisfying:

$$\begin{aligned} \mathrm{d}p[P](t) &= -r(t)p[P](t)\mathrm{d}t + q[P](t)\mathrm{d}W(t) \quad \text{ for all } t \in]0,T[, \\ p[P](T) &= 2[X[P](T) - A] + \lambda_E[P] \quad \textit{a.s. in } \Omega, \\ p[P](t,\omega)(\mu(t,\omega) - r(t)\mathbf{1}) &= -\sigma(t,\omega)(q[P](t,\omega))^\top \quad \text{ a.s. in } [0,T] \times \Omega. \end{aligned} \tag{4.25}$$

Proof. For the finiteness of $v(P)$ it suffices to prove that the feasible set is non-empty. Indeed, by (4.19) there is an i such that $\mathbb{E}[\int_0^T (\mu^i(t) - r(t))\mathrm{d}t] \neq 0$. Therefore, as in the proof of Lemma 4.12, we may build the portfolio π having 0 in every coordinate except for the ith one, which is set to

$$\left(\frac{A \exp\{-\int_0^T r(t)\mathrm{d}t\} - x}{\mathbb{E}[\int_0^T (\mu^i(t) - r(t))\mathrm{d}t]} \right) \mathrm{e}^{\int_0^{\cdot} r(t)\mathrm{d}t}.$$

We easily see that the corresponding wealth process has expected return equal to A at time T and so it is feasible. Suppose now that (X^1, π^1) and (X^2, π^2) attain $v(P)$. This implies that $\mathbb{E}\left[(X^1(T))^2\right] = \mathbb{E}\left[(X^2(T))^2\right]$.

If $X^1(T)$ were not almost surely equal to $X^2(T)$, by strict convexity of $Z \in L^2_{\mathcal{F}_T} \mapsto \mathbb{E}[Z^2]$ we would get that the pair $\frac{1}{2}(X^1 + X^2, \pi^1 + \pi^2)$ is feasible and induces a strictly smaller value of the objective function, yielding a contradiction. Calling now

$$\hat{X}(\cdot) := X^1(\cdot) - X^2(\cdot) = \int_0^{\cdot} \{r(X^1 - X^2) + (\pi^1 - \pi^2)^\top(\mu - r\mathbf{1})\}\mathrm{d}t + \int_0^{\cdot} (\pi^1 - \pi^2)^\top \sigma \mathrm{d}W(t),$$

we see that $\hat{X}(T) = 0$ and from Lemma 4.11 that $\pi^1 - \pi^2 \equiv 0$ and thus $\hat{X}(\cdot) \equiv 0$, and so that X^1 and X^2 are indistinguishable. For attainability, suppose first that (X^k, π^k) is a feasible optimizing sequence. We then know that $\mathbb{E}\left([X^k(T)]^2\right)$ is bounded. By Lemma 4.11 we get that π^k is bounded in $(L^{2,2}_{\mathbb{F}})^d$ and X^k is bounded in $L^{2,2}_{\mathbb{F}}$. Therefore, there exist $\pi \in (L^{2,2}_{\mathbb{F}})^d$, $\hat{X} \in L^{2,2}_{\mathbb{F}}$ such that, up to some subsequence, (X^k, π^k) converges weakly to $(\hat{X}, \pi)$ in $L^{2,2}_{\mathbb{F}} \times (L^{2,2}_{\mathbb{F}})^d$. Moreover, since in $L^{2,2}_{\mathbb{F}}$, we have that X^k_1 converges weakly to $r\hat{X} + \pi^\top(\mu - r\mathbf{1})$ and X^k_2 converges weakly to $\pi^\top \sigma$, we obtain that X^k converges weakly in $\mathcal{I}^1$ to

$$X(\cdot) := x + \int_0^{\cdot} \left[r\hat{X} + \pi^\top(\mu - r\mathbf{1})\right] \mathrm{d}t + \int_0^{\cdot} \pi^\top \sigma \mathrm{d}W(t).$$

Therefore, using that $\mathcal{I}$ is injected continuously in $L^{2,2}_{\mathbb{F}}$ by Proposition 4.1(i), uniqueness of the weak limit implies that $\hat{X} = X$. Moreover, using Proposition 4.1(i) again we see that $\mathbb{E}[X^k(T)] = A$ passes to the limit and we obtain that (X, π) is a feasible pair. Since the cost function is convex and strongly continuous we have that it is l.s.c. with respect to the weak convergence in $\mathcal{I}^1$, which implies that (X, π) is the optimal pair. Finally, the existence and uniqueness of the weak-Pontryagin multiplier $(p[P], q[P], \lambda_E[P])$ is a direct consequence of Theorem 4.4, Remark 4.5 and Lemma 4.12. Using (3.9), it is straightforward to see that $(p[P], q[P], \lambda_E[P])$ satisfies (4.25). □

In order to simplify the sensitivity analysis, we use a change of variables that reduces the number of parameters. We let $X'(\cdot) := \mathrm{e}^{-\int_0^{\cdot} r\mathrm{d}t} X(\cdot) - A\mathrm{e}^{-\int_0^T r(t)\mathrm{d}t}$ and for the portfolio variables we define the new ones by $\pi'(\cdot) = \mathrm{e}^{-\int_0^{\cdot} r\mathrm{d}s}\pi(\cdot)$. With this change of variables, we easily see that for $P' = (x - A\mathrm{e}^{-\int_0^T r(t)\mathrm{d}t}, 0, 0, \mu - r\mathbf{1}, \sigma)$ we have the identity

$$v(P) = \mathrm{e}^{2\int_0^T r\mathrm{d}s} v(P'). \tag{4.26}$$

Moreover, $(\bar{X}, \bar{\pi}, \bar{p}, \bar{q}, \bar{\lambda}_E) = (X[P], \pi[P], p[P], q[P], \lambda_E[P])$ if and only if

$$\begin{aligned}
(X[P'], \pi[P']) &= (\mathrm{e}^{-\int_0^{\cdot} r\mathrm{d}t}\bar{X}(\cdot) - A\mathrm{e}^{-\int_0^T r(t)\mathrm{d}t}, \mathrm{e}^{-\int_0^{\cdot} r\mathrm{d}t}\bar{\pi}) \\
(p[P'], q[P'], \lambda_E[P']) &= \left(\mathrm{e}^{\int_0^{\cdot} r\mathrm{d}t - 2\int_0^T r\mathrm{d}t}\bar{p},\ \mathrm{e}^{\int_0^{\cdot} r\mathrm{d}t - 2\int_0^T r\mathrm{d}t}\bar{q},\ \mathrm{e}^{-\int_0^T r\mathrm{d}t}\bar{\lambda}_E\right).
\end{aligned} \tag{4.27}$$

Therefore, in the following we will consider general perturbations with respect to the initial condition, the drift and diffusion coefficients, and for ease of notation we will write the value function only in terms of these parameters. That is, we shall assume that $r \equiv 0$, $A = 0$ and consider perturbed parameters of the form $P(k) := (x^k, \mu^k, \sigma^k)$. In the end of this section we shall undo the above change of variables and analyze the full original problem.

We will repeatedly use the notation

$$\begin{aligned}
(X^k, \pi^k, p^k, q^k, \lambda^k) &:= (X[P(k)], \pi[P(k)], p[P(k)], q[P(k)], \lambda_E[P(k)]), \\
(\bar{X}, \bar{\pi}, \bar{p}, \bar{q}, \bar{\lambda}_E) &:= (X[P], \pi[P], p[P], q[P], \lambda_E[P]),
\end{aligned}$$

We now prove a stability result, essential to our analysis.

Proposition 4.14. *For any sequence $P(k) \to P$ we have:*

(i) *Convergence of value functions holds: $v(P(k)) \to v(P)$.*
(ii) *The sequence $\left(X^k, \pi^k, p^k, q^k, \lambda^k\right)$ converges strongly in $\mathcal{I}^1 \times (L^{2,2}_{\mathbb{F}})^d \times \mathcal{I}^1 \times (L^{2,2}_{\mathbb{F}})^{1,d} \times \mathbb{R}$ to $\left(\bar{X}, \bar{\pi}, \bar{p}, \bar{q}, \bar{\lambda}_E\right)$.*

Proof of (i).
Begin by noticing that, since $P \in \hat{\mathcal{P}}_4$, there is a coordinate i (which we fix) such that $\left[\mathbb{E}(\int_0^T \mu^i(t)\mathrm{d}t)\right]^2 > 0$. This implies that, for k large enough,

$$\left[\mathbb{E}\left(\int_0^T (\mu^i)^k(t)\mathrm{d}t\right)\right]^2 \geq \frac{1}{2}\left[\mathbb{E}\left(\int_0^T \mu^i(t)\mathrm{d}t\right)\right]^2 > 0$$

and so the portfolios with ith component equal to $-x^k/\mathbb{E}(\int_0^T (\mu^i)^k \mathrm{d}t)$ (and zero in the remaining ones) are feasible for $(MVP(k))$. Using these feasible portfolios, we easily get the existence of $K > 0$ (independent of k) such that $v(P(k)) = \mathbb{E}[X^k(T)^2] \leq K$ and thus by Lemma 4.11 we obtain that π^k is bounded in $(L^{2,2}_{\mathbb{F}})^d$.

Now, consider first only those k such that $v(P(k)) \geq v(P)$ and define a portfolio ν^k equals to $\bar{\pi}$ except for the ith coordinate where it equals $\bar{\pi}^i + z^k$, with

$$z^k := \frac{-x^k - \mathbb{E}(\int_0^T \bar{\pi}^\top \mu^k \mathrm{d}t)}{\mathbb{E}(\int_0^T (\mu^i)^k \mathrm{d}t)}.$$

Calling $Z^k(\cdot) = x^k + \int_0^{\cdot} (\nu^k)^\top \mu^k \mathrm{d}t + \int_0^{\cdot} (\nu^k)^\top \sigma^k \mathrm{d}W(t)$, we easily check that (Z^k, ν^k) is feasible for $(MVP(k))$ and, since $z^k \to 0$, we have that $\mathbb{E}[(Z^k(T))^2] \to \mathbb{E}[(\bar{X}(T))^2]$. Hence, for any $\epsilon > 0$ and k large enough we obtain that $|v(P(k)) - v(P)| \leq v(P(k)) + \epsilon - \mathbb{E}[(Z^k(T))^2] \leq \epsilon$. On the other hand, by considering those k such that $v(P(k)) \leq v(P)$, with a similar manner we can construct out of π^k a new portfolio ξ^k obtained by modification of π^k's ith component in a way that it becomes feasible for the unperturbed problem. More precisely, it suffices to set $(\xi^j)^k = (\pi^j)^k$ for $j \neq i$ and $(\xi^i)^k := (\pi^i)^k + \hat{z}^k$, where

$$\hat{z}^k := \frac{-x - \mathbb{E}(\int_0^T (\pi^k)^\top \mu \mathrm{d}t)}{\mathbb{E}(\int_0^T \mu^i \mathrm{d}t)} = \frac{x^k - x - \mathbb{E}(\int_0^T (\pi^k)^\top [\mu - \mu^k] \mathrm{d}t)}{\mathbb{E}(\int_0^T \mu^i \mathrm{d}t)}.$$

Since π^k is bounded in $(L^{2,2}_{\mathbb{F}})^d$ we obtain that $\hat{z}^k \to 0$ and, as before, we get that for every $\epsilon > 0$ and k large enough, $|v(P(k)) - v(P)| \leq \epsilon$, which finally proves convergence of the value functions. □

Proof of (ii).
Let π be any weak limit point of π^k in $(L^{2,2}_{\mathbb{F}})^d$. Since, for $(y, Y, Z) \in \mathbb{R} \times L^{2,2}_{\mathbb{F}} \times (L^{2,2}_{\mathbb{F}})^d$

$$\langle X^k, (y, Y, Z)\rangle_{\mathcal{I}} = x^k y + \mathbb{E}\left(\int_0^T Y (\pi^k)^\top \mu^k \mathrm{d}t\right) + \mathbb{E}\left(\int_0^T (\pi^k)^\top \sigma^k Z \mathrm{d}t\right),$$

we get that, except for some subsequence, $\langle X^k, (y, Y, Z)\rangle_{\mathcal{I}} \to \langle X, (y, Y, Z)\rangle_{\mathcal{I}}$, where $X(\cdot) = x + \int_0^{\cdot} \pi^\top \mu \mathrm{d}t + \int_0^{\cdot} \pi^\top \sigma \mathrm{d}W(t)$, and thus $X^k \to X$ weakly in $\mathcal{I}^1$. Noticing that $\mathbb{E}(X(T)) = 0$, and by virtue of convergence of the value functions, we have similarly as in Lemma 4.13 that $(X, \pi) = (\bar{X}, \bar{\pi})$. By Proposition 4.1(i) we see that $X^k(T)$ converges weakly in $L^2_{\mathcal{F}_T}$ to $\bar{X}(T)$ and using that $\mathbb{E}([X^k(T)]^2) \to \mathbb{E}([\bar{X}(T)]^2)$ we obtain that $X^k(T) \to \bar{X}(T)$ strongly. Let us write $\hat{X}^k = x + \int_0^{\cdot} (\pi^k)^\top \mu \mathrm{d}t + \int_0^{\cdot} (\pi^k)^\top \sigma \mathrm{d}W(t)$. Then by Lemma 4.11:

$$\|\pi^k - \bar{\pi}\|^2_{2,2} \leq C\mathbb{E}[(\hat{X}^k(T) - \bar{X}(T))^2],$$

where $C = C(\mu, \sigma) > 0$ is some positive constant. Now, we have that $\mathbb{E}[(\bar{X}(T) - X^k(T))^2]$ tends to zero, and

$$\mathbb{E}[(\hat{X}^k(T) - X^k(T))^2] \leq |x - x^k|^2 + T\|\pi^k\|^2_{2,2}\left[\|\mu - \mu^k\|^2_{\infty,\infty} + \|\sigma - \sigma^k\|^2_{\infty,\infty}\right],$$

which also tends to zero. We conclude with the triangle inequality that $\pi^k \to \bar{\pi}$ strongly in $(L^{2,2}_{\mathbb{F}})^d$. Finally, since

$$\|X^k - \bar{X}\|^2_{\mathcal{I}} = |x - x^k|^2 + \|(\pi^k)^\top \mu^k - \bar{\pi}^\top \mu\|^2_{2,2} + \|(\pi^k)^\top \sigma^k - \bar{\pi}^\top \sigma\|^2_{2,2},$$

we conclude that $X^k \to \bar{X}$ strongly in $\mathcal{I}^1$. Now, for the weak-Pontryagin multipliers $(p^k, q^k, \lambda^k) \in \mathcal{I}^1 \times (L^{2,2}_{\mathcal{F}})^{1\times d} \times \mathbb{R}$, by (4.25) we have that:

$$
\begin{aligned}
\mathrm{d}p^k &= q^k \mathrm{d}W(t) \ \text{ for all } t \in]0,T[, \quad p^k(T) = 2X^k(T) + \lambda^k, \\
0 &= p^k(t,\omega)\mu^k(t,\omega) + \sigma^k(t,\omega)(q^k(t,\omega))^\top, \quad \text{for a.a. } (t,\omega) \in [0,T]\times\Omega.
\end{aligned}
$$

We will show now that the λ^k are bounded uniformly in k. Define $P^k = p^k - \lambda^k$. Then we know that $P^k(T) = 2X^k(T)$ and $\mathrm{d}P^k(t) = q^k \mathrm{d}W(t)$. Since the $X^k(T)$ are $L^2_{\mathcal{F}_T}$-bounded, classical estimates for linear BSDEs imply that both q^k and P^k are bounded in $(L^{2,2}_{\mathbb{F}})^{1\times d}$ and $L^{2,2}_{\mathbb{F}}$, respectively. On the other hand, we have that $(P^k + \lambda^k)\mu^k + \sigma^k(q^k)^\top = 0$, which proves that $\lambda^k\mu^k = -P^k\mu^k - \sigma^k(q^k)^\top$ and thus:

$$
|\lambda^k|\|\mu^k\|_{2,2} = \|\lambda^k\mu^k\|_{2,2} \leq \|\mu^k\|_{\infty,\infty}\|P^k\|_{2,2} + \|\sigma^k\|_{\infty,\infty}\|q^k\|_{2,2}.
$$

The right hand-side of the above expression is uniformly bounded by the nature of the perturbations we have, and the estimates we already had. Further, we check that $\|\mu^k\|_{2,2}$ is bounded away from zero since $\mu \not\equiv 0$ and thus λ^k is bounded. Take now any subsequence of (X^k, π^k, λ^k). Then, there exists $\hat{\lambda} \in \mathbb{R}$ such that, except for some subsequence, (X^k, π^k, λ^k) converges strongly to $(\bar{X}, \bar{\pi}, \hat{\lambda})$. This implies, by the classical estimates for linear BSDEs, that the corresponding (p^k, q^k) converge strongly in $L^{2,\infty}_{\mathbb{F}} \times (L^{2,2}_{\mathbb{F}})^{1\times d}$ to the solution (p,q) of

$$
\mathrm{d}p(t) = q(t)\mathrm{d}W(t) \ \text{ for } t \in]0,T[, \quad p(T) = \bar{X}(T) + \hat{\lambda}.
$$

Further, since $p^k(\cdot) = \lambda^k + \int_0^\cdot q^k \mathrm{d}W$, we have that $p^k \to p$ strongly in $\mathcal{I}^1$. Moreover, since $(q^k)^\top = -p^k(\sigma^k)^{-1}\mu^k$ converges in $(L^{2,2}_{\mathbb{F}})^{1,d}$ to $-p\sigma^{-1}\mu$ we conclude that $p\mu + \sigma q^\top = 0$. Therefore, by the uniqueness of the weak-Pontryagin multiplier in Lemma 4.13, we deduce that $(p,q,\hat{\lambda}) = (\bar{p}, \bar{q}, \bar{\lambda}_E)$. This proves that the whole sequence (p^k, q^k, λ^k) converges to $(\bar{p}, \bar{q}, \bar{\lambda}_E)$ strongly in $\mathcal{I}^1 \times (L^{2,2}_{\mathbb{F}})^{1,d} \times \mathbb{R}$. □

By Lemma 4.12, Proposition 4.14 and arguing exactly as in the proof of Theorem 4.9, we have the following:

Proposition 4.15. *The function $(x,\mu,\sigma) \mapsto v(x,0,0,\mu,\sigma)$ is of class C^1 on $\{(x,\mu,\sigma) : (x,0,0,\mu,\sigma) \in \hat{\mathcal{P}}_4\}$. Moreover, for every point $P' = (x,\mu,\sigma)$ in the latter set, we have*

$$
D_{(x,\mu,\sigma)}v(P';\Delta P') = \bar{p}(0)\Delta x + \mathbb{E}\left[\int_0^T \bar{\pi}(t)^\top \Delta\mu(t)\bar{p}(t)\mathrm{d}t\right] + \mathbb{E}\left[\int_0^T \bar{\pi}(t)^\top \Delta\sigma(t)\bar{q}(t)^\top\mathrm{d}t\right], \tag{4.28}
$$

where $\Delta P' = (\Delta x, \Delta\mu, \Delta\sigma)$ and $(\bar{X}, \bar{\pi}, \bar{p}, \bar{q}, \bar{\lambda}_E) = (X[P'], \pi[P'], p[P'], q[P'], \lambda_E[P'])$ is given by Lemma 4.13 with the identification $P' \leftrightarrow (x,0,0,\mu,\sigma)$.

We now unwind the change of variables done in order to reduce the size of the parameter space. In this way we obtain sensitivities with respect to the initial capital, deterministic interest/saving rates, the desired return, the drift and the diffusion coefficients.

Theorem 4.16. *The value function $v : \mathcal{P}_4 \mapsto \mathbb{R}$ is C^1 on $\hat{\mathcal{P}}_4$. Moreover, at every $P = (x,r,A,\mu,\sigma) \in \hat{\mathcal{P}}_4$ we have that*

$$
\begin{aligned}
D_x v(P;\Delta x) &= \bar{p}(0)\Delta x, \\
D_r v(P;\Delta r) &= \mathbb{E}\left(\int_0^T \bar{p}(t)(\bar{X}(t) - \bar{\pi}^\top \mathbf{1})\Delta r \mathrm{d}t\right), \\
D_A v(P;\Delta A) &= -\bar{\lambda}_E \Delta A, \\
D_\mu v(P;\Delta\mu) &= \mathbb{E}\left[\int_0^T \bar{\pi}(t)^\top \Delta\mu(t)\bar{p}(t)\mathrm{d}t\right], \\
D_\sigma v(P;\Delta\sigma) &= \mathbb{E}\left[\int_0^T \bar{\pi}(t)^\top \Delta\sigma(t)\bar{q}(t)^\top \mathrm{d}t\right],
\end{aligned} \tag{4.29}
$$

where $(\bar{X}, \bar{\pi}, \bar{p}, \bar{q}, \bar{\lambda}_E) = (X[P], \pi[P], p[P], q[P], \lambda_E[P])$ *is given by Lemma* 4.13.

Proof. Since $(x, r, A, \mu, \sigma) \to (x - A\mathrm{e}^{-\int_0^T r(t)\mathrm{d}t}, 0, 0, \mu - r\mathbf{1}, \sigma)$ is C^1, we can apply the chain rule in (4.26). Therefore, by (4.27) and Proposition 4.15 we have that

$$
\begin{aligned}
D_x v(P) &= \mathrm{e}^{2\int_0^T r(t)\mathrm{d}t}\, p[P'](0) = \bar{p}(0),\\
D_A v(P) &= \mathrm{e}^{2\int_0^T r(t)\mathrm{d}t}\,(-\mathrm{e}^{-\int_0^T r(t)\mathrm{d}t}) p[P'](0) = -\mathrm{e}^{\int_0^T r(t)\mathrm{d}t}\lambda_E[P'] = -\bar{\lambda}_E,\\
D_\mu v(P)\Delta\mu &= \mathrm{e}^{2\int_0^T r(t)\mathrm{d}t}\,\mathbb{E}\left(\int_0^T \mathrm{e}^{-\int_0^t r(s)\mathrm{d}s}\bar{\pi}(t)^\top \Delta\mu(t)\mathrm{e}^{\int_0^t r(s)\mathrm{d}s - 2\int_0^T r(s)\mathrm{d}s}\bar{p}(t)\mathrm{d}t\right),\\
&= \mathbb{E}\left[\int_0^T \bar{\pi}(t)^\top \Delta\mu(t)\bar{p}(t)\mathrm{d}t\right],\\
D_\sigma v(P)\Delta\sigma &= \mathrm{e}^{2\int_0^T r(t)\mathrm{d}t}\,\mathbb{E}\left(\int_0^T \mathrm{e}^{-\int_0^t r(s)\mathrm{d}s}\bar{\pi}(t)^\top \Delta\sigma(t)\mathrm{e}^{\int_0^t r(s)\mathrm{d}s - 2\int_0^T r(s)\mathrm{d}s}\bar{q}(t)^\top\mathrm{d}t\right),\\
&= \mathbb{E}\left[\int_0^T \bar{\pi}(t)^\top \Delta\sigma(t)\bar{q}(t)^\top\mathrm{d}t\right].
\end{aligned}
$$

Observe that by (4.25) applied at P' we have $p[P'](0) = \lambda_E[P']$, since $2X[P'](T) + \lambda_E[P'] = p[P'](0) + \int_0^T q[P'](t)\mathrm{d}W_t$ must hold. Thus setting $R(\cdot) := \int_0^{\cdot} \Delta r(t)\mathrm{d}t$ we obtain

$$
\begin{aligned}
D_r v(P; \Delta r) =& 2R(T)v(P) + \mathrm{e}^{\int_0^T r(t)\mathrm{d}t} p[P'](0)R(T)A\\
&- \mathrm{e}^{2\int_0^T r(t)\mathrm{d}t}\mathbb{E}\left[\int_0^T \mathrm{e}^{-\int_0^t r(s)\mathrm{d}s}\bar{\pi}^\top \Delta r(t)\mathbf{1}\mathrm{e}^{\int_0^t r(s)\mathrm{d}s - 2\int_0^T r(s)\mathrm{d}s}\bar{p}(t)\mathrm{d}t\right)\\
=& 2R(T)v(P) + \bar{\lambda}_E R(T)A - \mathbb{E}\left(\int_0^T \bar{\pi}^\top \Delta r(t)\mathbf{1}\bar{p}(t)\mathrm{d}t\right).
\end{aligned}
\tag{4.30}
$$

On the other hand,

$$
\mathbb{E}\left(\int_0^T \bar{X}(t)\bar{p}(t)\Delta r(t)\mathrm{d}t\right) = \mathbb{E}\left(R(T)\bar{X}(T)\bar{p}(T) - \int_0^T R(t)\mathrm{d}(\bar{X}(t)\bar{p}(t))\right). \tag{4.31}
$$

By Itô's formula, we can write

$$
\begin{aligned}
\mathrm{d}(\bar{X}(t)\bar{p}(t)) = &\left[r(t)\bar{X}(t)p(t) + \bar{\pi}^\top(\mu(t) - r(t)\mathbf{1})\bar{p}(t) - \bar{X}(t)r(t)\bar{p}(t) + \bar{\pi}(t)^\top\sigma(t)\bar{q}(t)^\top\right]\mathrm{d}t\\
&+ \left[\bar{X}(t)\bar{q} + \bar{p}(t)\bar{\pi}(t)^\top\sigma(t)\right]\mathrm{d}W(t).
\end{aligned}
$$

Since, by the third line in (4.25), $(\mu(t) - r(t)\mathbf{1})\bar{p}(t) = -\sigma(t)\bar{q}(t)^\top$ we obtain with Lemma A.1 that

$$
\mathbb{E}\left(\int_0^T R(t)\mathrm{d}(\bar{X}(t)\bar{p}(t))\right) = 0.
$$

Therefore, by (4.31) and the second line in (4.25), we get

$$
\begin{aligned}
\mathbb{E}\left(\int_0^T \bar{X}(t)\bar{p}(t)\Delta r(t)\mathrm{d}t\right) = \mathbb{E}\left(R(T)\bar{X}(T)\bar{p}(T)\right) &= R(T)\mathbb{E}\left(\bar{X}(T)[2(\bar{X}(T) - A) + \bar{\lambda}_E]\right),\\
&= 2R(T)v(P) + \bar{\lambda}_E R(T)A.
\end{aligned}
\tag{4.32}
$$

The conclusion follows from (4.30) and (4.32). ∎

4.3.1. Comparison with a known explicit result

We want to compare the theoretical sensitivities we obtained with those coming from a simplified model where an explicit solution is known. We choose to compare our results with the model in ([23], Example 4.1) (with null jump component). More precisely, we consider the (MVP) problem with $d = 1$, $r \equiv 0$ and $\mu(\cdot) : [0, T] \to \mathbb{R}$, $\sigma : [0, T] \to \mathbb{R}$ being deterministic bounded functions. Assuming that $\int_0^T \mu(t)\mathrm{d}t \neq 0$ and that σ is uniformly positive, problem (MVP) can be explicitly solved (see [23] for the details). Setting $\Sigma := \mu/\sigma$, the optimal portfolio, optimal states and adjoint states are given by

$$\bar{X}(t) = A + \frac{x - A}{\mathrm{e}^{\int_0^T \Sigma_s^2 ds} - 1}\left[\mathrm{e}^{\int_0^T \Sigma_s^2 \mathrm{d}s}\mathrm{e}^{-\int_0^t\left(\Sigma_s \mathrm{d}W_s + \frac{3}{2}\Sigma_s^2 \mathrm{d}s\right)} - 1\right],$$

$$\bar{\pi}(t) = \frac{(A - x)\mu(t)\mathrm{e}^{\int_0^T \Sigma_s^2 \mathrm{d}s}}{\sigma(t)^2\left(\mathrm{e}^{\int_0^T \Sigma_s^2 \mathrm{d}s} - 1\right)}\mathrm{e}^{-\int_0^t\left(\Sigma_s \mathrm{d}W_s + \frac{3}{2}\Sigma_s^2 \mathrm{d}s\right)}, \tag{4.33}$$

$$\bar{p}(t) = \frac{2(x - A)}{\mathrm{e}^{\int_0^T \Sigma_s^2 \mathrm{d}s} - 1}\mathrm{e}^{-\int_0^t\left(\Sigma_s \mathrm{d}W_s + \frac{1}{2}\Sigma_s^2 \mathrm{d}s\right)}, \tag{4.34}$$

$$\bar{q}(t) = \frac{2(A - x)}{\mathrm{e}^{\int_0^T \Sigma_s^2 \mathrm{d}s} - 1}\Sigma(t)\mathrm{e}^{-\int_0^t\left(\Sigma_s \mathrm{d}W_s + \frac{1}{2}\Sigma_s^2 \mathrm{d}s\right)}. \tag{4.35}$$

Thus, setting $P = (x, 0, A, \mu, \sigma)$ and $\Delta P = (\Delta x, 0, \Delta A, \Delta\mu, \Delta\sigma)$ from Theorem 4.16 we have

$$Dv(P; \Delta P) = \bar{p}(0)[\Delta x - \Delta A] + \mathbb{E}\left[\int_0^T \bar{p}(t)\bar{\pi}(t)\Delta\mu(t)\right] + \mathbb{E}\left[\int_0^T \bar{q}(t)\bar{\pi}(t)\Delta\sigma(t)\right].$$

If we assume that $\Delta\mu$ and $\Delta\sigma$ are deterministic, a brief computation then yields to:

$$D_x v(P; \Delta x) = \frac{2(x - A)\Delta x}{\mathrm{e}^{\int_0^T \Sigma_s^2 \mathrm{d}s} - 1}, \tag{4.36}$$

$$D_A v(P; \Delta A) = -\frac{2(x - A)\Delta A}{\mathrm{e}^{\int_0^T \Sigma_s^2 \mathrm{d}s} - 1}, \tag{4.37}$$

$$D_\mu v(P; \Delta\mu) = -\frac{2(A - x)^2\mathrm{e}^{\int_0^T \Sigma_s^2 \mathrm{d}s}}{\left(\mathrm{e}^{\int_0^T \Sigma_s^2 \mathrm{d}s} - 1\right)^2}\int_0^T \frac{\mu(t)\Delta\mu(t)}{\sigma(t)^2}\mathrm{d}t, \tag{4.38}$$

$$D_\sigma v(P; \Delta\sigma) = \frac{2(A - x)^2\mathrm{e}^{\int_0^T \Sigma_s^2 \mathrm{d}s}}{\left(\mathrm{e}^{\int_0^T \Sigma_s^2 \mathrm{d}s} - 1\right)^2}\int_0^T \frac{\mu(t)^2\Delta\sigma(t)}{\sigma(t)^3}\mathrm{d}t. \tag{4.39}$$

Since we know explicitly the solution, we can actually verify that

$$v(P) = \frac{(x - A)^2}{\mathrm{e}^{\int_0^T \Sigma_s^2 \mathrm{d}s} - 1},$$

and thus computing its derivatives we easily recover (4.36)–(4.39).

Appendix A.

In the first section of this Appendix we explore the relations between the adjoint operators of some simple bounded linear operators defined on $\mathcal{I}^n$ and some linear BSDEs. Next, in Section A.2, we provide the proofs of some technical results stated in Section 3.

A.1. Adjoint operators and backward stochastic differential equations

We start with two basic well-known results. However, since the proofs are short, we provide the details for the reader's convenience.

Lemma A.1. *Let $x \in (L_{\mathbb{F}}^{2,\infty})^n$ and $r \in (L_{\mathbb{F}}^{2,2})^n$. Then, for every $j = 1, \dots, d$,*

$$M^j(\cdot) := \int_0^{\cdot} x(s)^\top r(s) \mathrm{d}W^j(s) \quad \textit{is a martingale.}$$

Proof. Since $x \in (L_{\mathbb{F}}^{2,\infty})^n$ and $r \in (L_{\mathbb{F}}^{2,2})^n$ we have that the stochastic integral M^j is well-defined and is a local-martingale. By the Burkholder−Davis−Gundy inequality (see e.g [17]) we have the existence of a constant $K > 0$ such that

$$\mathbb{E}\left(\sup_{t\in[0,T]} |M^j(t)| \right) \leq K\mathbb{E}\left[\left(\int_0^T |x(s)^\top r(s)|^2 \mathrm{d}t \right)^{\frac{1}{2}} \right] \leq K\|x\|_{2,\infty}\|r\|_{2,2},$$

where the last inequality follows from the Cauchy−Schwarz inequality. Therefore, by ([25], Thm. 51), we have that $M^j(\cdot)$ is a martingale with null expectation. □

Using the above result, the following one is a straightforward consequence of Itô's Lemma and Lemma 2.2.

Lemma A.2. *Let $x, y \in \mathcal{I}^n$. Then*

$$\mathbb{E}\left(x(T)^\top y(T) \right) = x_0^\top y_0 + \mathbb{E}\left(\int_0^T \left[x(t)^\top y_1(t) + y(t)^\top x_1(t) + \sum_{j=1}^d (x_2^j(t))^\top y_2^j(t) \right] \mathrm{d}t \right).$$

Given a sigma-algebra $\mathcal{G} \subseteq \mathcal{F}$ we write $L_{\mathcal{G}}^p := L^p(\Omega, \mathcal{G}, \mathbb{P})$. The following Proposition will be useful.

Proposition A.3. *Let $g \in (L_{\mathcal{F}_T}^2)^n$ and $a \in (L_{\mathbb{F}}^{2,2})^n$. Then, for every $z \in \mathcal{I}^n$ we have that*

$$\begin{aligned} \mathbb{E}\left(g^\top z(T) \right) &= \left\langle \mathbb{E}\left(g|\mathcal{F}_{(\cdot)} \right) + \int_0^{\cdot} \mathbb{E}(g|\mathcal{F}_t)\mathrm{d}t, z \right\rangle_{\mathcal{I}}, \\ \mathbb{E}\left(\int_0^T a(t)^\top z(t)\mathrm{d}t \right) &= \left\langle \mathbb{E}\left(\int_0^T a(t)\mathrm{d}t|\mathcal{F}_{(\cdot)} \right) + \int_0^{\cdot} \mathbb{E}\left(\int_t^T a(s)\mathrm{d}s|\mathcal{F}_t \right) \mathrm{d}t, z \right\rangle_{\mathcal{I}}. \end{aligned} \tag{A.1}$$

In particular,

$$\mathbb{E}\left(g^\top z(T) + \int_0^T a(t)^\top z(t)\mathrm{d}t \right) = \left\langle p(0) + \int_0^{\cdot} p(t)\mathrm{d}t + \int_0^{\cdot} q(t)\mathrm{d}W(t), z \right\rangle_{\mathcal{I}}, \tag{A.2}$$

where $(p, q) \in \mathcal{I}^n \times (L_{\mathbb{F}}^{2,2})^{n\times d}$ is the unique solution of the BSDE

$$\begin{aligned} \mathrm{d}p &= -a(t)\mathrm{d}t + q(t)\mathrm{d}W(t), \\ p(T) &= g. \end{aligned}$$

Proof. Let us first prove (A.1). Let us denote by r_g for the unique element in $(L_{\mathbb{F}}^{2,2})^{n\times d}$ (see (2.1)) such that

$$\mathbb{E}\left(g|\mathcal{F}_{(\cdot)} \right) = \mathbb{E}(g) + \int_0^{\cdot} r_g(t)\mathrm{d}W(t). \tag{A.3}$$

Lemma A.2 implies that

$$\begin{aligned}\mathbb{E}\left(g^\top z(T)\right) &= \mathbb{E}\left(\mathbb{E}(g|\mathcal{F}_T)^\top z(T)\right),\\ &= \mathbb{E}\left(\mathbb{E}(g|\mathcal{F}_0)^\top z_0 + \int_0^T \left[\mathbb{E}(g|\mathcal{F}_t)^\top z_1 + \sum_{j=1}^d (r_g^j)^\top z_2^j\right]\mathrm{d}t\right),\\ &= \left\langle \mathbb{E}(g) + \int_0^\cdot \mathbb{E}(g|\mathcal{F}_t)\mathrm{d}t + \int_0^\cdot r_g(t)\mathrm{d}W(t), z\right\rangle_{\mathcal{I}},\end{aligned}$$

which, together with (A.3), yields to the first identity in (A.1). On the other hand, setting $y(\cdot) := \int_0^\cdot a(t)\mathrm{d}t$,

$$\mathbb{E}\left(\int_0^T a(t)^\top z(t)\mathrm{d}t\right) = \mathbb{E}\left(\int_0^T z(t)^\top \mathrm{d}y(t)\right) = \mathbb{E}\left(y(T)^\top z(T) - \int_0^T y(t)^\top z_1(t)\mathrm{d}t\right),$$

and the second identity in (A.1) follows from the first one. To establish (A.2), let $q \in (L^{2,2}_{\mathbb{F}})^{n\times d}$ be such that

$$\mathbb{E}\left(g + \int_0^T a(t)\mathrm{d}t\big|\mathcal{F}_{(\cdot)}\right) = \mathbb{E}\left(g + \int_0^T a(t)\mathrm{d}t\right) + \int_0^\cdot q(t)\mathrm{d}W(t),$$

and define

$$p(t) := \mathbb{E}\left(g + \int_t^T a(s)\mathrm{d}s\big|\mathcal{F}_t\right).$$

Then

$$p(t) = \mathbb{E}\left(g + \int_0^T a(s)\mathrm{d}t\big|\mathcal{F}_t\right) - \int_0^t a(s)\mathrm{d}s = p(0) - \int_0^t a(s)\mathrm{d}s + \int_0^t q(s)\mathrm{d}W(s),$$

from which the result follows. □

For $g \in (L^{\infty,\infty}_{\mathbb{F}})^{n\times n}$ and $h = (h^j)_{j=1}^d$ with $h^j \in (L^{\infty,\infty}_{\mathbb{F}})^{n\times n}$, let us define the operators A_g, $B_h : \mathcal{I}^n \to \mathcal{I}^n$ as

$$A_g z := \int_0^\cdot g(s)z(s)\mathrm{d}s, \qquad B_h z := \sum_{j=1}^d \int_0^\cdot h^j(s)z(s)\mathrm{d}W^j(s). \tag{A.4}$$

Proposition A.3 has the following consequence:

Corollary A.4. *The following assertions hold:*
(i) *The operator A_g is continuous and its adjoint $A_g^*: \mathcal{I}^n \to \mathcal{I}^n$ is given by*

$$A_g^* r(\cdot) = \mathbb{E}\left(\int_0^T g(t)^\top r_1(t)\mathrm{d}t\big|\mathcal{F}_{(\cdot)}\right) + \int_0^\cdot \mathbb{E}\left(\int_t^T g(s)^\top r_1(s)\mathrm{d}s\big|\mathcal{F}_t\right)\mathrm{d}t, \qquad \forall\, r \in \mathcal{I}^n. \tag{A.5}$$

Moreover,

$$A_g^* r(\cdot) = p_{g,r}(0) + \int_0^\cdot p_{g,r}(t)\mathrm{d}t + \int_0^\cdot q_{g,r}(t)\mathrm{d}W(t) \qquad \forall\, r \in \mathcal{I}^n,$$

where $(p_{g,r}, q_{g,r}) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$ is the unique solution of the following BSDE

$$\begin{aligned}\mathrm{d}p(t) &= -g(t)^\top r_1(t)\mathrm{d}t + q(t)\mathrm{d}W(t),\\ p(T) &= 0.\end{aligned}$$

(ii) *The operator B_h is continuous and its adjoint $B_h^*: \mathcal{I}^n \to \mathcal{I}^n$ is given by*

$$B_h^* r(\cdot) = \sum_{j=1}^d \mathbb{E}\left(\int_0^T h^j(t)^\top r_2^j(t)\mathrm{d}t\big|\mathcal{F}_{(\cdot)}\right) + \sum_{j=1}^d \int_0^\cdot \mathbb{E}\left(\int_t^T h^j(s)^\top r_2^j(s)\mathrm{d}s\big|\mathcal{F}_t\right)\mathrm{d}t, \qquad \forall\, r \in \mathcal{I}^n. \tag{A.6}$$

Moreover,

$$B_h^* r(\cdot) = p_{h,r}(0) + \int_0^{\cdot} p_{h,r}(t)\mathrm{d}t + \int_0^{\cdot} q_{h,r}(t)\mathrm{d}W(t) \qquad \forall\, r \in \mathcal{I}^n,$$

where $(p_{h,r}, q_{h.r}) \in \mathcal{I}^n \times (L_{\mathbb{F}}^{2,2})^{n\times d}$ *is the unique solution of the following BSDE*

$$\begin{aligned} \mathrm{d}p(t) &= -\textstyle\sum_{j=1}^d h^j(t)^\top r_2^j(t)\mathrm{d}t + q(t)\mathrm{d}W(t),\\ p(T) &= 0. \end{aligned}$$

Consequently, the adjoint of $A_g + B_h$ *is given by*

$$(A_g + B_h)^* r(\cdot) = p_r(0) + \int_0^{\cdot} p_r(t)\mathrm{d}t + \int_0^{\cdot} q_r(t)\mathrm{d}W(t) \qquad \forall\, r \in \mathcal{I}^n,$$

where $(p_r, q_r) \in \mathcal{I}^n \times (L_{\mathbb{F}}^{2,2})^{n\times d}$ *is the unique solution of the following BSDE*

$$\begin{aligned} \mathrm{d}p(t) &= -\left[g(t)^\top r_1(t) + \textstyle\sum_{j=1}^d h^j(t)^\top r_2^j(t)\right]\mathrm{d}t + q(t)\mathrm{d}W(t),\\ p(T) &= 0. \end{aligned}$$

Proof. For all $z \in \mathcal{I}^n$, we have that

$$\begin{aligned} \|A_g z\|_{\mathcal{I}} &= \left[\mathbb{E}\left(\textstyle\int_0^T |g(t)z(t)|^2\mathrm{d}t\right)\right]^{\frac{1}{2}} \leq n\|g\|_{\infty,\infty}\|z\|_{2,2} \leq n\sqrt{T}\|g\|_{\infty,\infty}\|z\|_{2,\infty},\\ \|B_h z\|_{\mathcal{I}} &= \textstyle\sum_{j=1}^d \left[\mathbb{E}\left(\int_0^T |h^j(t)z(t)|^2\mathrm{d}t\right)\right]^{\frac{1}{2}} \leq nd\|g\|_{\infty,\infty}\|z\|_{2,2} \leq nd\sqrt{T}\|g\|_{\infty,\infty}\|z\|_{2,\infty}. \end{aligned}$$

Therefore, Lemma 2.2 implies that the linear operators are indeed continuous. We also have that, by Lemma A.3:

$$\begin{aligned} \langle r, A_g z\rangle_I &= \mathbb{E}\left(\textstyle\int_0^T r_1(t)^\top g(t)z(t)\mathrm{d}t\right),\\ &= \left\langle \mathbb{E}(\textstyle\int_0^T g(t)^\top r_1(t)\mathrm{d}t|\mathcal{F}_{(\cdot)}) + \int_0^{\cdot}\mathbb{E}(\int_t^T g(s)^\top r_1(s)\mathrm{d}s|\mathcal{F}_t)\mathrm{d}t, z\right\rangle_{\mathcal{I}}, \end{aligned}$$

which, by Lemma A.3, implies the expression for A_g^* in (i). The corresponding identity for B_g^* in (ii) is obtained by an analogous argument, while assertion (iii) is a direct consequence of (i)–(ii). □

A.2. Proof of some technical results

Proof of Lemma 3.2. Given $z \in \mathcal{I}^n$, $v \in (L_{\mathbb{F}}^{2,2})^m$ and $\tau > 0$, by a first order Taylor expansion of f and σ we obtain

$$\begin{aligned} G(x+\tau z, u+\tau v) - G(x,u) = {} & \tau\textstyle\int_0^{\cdot}\left[f_x(t,x(t),u(t))z(t) + f_u(t,x(t),u(t))v(t) + r_1(t,\tau)\right]\mathrm{d}t\\ & +\tau\textstyle\int_0^{\cdot}\left[\sigma_x(t,x(t),u(t))z(t) + \sigma_u(t,x(t),u(t))v(t) + r_2(t,\tau)\right]\mathrm{d}W(t)\\ & -\tau z(\cdot), \end{aligned} \tag{A.7}$$

where

$$\begin{aligned} r_1(\omega,t,\tau) &:= \textstyle\int_0^1\left[Df(t,x(t)+\theta\tau z(t), u(t)+\theta\tau v(t)) - Df(t,x(t),u(t))\right](z,v)\mathrm{d}\theta,\\ r_2(\omega,t,\tau) &:= \textstyle\int_0^1\left[D\sigma(t,x(t)+\theta\tau z(t), u(t)+\theta\tau v(t)) - D\sigma(t,x(t),u(t))\right](z,v)\mathrm{d}\theta. \end{aligned}$$

By **(H1)**(ii), we have that

$$|r_1(\omega,t,\tau)|^2 + |r_2(\omega,t,\tau)|^2 \leq c'\left(|z(\omega,t)|^2 + |v(\omega,t)|^2\right) \quad \text{for a.a. } (\omega,t) \in \Omega\times[0,T]. \tag{A.8}$$

Since the left-hand side of (A.8) converges a.s. to 0 as $\tau \downarrow 0$, we deduce with Lemma 2.2 and the dominated convergence theorem that

$$\mathbb{E}\left(\int_0^T |r_1(t,\tau)|^2\mathrm{d}t\right) + \mathbb{E}\left(\int_0^T |r_2(t,\tau)|^2\mathrm{d}t\right) \to 0 \quad \text{as } \tau\downarrow 0,$$

and thus (3.3) follows from dividing by τ in (A.7), taking the limit $\tau \downarrow 0$ and the definition of convergence in $\mathcal{I}^n$. Now, fix $v \in (L^{2,2}_{\mathbb{F}})^m$ and $\xi \in \mathcal{I}^n$. Let us prove that there exists $z \in \mathcal{I}^n$ such that $DG(x,u)(z,v) = \xi$. By definition, this is equivalent to solving the SDE

$$\begin{aligned} \mathrm{d}z &= [f_x(t,x(t),u(t))z(t) + f_u(t,x(t),u(t))v(t) - \xi_1]\,\mathrm{d}t \\ &\quad + [\sigma_x(t,x(t),u(t))z(t) + \sigma_u(t,x(t),u(t))v(t) - \xi_2]\,\mathrm{d}W(t) \\ z(0) &= -\xi_0. \end{aligned}$$

Since $(\xi_1,\xi_2) \in (L^{2,2}_{\mathbb{F}})^n \times (L^{2,2}_{\mathbb{F}})^{n\times d}$, under **(H1)** classical results for solvability of linear SDEs (see *e.g.* [4], Thm. 2.1) imply that the above equation has a unique solution. □

Proof of Lemma 3.4. The proof that F is Gâteaux differentiable and that its Gâteaux derivative satisfies the first equation in (3.6) follows the same lines as the proof of Lemma 3.2. Now, note that given $(z,v) \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m$, with $\|z\|_{\mathcal{I}} = \|v\|_{2,2} = 1$, for all (x,u), $(x',u') \in \mathcal{I}^n \times (L^{2,2}_{\mathbb{F}})^m$

$$\begin{aligned} |DF(x,u)(z,v) - DF(x',u')(z,v)| \leq \|z\|_{2,\infty} &\left(\mathbb{E}\left[\int_0^T |\ell_x(t,x(t),u(t)) - \ell_x(t,x'(t),u'(t))|\,\mathrm{d}t\right]^2\right)^{\frac{1}{2}} \\ &+\|z\|_{2,\infty}\left(\mathbb{E}\left[|\Phi_x(x(T)) - \Phi_x(x'(T))|^2\right]\right)^{\frac{1}{2}} \\ &+\|v\|_{2,2}\left(\mathbb{E}\left[\int_0^T |\ell_u(t,x(t),u(t)) - \ell_u(t,x'(t),u'(t))|^2\,\mathrm{d}t\right]\right)^{\frac{1}{2}}. \end{aligned}$$

Therefore, by Lemma 2.1 we get that

$$\sup_{(z,v)\in\mathcal{I}^n\times(L^{2,2}_{\mathbb{F}})^m\ ;\ \|z\|_{\mathcal{I}}=\|v\|_{2,2}=1} |DF(x,u)(z,v) - DF(x',u')(z,v)|^2 \leq c w(x',u'),$$

where

$$\begin{aligned} w(x',u') := \mathbb{E}\Big(\int_0^T &\left[|\ell_x(t,x(t),u(t)) - \ell_x(t,x'(t),u'(t))|^2 + |\ell_u(t,x(t),u(t)) - \ell_u(t,x'(t),u'(t))|^2\right]\mathrm{d}t \\ &+ |\Phi_x(x(T)) - \Phi_x(x'(T))|^2\Big). \end{aligned}$$

Since ℓ_x, ℓ_u and Φ_x satisfy the linear growth property in (3.4), we have by dominated convergence that $w(x',u') \to 0$ as $\|x'-x\|_{\mathcal{I}} + \|u'-u\|_{2,2} \to 0$. Thus DF is continuous and therefore F is Fréchet differentiable. The proof of the analogous result for G_E and G_I follows the same lines. □

Acknowledgements. The first author is grateful for full support from the Berlin Mathematical School and Humboldt-Universität zu Berlin, and partial support by the Austrian Science Fund (FWF) under grant Y782-N25. The second author acknowledges partial supported by the Gaspar Monge Program for Optimization and Operation Research (PGMO). This work was initiated during a one month visit to the Hausdorff Research Institute for Mathematics at the University of Bonn in the framework of the Trimester Program Stochastic Dynamics in Economics and Finance. The authors likewise express their gratitude to J. Fontbona for constructive discussions.

References

[1] A. Bensoussan, Lectures on stochastic control. In Nonlinear filtering and stochastic control (Cortona, 1981). Vol. 972 of *Lect. Notes Math.* Springer, Berlin, New York (1982) 1–62.

[2] A. Bensoussan, Stochastic maximum principle for distributed parameter system. *J. Franklin Inst.* **315** (1983) 387–406.

[3] J.-M Bismut, Conjugate convex functions in optimal stochastic control. *J. Math. Anal. Appl.* **44** (1973) 384–404.

[4] J.-M. Bismut, Linear quadratic optimal stochastic control with random coefficients. *SIAM J. Control Optim.* **14** (1976) 419–444.

[5] J.-M Bismut, Théorie probabiliste du contrôle des diffusions. *Mem. Amer. Math. Soc.* **4** (167):xiii+130 (1976).

[6] J.F. Bonnans, Optimisation Continue. Dunod, Paris (2006).

[7] J.F. Bonnans and A. Shapiro, Optimization problems with perturbations: a guided tour. *SIAM Rev.* **40** (1998) 228–264.

[8] J.F. Bonnans and A. Shapiro, Perturbation analysis of optimization problems. *Springer Ser. Oper. Research*. Springer-Verlag, New York (2000).

[9] J.F. Bonnans and F.J. Silva, First and second order necessary conditions for stochastic optimal control problems. *Appl. Math. Optim.* **65** (2012) 403–439.

[10] A. Cadenillas and I. Karatzas, The stochastic maximum principle for linear convex optimal control with random coefficients. *SIAM J. Control Optim.* **33** (1995) 590–624.

[11] S. Chen, X. Li, and X.Y. Zhou, Stochastic linear quadratic regulators with indefinite control weight costs. *SIAM J. Control Optim.* **36** (1998) 1685–1702 (electronic).

[12] S. Chen and J. Yong, Stochastic linear quadratic optimal control problems. *Appl. Math. Optim.* **43** (2001) 21–45.

[13] X. Cheng and J. Yan, A new look at the Lagrange method for continuous-time stochastic optimization. *Sci. China Math.* **55** (2012) 2247–2258.

[14] D. Duffie and H. R. Richardson, Mean-variance hedging in continuous time. *Ann. Appl. Probab.* **1** (1991) 1–15.

[15] U.G. Haussmann, A stochastic maximum principle for optimal control of diffusions. *Pitman Research Notes Math. Series.* Longman, Scientific & Technical (1986).

[16] N. Ikeda and S. Watanabe, Stochastic differential equations and diffusion processes. Vol. 24 of *North-Holland Mathematical Library.* North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 2nd edition (1989).

[17] I. Karatzas and S.E. Shreve, Brownian motion and stochastic calculus. Vol. 113 of *Grad. Texts Math.* 2nd edition. Springer-Verlag, New York (1991).

[18] I. Karatzas and S.E. Shreve, Methods of mathematical finance. Vol. 39 of *Appl. Math.* Springer-Verlag, New York (1998).

[19] P. Kosmol and M. Pavon, Solving optimal control problems by means of general Lagrange functionals. *Automatica J. IFAC* **37** (2001) 907–913.

[20] H.J. Kushner, On the stochastic maximum principle: Fixed time of control. *J. Math. Anal. Appl.* **11** (1965) 78–92.

[21] P.D. Loewen, Parameter sensitivity in stochastic optimal control. *Stochastics* **22** (1987) 1–40.

[22] H. Maurer and J. Zowe, First and second order necessary and sufficient optimality conditions for infinite-dimensional programming problems. *Math. Program.* **16** (1979) 98–110,.

[23] B. Øksendal N.C. Framstad and A. Sulem, Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance. *J. Optim. Theory Appl.* **121** (2004) 77–98.

[24] S. Peng, A general stochastic maximum principle for optimal control problems. *SIAM J. Control Optim.* **28** (1990) 966–979.

[25] P.E. Protter, Stochastic integration and differential equations. Vol. 21 of *Stoch. Model. Appl. Probab.* 2nd edition. Version 2.1, Corrected third printing. Springer-Verlag, Berlin (2005).

[26] D. Revuz and M. Yor, Continuous martingales and Brownian motion. Vol. 293 of *Grundlehren der Math. Wissens. [Fundamental Principles of Mathematical Sciences]*, 3rd edition. Springer-Verlag, Berlin (1999).

[27] S.M. Robinson, First order conditions for general nonlinear optimization. *SIAM J. Appl. Math.* **30** (1976) 597–607.

[28] R. T. Rockafellar, Conjugate convex functions in optimal control and the calculus of variations. *J. Math. Anal. Appl.* **32** (1970) 174–222.

[29] R.T. Rockafellar, Conjugate duality and optimization. Lectures given at the Johns Hopkins University, Baltimore, Md., June (1973). *Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics*, No. 16. Society for Industrial and Applied Mathematics, Philadelphia, Pa. (1974).

[30] S. Tang, General linear quadratic optimal stochastic control problems with random coefficients: linear stochastic Hamilton systems and backward stochastic Riccati equations. *SIAM J. Control Optim.* **42** (2003) 53–75 (electronic).

[31] J. Yong and X. Y. Zhou, Stochastic controls. Hamiltonian systems and HJB equations. Vol. 43 of *Appl. Math.* Springer-Verlag, New York (1999).

[32] X.Y. Zhou, Maximum principle, dynamic programming, and their connection in deterministic control. *J. Optim. Theory Appl.* **65** (1990) 363–373.

[33] X.Y. Zhou, A unified treatment of maximum principle and dynamic programming in stochastic controls. *Stochastics Stochastics Rep.* **36** (1991) 137–161.

[34] X.Y. Zhou and D. Li, Continuous-time mean-variance portfolio selection: a stochastic LQ framework. *Appl. Math. Optim.* **42** (2000) 19–33.

[35] J. Zowe and S. Kurcyusz, Regularity and stability for the mathematical programming problem in banach spaces. *Appl. Math. Optim.* **5** (1979) 49–62.