



**HAL**  
open science

# Optimality conditions in variational form for non-linear constrained stochastic control problems

Laurent Pfeiffer

► **To cite this version:**

Laurent Pfeiffer. Optimality conditions in variational form for non-linear constrained stochastic control problems. *Mathematical Control & Related Fields*, 2020, 10 (3), pp.493-526. 10.3934/mcrf.2020008 . hal-03113195

**HAL Id: hal-03113195**

**<https://hal.inria.fr/hal-03113195>**

Submitted on 18 Jan 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## OPTIMALITY CONDITIONS IN VARIATIONAL FORM FOR NON-LINEAR CONSTRAINED STOCHASTIC CONTROL PROBLEMS

LAURENT PFEIFFER

Institute of Mathematics and Scientific Computing, University of Graz  
Heinrichstrasse 36, 8010 Graz, Austria

(Communicated by the associate editor name)

**ABSTRACT.** Optimality conditions in the form of a variational inequality are proved for a class of constrained optimal control problems of stochastic differential equations. The cost function and the inequality constraints are functions of the probability distribution of the state variable at the final time. The analysis uses in an essential manner a convexity property of the set of reachable probability distributions. An augmented Lagrangian method based on the obtained optimality conditions is proposed and analyzed for solving iteratively the problem. At each iteration of the method, a standard stochastic optimal control problem is solved by dynamic programming. Two academical examples are investigated.

**1. Introduction.** This article is devoted to the derivation of optimality conditions for a class of control problems of stochastic differential equations (SDE). For a given adapted control process  $u$ , let  $(X_t^{0,Y_0,u})_{t \in [0,T]}$  be the solution to the following SDE:

$$\begin{cases} dX_t^{0,Y_0,u} = b(X_t^{0,Y_0,u}, u_t) dt + \sigma(X_t^{0,Y_0,u}, u_t) dW_t, & \text{for a. e. } t \in [0, T], \\ X_0^{0,Y_0,u} = Y_0, \end{cases}$$

where the drift  $b$ , the volatility  $\sigma$ , and the random variable  $Y_0$  are given. For all  $t \in [0, T]$ , we denote by  $m_t^{0,Y_0,u}$  the probability distribution of  $X_t^{0,Y_0,u}$ . Along the article, constrained problems of the following form are considered:

$$\inf_{u \in \mathcal{U}_0(Y_0)} F(m_T^{0,Y_0,u}), \quad \text{subject to: } G(m_T^{0,Y_0,u}) \leq 0, \quad (1)$$

where the mappings  $F$  and  $G$  are given and satisfy differentiability assumptions. The control set  $\mathcal{U}_0(Y_0)$  is a set of adapted stochastic processes. In the article,  $n$  denotes the dimension of the state variable and  $N$  the number of constraints. A precise description of problem (1) will be given in Section 2. In this paper, we call the mappings  $F$  and  $G$  linear if they can be written in the form  $F(m) = \int_{\mathbb{R}^n} f(x) dm(x)$  and  $G(m) = \int_{\mathbb{R}^n} g(x) dm(x)$ , where  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  and  $g : \mathbb{R}^n \rightarrow \mathbb{R}^N$ . In this specific

---

2010 *Mathematics Subject Classification.* Primary: 90C15, 93E20; Secondary: 49J53, 49K99.

*Key words and phrases.* stochastic optimal control, optimality conditions, augmented Lagrangian method, mean-field-type control, relaxation.

This article is dedicated to Frédéric Bonnans in his 60th birthday.

case, problem (1) is equivalent to the following stochastic optimal control problem with an expectation constraint:

$$\inf_{u \in \mathcal{U}_0(Y_0)} \mathbb{E}[f(X_T^{0,Y_0,u})], \quad \text{subject to: } \mathbb{E}[g(X_T^{0,Y_0,u})] \leq 0. \quad (2)$$

The terminology *non-linear* used in the title refers to the fact that the functions  $F$  and  $G$  for which our result applies are not necessarily linear. In other words, the cost function and the constraints are not necessarily formulated as expectations of functions of  $X_T^{0,Y_0,u}$ . This is the specificity of the present article.

Stochastic optimal control problems with a non-linear cost function are mainly motivated by applications in economy and finance. In many situations, minimizing the expectation of a random cost may be unsatisfactory and one may prefer to take into account the “risk” induced by the dispersion of the cost. In the literature, there exist many different models of risk and some of them can be formulated as functions of the probability distribution of the state variable, as explained for example in [31, Chapter 6]. Some portfolio problems with risk-averse cost functions are studied in [26]. In [3, Section 5], a mean-variance portfolio selection problem is considered as well as in [22]. In [19], a gradient-based method is developed for minimization problems of the conditional value at risk, a popular risk-averse cost function. The risk can also be taken into account by considering a constraint of the form  $G(m_T^{0,Y_0,u}) \leq 0$ . For example, one can try to keep the probability of bankruptcy under a given threshold with a probability constraint. If  $G$  models the variance of the outcome of some industrial process, then a constraint on the variance can guarantee some uniformity in the outcome, which can be a desirable property. Final-time constraints have several applications in finance, see for example [33], where probability constraints are used for solving a super-replication problem or [32], where the final probability distribution is fixed for solving a determination problem of no-arbitrage bounds of exotic options. Finally, let us mention that a problem of the form (1) can be seen as a simple model for an optimization problem of a multi-agent system, where a government, that is to say a centralized controller, influences a (very large) population of agents, whose behavior is described by a SDE (see for example [2] and the references therein on this topic). Let us mention however that for general multi-agent models, the coefficients of the SDE at time  $t$  depend on the current probability distribution  $m_t^{0,Y_0,u}$ .

Let us describe the available results in the literature related to optimality conditions for problems similar to Problem (1). Problem (1), without constraints, is a specific case of an optimal control problem of a McKean-Vlasov process (or mean-field-type control problem). For this more general class of problems, the drift and volatility of the SDE possibly depend at any time  $t$  on the current distribution  $m_t^{0,Y_0,u}$ . Optimality conditions usually take the form of a stochastic maximum principle: for a solution  $\bar{u}$ ,  $\bar{u}_t$  minimizes almost surely and for almost every time  $t$  a Hamiltonian involving a costate which is obtained as a solution to a backward stochastic differential equation (see for example [3], [4], [9], [13]). In a different but related approach, one can consider control processes in a feedback form, that is to say in the form  $u_t = \mathbf{u}(t, X_t^{0,Y_0,u})$ , where the mapping  $\mathbf{u}: [0, T] \times \mathbb{R}^n \rightarrow U$  has to be optimized. Under regularity assumptions, the probability distribution  $m_t^{0,Y_0,u}$  has a density, say  $\mu(t, x)$ , which is the solution to the Fokker-Planck equation:

$$\partial_t \mu = \frac{1}{2} \nabla^2 : [\mu(t, x) a(x, \mathbf{u}(t, x))] - \nabla \cdot [\mu(t, x) b(x, \mathbf{u}(t, x))],$$

where  $a = \sigma\sigma^\top$  and where the operators  $\nabla \cdot$  and  $\nabla^2$  : are defined by

$$\nabla \cdot f(x) = \sum_{i=1}^n \partial_{x_i} f_i(x) \quad \text{and} \quad \nabla^2 : g(x) = \sum_{i,j=1}^n \partial_{x_i x_j}^2 g_{ij}(x),$$

respectively. A derivation of the Fokker-Planck can be found in [12, Lemma 3.3], for example. Note that if  $b$  and  $\sigma$  do not depend on  $\mu$ , then the Fokker-Planck equation is a linear partial differential equation. In this approach, the problem is an optimal control problem of the Fokker-Planck equation and optimality conditions take the form of a Pontryagin's maximum principle. The adjoint equation is in this setting an HJB equation. We refer the reader to [1], [2], [4], and [15] for this approach.

In this article, optimal control problems of the following form:

$$\inf_{u \in \mathcal{U}_0(Y_0)} \mathbb{E}[\phi(X_T^{0,Y_0,u})] \quad (3)$$

are called *standard problems*, considering the fact that they have been extensively studied in the last decades. They can be solved by dynamic programming, by computing the solution to the associated Hamilton-Jacobi-Bellman equation (see the textbooks [16], [28], [35] on this field). Since standard problems are of form (2) (without constraints), they fall into the general class of problems investigated in the paper. The optimality conditions provided in the present article can be shortly formulated as follows: if  $\bar{u}$  is a solution to (1) and satisfies a qualification condition, then it is also the solution to a standard problem of the form (3), where the involved function  $\phi$  is the derivative at  $m_T^{0,Y_0,\bar{u}}$  (in a specific sense) of the Lagrangian of the problem  $L := F + \langle \lambda, G \rangle$ , for some non-negative Lagrange multiplier  $\lambda$  satisfying a complementarity condition. Our optimality conditions therefore take the form of a variational inequality. Our analysis relies on the following technical result: the closure (for the Wasserstein distance associated with the  $L^1$ -distance) of the set of reachable probability distributions at time  $T$  is convex. This property is proved by constructing controls imitating the behaviour of relaxed controls. To the best of our knowledge, the optimality conditions for problem (1) are new, as well as the convexity property satisfied by the reachable set<sup>1</sup>. They differ from the maximum principle mentioned above and are, to a certain extent, related to the optimality conditions obtained for mean-field-type control problems formulated with feedback laws. The presence of non-linear constraints is another novelty of the article; in the literature, most constraints of the form  $G(m_T^{0,Y_0,\bar{u}}) \leq 0$  are expectation constraints, see for example [8], [27]. The existence of a Lagrange multiplier, even in a linear setting, is not often considered, see [7, Section 5] or [23, Section 6]. Finally, we mention that we also prove the sufficiency of the optimality conditions in variational form, when  $F$  and  $G$  are convex.

It is well-known that problems of the form (1) are time-inconsistent, i.e. it is (in general) not possible to write a dynamic programming principle by parameterizing problem (1) by its initial time and initial condition, as is customary for standard problems of the form (3). A dynamic programming principle can be written if one considers the whole initial probability distribution as a state variable, see [18], [29], [30]. However, in practice, this approach does not allow, in general, to solve the problem, because the complexity of the method grows exponentially with the dimension of the (discretized) space of probability distributions. The optimality

<sup>1</sup>A proof of the convexity property as well as optimality conditions in variational form for unconstrained problems can be found in the unpublished research report [24].

conditions in variational form and the convexity property proved in this article naturally lead to iterative methods for solving problem (1), based on successive resolutions of standard problems and thus overcoming the difficulty related to time-inconsistency. We propose, analyse, and test such a method in the article. The cost function of the standard problem to be solved at each iteration is the derivative, in a certain sense, of an augmented Lagrangian.

We give a precise formulation of the problem under study in Section 2. We also discuss the notion of differentiability which is used. In Section 3, we prove the convexity of the closure of the reachable set of probability distributions. Optimality conditions in variational form are proved in Section 4. The case of convex problems is discussed. Our numerical method for solving the problem is described and analyzed in Section 5. We provide results for two academical examples. Elements on optimal transportation theory are given in the appendix.

## 2. Formulation of the problem and assumptions.

### 2.1. Notation.

- The set of probability measures on  $\mathbb{R}^n$  is denoted by  $\mathcal{P}(\mathbb{R}^n)$ . For a function  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ , its integral (if well-defined) with respect to the measure  $m \in \mathcal{P}(\mathbb{R}^n)$  is denoted by

$$\int_{\mathbb{R}^n} \phi(x) dm(x) \quad \text{or} \quad \int_{\mathbb{R}^n} \phi dm.$$

Given two measures  $m_1$  and  $m_2 \in \mathcal{P}(\mathbb{R}^n)$ , we denote:

$$\int_{\mathbb{R}^n} \phi(x) d(m_2(x) - m_1(x)) := \int_{\mathbb{R}^n} \phi(x) dm_2(x) - \int_{\mathbb{R}^n} \phi(x) dm_1(x).$$

- For a given random variable  $X$  with values in  $\mathbb{R}^n$ , we denote by  $\sigma(X)$  the  $\sigma$ -algebra generated by  $X$ .
- For a given vector  $x \in \mathbb{R}^q$ , we denote by  $|x|$  its Euclidean norm and by  $|x|_\infty$  its supremum norm.
- For  $p \geq 1$ , we denote by  $\mathcal{P}_p(\mathbb{R}^n)$  the set of probability measures having a finite  $p$ -th moment:

$$\mathcal{P}_p(\mathbb{R}^n) := \left\{ m \in \mathcal{P}(\mathbb{R}^n) \mid \int_{\mathbb{R}^n} |x|^p dm(x) < +\infty \right\}.$$

We recall that for  $1 \leq p \leq q$ , the space  $\mathcal{P}_q(\mathbb{R}^n)$  is included into  $\mathcal{P}_p(\mathbb{R}^n)$ . We equip  $\mathcal{P}_1(\mathbb{R}^n)$  with the Wasserstein distance  $d_1$ . The definition of the Wasserstein distance is recalled in the appendix. In the article, we only make use of the dual representation of  $d_1$  [34, Remark 6.5]: for all  $m_1, m_2 \in \mathcal{P}_1(\mathbb{R}^n)$ ,

$$d_1(m_1, m_2) = \sup_{\phi \in 1\text{-Lip}(\mathbb{R}^n)} \int_{\mathbb{R}^n} \phi d(m_2 - m_1), \quad (4)$$

where  $1\text{-Lip}(\mathbb{R}^n) := \{ \phi : \mathbb{R}^n \rightarrow \mathbb{R} \mid |\phi(y) - \phi(x)| \leq |y - x|, \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n \}$  is the set of real-valued Lipschitz continuous functions of modulus 1.

- For all  $R \geq 0$ , we define

$$\bar{B}_p(R) := \left\{ m \in \mathcal{P}_p(\mathbb{R}^n) \mid \int_{\mathbb{R}^n} |x|^p dm(x) \leq R \right\}. \quad (5)$$

- The open (resp. closed) ball in  $\mathbb{R}^n$  of radius  $r \geq 0$  and center 0 is denoted by  $B_r$  (resp.  $\bar{B}_r$ ), its complement  $B_r^c$  (resp.  $\bar{B}_r^c$ ), for the Euclidean norm.

- For a given  $p \geq 1$ , we say that a function  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^q$  is dominated by  $|x|^p$  if for all  $\varepsilon > 0$ , there exists  $r > 0$  such that for all  $x \in B_r^c$ ,

$$|\phi(x)| \leq \varepsilon |x|^p. \quad (6)$$

- The convex envelope of a set  $\mathcal{R}$  is denoted  $\text{conv}(\mathcal{R})$ . When  $\mathcal{R}$  is a subset of  $\mathcal{P}_1(\mathbb{R}^n)$ , its closure for the  $d_1$ -distance is denoted  $\text{cl}(\mathcal{R})$ .

**2.2. State equation.** We fix a final time  $T > 0$  and a Brownian motion  $(W_t)_{t \in [0, T]}$  of dimension  $d$ . For all  $0 \leq t \leq T$ ,  $(W_s - W_t)_{s \in [t, T]}$  is a standard Brownian motion. For all  $s \in [t, T]$ , we denote by  $\mathcal{F}_{t,s}$  the  $\sigma$ -algebra generated by  $(W_\theta - W_t)_{\theta \in [t, s]}$ .

Let  $U$  be a compact subset of  $\mathbb{R}^k$ . Note that we do not make any other assumption on  $U$ : it can be non-convex, for example, or can be a discrete set. For a given random variable  $Y_t$  independent of  $\mathcal{F}_{t,T}$  with values in  $\mathbb{R}^n$ , we define the set  $\mathcal{U}_t^0$  as the set of control processes  $(u_s)_{s \in [t, T]}$  taking values in  $U$  such that for all  $s \in [t, T]$ ,  $u_s$  is  $\mathcal{F}_{t,s}$ -measurable and we define the set  $\mathcal{U}_t(Y_t)$  as the set of control processes  $(u_s)_{s \in [t, T]}$  taking values in  $U$  such that for all  $s \in [t, T]$ ,  $u_s$  is  $(\sigma(Y_t) \times \mathcal{F}_{t,s})$ -measurable.

The drift  $b : \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$  and the volatility  $\sigma : \mathbb{R}^n \times U \rightarrow \mathbb{R}^{n \times d}$  are given. For all  $u \in \mathcal{U}_t(Y_t)$ , we denote by  $(X_s^{t, Y_t, u})_{s \in [t, T]}$  the solution to the SDE

$$dX_s^{t, Y_t, u} = b(X_s^{t, Y_t, u}, u_s) ds + \sigma(X_s^{t, Y_t, u}, u_s) dW_s, \quad \forall s \in [t, T], \quad X_t^{t, Y_t, u} = Y_t. \quad (7)$$

The well-posedness of this SDE is ensured by Assumption A1 [21, Section 5] below. We also denote by  $m_s^{t, Y_t, u}$  the probability distribution of  $X_s^{t, Y_t, u}$ .

All along the article, we assume that the following assumption holds true. From now on, the initial condition  $Y_0$  and the real number  $p \geq 2$  introduced below are fixed.

**Assumption A1.** *There exists  $K > 0$  such that for all  $x, y \in \mathbb{R}^n$ , for all  $u, v \in U$ ,*

$$\begin{aligned} |b(x, u)| + |\sigma(x, u)| &\leq K(1 + |x|), \\ |b(x, u) - b(y, v)| + |\sigma(x, u) - \sigma(y, v)| &\leq K(|y - x| + |v - u|). \end{aligned}$$

*There exists  $p \geq 2$  such that  $\mathbb{E}[|Y_0|^p] < \infty$ .*

The following lemma is classical, see for example [17, Section 2.5].

**Lemma 2.1.** *There exist three constants  $C_1, C_2$ , and  $C_3 > 0$  depending on  $K, T, U$ , and  $p$  such that for all  $t \in [0, T]$ , for all random variables  $Y_t$  and  $\tilde{Y}_t$  independent of  $\mathcal{F}_{t,T}$  and taking values in  $\mathbb{R}^n$ , for all  $u \in \mathcal{U}_t(Y_t)$ , for all  $0 \leq h \leq T - t$ , for all  $t \leq s \leq T - h$ , the following estimates hold:*

1.  $\mathbb{E} \left[ \sup_{t \leq \theta \leq T} |X_\theta^{t, Y_t, u}|^p \right] \leq C_1 (1 + \mathbb{E}[|Y_t|^p])$
2.  $\mathbb{E} \left[ \sup_{s \leq \theta \leq s+h} |X_\theta^{t, Y_t, u} - X_s^{t, Y_t, u}|^p \right] \leq C_2 h^{p/2} (1 + \mathbb{E}[|Y_t|^p])$
3.  $\mathbb{E} \left[ \sup_{t \leq \theta \leq T} |X_\theta^{t, \tilde{Y}_t, u} - X_\theta^{t, Y_t, u}|^p \right] \leq C_3 \mathbb{E}[|\tilde{Y}_t - Y_t|^p].$

**Definition 2.2.** We denote by  $\mathcal{R}$  the set of reachable probability distributions at time  $T$ , given by

$$\mathcal{R} = \{m_T^{0, Y_0, u} \mid u \in \mathcal{U}_0(Y_0)\}. \quad (8)$$

By Lemma 2.1, there exists  $R > 0$  such that

$$\mathcal{R} \subseteq \bar{B}_p(R). \quad (9)$$

By Lemma A.2 (in the appendix),  $\bar{B}_p(R)$  is compact for the  $d_1$ -distance, thus it is bounded. It follows that  $\mathcal{R}$  and  $\text{cl}(\mathcal{R})$  are bounded. We can therefore consider, for future reference, the diameter of  $\text{cl}(\mathcal{R})$ , defined by

$$D := \sup_{m_1, m_2 \in \text{cl}(\mathcal{R})} d_1(m_1, m_2). \quad (10)$$

**2.3. Formulation of the problem and regularity assumptions.** Consider two mappings  $F: \mathcal{P}_p(\mathbb{R}^n) \rightarrow \mathbb{R}$  and  $G: \mathcal{P}_p(\mathbb{R}^n) \rightarrow \mathbb{R}^N$ . We aim at studying the following problem:

$$\inf_{u \in \mathcal{U}_0(Y_0)} F(m_T^{0, Y_0, u}) \quad \text{subject to: } G(m_T^{0, Y_0, u}) \leq 0. \quad (P)$$

Throughout the article, we assume that on top of Assumption A1, Assumptions A2 and A3 below, dealing with the continuity and the differentiability of  $F$  and  $G$ , are satisfied. The constant  $R$  used in these assumptions is given by (9).

**Assumption A2.** *The restrictions of  $F$  and  $G$  to  $\bar{B}_p(R)$  are continuous for the  $d_1$ -distance.*

In order to state optimality conditions, we need a notion of derivative for the mappings  $F$  and  $G$ . Denoting by  $\mathcal{M}(\mathbb{R}^n)$  the set of finite signed measures on  $\mathbb{R}^n$ , we define:

$$\widehat{\mathcal{M}}_p(\mathbb{R}^n) = \left\{ m \in \mathcal{M}(\mathbb{R}^n) \mid \int_{\mathbb{R}^n} |x|^p d|m|(x) < +\infty, \int_{\mathbb{R}^n} 1 dm(x) = 0 \right\}.$$

Let  $A: \widehat{\mathcal{M}}_p(\mathbb{R}^n) \rightarrow \mathbb{R}^q$  be a linear mapping. We say that the function  $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^q$  is a representative of  $A$  if for all  $m \in \widehat{\mathcal{M}}_p(\mathbb{R}^n)$ , the integral  $\int_{\mathbb{R}^n} \phi(x) dm(x)$  is well-defined and equal to  $Am$ :

$$Am = \int_{\mathbb{R}^n} \phi(x) dm(x).$$

If  $\phi$  is a representative of  $A$ , then for any constant  $c \in \mathbb{R}^q$ ,  $\phi + c$  is also a representative of  $A$ , since

$$\int_{\mathbb{R}^n} c dm(x) = 0, \quad \forall m \in \widehat{\mathcal{M}}_p(\mathbb{R}^n).$$

Conversely, if the value of a representative  $\phi$  is fixed for a given point  $x_0 \in \mathbb{R}^n$ , then for all  $x \in \mathbb{R}^n$ , the value of  $\phi(x)$  is determined by

$$\phi(x) = \phi(x_0) + A(\delta_x - \delta_{x_0}),$$

where  $\delta_x$  and  $\delta_{x_0}$  are the Dirac measures centered at  $x$  and  $x_0$ , respectively. Therefore, the representative, if it exists, is uniquely defined up to a constant.

**Definition 2.3.** Let  $f: [0, 1] \times [0, 1] \rightarrow \mathbb{R}^\ell$ .

1. We say that  $f$  is Gâteaux-differentiable at 0 if there exists  $Df(0) \in \mathbb{R}^{\ell \times 2}$  such that

$$|f(z) - f(0) - Df(0)z| = o(|z|), \quad \text{as } z \rightarrow 0.$$

2. We say that  $f$  is strictly differentiable at 0 if there exists  $Df(0) \in \mathbb{R}^{\ell \times 2}$  such that

$$|f(z_2) - f(z_1) - Df(0)(z_2 - z_1)| = o(|z_2 - z_1|), \quad \text{as } z_1 \rightarrow 0 \text{ and } z_2 \rightarrow 0.$$

**Assumption A3.** 1. For all  $m, m_1$ , and  $m_2 \in \bar{B}_p(R)$ , there exists a linear form  $DF(m): \widehat{\mathcal{M}}_p(\mathbb{R}^n) \rightarrow \mathbb{R}$  such that the mapping

$$\theta \in [0, 1]^2 \mapsto F((1 - \theta_1 - \theta_2)m + \theta_1 m_1 + \theta_2 m_2)$$

is Gâteaux-differentiable at 0, with derivative  $(DF(m)(m_2 - m), DF(m)(m_1 - m))$ . Moreover,  $DF(m)$  possesses a continuous representative, dominated by  $|x|^p$ , and denoted  $x \in \mathbb{R}^n \mapsto DF(m, x) \in \mathbb{R}$ .

2. For all  $m, m_1$ , and  $m_2 \in \bar{B}_p(R)$ , there exists a linear mapping  $DG(m): \widehat{\mathcal{M}}_p(\mathbb{R}^n) \rightarrow \mathbb{R}^N$  such that the mapping

$$\theta \in [0, 1]^2 \mapsto G((1 - \theta_1 - \theta_2)m + \theta_1 m_1 + \theta_2 m_2)$$

is strictly differentiable at 0, with derivative  $(DG(m)(m_2 - m), DG(m)(m_1 - m))$ . Moreover,  $DG(m)$  possesses a continuous representative, dominated by  $|x|^p$ , and denoted  $x \in \mathbb{R}^n \mapsto DG(m, x) \in \mathbb{R}$ .

In the article, we make use of the derivative  $DF(m)$  (a linear form from  $\widehat{\mathcal{M}}_p(\mathbb{R}^n)$  to  $\mathbb{R}$ ) and its representative. The two notions can be distinguished according to the presence (or not) of the variable  $x$ .

**2.4. Discussion of the notion of derivative.** A general class of cost functions satisfying Assumptions A2 and A3 can be described as follows. Let  $K \in \mathbb{N}$ , let  $\Psi: \mathbb{R}^K \rightarrow \mathbb{R}^q$  be differentiable, let  $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^K$  be a continuous function dominated by  $|x|^p$ . We define then on  $\mathcal{P}_p(\mathbb{R}^n)$ :

$$H(m) = \Psi\left(\int_{\mathbb{R}^n} \phi(x) dm(x)\right). \quad (11)$$

Note that for all control processes  $u \in \mathcal{U}_0(Y_0)$ ,  $H(m_T^{0, Y_0, u}) = \Psi(\mathbb{E}[\phi(X_T^{0, Y_0, u})])$ . For all  $R \geq 0$ , the continuity of  $H$  on  $\bar{B}_p(R)$  follows from Lemma A.3 (in the appendix). One can easily check that the mapping  $H$  is differentiable in the sense of Assumption A3.1. The representative of its derivative is given by

$$DH(m, x) = D\Psi\left(\int_{\mathbb{R}^n} \phi(z) dm(z)\right)\phi(x), \quad (12)$$

up to a constant. Furthermore, if  $\Psi$  is continuously differentiable, then  $H$  is differentiable in the sense of Assumption A3.2. Note that the function  $\phi$  does not need to be differentiable. Let us mention that the variance of a probability distribution  $m$  can be described with a function of the form (11), taking  $\phi_1(x) = x^2$ ,  $\phi_2(x) = x$ , and  $\Psi(z_1, z_2) = z_1 - z_2^2$ . Further examples are discussed in detail in [25, Section 4]. We finish this subsection with two remarks.

**Remark 1.** The fact that  $F$  should be defined on the whole space  $\mathcal{P}_p(\mathbb{R}^n)$  discards cost functions whose formulation is based on the density of the probability measure (since a density does not always exist, for probability distributions in  $\mathcal{P}_p(\mathbb{R}^n)$ ). For example, the following problem does not fit in the proposed framework:

$$\inf_{u \in \mathcal{U}_0(Y_0)} \frac{1}{2} \|f - f_{\text{ref}}\|_{L^2(R^n)}, \quad \text{subject to: } f = \text{PDF}(X_T^{0, Y_0, u}),$$

where PDF stands for probability density function and where  $f_{\text{ref}}$  is a given probability density function.



**Remark 2.** The notion of derivative provided in [12, Section 6] and the one introduced in Assumption A3 are of different nature, because they aim at evaluating the variation of functions from  $\bar{B}_p(R)$  to  $\mathbb{R}$  on different kinds of paths. While our derivative is represented by a function from  $\mathbb{R}^n$  to  $\mathbb{R}$ , the one of [12] is represented by a function from  $\mathbb{R}^n$  to  $\mathbb{R}^n$  (see [12, Theorem 6.5]). This difference of nature can be better understood by considering the mapping:  $m \in \mathcal{P}_p(\mathbb{R}^n) \mapsto \int_{\mathbb{R}^n} \phi dm$ . This mapping is a monomial, according to the terminology given in [12, Example, page 43]. Its derivative (in the sense of [12]) is represented by the mapping:  $x \in \mathbb{R}^n \mapsto D\phi(x) \in \mathbb{R}^n$  (see [12, Example, page 44]). In the current framework, the derivative of  $m \in \mathcal{P}_p(\mathbb{R}^n) \mapsto \int_{\mathbb{R}^n} \phi dm$  is the real-valued mapping  $x \in \mathbb{R}^n \mapsto \phi(x) \in \mathbb{R}$ , up to a constant.

**2.5. Existence of a solution.** Observe that problem (P) can be equivalently formulated as follows:

$$\inf_{m \in \mathcal{R}} F(m), \quad \text{subject to: } G(m) \leq 0, \quad (P')$$

where  $\mathcal{R}$  is defined by (8). Indeed, if  $\bar{u} \in \mathcal{U}_0(Y_0)$  is a solution to (P), then  $\bar{m} := m_T^{0, Y_0, \bar{u}}$  is a solution to (P') and conversely, if  $\bar{m} \in \mathcal{R}$  is a solution to (P'), then any  $\bar{u} \in \mathcal{U}_0(Y_0)$  such that  $\bar{m} = m_T^{0, Y_0, \bar{u}}$  is a solution to (P). The feasible set  $\mathcal{R}_{\text{ad}}$  of Problem (P') is defined by

$$\mathcal{R}_{\text{ad}} = \{m \in \mathcal{R} \mid G(m) \leq 0\}. \quad (13)$$

By continuity of  $F$  for the  $d_1$ -distance, we have

$$\inf_{\substack{m \in \mathcal{R} \\ \text{s.t.: } G(m) \leq 0}} F(m) = \inf_{m \in \mathcal{R}_{\text{ad}}} F(m) = \inf_{m \in \text{cl}(\mathcal{R}_{\text{ad}})} F(m). \quad (P'')$$

The notation (P'') refers to the problem on the right. Lemma 2.1 enables us to prove the existence of a solution to problem (P'').

**Lemma 2.4.** *If  $\mathcal{R}_{\text{ad}}$  is non-empty, then problem (P'') has a solution.*

*Proof.* It is proved in Lemma A.2 that the set  $\bar{B}_p(R)$  is compact for the  $d_1$ -distance. By Lemma 2.1,  $\mathcal{R} \subseteq \bar{B}_p(R)$  and therefore,  $\mathcal{R}_{\text{ad}} \subseteq \bar{B}_p(R)$ . Since  $\bar{B}_p(R)$  is closed,  $\text{cl}(\mathcal{R}_{\text{ad}}) \subseteq \bar{B}_p(R)$  and therefore,  $\text{cl}(\mathcal{R}_{\text{ad}})$  is compact, since it is a closed set of a compact set. The existence of a solution follows, since  $F$  is continuous.  $\square$

Although we have an existence result for (P''), it is in general difficult to prove the existence of a solution to (P).

**3. Convexity of the reachable set.** This subsection is dedicated to the proof of the convexity of the closure of  $\mathcal{R}$  (the set of reachable probability distributions at time  $T$ ). This result is an important tool for the proof of the optimality conditions in Section 4 and for the numerical method developed in Section 5. Let us explain the underlying purpose with a simple example. Consider two processes  $u_1$  and  $u_2 \in \mathcal{U}_0(Y_0)$  and the corresponding final probability distributions  $m_T^{0, Y_0, u_1}$  and  $m_T^{0, Y_0, u_2}$ . We aim at building a control process  $u$  such that

$$m_T^{0, Y_0, u} = \frac{1}{2} (m_T^{0, Y_0, u_1} + m_T^{0, Y_0, u_2}). \quad (14)$$

A very simple way of building such a control process is to define a random variable  $S$  independent of  $Y_0$  and  $\mathcal{F}_{0, T}$  taking two different values with probability 1/2. A control  $u$  realizing (14) can then be constructed in  $\mathcal{U}_0((S, Y_0))$ : it suffices that

$u = u_1$  for one value of  $S$  and that  $u = u_2$  for the other. The obtained controlled process can be seen as a relaxed control, since it is now measurable with respect to a larger filtration. The main idea of Lemma 3.1 is to construct control processes in  $\mathcal{U}_0(Y_0)$  imitating the behaviour of the relaxed control process  $u$ .

**Lemma 3.1.** *The closure of the set of reachable probability measures for the  $d_1$ -distance, denoted  $\text{cl}(\mathcal{R})$ , is convex.*

*Proof.* Let us set  $\tilde{\mathcal{R}} = \{\frac{1}{2}m_1 + \frac{1}{2}m_2 \mid m_1 \in \mathcal{R}, m_2 \in \mathcal{R}\}$ . Our approach mainly consists in proving that

$$\tilde{\mathcal{R}} \subseteq \text{cl}(\mathcal{R}). \quad (15)$$

Let  $u^1$  and  $u^2$  in  $\mathcal{U}_0(Y_0)$ . To prove (15), it suffices to prove that there exists a sequence  $(u^\varepsilon)_{\varepsilon \geq 0}$  in  $\mathcal{U}_0(Y_0)$  such that

$$d_1\left(m_T^{0, Y_0, u^\varepsilon}, \frac{1}{2} \sum_{k=1}^2 m_T^{0, Y_0, u^k}\right) \xrightarrow{\varepsilon \rightarrow 0} 0. \quad (16)$$

Let  $0 < \varepsilon < T$ , let  $\tilde{u}^1$  and  $\tilde{u}^2$  be two processes in  $\mathcal{U}_\varepsilon(Y_0)$  such that for all  $k = 1$  and  $k = 2$ , the processes  $(u_s^k)_{s \in [0, T-\varepsilon]}$  and  $(\tilde{u}_s^k)_{s \in [\varepsilon, T]}$  can be seen as the same measurable function of respectively

$$(Y_0, (W_s - W_0)_{s \in [0, T-\varepsilon]}) \quad \text{and} \quad (Y_0, (W_s - W_\varepsilon)_{s \in [\varepsilon, T]}).$$

In other words, we simply delay the observation of the variation of the Brownian motion of a time  $\varepsilon$ . Let us denote by  $A_1$  the event  $(W_\varepsilon^1 - W_0^1) \in (-\infty, 0)$  and by  $A_2$  the event  $(W_\varepsilon^1 - W_0^1) \in (0, \infty)$ , where  $W^1$  is the first coordinate of the Brownian motion. Of course,  $\mathbb{P}[A_1] = \mathbb{P}[A_2] = \frac{1}{2}$ . Fixing  $u^0 \in U$ , we define  $u^\varepsilon \in \mathcal{U}_0(Y_0)$  as follows:

$$\begin{aligned} u_t^\varepsilon &= u^0, & \text{for a. e. } t \in (0, \varepsilon), \\ u_t^\varepsilon &= \tilde{u}_t^k, & \text{for a. e. } t \in (\varepsilon, T), \text{ when } A_k \text{ is realized.} \end{aligned}$$

For all  $\phi \in 1\text{-Lip}(\mathbb{R}^n)$ ,

$$\begin{aligned} \mathbb{E}[\phi(X_T^{0, Y_0, u^\varepsilon})] &= \frac{1}{2} \sum_{k=1}^2 \mathbb{E}\left[\phi\left(X_T^{\varepsilon, X_\varepsilon^{0, Y_0, u^\varepsilon}, u^\varepsilon}\right) \mid A_k\right] \\ &= \frac{1}{2} \sum_{k=1}^2 \mathbb{E}\left[\phi\left(X_T^{\varepsilon, X_\varepsilon^{0, Y_0, u^0}, \tilde{u}^k}\right) \mid A_k\right] = \frac{1}{2} \sum_{k=1}^2 \left[a_k + b_k + \mathbb{E}[\phi(X_T^{0, Y_0, u^k})]\right], \end{aligned} \quad (17)$$

where  $a_k$  and  $b_k$  are given by

$$\begin{aligned} a_k &= \mathbb{E}\left[\phi\left(X_T^{\varepsilon, X_\varepsilon^{0, Y_0, u^0}, \tilde{u}^k}\right) - \phi\left(X_T^{\varepsilon, Y_0, \tilde{u}^k}\right) \mid A_k\right], \\ b_k &= \mathbb{E}\left[\phi\left(X_T^{\varepsilon, Y_0, \tilde{u}^k}\right) \mid A_k\right] - \mathbb{E}[\phi(X_T^{0, Y_0, u^k})]. \end{aligned}$$

Let us first estimate  $a_k$ . Using the Lipschitz-continuity of  $\phi$ , we obtain that

$$\frac{1}{2}|a_k| \leq \mathbb{P}[A_k] \mathbb{E}\left[\left|X_T^{\varepsilon, X_\varepsilon^{0, Y_0, u^0}, \tilde{u}^k} - X_T^{\varepsilon, Y_0, \tilde{u}^k}\right| \mid A_k\right] \leq \mathbb{E}\left[\left|X_T^{\varepsilon, X_\varepsilon^{0, Y_0, u^0}, \tilde{u}^k} - X_T^{\varepsilon, Y_0, \tilde{u}^k}\right|\right].$$

We deduce from the Cauchy-Schwarz inequality and Lemma 2.1 that

$$\begin{aligned} \frac{1}{2}|a_k| &\leq \mathbb{E}\left[\left|X_T^{\varepsilon, X_\varepsilon^0, Y_0, u^0, \tilde{u}_k} - X_T^{\varepsilon, Y_0, \tilde{u}_k}\right|^2\right]^{1/2} \\ &\leq \sqrt{C_3}\mathbb{E}\left[|X_\varepsilon^0, Y_0, u^0 - Y_0|^2\right]^{1/2} \\ &\leq \sqrt{C_2 C_3 \varepsilon} (1 + \mathbb{E}[|Y_0|^2])^{1/2}. \end{aligned} \quad (18)$$

Let us estimate  $b_k$ . Since  $\tilde{u}_k$  and  $Y_0$  are independent of  $A_k$  and using the definition of  $\tilde{u}_k$ , we obtain that

$$\mathbb{E}[\phi(X_T^{\varepsilon, Y_0, \tilde{u}_k}) | A_k] = \mathbb{E}[\phi(X_T^{\varepsilon, Y_0, \tilde{u}_k})] = \mathbb{E}[\phi(X_{T-\varepsilon}^{0, Y_0, u^k})].$$

Therefore,

$$b_k = \mathbb{E}[\phi(X_{T-\varepsilon}^{0, Y_0, u^k}) - \phi(X_T^{0, Y_0, u^k})].$$

We obtain with the Lipschitz-continuity of  $\phi$ , the Cauchy-Schwarz inequality, and Lemma 2.1 that

$$\begin{aligned} |b_k| &\leq \mathbb{E}[|X_T^{0, Y_0, u^k} - X_{T-\varepsilon}^{0, Y_0, u^k}|] \\ &\leq \mathbb{E}[|X_T^{0, Y_0, u^k} - X_{T-\varepsilon}^{0, Y_0, u^k}|^2]^{1/2} \\ &\leq \sqrt{C_2 \varepsilon} (1 + \mathbb{E}[|Y_0|^2])^{1/2}. \end{aligned} \quad (19)$$

Combining (17), (18), and (19), we obtain that

$$\left| \mathbb{E}[\phi(X_T^{0, Y_0, u^\varepsilon})] - \frac{1}{2} \sum_{k=1}^2 \mathbb{E}[\phi(X_T^{0, Y_0, u^k})] \right| \leq \frac{1}{2} \sum_{k=1}^2 (|a_k| + |b_k|) \leq M\sqrt{\varepsilon},$$

where  $M$  is a constant independent of  $\phi$  and  $\varepsilon$ . Using the dual representation of  $d_1$  (given by (4)), we deduce that

$$\begin{aligned} d_1\left(m_T^{0, Y_0, u^\varepsilon}, \frac{1}{2} \sum_{k=1}^2 m_T^{0, Y_0, u^k}\right) \\ \leq \sup_{\phi \in 1\text{-Lip}(\mathbb{R}^n)} \left\{ \mathbb{E}[\phi(X_T^{0, Y_0, u^\varepsilon})] - \frac{1}{2} \sum_{k=1}^2 \mathbb{E}[\phi(X_T^{0, Y_0, u^k})] \right\} = M\sqrt{\varepsilon}. \end{aligned}$$

This proves (16) and thus justifies (15). We can now conclude the proof. It follows from (15) that

$$\text{cl}(\tilde{\mathcal{R}}) \subseteq \text{cl}(\mathcal{R}). \quad (20)$$

Since  $\mathcal{R} \subseteq \tilde{\mathcal{R}}$ , we have  $\text{cl}(\mathcal{R}) \subseteq \text{cl}(\tilde{\mathcal{R}})$ , and therefore by (20),  $\text{cl}(\mathcal{R}) = \text{cl}(\tilde{\mathcal{R}})$ . It remains to prove that  $\text{cl}(\tilde{\mathcal{R}})$  is convex, which is an easy task.  $\square$

On the feasibility of the problem. Proving the feasibility (or infeasibility) of the problem is in general a difficult task. Still, Lemma 3.1 can be used to prove the existence of  $m \in \text{cl}(\mathcal{R})$  such that  $G(m) < 0$  (this ensures the feasibility of the problem), or can be used to prove that for all  $m \in \text{cl}(\mathcal{R})$ ,  $G(m) > 0$  (in this case the problem is infeasible).

Given controls  $u_1, u_2, \dots, u_K \in \mathcal{U}_0(Y_0)$ , we obtain with Lemma 3.1 that the set  $\text{conv}(\{m_T^{0, Y_0, u_k}\}_{k=1, \dots, K})$  is included into  $\text{cl}(\mathcal{R})$ . Thus, if there exists  $\lambda \in \mathbb{R}^K$  such that  $\lambda \geq 0$ ,  $\sum_{k=1}^K \lambda_k = 1$ , and  $G(\sum_{k=1}^K \lambda_k m_T^{0, Y_0, u_k}) < 0$ , then the problem is feasible.

Let us consider now the situation where  $G$  is of the form  $G(m) = \Psi(\int \phi(x) dm(x))$  with  $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^K$  and  $\Psi: \mathbb{R}^K \rightarrow \mathbb{R}^N$ , as proposed in subsection 2.4. Let us set

$$\mathcal{Z} = \left\{ \int \phi(x) dm(x) \mid m \in \text{cl}(\mathcal{R}) \right\} \subseteq \mathbb{R}^K.$$

If there exists  $z \in \mathcal{Z}$  such that  $G(z) < 0$ , then the problem is feasible. If for all  $z \in \mathcal{Z}$ ,  $G(z) > 0$ , then the problem is infeasible. Thus, if we can describe in a convenient way the set  $\mathcal{Z}$ , then the issue of feasibility is reduced to a feasibility problem of finite dimension and is likely to be easier.

Regarding the description of  $\mathcal{Z}$ , a dual approach can be considered. The set  $\mathcal{Z}$  is indeed convex, by Lemma 3.1. It is moreover compact, since  $\text{cl}(\mathcal{R})$  is compact and the mapping  $\int \phi dm(x)$  continuous for the  $d_1$ -distance. Given  $\lambda \in \mathbb{R}^K$ , we define  $Z(\lambda) = \inf_{z \in \mathcal{Z}} \langle \lambda, z \rangle$ . We have

$$Z(\lambda) = \inf_{m \in \mathcal{R}} \int \langle \lambda, \phi(x) \rangle dm(x) = \inf_{u \in \mathcal{U}_0(Y_0)} \mathbb{E}[\langle \lambda, \phi(X_T^{0, Y_0, u}) \rangle],$$

thus the quantity  $Z(\lambda)$  can be computed numerically by solving a standard problem. By the Hahn-Banach theorem, we have the following relation:

$$\mathcal{Z} = \{z \in \mathbb{R}^K \mid \langle \lambda, z \rangle \geq Z(\lambda), \forall \lambda \text{ such that } |\lambda| = 1\}.$$

Thus an outer approximation of  $\mathcal{Z}$  can be computed by replacing the sphere of radius 1 (in  $\mathbb{R}^K$ ) by a discrete subset, in the right-hand side of the above relation.

In the situation of expectation constraints, that is  $G(m) = \Psi(\int \phi(x) dm(x))$  with  $\Psi(z) = z$ , the following problem can be considered:

$$\sup_{\lambda \in \mathbb{R}^K} Z(\lambda), \quad \text{subject to: } |\lambda| = 1, \lambda \geq 0. \quad (21)$$

If the value of problem (21) is negative, then Problem  $P$  is feasible, if the value of Problem (21) is positive, then Problem  $(P)$  is infeasible. For proving this result, one has to prove that (21) is the dual of:  $\inf_{z \in \mathcal{Z}} \sup_{\lambda \in \mathbb{R}^K} \langle \lambda, z \rangle$ , subject to  $|\lambda| = 1$ ,  $\lambda \geq 0$ , and one has to observe that for all  $z \in \mathbb{R}^K$ ,  $z < 0$  if and only if the following problem has a negative value:  $\sup_{\lambda \in \mathbb{R}^K} \langle \lambda, z \rangle$  subject to  $|\lambda| = 1$ ,  $\lambda \geq 0$ .

**4. Optimality conditions.** We prove in this section the main result: if a control  $\bar{u}$  is a solution to  $(P)$  and satisfies a qualification condition, then it is the solution to a standard problem of the form (3). Before proving our result, we recall in Subsection 4.1 some well-known properties of the value function associated with a standard problem.

**4.1. Standard problems.** Let  $\phi: \mathbb{R}^n \rightarrow \mathbb{R}$  be a continuous function dominated by  $|x|^p$ . Let us define:

$$\Phi: m \in \mathcal{P}_p(\mathbb{R}^n) \mapsto \int_{\mathbb{R}^n} \phi(x) dm(x).$$

The mapping  $\Phi$  is linear, in so far as for all  $m_1$  and  $m_2 \in \mathcal{P}_p(\mathbb{R}^n)$ , for all  $\theta \in [0, 1]$ ,

$$\Phi(\theta m_1 + (1 - \theta)m_2) = \theta \Phi(m_1) + (1 - \theta)\Phi(m_2).$$

It is also continuous for the  $d_1$ -distance, see Lemma A.3 (in the appendix). We denote by  $(P(\phi))$  the following standard problem:

$$\inf_{u \in \mathcal{U}_0(Y_0)} \mathbb{E}[\phi(X_T^{0, Y_0, u})]. \quad (P(\phi))$$

Let  $\hat{u} \in \mathcal{U}_0(Y_0)$  and let  $\hat{m} = m_T^{0, Y_0, \hat{u}}$ . By continuity of  $\Phi$ , the control process  $\hat{u}$  is a solution to  $(P(\phi))$  if and only if

$$\Phi(\hat{m}) = \inf_{m \in \text{cl}(\mathcal{R})} \Phi(m).$$

We recall, for future reference, some well-known results concerning the value function associated with the standard problem  $(P(\phi))$ . We refer to the textbooks [16], [28], [35] on this topic. The value function  $V$  associated with  $(P(\phi))$  is defined for all  $t \in [0, T]$  and for all  $x \in \mathbb{R}^n$  by

$$V(t, x) = \inf_{u \in U_t^0} \mathbb{E}[\phi(X_T^{t, x, u})].$$

It can be characterized as the unique viscosity solution to the following Hamilton-Jacobi-Bellman (HJB) equation:

$$-\partial_t V(t, x) = \inf_{u \in U} \left\{ H(x, u, \partial_x V(t, x), \partial_{xx}^2 V(t, x)) \right\}, \quad V(T, x) = \phi(x), \quad (22)$$

where the Hamiltonian  $H$  is defined for  $x \in \mathbb{R}^n$ ,  $u \in U$ ,  $p \in \mathbb{R}^n$ , and  $Q \in \mathbb{R}^{n \times n}$  by

$$H(x, u, p, Q) = \langle p, b(x, u) \rangle + \frac{1}{2} \text{tr}(\sigma(x, u) \sigma(x, u)^\top Q).$$

If  $V$  is sufficiently smooth, one can prove with a verification argument that any control process  $u$  is a global solution to  $(P(\phi))$  if almost surely,

$$u_t \in \arg \min_{v \in U} \left\{ H(X_t^{0, Y_0, u}, v, \partial_x V(t, X_t^{0, Y_0, u}), \partial_{xx}^2 V(t, X_t^{0, Y_0, u})) \right\}, \quad \text{for a.e. } t \in (0, T).$$

Finally, note that the value of problem  $(P(\phi))$  is given by

$$\inf_{u \in \mathcal{U}_0(Y_0)} \mathbb{E}[\phi(X_T^{0, Y_0, u})] = \int_{\mathbb{R}^n} V(0, x) dm(x),$$

where  $m$  denotes the probability distribution of  $Y_0$ .

**4.2. Main result.** In this subsection, we give first-order optimality conditions in variational form for problem  $(P)$  (defined in the introduction, page 6). Along this subsection, a solution  $\bar{u} \in \mathcal{U}_0(Y_0)$  to problem  $(P)$  is fixed. We also set

$$\bar{m} = m_T^{0, Y_0, \bar{u}}.$$

We first give a metric regularity result (Theorem 4.1), which is a key tool for the proof of the optimality conditions (Theorem 4.3).

Let us consider the sets  $A$  and  $I$  of active and inactive constraints at  $\bar{m}$ , defined by

$$A = \{j = 1, \dots, N \mid G_j(\bar{m}) = 0\} \quad \text{and} \quad I = \{j = 1, \dots, N \mid G_j(\bar{m}) < 0\}.$$

Let  $N_A$  be the cardinality of  $A$ . We define

$$G_A(m) = (G_j(m))_{j \in A} \in \mathbb{R}^{N_A} \quad \text{and} \quad G_I(m) = (G_j(m))_{j \in I} \in \mathbb{R}^{N - N_A}.$$

We have

$$\mathcal{R}_{\text{ad}} = \{m \in \mathcal{R} \mid G_A(m) \leq 0, G_I(m) \leq 0\}, \quad G_A(\bar{m}) = 0, \quad \text{and} \quad G_I(\bar{m}) < 0.$$

The following assumption is a qualification condition.

**Assumption A4.** *There exists  $m_0 \in \text{cl}(\mathcal{R})$  such that  $DG_A(\bar{m})(m_0 - \bar{m}) < 0$ .*

For all  $z \in \mathbb{R}^{N_A}$ , we denote by  $z_+$  the vector defined by  $(z_+)_j = \max(z_j, 0)$  for  $j \in A$ .

**Theorem 4.1.** *If Assumption A4 holds, then for all  $m \in \text{cl}(\mathcal{R})$ , there exist two constants  $\bar{\theta} \in (0, 1]$  and  $C > 0$  such that for all  $\theta \in [0, \bar{\theta}]$ , for all  $\varepsilon > 0$ , there exists  $\eta \in [0, 1]$  such that*

$$\eta \leq C |G_A(m_\theta)_+|_\infty + \varepsilon \quad (23)$$

and such that  $G((1 - \eta)m_\theta + \eta m_0) < 0$ , where  $m_\theta = (1 - \theta)\bar{m} + \theta m$  and where  $m_0$  is given by Assumption A4.

The estimate (23) is basically an estimate of the distance of  $m_\theta$  to  $\text{cl}(\mathcal{R}_{\text{ad}})$ . Indeed, by the convexity of  $\text{cl}(\mathcal{R})$ , the probability measure  $(1 - \eta)m_\theta + \eta m_0$  lies in  $\text{cl}(\mathcal{R})$ . Since  $G$  is continuous and since  $G((1 - \eta)m_\theta + \eta m_0) < 0$ , the probability measure  $(1 - \eta)m_\theta + \eta m_0$  lies in  $\text{cl}(\mathcal{R}_{\text{ad}})$ . It is at a distance  $\eta d_1(m_\theta, m_0)$  of  $m_\theta$ . The real number  $\eta \geq 0$  is of same order as the quantity  $|G_A(m_\theta)_+|_\infty$ , which indicates how much the constraints are violated.

*Proof of Theorem 4.1.* For all  $\theta \in [0, 1]$ , we define

$$\mathcal{G}_\theta: \eta \in [0, 1] \mapsto \mathcal{G}_\theta(\eta) = G_A((1 - \eta)m_\theta + \eta m_0) \in \mathbb{R}^{N_A}.$$

By Assumption A4,  $DG_A(\bar{m})(m_0 - \bar{m}) < 0$ . Let  $\alpha > 0$  be such that

$$DG_A(\bar{m})(m_0 - \bar{m}) \leq -\alpha \mathbf{1}, \quad (24)$$

where  $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^{N_A}$ . The above inequality (as well as all those involving vectors) must be understood coordinatewise.

*Claim 1.* There exist  $\bar{\eta} \in (0, 1]$  and  $\bar{\theta} \in (0, 1]$  such that for all  $\eta \in [0, \bar{\eta}]$  and for all  $\theta \in [0, \bar{\theta}]$ ,

$$\mathcal{G}_\theta(\eta) - \mathcal{G}_\theta(0) \leq -\frac{\alpha\eta}{2} \mathbf{1}. \quad (25)$$

Let us prove this claim. The reader can check that as a consequence of the differentiability assumption A3, the mapping  $(\theta, \eta) \in [0, 1]^2 \mapsto \mathcal{G}_\theta(\eta)$  is strictly differentiable at 0 (in the sense of Definition 2.3). Moreover, the partial derivative with respect to  $\eta$  is given by

$$D\mathcal{G}_0(0) = DG_A(\bar{m})(m_0 - \bar{m}).$$

As a consequence, there exist  $\bar{\theta}$  and  $\bar{\eta} \in [0, 1]$  such that for all  $\theta \in [0, \bar{\theta}]$  and for all  $\eta \in [0, \bar{\eta}]$ ,

$$|\mathcal{G}_\theta(\eta) - \mathcal{G}_\theta(0) - D\mathcal{G}_0(0)\eta| \leq \frac{\alpha}{2}\eta.$$

Combining the above inequality with (24), we obtain that

$$\mathcal{G}_\theta(\eta) - \mathcal{G}_\theta(0) \leq D\mathcal{G}_0(0)\eta + \frac{\alpha}{2}\eta \mathbf{1} \leq -\frac{\alpha}{2}\eta \mathbf{1},$$

which concludes the proof of the claim.

Now, let  $\eta_0 \in (0, 1]$  and  $\theta_0 \in (0, 1]$  be sufficiently small, so that for all  $\eta \in [0, \eta_0]$  and for all  $\theta \in [0, \theta_0]$ ,

$$G_I((1 - \eta)m_\theta + \eta m_0) < 0.$$

Let us set

$$\gamma = \frac{\alpha}{4} \min(\bar{\eta}, \eta_0). \quad (26)$$

Recall that  $\mathcal{G}_0(0) = G_A(\bar{m}) = 0$ . Once again, we reduce the value of  $\bar{\theta}$ , if necessary, so that  $\bar{\theta} \leq \theta_0$  and so that for all  $\theta \in [0, \bar{\theta}]$ ,

$$|\mathcal{G}_\theta(0)|_\infty \leq \gamma. \quad (27)$$

*Claim 2.* There exists  $C > 0$  such that for all  $y \in \mathbb{R}^{N_A}$  with  $|y|_\infty \leq \gamma$ , for all  $\theta \in [0, \bar{\theta}]$ , there exists  $\eta \in [0, \bar{\eta}]$  such that

$$\mathcal{G}_\theta(\eta) + y \leq 0 \quad \text{and} \quad \eta \leq C |(\mathcal{G}_\theta(0) + y)_+|_\infty.$$

Let us prove the claim. We set  $C = \frac{2}{\alpha}$ . For a given  $\theta \in [0, \bar{\theta}]$  and for a given  $y$  such that  $|y|_\infty \leq \gamma$ , we set  $\eta = C |(\mathcal{G}_\theta(0) + y)_+|_\infty$ . Using (26), (27), and the definition of  $C$ , we obtain that

$$|(\mathcal{G}_\theta(0) + y)_+|_\infty \leq |\mathcal{G}_\theta(0) + y|_\infty \leq |\mathcal{G}_\theta(0)|_\infty + |y|_\infty \leq 2\gamma \leq \frac{\alpha \bar{\eta}}{2} = \frac{\bar{\eta}}{C}.$$

Therefore,  $\eta = C |(\mathcal{G}_\theta(0) + y)_+|_\infty \leq \bar{\eta}$ . Using the first claim, we obtain that

$$\mathcal{G}_\theta(\eta) + y \leq \mathcal{G}_\theta(0) - \frac{\alpha \eta}{2} \mathbf{1} + y. \quad (28)$$

It directly follows from the definitions of  $C$  and  $\eta$  that

$$-\frac{\alpha \eta}{2} \mathbf{1} = -\frac{\eta}{C} \mathbf{1} \leq -|(\mathcal{G}_\theta(0) + y)_+|_\infty \mathbf{1} \leq -(\mathcal{G}_\theta(0) + y)_+.$$

Combined with (28), we obtain that

$$\mathcal{G}_\theta(\eta) + y \leq (\mathcal{G}_\theta(0) + y) - (\mathcal{G}_\theta(0) + y)_+ \leq 0.$$

This proves the second claim.

*Conclusion.* We can finally prove the theorem. Let  $\theta \in [0, \bar{\theta}]$  and let  $\varepsilon > 0$ . We set

$$y = \min\left(\gamma, \frac{\varepsilon}{C}\right) \mathbf{1}.$$

Since  $|y|_\infty \leq \gamma$ , there exists  $\eta \in [0, \bar{\eta}]$  such that  $\mathcal{G}_\theta(\eta) + y \leq 0$  and such that  $\eta \leq C |(\mathcal{G}_\theta(0) + y)_+|_\infty$ , by the second claim. Using the inequality  $||b|_\infty - |a|_\infty| \leq |b - a|_\infty$ , which is easy to establish, we obtain that

$$\eta \leq C |(\mathcal{G}_\theta(0) + y)_+|_\infty \leq C (|\mathcal{G}_\theta(0)|_\infty + |y|_\infty). \quad (29)$$

Since  $\mathcal{G}_\theta(0) = G_A(m_\theta)$  and  $|y|_\infty \leq \varepsilon/C$ , we obtain that

$$\eta \leq C |G_A(m_\theta)|_\infty + \varepsilon.$$

The estimate (23) is therefore satisfied. Since  $|\mathcal{G}_\theta(0)|_\infty \leq \gamma$  and  $|y|_\infty \leq \gamma$ , we deduce from (26) and from (29) that

$$\eta \leq 2C\gamma = \frac{4\gamma}{\alpha} \leq \eta_0. \quad (30)$$

Let us set  $\hat{m} = (1 - \eta)m + \eta m_0$ . By construction,  $\mathcal{G}_\theta(\eta) + y = G_A(\hat{m}) + y \leq 0$  and thus,  $G_A(\hat{m}) < 0$ , since  $y > 0$ . Moreover,  $G_I(\hat{m}) < 0$ , since  $\theta \in [0, \theta_0]$  and  $\eta \in [0, \eta_0]$ . Therefore,  $G(\hat{m}) < 0$ . The theorem is proved.  $\square$

In the following theorem, we prove first-order optimality conditions in variational form for problem (P).

**Definition 4.2.** The Lagrangian  $L$  is given by

$$L: (m, \lambda) \in \bar{B}_p(R) \times \mathbb{R}^N \mapsto F(m) + \langle \lambda, G(m) \rangle.$$

The Lagrangian is differentiable (with respect to  $m$ ) in the sense of Assumption A3.1 with

$$DL(m, \lambda) \tilde{m} = DF(m) \tilde{m} + \langle \lambda, DG(m) \tilde{m} \rangle, \quad \forall \tilde{m} \in \widehat{\mathcal{M}}_p(\mathbb{R}^n).$$

A representative of  $\tilde{m} \mapsto DL(m, \lambda) \tilde{m}$  is given by

$$DL(m, \lambda, x) = DF(m, x) + \langle \lambda, DG(m, x) \rangle,$$

up to a constant. Note that the mapping  $\tilde{m} \in \bar{B}_p(R) \mapsto DL(m, \lambda)\tilde{m}$  is continuous, by Lemma A.3.

In the sequel, we say that a non-negative Lagrange multiplier  $\lambda \in \mathbb{R}^N$  satisfies the complementarity condition at  $\bar{m}$  if

$$G_i(\bar{m}) < 0 \implies \lambda_i = 0, \quad \text{for all } i = 1, \dots, N. \quad (31)$$

**Theorem 4.3.** *Assume that  $\bar{u}$  is a solution to problem (P). Let  $\bar{m} = m_T^{0, Y_0, \bar{u}}$ . If Assumption A4 holds, then there exists a non-negative Lagrange multiplier  $\lambda \in \mathbb{R}^N$  satisfying the complementarity condition (31) at  $\bar{m}$  which is such that  $\bar{u}$  is a solution to the standard problem (P( $\phi$ )) with*

$$\phi(x) = DL(\bar{m}, \lambda, x).$$

**Remark 3.** We say then that the control process  $\bar{u}$  satisfies the optimality conditions in variational form. The optimality of  $\bar{u}$  for the standard problem with  $\phi(x) = DL(\bar{m}, \lambda, x)$  is equivalent to the following variational inequality:

$$\mathbb{E}[DL(\bar{m}, \lambda, X_T^{0, Y_0, u})] \geq \mathbb{E}[DL(\bar{m}, \lambda, X_T^{0, Y_0, \bar{u}})], \quad \forall u \in \mathcal{U}_0(Y_0)$$

and also equivalent to the following inequality:

$$DL(\bar{m}, \lambda)(m - \bar{m}) \geq 0, \quad \forall m \in \text{cl}(\mathcal{R}). \quad (32)$$

In the sequel, we say that a probability measure  $\bar{m} \in \text{cl}(\mathcal{R})$  such that  $G(\bar{m}) \leq 0$  satisfies the optimality conditions in variational form if there exists a multiplier  $\lambda \geq 0$  satisfying the complementarity condition (31) and such that (32) holds. We mention that for a solution  $\bar{m}$  to (P''), not necessarily lying in  $\mathcal{R}_{\text{ad}}$ , one can extend the proof given below and demonstrate that the optimality conditions in variational form are also satisfied.

*Proof of Theorem 4.3.* In view of the complementarity condition, it suffices to prove the existence of  $\lambda_A \in \mathbb{R}^{N_A}$  such that  $\lambda_A \geq 0$  and such that

$$\inf_{m \in \text{cl}(\mathcal{R})} D\tilde{L}(\bar{m}, \lambda_A)(m - \bar{m}) = 0,$$

where  $\tilde{L}(m, \lambda_A) = F(m) + \langle \lambda_A, G_A(m) \rangle$ . For all  $m \in \text{cl}(\mathcal{R})$ ,

$$D\tilde{L}(\bar{m}, \lambda_A)(m - \bar{m}) = DF(\bar{m})(m - \bar{m}) + \langle \lambda_A, DG_A(\bar{m})(m - \bar{m}) \rangle.$$

For all  $y \in \mathbb{R}^{N_A}$ , we consider the following optimization problem, denoted (LP(y)):

$$V(y) = \inf_{m \in \text{cl}(\mathcal{R})} DF(\bar{m})(m - \bar{m}), \quad \text{subject to: } DG_A(\bar{m})(m - \bar{m}) + y \leq 0. \quad (LP(y))$$

*Step 1.* We first prove that  $V(0) = 0$ . For  $y = 0$ ,  $\bar{m}$  is feasible (for problem (LP(y)) with  $y = 0$ ), thus  $V(0) \leq DF(\bar{m})(\bar{m} - \bar{m}) = 0$ . Now, let  $m \in \text{cl}(\mathcal{R})$  be such that  $DG_A(\bar{m})(m - \bar{m}) \leq 0$ . Let  $\bar{\theta} > 0$  and  $C > 0$  be given by Theorem 4.1. Let  $(\theta_k)_{k \in \mathbb{N}}$  be a convergent sequence with limit 0 taking values in  $(0, \bar{\theta}]$ . For all  $k$ , we set

$$m_k = (1 - \theta_k)\bar{m} + \theta_k m.$$

By Assumption A3, we have

$$G_A(m_k) = G_A(\bar{m}) + \theta_k DG_A(\bar{m})(m - \bar{m}) + o(\theta_k).$$

Since  $G_A(\bar{m}) = 0$  and  $DG_A(\bar{m})(m - \bar{m}) \leq 0$ , we have

$$|G_A(m_k)_+|_\infty = o(\theta_k).$$



By Theorem 4.1, there exists for all  $k \in \mathbb{N}$  a real number  $\eta_k \in [0, 1]$  such that

$$\eta_k \leq C|G_A(m_k)_+|_\infty + \theta_k^2 = o(\theta_k)$$

and such that  $G(\hat{m}_k) < 0$ , where

$$\begin{aligned} \hat{m}_k &= (1 - \eta_k)m_k + \eta_k m_0 \\ &= (1 - \eta_k)(1 - \theta_k)\bar{m} + (1 - \eta_k)\theta_k m + \eta_k m_0. \end{aligned}$$

Thus,

$$\hat{m}_k - \bar{m} = (1 - \eta_k)\theta_k(m - \bar{m}) + \eta_k(m_0 - \bar{m}). \quad (33)$$

Since  $\text{cl}(\mathcal{R})$  is convex (Lemma 3.1),  $\hat{m}_k \in \text{cl}(\mathcal{R})$ . Therefore, for all  $k$ , by continuity of  $F$  and  $G$ , there exists  $\tilde{m}_k \in \mathcal{R}$  such that  $G(\tilde{m}_k) \leq 0$  and such that  $F(\tilde{m}_k) - F(\hat{m}_k) \leq \theta_k^2$ . Using the differentiability assumption on  $F$  (Assumption A3), the feasibility of  $\tilde{m}_k$ , the fact that  $\eta_k = o(\theta_k)$  and (33), we obtain that

$$\begin{aligned} 0 &\leq \frac{F(\tilde{m}_k) - F(\bar{m})}{\theta_k} \\ &\leq \frac{F(\hat{m}_k) - F(\bar{m})}{\theta_k} + \frac{F(\tilde{m}_k) - F(\hat{m}_k)}{\theta_k} \\ &= (1 - \eta_k)DF(\bar{m})(m - \bar{m}) + \frac{\eta_k}{\theta_k}DF(\bar{m})(m_0 - \bar{m}) + o(1) \\ &= DF(\bar{m})(m - \bar{m}) + o(1). \end{aligned}$$

It follows that  $DF(\bar{m})(m - \bar{m}) \geq 0$  and finally proves that  $V(0) = 0$ .

*Step 2.* We compute now the Legendre-Fenchel transform (see [6, Relation 2.210] for a definition) of  $V$ . For all  $\lambda_A \in \mathbb{R}^{N_A}$ , we have

$$\begin{aligned} V^*(\lambda_A) &= \sup_{y \in \mathbb{R}^{N_A}} (\langle \lambda_A, y \rangle - V(y)) \\ &= \sup_{y \in \mathbb{R}^{N_A}} \left( \langle \lambda_A, y \rangle - \left( \inf_{m \in \text{cl}(\mathcal{R})} DF(\bar{m})(m - \bar{m}), \text{ s.t.: } DG_A(\bar{m})(m - \bar{m}) + y \leq 0 \right) \right) \\ &= \sup_{\substack{y \in \mathbb{R}^{N_A} \\ m \in \text{cl}(\mathcal{R})}} \left( \langle \lambda_A, y \rangle - DF(\bar{m})(m - \bar{m}), \text{ s.t.: } DG_A(\bar{m})(m - \bar{m}) + y \leq 0 \right). \end{aligned}$$

Using the change of variable  $z = DG_A(\bar{m})(m - \bar{m}) + y$ , we obtain:

$$\begin{aligned} V^*(\lambda_A) &= \sup_{\substack{z \in \mathbb{R}^{N_A} \\ m \in \text{cl}(\mathcal{R})}} \left( \langle \lambda_A, z - DG_A(\bar{m})(m - \bar{m}) \rangle - DF(\bar{m})(m - \bar{m}), \text{ s.t.: } z \leq 0 \right) \\ &= \sup_{m \in \text{cl}(\mathcal{R})} \left( \left( \sup_{z \in \mathbb{R}^{N_A}} \langle \lambda_A, z \rangle, \text{ s.t.: } z \leq 0 \right) - D\tilde{L}(\bar{m}, \lambda_A)(m - \bar{m}) \right). \end{aligned}$$

Observing that

$$\left( \sup_{z \in \mathbb{R}^{N_A}} \langle \lambda_A, z \rangle, \text{ subject to: } z \leq 0 \right) = \begin{cases} 0 & \text{if } \lambda_A \geq 0 \\ +\infty & \text{otherwise,} \end{cases}$$

we deduce that

$$V^*(\lambda_A) = \begin{cases} -\inf_{m \in \text{cl}(\mathcal{R})} D\tilde{L}(\bar{m}, \lambda_A)(m - \bar{m}) & \text{if } \lambda_A \geq 0 \\ +\infty & \text{otherwise.} \end{cases} \quad (34)$$

*Step 3.* Using the convexity of  $\text{cl}(\mathcal{R})$  (Lemma 3.1), one can easily show that  $V$  is a convex function. Let  $\alpha > 0$  be such that (24) holds. Then, for any  $y \in \mathbb{R}^{N_A}$  with  $|y|_\infty \leq \alpha$ ,

$$DG_A(\bar{m})(m_0 - \bar{m}) + y \leq 0$$

and therefore, problem  $(LP(y))$  is feasible and  $V(y) \leq DF(\bar{m})(m_0 - \bar{m})$ . It follows from [6, Proposition 2.018, Proposition 2.126] that  $V$  is continuous in the neighbourhood of 0 and has a non-empty subdifferential  $\partial V(0)$  at 0. Let  $\lambda_A \in \partial V(0)$ , by [6, Relation 2.232], we have

$$V^*(\lambda_A) = V(0) + V^*(\lambda) = \langle 0, \lambda \rangle = 0.$$

Thus, by (34),  $\lambda_A \geq 0$  and

$$\inf_{m \in \text{cl}(\mathcal{R})} D\tilde{L}(\bar{m}, \lambda_A)(m - \bar{m}) = 0.$$

The theorem is proved.  $\square$

The approach which has been employed to prove Theorem 4.3 is similar to the one used in [5] for proving Pontryagin's principle in a deterministic framework. In this reference, an auxiliary convex optimization problem (denoted by  $(PL)$ , introduced at the end of Section 3, and involving the final-state variable) is introduced. Pontryagin's principle is obtained by analyzing the dual of that problem, which plays a similar role as Problem  $(LP(y))$  in the proof of Theorem 4.3.

The following lemma shows that the value of the standard problem can be used to estimate the loss of optimality of a given probability measure  $\hat{m}$  in  $\mathcal{R}_{\text{ad}}$  (defined by (13)), when the mappings  $F, G_1, \dots, G_N$  are convex. The calculations involved in the proof are rather standard, see for example [20, Theorem 12.14].

**Lemma 4.4.** *Denote by  $\text{Val}(P)$  the value of Problem  $(P)$ . Assume that the mappings  $F, G_1, \dots, G_N$  are convex. Then, for all  $\hat{m} \in \mathcal{R}_{\text{ad}}$ , for all non-negative  $\lambda \in \mathbb{R}^N$  such that the complementarity condition holds at  $\hat{m}$ , the following upper estimate holds:*

$$F(\hat{m}) - \text{Val}(P) \leq - \inf_{m \in \text{cl}(\mathcal{R})} DL(\hat{m}, \lambda)(m - \hat{m}). \quad (35)$$

*Proof.* Let  $m \in \mathcal{R}_{\text{ad}}$ . Since  $F$  is convex, we have

$$F(m) - F(\hat{m}) \geq DF(\hat{m})(m - \hat{m}). \quad (36)$$

Denoting by  $A$  the active set at  $\hat{m}$  and setting

$$G_A(m) = (G_j(m))_{j \in A} \quad \text{and} \quad \lambda_A = (\lambda_j)_{j \in A},$$

we obtain, using the feasibility of  $m$  and the convexity of  $G_1, \dots, G_m$  that

$$0 \geq G_A(m) = G_A(m) - G_A(\hat{m}) \geq DG_A(\hat{m})(m - \hat{m}).$$

Since  $\lambda_A \geq 0$ , we deduce that

$$0 \geq \langle \lambda_A, DG_A(\hat{m})(m - \hat{m}) \rangle. \quad (37)$$

By the complementarity condition,

$$\langle \lambda_A, DG_A(\hat{m})(m - \hat{m}) \rangle = \langle \lambda, DG(\hat{m})(m - \hat{m}) \rangle. \quad (38)$$

Adding (36), (37), and (38) together, we obtain that

$$F(m) - F(\hat{m}) \geq DF(\hat{m})(m - \hat{m}) + \langle \lambda, DG(\hat{m})(m - \hat{m}) \rangle = DL(\hat{m}, \lambda)(m - \hat{m}).$$

Minimizing successively both sides with respect to  $m$ , we obtain that

$$\text{Val}(P) - F(\hat{m}) \geq \inf_{m \in \mathcal{R}_{\text{ad}}} DL(\hat{m}, \lambda)(m - \hat{m}).$$

Since  $\mathcal{R}_{\text{ad}} \subseteq \text{cl}(\mathcal{R})$ , we finally obtain that

$$\text{Val}(P) - F(\hat{m}) \geq \inf_{m \in \text{cl}(\mathcal{R})} DL(\hat{m}, \lambda)(m - \hat{m}),$$

which concludes the proof.  $\square$

As a corollary, we obtain that the optimality conditions in variational form are sufficient optimality conditions, in the convex case.

**Corollary 1.** *Assume that  $F, G_1, \dots, G_N$  are convex. Let  $\hat{u}$  be a feasible control process satisfying the optimality conditions in variational form. Then,  $\hat{u}$  is a solution to (P).*

*Proof.* In this situation, the right-hand side of inequality (35) is equal to 0, which directly proves the optimality of  $\hat{u}$ .  $\square$

We finish this section with a corollary dealing with stochastic optimal control problems with an expectation constraint.

**Corollary 2.** *Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $g: \mathbb{R}^n \rightarrow \mathbb{R}^N$  be two continuous functions, dominated by  $|x|^p$ . Assume that there exists  $u_0 \in \mathcal{U}_0(Y_0)$  such that  $\mathbb{E}[g(X_T^{0, Y_0, u_0})] < 0$ . Then, any  $u \in \mathcal{U}_0(Y_0)$  is a solution to the following problem:*

$$\inf_{u \in \mathcal{U}_0(Y_0)} \mathbb{E}[f(X_T^{0, Y_0, u})] \quad \text{subject to: } \mathbb{E}[g(X_T^{0, Y_0, u})] \leq 0 \quad (39)$$

*if and only if  $u$  is feasible and there exists  $\lambda \geq 0$  such that  $\mathbb{E}[g_i(X_T^{0, Y_0, u})] < 0 \implies \lambda_i = 0$ , for all  $i = 1, \dots, N$ , and such that  $u$  is a solution to (P( $\phi$ )) with*

$$\phi(x) = f(x) + \langle \lambda, g(x) \rangle.$$

*Proof.* Setting  $F(m) = \int_{\mathbb{R}^n} f(x) dm(x)$  and  $G(m) = \int_{\mathbb{R}^n} g(x) dm(x)$ , we obtain that problem (39) falls into the general class of problems studied in the article. The functions  $F$  and  $G$  satisfy the required regularity assumptions (see Subsection 2.4). Note that  $DF(m, x) = f(x)$  and that  $DG(m, x) = g(x)$  (up to a constant). The existence of  $u_0$  ensures that the qualification condition is satisfied. The mappings  $F$  and  $G$  are clearly convex, therefore, the optimality conditions in variational form are necessary and sufficient, by Theorem 4.3 and Corollary 1.  $\square$

## 5. Numerical method and results.

**5.1. Augmented Lagrangian Method.** We provide in this section a numerical method for solving problem (P) and give results for two academical problems. The method is an augmented Lagrangian method combined with a projected-gradient-type algorithm.

Let us begin with a rough description of the method, consisting of Algorithms 1 and 2 (page 21). The second algorithm is a building block of the first one. The augmented Lagrangian method is used to solve the following problem:

$$\inf_{m \in \text{cl}(\mathcal{R}), s \in \mathbb{R}^N} F(m), \quad \text{subject to: } G(m) + s = 0, \quad s \geq 0. \quad (40)$$

At the end of the while loop of Algorithm 1 (line 19), the method provides a probability measure  $m_k \in \text{cl}(\mathcal{R})$  satisfying approximately the optimality conditions in variational form for some Lagrange multiplier  $\lambda_k$ . At this stage, the method

has not computed a control process  $\bar{u}$  such that  $m_k = m_T^{0, Y_0, \bar{u}}$ . The ultimate step of the algorithm (line 20) aims at recovering such a control  $\bar{u}$  by solving the standard problem  $(P(\phi))$  with  $\phi = DL(m_k, \lambda_k)$ . One has to check a posteriori that  $\bar{u}$  approximately satisfies the optimality conditions in variational form with associated Lagrange multiplier  $\bar{\lambda} = \lambda_k$ .

Let us go into the details of the method. The augmented Lagrangian  $L_A$  associated with (40) is given by

$$L_A: (m, s, \lambda, c) \in \text{cl}(\mathcal{R}) \times \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}_+ \mapsto F(m) + \langle \lambda, G(m) + s \rangle + \frac{c}{2} |G(m) + s|^2.$$

The employed norm in the above definition is the Euclidean norm. Note that the constraints  $s \geq 0$  and  $m \in \text{cl}(\mathcal{R})$  are not dualized, since they will be ensured by the projected gradient method. The mapping  $L_A(\cdot, s, \lambda, c)$  is differentiable (in the sense of Assumption A3.1) with respect to  $m$ , with

$$\begin{aligned} DL_A(m, s, \lambda, c)\hat{m} &= DF(m)\hat{m} + \langle \lambda, DG(m)\hat{m} \rangle + \langle c(G(m) + s), DG(m)\hat{m} \rangle \\ &= DL(m, \lambda + c(G(m) + s))\hat{m}, \quad \forall \hat{m} \in \widehat{\mathcal{M}}_p(\mathbb{R}^n). \end{aligned}$$

A representative of  $\hat{m} \mapsto DL_A(m, s, \lambda, c)\hat{m}$  is therefore given by

$$\begin{aligned} DL_A(m, s, \lambda, c, x) &= DF(m) + \langle \lambda + c(G(m) + s), DG(m, x) \rangle \\ &= DL(m, \lambda + c(G(m) + s), x), \end{aligned} \quad (41)$$

up to a constant. The partial gradient of  $L_A$  with respect to  $s$  is given by

$$\nabla_s L_A(m, s, \lambda, c) = \lambda + c(G(m) + s).$$

Let us first focus on Algorithm 2. It aims at solving the following problem:

$$\inf_{m \in \text{cl}(\mathcal{R}), s \geq 0} L_A(m, s, \lambda, c), \quad (42)$$

for values of the Lagrange multiplier  $\lambda$  and the penalty parameter  $c > 0$  given as input variables. The algorithm constructs a sequence  $(m_\ell)_{\ell=0,1,\dots}$  of probability distributions in  $\text{cl}(\mathcal{R})$  and a sequence  $(s_\ell)_{\ell=0,1,\dots}$  of non-negative slack variables in the while loop. At each iteration  $\ell$  of the while loop, a kind of line-search (line 10) is performed: the mapping  $L_A(\cdot, \cdot, \lambda, c)$  is minimized over the set

$$\{(m_\ell(\theta), s_\ell(\theta)) \mid \theta \in [0, 1]\}, \quad (43)$$

where

$$m_\ell(\theta) = (1 - \theta)m_\ell + \theta\tilde{m}_\ell, \quad s_\ell(\theta) = \max(s_\ell + \theta\delta s_\ell, 0)$$

and where  $\delta s_\ell = -\nabla_s L_A(m_\ell, s_\ell, \lambda, c)$ . The max operator in the above expression must be understood coordinatewise, it is nothing but a projection of  $s_\ell + \theta\delta s_\ell$  on  $\mathbb{R}_{\geq 0}^N$ . The probability distribution  $\tilde{m}_\ell$  is chosen as a solution to

$$\inf_{m \in \text{cl}(\mathcal{R})} DL_A(m_\ell, s_\ell, \lambda, c)(m - m_\ell).$$

The value of the above problem is non-positive. The measure  $\tilde{m}_\ell - m_\ell$  can therefore be seen as a descent direction for the variable  $m$ . The next iterate of the algorithm is given by

$$m_{\ell+1} = (1 - \theta_\ell)m_\ell + \theta_\ell\tilde{m}_\ell \quad \text{and} \quad s_{\ell+1} = \max(s_\ell + \theta_\ell\delta s_\ell, 0),$$

where  $\theta_\ell$  minimizes  $L_A(m_\ell(\theta), s_\ell(\theta), \lambda, c)$  over  $[0, 1]$ . Note that Lemma 3.1 plays here a crucial role: it guarantees that  $m_{\ell+1} \in \text{cl}(\mathcal{R})$ . Let us note that Algorithm 2 is not, strictly speaking, a projected gradient method, since  $m_{\ell+1}$  is not obtained as a projection on  $\text{cl}(\mathcal{R})$ .

Let us discuss the criterion used in Algorithm 2. It is related to the optimality conditions associated with problem (42): if  $(\bar{m}, \bar{s})$  is a solution to (42), then, by Theorem 4.3,

$$\inf_{m \in \text{cl}(\mathcal{R})} DL_A(\bar{m}, \bar{s}, \lambda, c)(m - \bar{m}) = 0, \quad (44)$$

moreover, the optimality of  $\bar{s}$  implies that

$$\nabla_s L_A(\bar{m}, \bar{s}, \lambda, c) \geq 0 \quad \text{and} \quad \left[ s_i > 0 \implies \nabla_{s_i} L_A(\bar{m}, \bar{s}, \lambda, c) = 0, \forall i = 1, \dots, N \right]. \quad (45)$$

Algorithm 2 stops when the variable  $\varepsilon_\ell$  (defined line 6 in the algorithm) is smaller than  $\omega$ , i.e. when both

$$- \inf_{m \in \text{cl}(\mathcal{R})} DL_A(m_\ell, s_\ell, \lambda, c)(m - m_\ell) \leq \omega, \quad (46)$$

$$|s_\ell - \max(s_\ell + \delta s_\ell, 0)|_\infty \leq \omega. \quad (47)$$

The inequality (46) ensures that the optimality condition (44) is approximately satisfied (note that the left-hand side of (46) is always non-negative). The inequality (47) implies that

$$s_\ell - \max(s_\ell + \delta s_\ell, 0) \geq -\omega \mathbf{1},$$

thus that

$$s_\ell + \delta s_\ell \leq \max(s_\ell + \delta s_\ell, 0) \leq s_\ell + \omega \mathbf{1}$$

and finally that

$$\nabla_s L_A(m_\ell, s_\ell, \lambda, c) = -\delta s_\ell \geq -\omega \mathbf{1}. \quad (48)$$

Moreover, for all  $i = 1, \dots, N$ , if  $s_{\ell,i} > \omega$ , then  $\max(s_{\ell,i} + \delta s_{\ell,i}, 0) > 0$  (by (47)) and therefore,

$$s_{\ell,i} + \delta s_{\ell,i} = \max(s_{\ell,i} + \delta s_{\ell,i}, 0)$$

and finally,

$$\omega \geq |s_{\ell,i} - \max(s_{\ell,i} + \delta s_{\ell,i}, 0)| = |\delta s_{\ell,i}| = |\nabla_{s_i} L_A(m_\ell, s_\ell, \lambda, c)|. \quad (49)$$

Inequalities (48) and (49) therefore ensure that the optimality condition (45) is approximately satisfied.

Let us come back to Algorithm 1. It constructs a sequence  $(m_k)_{k=0,1,\dots}$  of probability distributions in  $\text{cl}(\mathcal{R})$ , a sequence of non-negative slack variables  $(s_k)_{k=0,1,\dots}$  and a sequence of Lagrange multipliers  $(\lambda_k)_{k=0,1,\dots}$ . Two sequences of tolerances are also constructed,  $(\eta_k)_{k=0,1,\dots}$  and  $(\omega_k)_{k=0,1,\dots}$ , as well as a sequence of penalty parameters  $(c_k)_{k=0,1,\dots}$ . At the iteration  $k$ , the augmented Lagrangian is minimized by using Algorithm 2 with  $\lambda = \lambda_k$  and the current tolerance  $\omega_k$ . The triplet  $(m_{k+1}, s_{k+1}, \varepsilon_{k+1})$  is the output of Algorithm 2. Three cases are then considered.

- If  $|G(m_{k+1}) + s_{k+1}| \leq \eta_k$ , then we consider that the penalty term  $c_k$  is large enough. The Lagrange multiplier is updated as follows:

$$\lambda_{k+1} = \nabla_s L_A(m_{k+1}, s_{k+1}, \lambda_k, c_k) = \lambda_k + c_k(G(m_k) + s_k).$$

This update rule is motivated by (41).

- If moreover  $|G(m_{k+1}) + s_{k+1}| \leq \eta_*$  and  $\varepsilon_{k+1} \leq \omega_*$ , then the algorithm stops and we have that

$$\begin{aligned} & \inf_{m \in \text{cl}(\mathcal{R})} DL_A(m_{k+1}, s_{k+1}, \lambda_k, c_k)(m - m_{k+1}) \\ &= \inf_{m \in \text{cl}(\mathcal{R})} DL(m_{k+1}, \lambda_{k+1})(m - m_{k+1}) \geq -\omega_*. \end{aligned}$$

Moreover, by (48),  $\lambda_{k+1} \geq -\omega_* \mathbf{1}$  and for all  $i = 1, \dots, N$ , if  $G_i(m_{k+1}) > \eta_* + \omega_*$ , then

$$s_{k+1,i} \geq |G_i(m_{k+1})| - |G_i(m_{k+1}) + s_{k+1,i}| \geq \omega_*$$

and therefore  $|\lambda_{k+1,i}| \leq \omega_*$  by (49). The probability distribution  $m_{k+1}$  satisfies approximately the optimality conditions in variational form, with associated Lagrange multiplier  $\lambda_{k+1}$ .

- Otherwise, the penalty term is unchanged and the tolerances  $\eta_k$  and  $\omega_k$  are tightened (line 12).
- If  $|G(m_{k+1}) + s_{k+1}| > \omega_k$ , then the penalty term  $c_k$  is regarded as too weak, it is therefore increased. The estimate of the Lagrange multiplier  $\lambda$  is unchanged and the tolerances are re-initialized (line 16).

**Remark 4.** In practice, the main difficulty in the method is the resolution of the standard problem. It consists of two phases: in a backward phase, the Hamilton-Jacobi-Bellman equation associated with the standard problem must be solved (see subsection 4.1). It provides an optimal control (for the standard problem) in a feedback form. One must then compute the probability distribution  $(m_t)_{t \in [0, T]}$  which is associated, in a forward phase.

**Remark 5.** Algorithm 1 is taken from [20, Algorithm 17.4] (see also [14]). Note that the update rules for  $\eta_k$  and  $\omega_k$  have been modified, in order to avoid too strong variations of the parameters  $\eta_k$  and  $\omega_k$ .

---

**Algorithm 1:** Augmented Lagrangian method for solving Problem (P)

---

```

1 Input:  $m_{\text{init}} \in \text{cl}(\mathcal{R})$ ,  $s_{\text{init}} \in \mathbb{R}_{\geq 0}^N$ ,  $\lambda_{\text{init}} \in \mathbb{R}^N$ ,  $\eta_* > 0$ ,  $\omega_* > 0$ ;
2 Set  $k = 0$ ,  $c_0 = 10$ ,  $\omega_0 = 1/c_0$ ,  $\eta_0 = 1/c_0^{0.1}$ ;
3 Set  $m_0 = m_{\text{init}}$ ,  $\lambda_0 = \lambda_{\text{init}}$ ,  $s_0 = s_{\text{init}}$ ;
4 Set T = true;
5 while T = true do
6   Compute  $(m_{k+1}, s_{k+1}, \varepsilon_{k+1})$  as an output of Algorithm 2 with input
    $(m_k, s_k, \lambda_k, c_k, \omega_k)$ ;
7   if  $|G(m_{k+1}) + s_{k+1}| \leq \eta_k$  then
8     Set  $\lambda_{k+1} = \lambda_k + c_k(G(m_{k+1}) + s_{k+1})$ ;
9     if  $|G(m_{k+1}) + s_{k+1}| \leq \eta_*$  and  $\varepsilon_{k+1} \leq \omega_*$  then
10      | Set T = false;
11    else
12      | Set  $c_{k+1} = c_k$ ,  $\eta_{k+1} = \eta_k/10^{0.1}$ ,  $\omega_{k+1} = \omega_k/10$ ;
13    end
14  else
15    | Set  $\lambda_{k+1} = \lambda_k$ ;
16    | Set  $c_{k+1} = 10c_k$ ,  $\eta_{k+1} = 1/c_{k+1}^{0.1}$ ,  $\omega_{k+1} = 1/c_{k+1}$ ;
17  end
18  Set  $k = k + 1$ ;
19 end
20 Compute a solution  $\bar{u}$  to  $(P(\phi))$  with  $\phi(x) = DL(m_k, \lambda_k, x)$ ;
21 Output:  $\bar{u}$ ,  $\bar{\lambda} = \lambda_k$ .

```

---

---

**Algorithm 2:** Projected gradient method for minimizing  $L_A(\cdot, \cdot, \lambda, c)$  on  $\text{cl}(\mathcal{R}) \times \mathbb{R}_{\geq 0}^N$

---

```

1 Input:  $m_{\text{init}} \in \text{cl}(\mathcal{R})$ ,  $s_{\text{init}} \in \mathbb{R}_{\geq 0}^N$ ,  $\lambda \in \mathbb{R}^N$ ,  $c > 0$ ,  $\omega > 0$ ;
2 Set  $\ell = 0$ ,  $m_0 = m_{\text{init}}$ ,  $s_0 = s_{\text{init}}$ ,  $\mathbf{T} = \text{true}$ ;
3 while  $\mathbf{T} = \text{true}$  do
4   Compute a solution  $\tilde{m}_\ell$  to:  $\inf_{m \in \text{cl}(\mathcal{R})} DL_A(m_\ell, s_\ell, \lambda, c)(m - m_\ell)$ ;
5   Set  $\delta s_\ell = -(\lambda + c(G(m_\ell) + s_\ell))$ ;
6   Set  $\varepsilon_\ell = \max(DL_A(m_\ell, s_\ell, \lambda, c)(m_\ell - \tilde{m}_\ell), |s_\ell - \max(s_\ell + \delta s_\ell, 0)|_\infty)$ ;
7   if  $\varepsilon_\ell \leq \omega$  then
8     | Set  $\mathbf{T} = \text{false}$ ;
9   else
10    Compute a solution  $\theta_\ell$  to:
11    |  $\inf_{\theta \in [0,1]} L_A((1 - \theta)m_\ell + \theta\tilde{m}_\ell, \max(s_\ell + \theta\delta s_\ell, 0), \lambda, c)$ ;
12    | Set  $m_{\ell+1} = (1 - \theta_\ell)m_\ell + \theta_\ell\tilde{m}_\ell$ ,  $s_{\ell+1} = \max(s_\ell + \theta\delta s_\ell, 0)$ ;
13    | Set  $\ell = \ell + 1$ ;
14  end
15 Output:  $m_\ell, s_\ell, \varepsilon_\ell$ .
```

---

**5.2. Convergence analysis.** We investigate in this subsection the termination of Algorithms 1 and 2. Our analysis follows the main lines of [20, Chapters 3 and 17]. Let us mention that we do not tackle in this subsection the issues related to discretization. In general, termination proofs for line-search methods require that the function to be minimized is differentiable with a Lipschitz-continuous gradient. A similar assumption is therefore considered below.

**Assumption A5.** *The mappings  $F$  and  $G$  are differentiable in the sense of Assumption A3. Moreover, there exist two constants  $K_1 > 0$  and  $K_2 > 0$  such that for all  $m_1, m_2, m_3$ , and  $m_4 \in \bar{B}_p(R)$ ,*

$$\begin{cases} |(DF(m_2) - DF(m_1))(m_4 - m_3)| \leq K_1 d_1(m_1, m_2) d_1(m_3, m_4), \\ |(DG(m_2) - DG(m_1))(m_4 - m_3)| \leq K_1 d_1(m_1, m_2) d_1(m_3, m_4), \end{cases} \quad (50)$$

and such that

$$\begin{aligned} |DF(m_1)(m_4 - m_3)| &\leq K_2 d_1(m_3, m_4), \\ |DG(m_1)(m_4 - m_3)| &\leq K_2 d_1(m_3, m_4). \end{aligned}$$

**Remark 6.** It can be easily proved that under Assumption A5,  $F$  and  $G$  are Lipschitz continuous for the distance  $d_1$ , with modulus  $K_2$ .

Before starting the convergence analysis, we give an example of a mapping satisfying Assumption A5.

**Lemma 5.1.** *If  $F$  and  $G$  are of the form  $m \in \bar{B}_p(R) \mapsto \Psi(\int \phi dm)$  where  $\Psi$  is differentiable with a Lipschitz continuous derivative on bounded sets, and where  $\phi$  is globally Lipschitz continuous, then Assumption A5 holds.*

*Proof.* Recall the expression of the derivative, given in this situation by (12). Let  $K_a$  be the Lipschitz modulus of  $\phi$ . For  $m \in \bar{B}_p(R)$ , by Hölder's inequality,

$$\begin{aligned} \left| \int_{\mathbb{R}^n} \phi(x) dm(x) \right| &\leq \int_{\mathbb{R}^n} |\phi(0)| + K_a|x| dm(x) \\ &\leq |\phi(0)| + K_a \int_{\mathbb{R}^n} |x|^p dm(x) \\ &\leq |\phi(0)| + K_a R =: K_b. \end{aligned}$$

Let  $K_c$  be the Lipschitz modulus of  $D\Psi$  on the ball of centre 0 and radius  $K_b$ . Let  $K_d$  be a bound of  $|D\Psi|$  on the same ball. Using the dual representation of the Wasserstein distance  $d_1$ , we obtain that

$$\left| \int_{\mathbb{R}^n} D\Psi(\int \phi dm_1) \phi(x) d(m_3(x) - m_4(x)) \right| \leq K_a K_d d_1(m_3, m_4),$$

since  $x \in \mathbb{R}^n \mapsto D\Psi(\int \phi dm_1) \phi(x)$  is Lipschitz continuous with modulus  $K_a K_d$ . Similarly, we also obtain that

$$\begin{aligned} \left| \int_{\mathbb{R}^n} (D\Psi(\int \phi dm_2) - D\Psi(\int \phi dm_1)) \phi(x) d(m_3(x) - m_4(x)) \right| \\ \leq K_a^2 K_c d_1(m_1, m_2) d_1(m_3, m_4), \end{aligned}$$

since

$$|D\Psi(\int \phi dm_2) - D\Psi(\int \phi dm_1)| \leq K_a K_c d_1(m_1, m_2).$$

Thus, Assumption A5 holds with  $K_1 = K_a^2 K_c$  and  $K_2 = K_a K_d$ .  $\square$

The following lemma provides some useful properties dealing with the Lipschitz-continuity of the derivatives of the augmented Lagrangian.

**Lemma 5.2.** *Under Assumption A5, for all  $\lambda \in \mathbb{R}^N$ , for all  $c > 0$ , for all bounded sets  $\mathcal{S}$ , there exist three constants  $K_3$ ,  $K_4$  and  $K_5 > 0$  such that for all  $m_1$  and  $m_2 \in \text{cl}(\mathcal{R})$ , for all  $s_1$  and  $s_2 \in \mathcal{S}$ ,*

- $|\nabla_s L_A(m_1, s_1, \lambda, c)| \leq K_3$ ,
- $|\nabla_s L_A(m_2, s_2, \lambda, c) - \nabla_s L_A(m_1, s_1, \lambda, c)| \leq K_4(d_1(m_1, m_2) + |s_2 - s_1|)$ ,
- $|(DL_A(m_2, s_2, \lambda, c) - DL_A(m_1, s_1, \lambda, c))(m_4 - m_3)| \\ \leq K_5(d_1(m_1, m_2) + |s_2 - s_1|)d_1(m_3, m_4)$ .

*Proof.* It is clear that  $|\nabla_s L_A(m_1, s_1, \lambda, c)| = |\lambda + c(G(m_1) + s_1)|$  is bounded, since  $G$  is Lipschitz continuous and since  $\text{cl}(\mathcal{R})$  and  $\mathcal{S}$  are bounded. The first inequality follows.

We obtain with the Lipschitz-continuity of  $G$  that

$$\begin{aligned} |\nabla_s L_A(m_2, s_2, \lambda, c) - \nabla_s L_A(m_1, s_1, \lambda, c)| &= c|(G(m_2) - G(m_1)) + (s_2 - s_1)| \\ &\leq c(K_2 d_1(m_1, m_2) + |s_2 - s_1|), \end{aligned}$$

which proves the second inequality.

For proving the third inequality, we focus on the Lipschitz continuity of the mapping  $(s, m) \mapsto (G(m) + s)^\top DG(m)$  (the other terms involved in  $DL_A(\cdot, \cdot, \lambda, c)$  can be easily treated). Let  $K$  and  $S > 0$  be such that for all  $m_1 \in \text{cl}(\mathcal{R})$ , for all



$s \in \mathcal{S}$ ,  $|G(m_1)| \leq K$  and  $|s_1| \leq S$ . We have

$$\begin{aligned}
& |((G(m_2) + s_2)^\top DG(m_2) - (G(m_1) + s_1)^\top DG(m_1))(m_4 - m_3)| \\
& \leq |(G(m_2) - G(m_1) + s_2 - s_1)^\top DG(m_2)(m_4 - m_3)| \\
& \quad + |(G(m_1) + s_1)^\top (DG(m_2) - DG(m_1))(m_4 - m_3)| \\
& \leq (K_2 d_1(m_1, m_2) + |s_2 - s_1|) K_2 d_1(m_3, m_4) \\
& \quad + (K + S) K_1 d_1(m_1, m_2) d_1(m_3, m_4) \\
& \leq \max(K_2^2 + (K + S) K_1, K_2) (d_1(m_1, m_2) + |s_2 - s_1|) d_1(m_3, m_4).
\end{aligned}$$

The third inequality follows.  $\square$

**Proposition 1.** *Under Assumption A5, Algorithm 2 terminates.*

*Proof.* We do a proof by contradiction and therefore assume that the algorithm never terminates. Therefore, it generates a sequence  $(m_\ell, s_\ell, \varepsilon_\ell)_{\ell \in \mathbb{N}}$  which is such that  $\varepsilon_\ell > \omega$ , for all  $\ell \in \mathbb{N}$ . One can easily prove that the following set is bounded:

$$\mathcal{S} := \{s \in \mathbb{R}^N \mid \exists m \in \text{cl}(\mathcal{R}), L_A(m, s, \lambda, c) \leq L_A(m_0, s_0, \lambda, c)\},$$

since  $G$  is bounded on  $\text{cl}(\mathcal{R})$  and since for a fixed  $m \in \text{cl}(\mathcal{R})$ ,  $s \in \mathbb{R}^N \mapsto L_A(m, s, \lambda, c)$  is linear-quadratic, with a dominant term  $c|s|^2$  independent of  $m$ . In a similar way, one can prove that  $L_A(\cdot, \cdot, \lambda, c)$  is bounded from below. By construction, the sequence  $(L_A(m_\ell, s_\ell, \lambda, c))_{\ell \in \mathbb{N}}$  is decreasing, therefore, for all  $\ell \in \mathbb{N}$ ,  $s_\ell \in \mathcal{S}$ . Let  $K_3$ ,  $K_4$ , and  $K_5$  be the three constants given by Lemma 5.2, for the set  $\mathcal{S}$ .

The proof mainly consists in finding an upper estimate of the decay

$$L_A(m_{\ell+1}, s_{\ell+1}, \lambda, c) - L_A(m_\ell, s_\ell, \lambda, c),$$

at a given iteration  $\ell$ . This is achieved with estimate (53) below. Let us introduce some notation, used only in this proof. For  $\theta \in [0, 1]$ , we denote

$$m(\theta) = (1 - \theta)m_\ell + \theta \tilde{m}_\ell \quad \text{and} \quad s(\theta) = \max(s_\ell + \theta \delta s_\ell, 0).$$

We also omit the arguments  $\lambda$  and  $c$  of the augmented Lagrangian (since they are fixed). Let  $\theta \in [0, 1]$ . Observe first that by Lemma A.1,

$$d_1(m(0), m(\theta)) \leq \theta d_1(m_\ell, \tilde{m}_\ell) \leq \theta D, \tag{51}$$

where  $D$  is the diameter of  $\text{cl}(\mathcal{R})$  (defined by (10)). Let us estimate  $|s(\theta) - s(0)|$ . By Lemma 5.2,  $|\delta s_\ell| = |\nabla_s L_A(m_\ell, s_\ell, \lambda, c)| \leq K_3$ . Since  $s_\ell \geq 0$  and since the mapping  $s \in \mathbb{R}^N \mapsto \max(s, 0)$  is Lipschitz-continuous with modulus 1 (it is a projection), we have

$$|s(\theta) - s(0)| = |\max(s_\ell + \theta \delta s_\ell, 0) - \max(s_\ell, 0)| \leq |s_\ell + \theta \delta s_\ell - s_\ell| = \theta |\delta s_\ell| \leq \theta K_3. \tag{52}$$

Now, we split the decay into two terms as follows:

$$\begin{aligned}
& L_A(m(\theta), s(\theta)) - L_A(m(0), s(0)) \\
& = \underbrace{L_A(m(\theta), s(\theta)) - L_A(m(\theta), s(0))}_{(a)} + \underbrace{L_A(m(\theta), s(0)) - L_A(m(0), s(0))}_{(b)}.
\end{aligned}$$

We split the first term as follows:

$$(a) = \underbrace{\int_0^1 \langle \nabla_s L_A(m(\theta), s(0) + \xi(s(\theta) - s(0))) - \nabla_s L_A(m(0), s(0)), s(\theta) - s(0) \rangle d\xi}_{(a_1)} + \underbrace{\langle \nabla_s L_A(m(0), s(0)), s(\theta) - s(0) \rangle}_{(a_2)}.$$

Combining Lemma 5.2 with estimates (51) and (52), we obtain that

$$(a_1) \leq K_4(D\theta + K_3\theta)K_3\theta = K_3K_4(D + K_3)\theta^2.$$

Since  $s(\theta)$  is the orthogonal projection of  $s_\ell + \theta\delta s_\ell$  on  $\mathbb{R}_{\geq 0}^N$  and since  $s_\ell = s(0) \in \mathbb{R}_{\geq 0}^N$ , we have

$$\langle s(0) - s(\theta), s(0) + \theta\delta s_\ell - s(\theta) \rangle \leq 0.$$

Using  $\delta s_\ell = -\nabla_s L_A(m_\ell, s_\ell, \lambda, c)$ , we deduce that

$$\theta \langle \nabla_s L_A(m_\ell, s_\ell, \lambda, c), s(\theta) - s_0 \rangle \leq -|s(\theta) - s(0)|^2.$$

It is proved in [10, Lemma 2.2] that

$$|s(\theta) - s(0)| \geq \theta|s(1) - s(0)|.$$

Combining the last two estimates, we obtain that

$$(a_2) = \langle \nabla_s L_A(m_\ell, s_\ell, \lambda, c), s(\theta) - s(0) \rangle \leq -\theta|s(1) - s(0)|^2 \\ = -\theta|s_\ell - \max(s_\ell + \delta s_\ell, 0)|^2.$$

Let us estimate (b). We have

$$(b) = \int_0^\theta (DL_A(m(\xi), s(0)) - DL_A(m(0), s(0)))(\tilde{m}_\ell - m_\ell) d\xi \\ + \theta DL_A(m(0), s(0))(\tilde{m}_\ell - m_\ell) \\ \leq \int_0^\theta K_5 D\xi D d\xi + \theta DL_A(m_\ell, s_\ell)(\tilde{m}_\ell - m_\ell) \\ = \frac{1}{2}K_5 D^2 \theta^2 + \theta DL_A(m_\ell, s_\ell)(\tilde{m}_\ell - m_\ell).$$

For all  $\ell \in \mathbb{N}$ , we denote

$$\tilde{\varepsilon}_\ell = DL_A(m_\ell, s_\ell, \lambda, c)(m_\ell - \tilde{m}_\ell) + |s_\ell - \max(s_\ell + \delta s_\ell, 0)|^2 \geq 0.$$

Combining the three obtained upper estimates of (a<sub>1</sub>), (a<sub>2</sub>), and (b), we obtain that there exists a constant  $K$ , independent of  $\ell$ , such that for all  $\theta \in [0, 1]$ ,

$$L_A(m(\theta), s(\theta)) - L_A(m(0), s(0)) \leq \frac{1}{2}K\theta^2 - \theta\tilde{\varepsilon}_\ell. \quad (53)$$

For all  $\ell \in \mathbb{N}$ , we define

$$\tilde{\theta}_\ell = \min\left(1, \frac{\tilde{\varepsilon}_\ell}{K}\right).$$

If  $\tilde{\theta}_\ell = \frac{\tilde{\varepsilon}_\ell}{K}$ , then

$$L_A(m(\tilde{\theta}_\ell), s(\tilde{\theta}_\ell)) - L_A(m(0), s(0)) \leq -\frac{1}{2K}\tilde{\varepsilon}_\ell^2.$$

Otherwise,  $\tilde{\theta}_\ell = 1$  and  $K \leq \tilde{\varepsilon}_\ell$ . Therefore

$$L_A(m(\tilde{\theta}_\ell), s(\tilde{\theta}_\ell)) - L_A(m(0), s(0)) \leq \frac{1}{2}K - \tilde{\varepsilon}_\ell \leq -\frac{1}{2}\tilde{\varepsilon}_\ell.$$

Therefore, for all  $\ell \in \mathbb{N}$ ,

$$L_A(m(\tilde{\theta}_\ell), s(\tilde{\theta}_\ell)) - L_A(m(0), s(0)) \leq -\frac{1}{2} \min\left(\frac{1}{K} \tilde{\varepsilon}_\ell^2, \tilde{\varepsilon}_\ell\right)$$

and thus

$$\begin{aligned} L_A(m_{\ell+1}, s_{\ell+1}) - L_A(m_\ell, s_\ell) &\leq L_A(m(\tilde{\theta}_\ell), s(\tilde{\theta}_\ell)) - L_A(m(0), s(0)) \\ &\leq -\frac{1}{2} \min\left(\frac{1}{K} \tilde{\varepsilon}_\ell^2, \tilde{\varepsilon}_\ell\right). \end{aligned}$$

Recall that the sequence  $(L_A(m_\ell, s_\ell))_{\ell \in \mathbb{N}}$  is bounded from below. Let  $\bar{L}_A$  be a lower bound. We deduce from the above estimate that for all  $q \in \mathbb{N}$ ,

$$\sum_{\ell=0}^q \min\left(\frac{1}{K} \tilde{\varepsilon}_\ell^2, \tilde{\varepsilon}_\ell\right) \leq 2(L_A(m_0, s_0) - L_A(m_{q+1}, s_{q+1})) \leq 2(L_A(m_0, s_0) - \bar{L}_A).$$

The sequence  $(\min(K \tilde{\varepsilon}_\ell^2, \tilde{\varepsilon}_\ell))_{\ell \in \mathbb{N}}$  is therefore summable and thus converges to 0. It follows that  $(\tilde{\varepsilon}_\ell)_{\ell \in \mathbb{N}}$  converges to 0. Since  $\tilde{\varepsilon}_\ell$  is the sum of two non-negative terms, they both converge to 0, i.e.

$$DL_A(m_\ell, s_\ell, \lambda, c)(m_\ell - \tilde{m}_\ell) \xrightarrow{\ell \rightarrow \infty} 0 \quad \text{and} \quad |s_\ell - \max(s_\ell + \delta s_\ell, 0)| \xrightarrow{\ell \rightarrow \infty} 0.$$

It follows that  $\varepsilon_\ell \xrightarrow{\ell \rightarrow \infty} 0$ , which is a contradiction.  $\square$

**Proposition 2.** *Under Assumption A5, if Algorithm 1 does not terminate, then any limit point  $(\bar{m}, \bar{s})$  of  $(m_k, s_k)_{k \in \mathbb{N}}$  — there exists at least one — is such that*

$$\inf_{m \in \text{cl}(\mathcal{R})} \langle G(\bar{m}) + \bar{s}, DG(\bar{m})(m - \bar{m}) \rangle = 0 \quad (54)$$

and such that

$$G(\bar{m}) + \bar{s} \geq 0 \quad \text{and} \quad \left[ \bar{s}_i > 0 \implies G_i(\bar{m}) + \bar{s}_i = 0, \quad \forall i = 1, \dots, N \right]. \quad (55)$$

Observe that the two conditions satisfied by  $(\bar{m}, \bar{s})$  are the optimality conditions for the problem

$$\inf_{m \in \text{cl}(\mathcal{R}), s \in \mathbb{R}^N} |G(m) + s|^2, \quad \text{subject to: } s \geq 0.$$

*Proof of Proposition 2.* Let us assume that the algorithm does not terminate. Let us first prove that there are infinitely many indices  $k$  such that  $|G(m_{k+1}) + s_{k+1}| > \eta_k$ . Suppose that it is not the case, then there exists  $k_0$  such that for all  $k \geq k_0$ ,  $|G(m_{k+1}) + s_{k+1}| \leq \eta_k$ . Considering the update formulas for  $\eta_k$  and  $\omega_k$  used in this situation (line 12), we obtain that  $\eta_k \rightarrow 0$  and  $\omega_k \rightarrow 0$  and thus for some  $k \geq k_0$  sufficiently large,  $\eta_k \leq \eta_*$  and  $\omega_k \leq \omega_*$ . The algorithm necessarily terminates when these two inequalities hold, which is a contradiction.

When  $|G(m_{k+1}) + s_{k+1}| \leq \eta_k$ ,  $c_k$  is unchanged and when  $|G(m_{k+1}) + s_{k+1}| > \eta_k$  (which occurs infinitely many times),  $c_{k+1} = 10c_k$ . Therefore,  $c_k \rightarrow \infty$ . We also have that for all  $k \in \mathbb{N}$ ,  $c_k \geq 1$ . It is easy to prove by induction that for all  $k \in \mathbb{N}$ ,  $\eta_k \leq 1/c_k^{0.1}$  and that  $\omega_k \leq 1/c_{k+1}$ . Therefore,  $\eta_k \rightarrow 0$  and  $\omega_k \rightarrow 0$ . For  $k$  large enough, say for  $k \geq k_1$ ,  $\eta_k \leq \eta_*$  and  $\omega_k \leq \omega_*$ . Therefore, for  $k \geq k_1$ ,  $|G(m_{k+1}) + s_{k+1}| > \eta_k$  (otherwise, the algorithm would terminate). It follows that for  $k \geq k_1$ , the Lagrange multiplier is not updated anymore:  $\lambda_k = \lambda_{k_1}$ . We denote this constant value of the Lagrange multiplier by  $\lambda$ , for simplicity.

We now prove that the sequence  $(s_k)_{k \in \mathbb{N}}$  is bounded. Let

$$K = \sup_{m \in \text{cl}(\mathcal{R})} |F(m)| < \infty.$$

The following inequalities hold true:

$$-K - |\lambda| \cdot |G(m) + s| + \frac{c}{2} |G(m) + s|^2 \leq L_A(m, s, \lambda, c), \quad (56)$$

$$L_A(m, s, \lambda, c) \leq K + |\lambda| \cdot |G(m) + s| + \frac{c}{2} |G(m) + s|^2. \quad (57)$$

The value of the augmented Lagrangian is decreasing along the iterations of Algorithm 2. Moreover, the pair  $(m_{k+1}, s_{k+1})$  is obtained as an output of Algorithm 2, with initial value  $(m_k, s_k)$ . Therefore,

$$L_A(m_{k+1}, s_{k+1}, \lambda, c_k) \leq L_A(m_k, s_k, \lambda, c_k), \quad \forall k \geq k_1.$$

Using (56) and denoting  $y_k = G(m_k) + s_k$ , we obtain that

$$-K - |\lambda| \cdot |y_{k+1}| + \frac{c_k}{2} |y_{k+1}|^2 \leq K + |\lambda| \cdot |y_k| + \frac{c_k}{2} |y_k|^2.$$

Dividing by  $c_k/2$  and adding  $|\lambda|^2/c_k^2$  on both sides and factorizing, we obtain that

$$-\frac{2K}{c_k} + \left( |y_{k+1}| - \frac{|\lambda|}{c_k} \right)^2 \leq \frac{2K}{c_k} + \left( |y_k| + \frac{|\lambda|}{c_k} \right)^2.$$

Using the inequality  $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ , we obtain

$$|y_{k+1}| - \frac{|\lambda|}{c_k} \leq |y_k| + \frac{|\lambda|}{c_k} + \frac{2\sqrt{K}}{\sqrt{c_k}}$$

and finally, by induction, for all  $q \geq k_1$ ,

$$|y_q| \leq |y_{k_1}| + \sum_{k=k_1}^{q-1} \left( \frac{2|\lambda|}{c_k} + \frac{2\sqrt{K}}{\sqrt{c_k}} \right). \quad (58)$$

Since the sequence  $(c_k)_{k \geq k_1}$  is a geometric sequence of ratio 10, the sequences  $(1/c_k)_{k \geq k_1}$  and  $(1/\sqrt{c_k})_{k \geq k_1}$  are also geometric with ratios 1/10 and 1/ $\sqrt{10}$ , respectively. These last two sequences are therefore summable and we deduce from (58) that  $(y_k)_{k \in \mathbb{N}}$  is bounded. Since  $(G(m_k))_{k \in \mathbb{N}}$  is bounded, we finally obtain that  $(s_k)_{k \in \mathbb{N}}$  is bounded.

Since  $\text{cl}(\mathcal{R})$  is compact and  $(s_k)_{k \in \mathbb{N}}$  bounded, the sequence  $(m_k, s_k)_{k \in \mathbb{N}}$  possesses at least one accumulation point. Let  $(\bar{m}, \bar{s}) \in \text{cl}(\mathcal{R}) \times \mathbb{R}_{\geq 0}^N$  be an accumulation point. To simplify, we assume that the whole sequence  $(m_k, s_k)_{k \in \mathbb{N}}$  converges to  $(\bar{m}, \bar{s})$ . The arguments which follow can be easily adapted if only a subsequence converge to  $(\bar{m}, \bar{s})$ . Let  $m \in \text{cl}(\mathcal{R})$ . For all  $k \in \mathbb{N}$ , we have

$$DF(m_k) + \langle \lambda + c_{k-1}(G(m_k) + s_k), DG(m_k)(m - m_k) \rangle \geq -\omega_{k-1}.$$

Dividing by  $c_{k-1}$  and passing to the limit, we obtain that

$$\langle G(\bar{m}) + \bar{s}, DG(\bar{m})(m - \bar{m}) \rangle \geq 0.$$

Minimizing the left-hand side with respect to  $m$ , we obtain (54). It remains to prove (55). For all  $k \geq k_1$ , we have

$$|s_k - \max(s_k - (\lambda + c_{k-1}(G(m_k) + s_k)), 0)|_\infty \leq \omega_{k-1}. \quad (59)$$

Dividing by  $c_{k-1}$  and passing to the limit, we obtain that

$$\max(-(G(\bar{m}) + \bar{s}), 0) = 0,$$

which proves that  $G(\bar{m}) + \bar{s} \geq 0$ . Let  $i \in \{1, \dots, N\}$  be such that  $\bar{s}_i > 0$ . For  $k$  large enough,  $\omega_{k-1} < s_{k,i}$  and therefore, as a consequence of (59),

$$\max(s_{k,i} - (\lambda_i + c_{k-1}(G_i(m_k) + s_{k,i})), 0) > 0,$$

meaning that

$$s_{k,i} - (\lambda_i + c_{k-1}(G_i(m_k) + s_{k,i})) > 0.$$

Dividing by  $c_{k-1}$  and passing to the limit, we obtain that  $G_i(\bar{m}) + \bar{s}_i \leq 0$  and therefore, we obtain that  $G_i(\bar{m}) + \bar{s}_i = 0$ , since  $G(\bar{m}) + \bar{s} \geq 0$ , which proves that (55) holds.  $\square$

**5.3. Results.** We present numerical results for two academical problems. The considered SDE is the following for the two of them:

$$dX_t = u_t dt + dW_t, \quad X_0 = 0, \quad U = [-2, 2]. \quad (60)$$

One can only act on the drift, the volatility is constant and equal to 1. All standard problems are solved by dynamic programming. The corresponding HJB equation is discretized with a semi-Lagrangian scheme (see [11]), which consists in approximating the SDE by a controlled Markov chain, defined at times  $\{0, \delta t, \dots, T\}$  with  $\delta t = 10^{-2}$  and taking values in  $\{-5, -5 + \delta x, \dots, 5 - \delta x, 5\}$ , with  $\delta x = 10^{-3}$  and using reflecting boundary conditions. At any mesh point, the minimization of the Hamiltonian is realized by enumeration, for a discretized set of controls  $\{-2, -2 + \delta u, \dots, 2\}$  with  $\delta u = 10^{-1}$ . As mentioned in Remark 4, the resolution of standard problems (as the one in Algorithm 2, line 4) is done in two phases. Once an optimal control  $u$  has been found by dynamic programming, the corresponding probability distribution is obtained by solving the Chapman-Kolmogorov equation associated with the discretized Markov chain.

The minimization with respect to  $\theta$  involved in the computation of a steplength (line 10, Algorithm 2) is done by enumeration. The considered discretized set is  $\{0, \delta\theta, 2\delta\theta, \dots, 1\}$  with  $\delta\theta = 10^{-6}$ . Note this step of the method is computationally inexpensive (at least for the considered test cases).

The Algorithm 1 is initialized with  $m_{\text{init}} = \delta_0$  (the Dirac distribution centered at 0),  $s_{\text{init}} = 0$ , and  $\lambda_{\text{init}} = 0$ .

Since the SDE (60) is linear with respect to  $u$ , the Hamiltonian  $H$  is itself linear with respect to  $u$  and one can expect that optimal controls only take the boundary values  $-2$  and  $2$  when the derivative (w.r.t.  $x$ ) of the solution to the HJB equation is positive (resp. negative). The optimal controls obtained below indeed take these values for most of the mesh points. This is why we worked with a rather coarse discretization of  $U$ .

Test case 1: bounded variance. For the first test case, we consider the following cost function and constraint:

$$F(m) = \int_{\mathbb{R}} x dm(x) \quad \text{and} \quad G(m) = \int_{\mathbb{R}} x^2 dm(x) - \left( \int_{\mathbb{R}} x dm(x) \right)^2 - \alpha, \quad (61)$$

where  $\alpha = 0.4$ . Observe that  $F(m_T^{0, Y_0, u}) = \mathbb{E}[X_T^{0, Y_0, u}]$ : the cost function is the expectation of the final state. The mapping  $G(m_T^{0, Y_0, u}) + \alpha$  is the variance of  $X_T^{0, Y_0, u}$ . The cost function  $F$  is linear and  $G$  is of the form (11). The derivative of

$G$  at any  $m$ , obtained with (12), is a linear-quadratic function (with respect to  $x$ ), given by

$$DG(m, x) = x^2 - 2 \left[ \int_{\mathbb{R}} y dm(y) \right] x.$$

Convergence results are shown on Figure 1. The tolerances  $\eta_*$  and  $\omega_*$  are chosen equal, for values ranging from  $10^{-3}$  to  $10^{-6}$ . For each of these tolerances, the values of  $G(\bar{u})$  and  $\bar{\lambda}$  are provided. The column ‘‘Var. Ineq.’’ contains the following value:

$$- \inf_{m \in \text{cl}(\mathcal{R})} DL(m_T^{0, Y_0, \bar{u}}, \bar{\lambda})(m - m_T^{0, Y_0, \bar{u}}),$$

which somehow indicates to what extent the variational inequality is satisfied. The column  $c$  shows the value of the penalty parameter at the last iteration. The last column shows the total number of standard problems which have to be solved.

It can be first observed that for tolerances below  $10^{-4}$ , the variational inequality is almost satisfied. The violation of the constraint is small and of the same order as the tolerances. The obtained Lagrange multipliers converge when the tolerance goes to 0. We also observe that the mechanism of Algorithm 1 avoids that the penalty term  $c$  becomes very high, for small tolerances.

| Tolerance | $G(\bar{u})$ | $\bar{\lambda}$ | Var. Ineq. | $c$  | Iterations |
|-----------|--------------|-----------------|------------|------|------------|
| 1e-3      | 3.72 e-3     | 1.285           | 1.92 e-5   | 100  | 29         |
| 1e-4      | 7.54 e-4     | 1.318           | 1.77 e-15  | 100  | 39         |
| 1e-5      | 1.87 e-5     | 1.324           | 2.66 e-15  | 1000 | 60         |
| 1e-6      | 1.87 e-5     | 1.324           | 2.66 e-15  | 1000 | 60         |

FIGURE 1. Convergence results for the Test Case 1

The optimal control generated by the algorithm is shown on Figure 2a (page 30), the associated probability measure (at any time  $t$ ) is shown on Figure 2b. The value function associated with the standard problem with cost function  $DL(m_T^{0, Y_0, \bar{u}}, \bar{\lambda}, \cdot)$  is represented on Figure 2c, we recall that it plays the role of an adjoint equation.

As expected, the optimal control has a kind of bang-bang structure. It is constant with respect to time, equal to  $-2$  for  $x \geq -1.6$  and to  $2$  for  $x \leq -1.6$ . If the same problem was solved without constraint, the optimal control would be equal to  $-2$ , in order to minimize the expectation of the final state. Here, the optimal control must be equal to  $2$  when  $x$  is smaller than  $-1.6$  in order to keep the variance sufficiently small and to satisfy the constraint.

Test case 2: expectation constraint. For this second test case, we consider the following cost function and constraint:

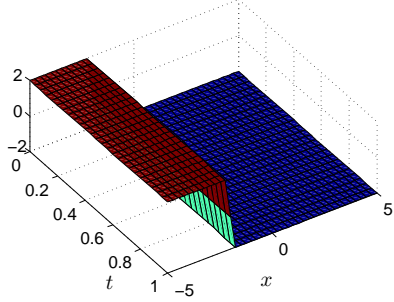
$$F(m) = \int_{\mathbb{R}} x dm(x) \quad \text{and} \quad G(m) = - \int_{\mathbb{R}} e^{-x^2} dm(x) + \alpha, \quad (62)$$

with  $\alpha = 0, 4$ . Note that both  $F$  and  $G$  are linear. Roughly speaking, the constraint  $G(m) \leq 0$  ensures that a proportion  $\alpha$  of the final probability measure remains around 0.

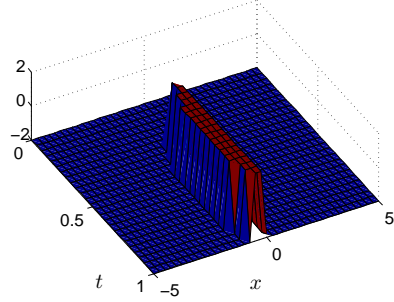
Convergence results are given in Figure 3. The value of  $G(\bar{u})$  converges to 0, suggesting that the constraint is active for the undiscretized problem. Convergence of the Lagrange multiplier is observed. The variational inequality is exactly satisfied, since the derivatives of  $F$  and  $G$  do not depend on  $m$ . The value of the penalty parameter does not increase much.

Test case 1 (problem (61))

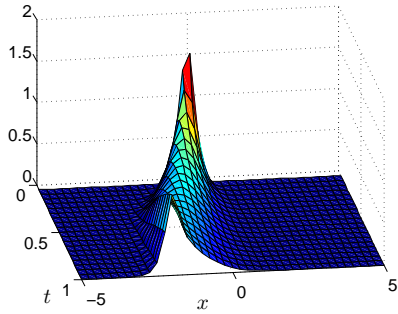
Test case 2 (problem (62))



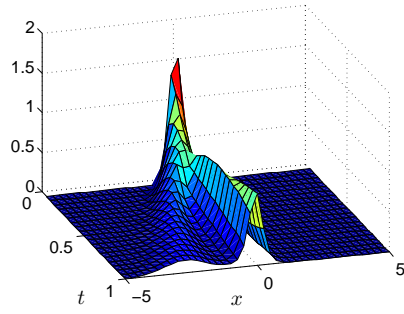
(A) Control



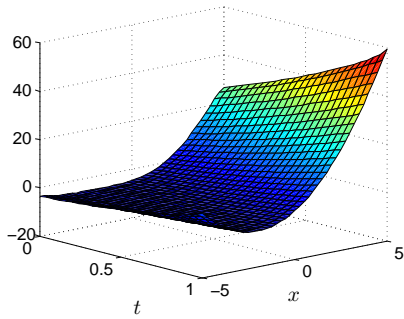
(D) Control



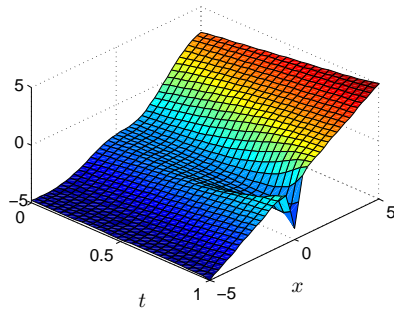
(B) Probability distribution



(E) Probability distribution



(C) Adjoint equation



(F) Adjoint equation

FIGURE 2. Numerical results

| Tolerance | $G(\bar{u})$ | $\bar{\lambda}$ | Var. Ineq. | $c$ | Iterations |
|-----------|--------------|-----------------|------------|-----|------------|
| 1e-3      | -1.93 e-2    | 4.119           | 0          | 100 | 37         |
| 1e-4      | 1.19 e-3     | 4.024           | 0          | 100 | 53         |
| 1e-5      | -8.22 e-5    | 4.026           | 0          | 100 | 64         |
| 1e-6      | -8.22 e-5    | 4.026           | 0          | 100 | 64         |

FIGURE 3. Convergence results for the Test Case 2

The optimal control, the probability distribution (at any time) and the adjoint are provided in Figures 2d, 2e, and 2f (page 30). As can be observed, the optimal control only takes the boundary values. The value 2 is taken in a small region around  $x = 0$ , after  $t \approx 0.4$ , which guarantees that a sufficiently large proportion of the distribution remains located around 0, as can be seen on the graph of the probability distribution.

**6. Conclusion.** We have proved optimality conditions for a class of constrained non-linear stochastic optimal control problems, using an appropriate concept of differentiability for the cost function and the constraints. The convexity of the closure of the reachable set of probability measures plays an essential role in the proof of these results. An augmented Lagrangian method, based on the convexity property and the optimality conditions has been proposed, demonstrating the relevance of these properties. Good convergence results have been obtained for examples with a one-dimensional state variable. Future work will focus on the extension of these results to more general problems, for example, for cost functions containing an integral cost depending on the current probability distribution.

#### Appendix A. Elements on optimal transportation.

Wasserstein distance. Let us recall the definition of the Wasserstein distance, denoted by  $d_1$  in this article. For all  $m_1$  and  $m_2$  in  $\mathcal{P}_1(\mathbb{R}^n)$ ,

$$d_1(m_1, m_2) = \inf_{\pi \in \Pi(m_1, m_2)} \int_{\mathbb{R}^n \times \mathbb{R}^n} |y - x| d\pi(x, y), \quad (63)$$

the set  $\Pi(m_1, m_2)$  being the set of transportation mappings from  $m_1$  to  $m_2$  defined as:

$$\left\{ \pi \in \mathcal{P}(\mathbb{R}^{2n}) \mid \begin{cases} \pi(A \times \mathbb{R}^n) = m_1(A), \\ \pi(\mathbb{R}^n \times A) = m_2(A), \end{cases} \text{ for all measurable } A \subset \mathbb{R}^n \right\}.$$

**Lemma A.1.** For all  $m_1$  and  $m_2 \in \mathcal{P}_1(\mathbb{R}^n)$ , for all  $\theta \in [0, 1]$ ,

$$d_1((1 - \theta)m_1 + \theta m_2) \leq \theta d_1(m_1, m_2). \quad (64)$$

*Proof.* Let  $\phi \in 1\text{-Lip}(\mathbb{R}^n)$ . Then,

$$\int_{\mathbb{R}^n} \phi d(((1 - \theta)m_1 + \theta m_2) - m_1) = \theta \int_{\mathbb{R}^n} \phi d(m_2 - m_1) \leq \theta d_1(m_1, m_2).$$

The last inequality follows from the dual representation (4). Maximizing the left-hand side with respect to  $\phi \in 1\text{-Lip}(\mathbb{R}^n)$ , we obtain inequality (64).  $\square$



A compactness property.

**Lemma A.2.** *For all  $p > 1$  and  $R \geq 0$ , the subset  $\bar{B}_p(R)$  of  $\mathcal{P}_1(\mathbb{R}^n)$  (defined in (5)) is compact for the  $d_1$ -distance.*

*Proof.* We first prove that  $\bar{B}_p(R)$  is compact for the weak topology of  $\mathcal{P}(\mathbb{R}^n)$ . For all  $r \geq 0$  and for all  $m \in \bar{B}_p(R)$ ,

$$R \geq \int_{\bar{B}_r^c} |x|^p dm(x) \geq r^p \int_{\bar{B}_r^c} 1 dm(x), \quad (65)$$

and thus,  $m(\bar{B}_r^c) \leq R/r^p \rightarrow 0$ , meaning that  $\bar{B}_p(R)$  is tight. By Prokhorov's theorem [34, Page 43],  $\bar{B}_p(R)$  is therefore precompact for the weak-topology. Now, let  $(m_k)_{k \in \mathbb{N}}$  be a sequence in  $\bar{B}_p(R)$  weakly converging to  $\bar{m} \in \mathcal{P}(\mathbb{R}^n)$ . For all  $r \geq 0$ , the function  $\min(|x|^p, r)$  is continuous and bounded, thus:

$$\int_{\mathbb{R}^n} \min(|x|^p, r) d\bar{m}(x) = \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} \min(|x|^p, r) dm_k(x) \leq R.$$

We obtain, using the monotone convergence theorem:

$$\int_{\mathbb{R}^n} |x|^p d\bar{m}(x) = \lim_{r \rightarrow \infty} \int_{\mathbb{R}^n} \min(|x|^p, r) d\bar{m}(x) \leq R,$$

thus  $\bar{m} \in \bar{B}_p(R)$ . Therefore,  $\bar{B}_p(R)$  is weakly closed, and thus weakly compact.

Finally, we need to prove that any weakly converging sequence  $(m_k)_{k \in \mathbb{N}}$  in  $\bar{B}_p(R)$  to some  $\bar{m} \in \bar{B}_p(R)$  also converges for the  $d_1$ -distance. By [34, Definition 6.8/Theorem 6.9], it suffices to prove that

$$\int_{\mathbb{R}^n} |x| dm_k(x) \xrightarrow{k \rightarrow \infty} \int_{\mathbb{R}^n} |x| d\bar{m}(x). \quad (66)$$

Observe that for all  $r \geq 0$ , for all  $m \in \bar{B}_p(R)$ , similarly to (65), we find that

$$\left| \int_{\mathbb{R}^n} |x| - \min(|x|, r) dm(x) \right| \leq \int_{\bar{B}_r^c} |x| dm(x) \leq \frac{R}{r^{p-1}}.$$

Therefore, for all  $r \geq 0$ ,

$$\begin{aligned} \limsup_{k \rightarrow \infty} \left| \int_{\mathbb{R}^n} |x| d(m_k(x) - \bar{m}(x)) \right| \\ \leq \underbrace{\left( \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} \min(|x|, r) d(m_k(x) - \bar{m}(x)) \right)}_{=0} + \frac{2R}{r^{p-1}}. \end{aligned}$$

We obtain (66), making  $r$  tend to  $+\infty$ .  $\square$

A continuity property. We prove in the following lemma the continuity of linear mappings for the  $d_1$ -distance on  $\bar{B}_p(R)$  under a growth condition.

**Lemma A.3.** *Let  $p > 1$ ,  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$  be dominated by  $|x|^p$  (in the sense of (6)). Then, for all  $R \geq 0$ , the following mapping:  $m \in \bar{B}_p(R) \mapsto \int_{\mathbb{R}^n} \phi(x) dm(x)$  is continuous for the  $d_1$ -distance.*

*Proof.* Let  $(m_k)_{k \in \mathbb{N}}$  be a sequence in  $\bar{B}_p(R)$  converging to  $\bar{m} \in \bar{B}_p(R)$  for the  $d_1$ -distance. Let  $\varepsilon > 0$  and let  $r$  be such that (6) holds. We define  $\hat{\phi} : x \in \mathbb{R}^n \mapsto \hat{\phi}(x) = \phi(P(x))$ , where  $P$  is the orthogonal projection on  $\bar{B}_r$ . For all  $x \in \mathbb{R}^n$ , if

$x \in \bar{B}_r$ , then  $\hat{\phi}(x) = \phi(x)$  and if  $x \in \bar{B}_r^c$ , then  $|\hat{\phi}(x)| = |\phi(rx/|x|)| \leq \varepsilon r^p \leq \varepsilon |x|^p$ . Thus, for all  $m \in \bar{B}_p(R)$ ,

$$\left| \int_{\mathbb{R}^n} \hat{\phi} - \phi \, dm \right| \leq \left| \int_{B_r} \hat{\phi} - \phi \, dm \right| + \int_{B_r^c} |\phi| \, dm + \int_{B_r^c} |\hat{\phi}| \, dm \leq 2\varepsilon R.$$

By [34, Definition 6.8/Theorem 6.9], the convergence for the  $d_1$ -distance implies the weak convergence, thus, since  $\hat{\phi}$  is continuous and bounded, we obtain:

$$\begin{aligned} & \limsup_{k \rightarrow \infty} \left| \int_{\mathbb{R}^n} \phi \, d(m_k - \bar{m}) \right| \\ & \leq \lim_{k \rightarrow \infty} \underbrace{\left| \int_{\mathbb{R}^n} \hat{\phi} \, d(m_k - \bar{m}) \right|}_{=0} + \limsup_{k \rightarrow \infty} \left| \int_{\mathbb{R}^n} (\hat{\phi} - \phi) \, d(m_k - \bar{m}) \right| \\ & \leq 4\varepsilon R. \end{aligned}$$

The result follows when  $\varepsilon$  tends to 0.  $\square$

## REFERENCES

- [1] Y. Achdou and M. Laurière. On the system of partial differential equations arising in mean field type control. *Discrete and Continuous Dynamical Systems*, 35(9):3879–3900, 2015.
- [2] G. Albi, Y.-P. Choi, M. Fornasier, and D. Kalise. Mean field control hierarchy. *Applied Mathematics & Optimization*, 76(1):93–135, 2017.
- [3] D. Andersson and B. Djehiche. A maximum principle for SDEs of mean-field type. *Applied Mathematics & Optimization*, 63(3):341–356, 2011.
- [4] A. Bensoussan, J. Frehse, and P. Yam. *Mean field games and mean field type control theory*. Springer Briefs in Mathematics. Springer, New York, 2013.
- [5] J. F. Bonnans, L. Pfeiffer, and O. S. Serea. Sensitivity analysis for relaxed optimal control problems with final-state constraints. *Nonlinear Anal.*, 89:55–80, 2013.
- [6] J. F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer Series in Operations Research. Springer-Verlag, New York, 2000.
- [7] J. F. Bonnans and F. J. Silva. First and second order necessary conditions for stochastic optimal control problems. *Applied Mathematics & Optimization*, 65(3):403–439, Jun 2012.
- [8] B. Bouchard, R. Elie, and N. Touzi. Stochastic target problems with controlled loss. *SIAM J. Control Optim.*, 48(5):3123–3150, 2009.
- [9] R. Buckdahn, B. Djehiche, and J. Li. A general stochastic maximum principle for SDEs of mean-field type. *Appl. Math. Optim.*, 64(2):197–216, 2011.
- [10] P. H. Calamai and J. J. Moré. Projected gradient methods for linearly constrained problems. *Mathematical Programming*, 39(1):93–116, 1987.
- [11] F. Camilli and M. Falcone. An approximation scheme for the optimal control of diffusion processes. *RAIRO Modél. Math. Anal. Numér.*, 29(1):97–122, 1995.
- [12] P. Cardaliaguet. Notes on Mean Field Games, 2012.
- [13] R. Carmona and F. Delarue. Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics. *Ann. Probab.*, 43(5):2647–2700, 2015.
- [14] A. R. Conn, N. I. M. Gould, and Ph. L. Toint. *LANCELOT*, volume 17 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1992. A Fortran package for large-scale nonlinear optimization (release A).
- [15] A. Fleig and R. Guglielmi. Optimal control of the Fokker–Planck equation with space-dependent controls. *Journal of Optimization Theory and Applications*, 174(2):408–427, 2017.
- [16] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer, second edition, 2006.
- [17] N. V. Krylov. *Controlled diffusion processes*, volume 14 of *Applications of Mathematics*. Springer-Verlag, New York-Berlin, 1980.
- [18] M. Laurière and O. Pironneau. Dynamic programming for mean-field type control. *Journal of Optimization Theory and Applications*, 169(3):902–924, 2016.
- [19] C. W. Miller and I. Yang. Optimal control of conditional value-at-risk in continuous time. *SIAM J. Control Optim.*, 55(2):856–884, 2017.

- [20] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2006.
- [21] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Hochschultext / Universitext. Springer, 2003.
- [22] J. L. Pedersen and G. Peskir. Optimal mean-variance portfolio selection. *Math. Financ. Econ.*, 11(2):137–160, 2017.
- [23] S. G. Peng. A general stochastic maximum principle for optimal control problems. *SIAM J. Control Optim.*, 28(4):966–979, 1990.
- [24] L. Pfeiffer. Optimality conditions for mean-field type optimal control problems. *SFB-report 2015-015*, 2015.
- [25] L. Pfeiffer. Numerical methods for mean-field type optimal control problems. *Pure and Applied Functional Analysis*, 1(4):629–655, 2016.
- [26] L. Pfeiffer. Risk-averse Merton’s portfolio problem. *IFAC-PapersOnLine*, 49(8):266 – 271, 2016.
- [27] L. Pfeiffer. Two approaches to stochastic optimal control problems with a final time expectation constraint. *Applied Mathematics & Optimization*, 77(2):377–404, 2018.
- [28] H. Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications*, volume 61 of *Stochastic Modelling and Applied Probability*. Springer, 2009.
- [29] H. Pham and X. Wei. Dynamic programming for optimal control of stochastic McKean–Vlasov dynamics. *SIAM Journal on Control and Optimization*, 55(2):1069–1101, 2017.
- [30] H. Pham and X. Wei. Bellman equation and viscosity solutions for mean-field stochastic control problem. *ESAIM-COCV*, to appear.
- [31] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming*. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, second edition, 2014.
- [32] X. Tan and N. Touzi. Optimal transportation under controlled stochastic dynamics. *Ann. Probab.*, 41(5):3201–3240, 2013.
- [33] N. Touzi. Direct characterization of the value of super-replication under stochastic volatility and portfolio constraints. *Stochastic Process. Appl.*, 88(2):305–328, 2000.
- [34] C. Villani. *Optimal transport. Old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Berlin, 2009.
- [35] J. Yong and X. Y. Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1999.

Received xxxx 20xx; revised xxxx 20xx.

*E-mail address:* [laurent.pfeiffer@uni-graz.at](mailto:laurent.pfeiffer@uni-graz.at)