Duality and approximation of stochastic optimal control problems under expectation constraints

Laurent Pfeiffer, Xiaolu Tan, Yulong Zhou

Duality and approximation of stochastic optimal control problems under expectation constraints

Laurent Pfeiffer* Xiaolu Tan† Yulong Zhou‡

July 1, 2020

Abstract

We consider a continuous time stochastic optimal control problem under both equality and inequality constraints on the expectation of some functionals of the controlled process. Under a qualification condition, we show that the problem is in duality with an optimization problem involving the Lagrange multiplier associated with the constraints. Then by convex analysis techniques, we provide a general existence result and some a priori estimation of the dual optimizers. We further provide a necessary and sufficient optimality condition for the initial constrained control problem. The same results are also obtained for a discrete time constrained control problem. Moreover, under additional regularity conditions, it is proved that the discrete time control problem converges to the continuous time problem, possibly with a convergence rate. This convergence result can be used to obtain numerical algorithms to approximate the continuous time control problem, which we illustrate by two simple numerical examples.

Keywords: Stochastic optimal control, convex duality, numerical approximation.

AMS Subject Classification(2010): 93E20; 49K45; 60H10.

1 Introduction

We study in this article a stochastic optimal control problem under expectation constraints, in a path-dependent framework. Concretely, let the controlled diffusion process be governed by the dynamics

$$dX_t^\alpha = \mu(t, X_{t\wedge\cdot}^\alpha, \alpha_t)dt + \sigma(t, X_{t\wedge\cdot}^\alpha, \alpha_t)dB_t,$$

where B is a Brownian motion. Given functionals Φ, Ψ_0, and Ψ_1, we consider the constrained optimization problem

$$\inf_\alpha \ \mathbb{E}\big[\Phi(X_\cdot^\alpha)\big], \quad \text{subject to: } \mathbb{E}[\Psi_0(X_\cdot^\alpha)] = 0 \text{ and } \mathbb{E}[\Psi_1(X_\cdot^\alpha)] \leq 0. \tag{1.1}$$

In the above problem, $X_{t\wedge\cdot}^\alpha$ denotes the stopped process $(X_{t\wedge s}^\alpha)_{s\in[0,T]}$ and the coefficient functions μ, σ, Φ, Ψ_0 and Ψ_1 may depend on the whole paths of $(X_s^\alpha)_{s\in[0,T]}$.

*Inria and CMAP (UMR 7641), CNRS, Ecole Polytechnique, Institut Polytechnique de Paris. Email: laurent.pfeiffer@inria.fr

†Department of Mathematics, The Chinese University of Hong Kong. Email: xiaolu.tan@cuhk.edu.hk

‡School of Mathematics, Sun Yat-Sen University, Guangzhou, China. Email: zhouyulong@mail.sysu.edu.cn

To illustrate the applications of (1.1), let us first mention the case of probability constraints. Such constraints can typically be written as follows:

$$\mathbb{P}\big[h(X_\cdot^\alpha) \leq 0\big] \leq \gamma, \tag{1.2}$$

where h is vector-valued and $\gamma \in (0,1)$. The constraint (1.2) can be put in the form $\mathbb{E}[\Psi_1(X_\cdot^\alpha)] \leq 0$ by setting $\Psi_1(X_\cdot^\alpha) = \mathbf{1}_{h(X_\cdot^\alpha)\leq 0} - \gamma$, where $\mathbf{1}$ denotes the indicator function. Note that such a function Ψ_1 is in general discontinuous. Probability constraints are typically used to take into account undesirable events, which may however arise with nonzero probability whatever the employed control process. We refer to Shapiro, Dentcheva and Ruszczyński [27, Chapter 4] for a general presentation. Another important application of problem (1.1) may concern the characterization of the Pareto front in a multi-objective setting, see e.g. Yong and Zhou [31, Chapter 3]. More examples of applications of problem (1.1), such as the control problem under state constraints, the semi-static hedging problem in finance, will also be discussed in Section 2.

In the literature, different approaches have been proposed to study the constrained control problem (1.1). The first one consists in introducing an additional process, denoted by $(Y_t^0, Y_t^1)_{t\in[0,T]}$, with initial value $Y_0^i = 0$ and final value $Y_T^i = \Psi_i(X_\cdot^\alpha)$, $i = 0, 1$, and having a martingale or submartingale dynamics. Then one can either reformulate the initial problem as a stochastic target problem which can be studied by the geometric dynamic programming approach (see e.g. Soner and Touzi [28], Bouchard, Elie and Imbert [7], Bouchard and Nutz [8], Chow, Yu and Zhou [10], etc.), or one can use the so-called *level-set approach* to reformulate it into an optimization problem over a family of (unconstrained) singular control problems (see e.g. Bokanowski, Picarelli and Zidani [3] and the references therein). In the second main approach, one looks for necessary optimality conditions in the form of Pontryagin's maximum principle, involving a Lagrange multiplier, see e.g. Yong and Zhou [31, Chapter 3] and Bonnans and Silva [6, Section 5]. Such optimality conditions are obtained with the variational technique, which consists in calculating a first-order Taylor expansion of the cost with respect to small perturbations of the control. The constraints are usually tackled with Ekeland's principle or with the Lyusternik-Graves theorem (under a qualification condition).

In the present work, we follow a *duality approach* by considering the following dual problem:

$$\sup_{\lambda_0\in\mathbb{R},\,\lambda_1\geq 0} \inf_{\alpha} \quad \mathbb{E}\big[\Phi(X_\cdot^\alpha) + \lambda_0\Psi_0(X_\cdot^\alpha) + \lambda_1\Psi_1(X_\cdot^\alpha)\big]. \tag{1.3}$$

The above problem can be formally obtained by writing (1.1) as a saddle-point problem and then by exchanging the inf and sup operators. Our duality result, Theorem 3.3, states that there is no duality gap between (1.1) and (1.3), i.e. the two problems have the same value. Such a duality result should be in line with Kantorovich's duality for the optimal transport (OT) problem, recalling that the marginal constraint in OT is equivalent to infinitely many expectation constraints. Let us refer in particular to Mikami and Thieullen [20], Tan and Touzi [30] for an optimal transport problem along controlled diffusion processes. In contrast to the duality theory in OT, which is generally based on the compactness of the control space satisfying the marginal constraints, together with the regularity of the reward/cost function, we will develop and explore here the Lagrange relaxation approach which was utilized in Pfeiffer [23] for a discrete time control problem.

In a first step, we will develop the Lagrange relaxation approach in an abstract framework, for a problem in the form:

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\xi], \quad \text{subject to:} \quad \mathbb{E}^{\mathbb{P}}[\zeta_0] = 0,\ \mathbb{E}^{\mathbb{P}}[\zeta_1] \leq 0,$$

where $\mathcal{P}$ is a family of probability measures on an abstract measurable space $(\Omega, \mathcal{F})$. By assuming that $\mathcal{P}$ is convex, and using a qualification condition, we obtain an abstract duality result. To reduce the constrained stochastic control problem (1.1) to the above abstract framework, we use a weak formulation of the control problem by considering the law of controlled processes on the canonical space. Such an approach allows to study the problem in a very general framework, in both discrete time and continuous time setting, Markovian or path-dependent, and requires no regularity on the reward/cost function.

In a second step, we explore properties of the initial constrained optimization problem as well as its dual problem. First, we provide a general existence result on the optimizers of the dual problem, together with some estimation. Next, using the dual optimizer, we obtain a necessary and sufficient optimality condition for the initial constrained problem. To the best of our knowledge, such an optimality condition in the non-Markovian setting should be new in the literature. Restricting to the Markovian setting, our optimality condition reduces to the *optimality conditions in variational form* in Pfeiffer [24, Section 4], and could lead naturally to the necessary condition in the maximum principle in [31, 6] described previously.

Finally, we also consider a discrete time control problem under constraints and study its convergence to the continuous time problem. Using the (Wasserstein) weak convergence technique, together with some a priori estimation on the dual problem, we are able to provide a very general approximation result. Moreover, with the a priori bound of the dual optimizer, and based on the approximation results for control problem without constraints, we can obtain some results on the convergence rate. This approximation result can be used to suggest numerical algorithms to approximate the continuous time problem. In particular, it can be considered as an extension of the classical weak convergence approach of Kushner and Dupuis [19] to approximate control problems without constraints. Notice also that similar numerical algorithms have been used in [30, Section 5] for a stochastic optimal transport problem. Our technique, based on the (Wasserstein) weak convergence arguments and a priori dual optimizer estimation, leads to a much more general convergence result, and with possible convergence rate, which is new in the literature.

The following of the paper is organized as follows. In Section 2, we provide a detailed weak formulation of the control problem under expectation constraints, in both continuous time version and discrete time version. Then in Section 3, we give the technical conditions as well as the main results, including the main duality and approximation results. In Section 4, we provide two simple numerical examples, by considering a linear quadratic control problem under constraints. The technical proofs of the main results are completed in Section 5.

2 Optimal control problems under expectation constraints

We provide here a weak formulation of the optimal control problem under expectation constraints, in both continuous time version and discrete time version, as well as the corresponding dual formulations.

2.1 An optimal control problem under constraints in continuous time

Let $n > 0$ be a positive integer, $T > 0$ and $\Omega := C([0,T], \mathbb{R}^n)$ denote the canonical space of all $\mathbb{R}^n$-valued continuous paths on $[0,T]$, with canonical process X. Let $\|\cdot\|$ denote the uniform convergence norm on Ω. Let (A, ρ) be a nonempty Polish space and $(\mu, \sigma)\colon [0,T] \times \Omega \times A \to \mathbb{R}^n \times \mathbb{S}^n$ be the coefficient functions for the controlled diffusion processes, where $\mathbb{S}^n$ denotes the collection of all $n \times n$-dimensional matrices. We assume that μ and σ are non-anticipative, in

the sense that $(\mu,\sigma)(t,\omega,a) = (\mu,\sigma)(t,\omega_{t\wedge\cdot},a)$ for all $(t,\omega,a) \in [0,T]\times\Omega\times A$. Moreover, for some fixed point $a_0 \in A$, there exists a constant $K > 0$ such that, for all $(t,\omega,a) \in [0,T]\times\Omega\times A$,

$$\big|(\mu,\sigma)(t,\omega,a)\big| \le K(1 + \|\omega_{t\wedge\cdot}\| + \rho(a,a_0)\big). \tag{2.4}$$

We also fix $x_0 \in \mathbb{R}^n$ as initial position of the controlled process.

Definition 2.1. A term $\gamma = (\Omega^\gamma, \mathcal{F}^\gamma, \mathbb{F}^\gamma = (\mathcal{F}^\gamma_t)_{t\in[0,T]}, \mathbb{P}^\gamma, X^\gamma, \alpha^\gamma, B^\gamma)$ is a weak control term if $(\Omega^\gamma, \mathcal{F}^\gamma, \mathbb{F}^\gamma, \mathbb{P}^\gamma)$ is a filtered probability space, in which B^γ is an n-dimensional standard Brownian motion, α^γ is an A-valued predictable process satisfying

$$\mathbb{E}^{\mathbb{P}^\gamma}\Big[\|X^\gamma\|^2 + \int_0^T \big(\rho(\alpha^\gamma_t, a_0)\big)^2 dt\Big] < \infty, \tag{2.5}$$

and X^γ is an $\mathbb{R}^n$–valued adapted continuous process such that

$$X^\gamma_t = x_0 + \int_0^t \mu(s, X^\gamma_{s\wedge\cdot}, \alpha^\gamma_s)ds + \int_0^t \sigma(s, X^\gamma_{s\wedge\cdot}, \alpha^\gamma_s)dB^\gamma_s, \quad t\in[0,T], \quad \mathbb{P}^\gamma\text{-a.s.} \tag{2.6}$$

Let us denote by Γ the collection of all weak control terms.

Remark 2.2. (i) The above control is called a weak control because the probability space and the Brownian motion are not a priori fixed, and more importantly, the control process may not be adapted to the filtration generated by the Brownian motion. Using the martingale problem, one can interpret the weak controls as probability measures on the canonical space, and the set of weak control (measures) is convex (but may not be closed). Such a convexity property will be essential in the proof of the duality result. The weak control may be different from the classical strong control, where a fixed probability space and a Brownian motion are given. See also Section 3.4 for a detailed comparaison between the strong and weak formulation of the constrained control problem.

(ii) The growth condition (2.4), together with the integrability condition (2.5), ensures that the (stochastic) integrals in (2.6) are well defined. In particular, one has

$$\int_0^T \big(\big|\mu(s, X^\gamma_{s\wedge\cdot}, \alpha^\gamma_s)\big| + \big|\sigma(s, X^\gamma_{s\wedge\cdot}, \alpha^\gamma_s)\big|^2\big)ds < \infty, \ \mathbb{P}^\gamma\text{–a.s.}$$

At the same time, there is a freedom to choose the metric ρ for the Polish space A. In particular, when (μ,σ) are uniformly bounded, one can choose a uniformly bounded metric ρ so that (2.5) holds true for any process α^γ.

Let $m, \ell \ge 0$ and let $\Phi\colon \Omega \to \mathbb{R}$ and $\Psi_i\colon \Omega \to \mathbb{R}$, $i = 1,\cdots,m+\ell$ be functionals defined on Ω, we consider the following optimization problem under expectation constraints:

$$V(z) := \inf_{\gamma\in\Gamma(z)} \mathbb{E}^{\mathbb{P}^\gamma}\big[\Phi\big(X^\gamma_\cdot\big)\big], \tag{2.7}$$

where for all $z \in \mathbb{R}^m \times \mathbb{R}^\ell$,

$$\begin{aligned}\Gamma(z) \ := \ \big\{\gamma\in\Gamma \ &: \mathbb{E}^{\mathbb{P}^\gamma}\big[\Psi_i\big(X^\gamma_\cdot\big)\big] = z_i, \ i = 1,\cdots,m,\\ &\text{and} \ \ \mathbb{E}^{\mathbb{P}^\gamma}\big[\Psi_{m+j}\big(X^\gamma_\cdot\big)\big] \le z_{m+j}, \ j = 1,\cdots,\ell\big\}.\end{aligned} \tag{2.8}$$

We are actually mostly interested in the case $z = 0$ but directly consider a parameterized version of the problem in (2.7), for conviniency. Integrability condtions on Φ and Ψ will be specified in Section 3.1.

Remark 2.3. We can consider more general costs and constraints, such as

$$\inf_{\gamma\in\Gamma(z)} \mathbb{E}^{\mathbb{P}^\gamma}\Big[\int_0^T L\big(s, X^\gamma_{s\wedge\cdot}, \alpha^\gamma_s\big)ds + \Phi\big(X^\gamma_\cdot\big)\Big],$$

for some function $L\colon [0,T]\times\Omega\times A\to\mathbb{R}$. By introducing an $n+1$ dimensional process $\widetilde{X}$ defined by $\widetilde{X}^i := X^i$ for $i = 1,\cdots,n$ and $\widetilde{X}^{n+1}_t := \int_0^t L(s, X^\gamma_{s\wedge\cdot}, \alpha^\gamma_s)ds$ and the corresponding new criteria function, one can reduce the problem to the formulation (2.7). From a numerical approximation point of view, it is always better to stay in the minimal dimension context. Here we would like to use this formulation (à la Mayer) to study the problem for ease of presentation and leave the dimensional reduction to the numerical implementation step.

A dual formulation By penalizing the constraints, it is clear that for $z = 0$,

$$V(0) = \inf_{\gamma\in\Gamma}\sup_{\lambda\in\mathbb{R}^m\times\mathbb{R}^\ell_+} \mathbb{E}^{\mathbb{P}^\gamma}\big[\Upsilon(X^\gamma_\cdot,\lambda)\big],$$

where

$$\Upsilon(\omega,\lambda) = \Phi(\omega) + \lambda\cdot\Psi(\omega).$$

It follows that a dual formulation of the optimization problem $V(0)$ can be given by

$$D(0) := \sup_{\lambda\in\mathbb{R}^m\times\mathbb{R}^\ell_+} d(\lambda), \quad \text{where } d(\lambda) := \inf_{\gamma\in\Gamma} \mathbb{E}^{\mathbb{P}^\gamma}\big[\Upsilon(X^\gamma_\cdot,\lambda)\big]. \tag{2.9}$$

Some examples We also provide some examples of applications of the above constrained optimization problem.

Example 2.4 (Optimal control under state constraint). The optimal control problem under state constraint has been studied by different papers (see e.g. [3] and the references therein), it can be formulated as follows:

$$\inf\big\{\mathbb{E}\big[\Phi(X^\gamma_T)\big] \ : \gamma\in\Gamma,\ X^\gamma_t\in E \text{ a.s. for all } t\in[0,T]\big\}, \tag{2.10}$$

where E is a closed convex subset of $\mathbb{R}^n$. Let us denote by $d(x,E)$ the distance between x and E, and let us define

$$\Psi(X^\gamma_\cdot) \ := \ \max_{0\le t\le T} d(X^\gamma_t, E),$$

then Problem (2.10) is equivalent to the following problem under expectation constraint (in a path-dependent version): $\inf\big\{\mathbb{E}\big[\Phi(X^\gamma_T)\big] \ : \gamma\in\Gamma,\ \mathbb{E}\big[\Psi(X^\gamma_\cdot)\big] = 0\big\}$.

Example 2.5 (Semi-static super-replication problem in finance). Let us consider an uncertain volatility financial model, where the interest rate $r = 0$ and a risky asset price follows the dynamics $dS^\gamma_t = \sigma^\gamma_t S^\gamma_t dW_t$ with an adapted process $\sigma^\gamma = (\sigma^\gamma_t)_{0\le t\le T}$. Denote by Γ the collection of all models $\gamma = (S^\gamma, \sigma^\gamma)$ such that σ^γ takes value in $[\underline{\sigma},\overline{\sigma}]$, and assume that the market model is uncertain (unknown), but the class of all possible market dynamics is given by $\{(S^\gamma_t)_{t\in[0,T]} \ : \gamma\in\Gamma\}$.

A dynamic trading strategy is an adapted process $H = (H_t)_{0\le t\le T}$ and $(H\cdot S)_T := \int_0^T H_t dS^\gamma_t$ is the Profit & Loss of the dynamic strategy. Denote by $\mathcal{H}$ the class of all dynamic strategies. On the market, there are two vanilla options of payoff $\Psi_1(S^\gamma_T)$ and $\Psi_2(S^\gamma_T)$ at maturity T. Assume that an agent is allowed to buy or to sell the first option with price z_1, and is allowed to buy the second option with price z_2 (but cannot short/sell it). There is another option with payoff

$\Phi(S^\gamma_\cdot)$, and the agent aims to super-replicate it with both dynamic trading strategy H and static strategy on options Ψ_1 and Ψ_2. Then the collection of all super-replication strategies is given by

$$\begin{aligned}\mathcal{D} :=\big\{&(x, H, \lambda_1, \lambda_2) \in \mathbb{R} \times \mathcal{H} \times \mathbb{R} \times \mathbb{R}_+ \ : \\ &x + (H \cdot S^\gamma)_T + \lambda_1(\Psi_1(S^\gamma_T) - z_1) + \lambda_2(\Psi_2(S^\gamma_T) - z_2) \geq \Phi(S^\gamma_\cdot), \text{a.s. } \forall \gamma \in \Gamma\big\},\end{aligned}$$

and the minimal super-replication cost of option Φ is given by

$$\inf\big\{x \ : (x, H, \lambda_1, \lambda_2) \in \mathcal{D}\big\}.$$

Next, by the duality theory (see e.g. Denis and Martini [11]) in finance, the above problem has a dual formulation:

$$\inf_{(\lambda_1,\lambda_2)\in\mathbb{R}\times\mathbb{R}_+} \ \sup_{\gamma\in\Gamma} \ \mathbb{E}\big[\Phi(S^\gamma_\cdot) - \lambda_1(\Psi_1(S^\gamma_T) - z_1) - \lambda_2(\Psi_2(S^\gamma_T) - z_2)\big],$$

which is equivalent to a constrained optimal control problem:

$$\sup_{\gamma\in\Gamma} \ \big\{\mathbb{E}[\Phi(S^\gamma_\cdot)] \ : \gamma \in \Gamma, \ \mathbb{E}[\Psi_1(S^\gamma_T)] = z_1, \ \mathbb{E}[\Psi_2(S^\gamma_T)] \leq z_2\big\}.$$

Notice that the above application in finance has also motivated the so-called optimal Skorokhod embedding problem, and the martingale optimal transport problem which has recently generated an important literature (see e.g. [2] among many others).

2.2 A discrete time version of the constrained control problem

Let us introduce a discrete time version of the constrained control problem, with the objective to approximate the continuous time problem (2.7).

Let $N > 0$ be an integer and $h := \frac{T}{N}$, denote $t_k := kh$ and $\mathbb{T}_h := \{t_0, t_1, \cdots, t_N\}$. Let $H_h : \mathbb{T}_h \times \Omega \times A \times [0,1] \to \mathbb{R}^n$ be a function satisfying

$$H_h(t_k, \omega, a, u) = H_h(t_k, \omega_{t_k\wedge\cdot}, a, u), \quad \text{for all } (t_k, \omega, a, u) \in \mathbb{T}_h \times \Omega \times A \times [0,1].$$

Definition 2.6. We say that

$$\gamma \ = \ \big(\Omega^\gamma, \mathcal{F}^\gamma, \mathbb{F}^\gamma, \mathbb{P}^\gamma, \alpha^\gamma, X^\gamma, U^\gamma\big)$$

is a weak discrete time control term if $(\Omega^\gamma, \mathcal{F}^\gamma, \mathbb{F}^\gamma, \mathbb{P}^\gamma)$ is a filtered probability space, equipped with an A-valued process α^γ, an $\mathbb{R}^n$–valued process X^γ, and a $[0,1]$–valued process U^γ, which are all adapted, and for each $k = 1, \cdots, N$, $U^\gamma_{t_k}$ is of uniform distribution on $[0,1]$ and independent of $\mathcal{F}^\gamma_{t_{k-1}}$, and finally, $X^\gamma_0 = x_0$,

$$X^\gamma_{t_k+1} \ = \ X^\gamma_{t_k} \ + \ H_h\big(t_k, \widehat{X}^\gamma_{t_k\wedge\cdot}, \alpha^\gamma_{t_k}, U^\gamma_{t_{k+1}}\big), \tag{2.11}$$

where $\widehat{X}^\gamma$ denotes the continuous time process on $[0,T]$ obtained by linear interpolation of $(X^\gamma_{t_i})_{t_i\in\mathbb{T}_h}$.

Let us denote by Γ_h the collection of all weak discrete time control rules, and by $\Gamma_h(z) \subset \Gamma_h$, for $z \in \mathbb{R}^{m+\ell}$, the subset of all controls γ such that

$$\mathbb{E}^{\mathbb{P}^\gamma}\big[\Psi_i\big(\widehat{X}^\gamma_\cdot\big)\big] = z_i, \ i = 1, \cdots, m, \quad \text{and} \ \ \mathbb{E}^{\mathbb{P}^\gamma}\big[\Psi_{m+j}\big(\widehat{X}^\gamma_\cdot\big)\big] \leq z_{m+j}, \ j = 1, \cdots, \ell.$$

Then a weak formulation of the discrete time constrained problem is given by

$$V_h(z) \ := \inf_{\gamma\in\Gamma_h(z)} \mathbb{E}^{\mathbb{P}^\gamma}\big[\Phi\big(\widehat{X}^\gamma_\cdot\big)\big]. \tag{2.12}$$

A dual formulation As in (2.9), we introduce the corresponding dual problem for the discrete time optimization problem, given by

$$D_h(0) := \sup_{\lambda\in\mathbb{R}^m\times\mathbb{R}^\ell_+} d_h(\lambda), \quad \text{where } d_h(\lambda) := \inf_{\gamma\in\Gamma_h} \mathbb{E}^{\mathbb{P}^\gamma}\big[\Upsilon(\widehat{X}^\gamma_\cdot,\lambda)\big]. \tag{2.13}$$

Remark 2.7. In the definition of the weak control term in Definition 2.6, if we fix the filtration to be the one generated by $(U^\gamma_k)_{k=1,\cdots,N}$, i.e. $\mathcal{F}^\gamma_k = \sigma(U^\gamma_i,\ i\le k)$, then it can be considered as a strong version of the control rule. For the problem $d_h(\lambda)$ in (2.13), it is equivalent to consider only strong control rules by the dynamic programming principle. However, it may change the value of control problem under constraints in (2.12).

3 Main results

We now provide the assumptions and main results of the paper, including our duality and approximation results. We then also discuss the numerical resolution of the discrete time (dual) constrained control problem.

3.1 Assumptions

The first assumption is on the integrability of the functionals Φ and Ψ, and a Robinson qualification condition on the constraints in both continuous and discrete time setting (see also the discussions in Section 3.4 on the validity of the qualification condition).

Assumption 3.1. (i) *The random variables $\Phi(X^\gamma_\cdot)$, $\Psi_i(X^\gamma_\cdot)$, $\Phi(\widehat{X}^{\gamma_h}_\cdot)$ and $\Psi_i(\widehat{X}^{\gamma_h}_\cdot)$ are integrable, for all $i=1,\cdots,m+\ell$, $\gamma\in\Gamma$ and $\gamma_h\in\Gamma_h$. Moreover, there is a constant $M>0$ such that*

$$\sup_{\gamma\in\Gamma}\Big|\mathbb{E}^{\mathbb{P}^\gamma}[\Phi(X^\gamma_\cdot)]\Big| \le M \quad \text{and} \quad \sup_{\gamma_h\in\Gamma_h}\Big|\mathbb{E}^{\mathbb{P}^{\gamma_h}}[\Phi(\widehat{X}^{\gamma_h}_\cdot)]\Big| \le M.$$

(ii) *There is a constant $\varepsilon>0$ such that*

$$B_\varepsilon(0)\subset \mathrm{Conv}\big\{z\in\mathbb{R}^{m+\ell} : \exists\gamma\in\Gamma,\ s.t.\ \mathbb{E}^{\mathbb{P}^\gamma}[\Psi(X^\gamma_\cdot)]+z\in\{0_m\}\times\mathbb{R}^\ell_-\big\}, \tag{3.14}$$

and

$$B_\varepsilon(0)\subset \mathrm{Conv}\big\{z\in\mathbb{R}^{m+\ell} : \exists\gamma_h\in\Gamma_h,\ s.t.\ \mathbb{E}^{\mathbb{P}^{\gamma_h}}[\Psi(\widehat{X}^{\gamma_h}_\cdot)]+z\in\{0_m\}\times\mathbb{R}^\ell_-\big\}, \tag{3.15}$$

where $0_m\in\mathbb{R}^m$ denotes the original point in $\mathbb{R}^m$ and $B_\varepsilon(0)$ denotes the closed ball in $\mathbb{R}^{m+\ell}$ with center 0 and radius ε for the supremum norm.

The next assumption ensures that the space of the laws of controlled process X^γ are relatively compact, and the discrete time control problem consists of a good approximation of the continuous time problem when the time step $h=T/N$ is small.

Assumption 3.2. (i) *The Polish space A is compact, and the associated metric ρ is uniformly bounded. The function μ and σ are continuous in all arguments and for some constant $K>0$,*

$$\|(\mu,\sigma)(t,\omega,a)-(\mu,\sigma)(t,\omega',a)\|\le K\|\omega_{t\wedge\cdot}-\omega'_{t\wedge\cdot}\|, \quad \textit{for all } (t,\omega,\omega',a). \tag{3.16}$$

Moreover, the functions $\Phi(\omega)$ and $\Psi_i(\omega)$, $i=1,\cdots,m+\ell$ are continuous and bounded by $K(1+\|\omega\|^2)$ for some constant $K>0$.

(ii) *For every* $(t_k, \omega, a) \in \mathbb{T}_h \times \Omega \times A$*, with a random variable* U *of uniform distribution* $\mathcal{U}[0,1]$*, one has*

$$\mathbb{E}[H_h(t_k, \omega, a, U)] = \mu(t_k, \omega, a)h, \quad \mathrm{Var}[H_h(t_k, \omega, a, U)] = \sigma\sigma^\top(t_k, \omega, a)h.$$

Moreover, there is some constant $C > 0$ *such that* $\mathbb{E}\big[\big|H_h(t_k, \omega, a, U)\big|^3\big] \le Ch^{3/2}$ *for all* $(t_k, \omega, a) \in \mathbb{T}_h \times \Omega \times A$ *and* $h > 0$*.*

3.2 Duality and approximation of the constrained control problems

Our first main result is on the duality of the two constrained problems, together with the existence of dual optimizers and some estimation. Let us denote

$$\|\lambda\|_1 := \sum_{i=1}^{m+\ell} |\lambda_i|, \quad \text{for } \lambda = (\lambda_1, \cdots, \lambda_{m+\ell}) \in \mathbb{R}^{m+\ell}.$$

Theorem 3.3. *Let Assumption 3.1 hold true.*
(i) *We have*

$$V(0) = D(0) \quad \textit{and} \quad V_h(0) = D_h(0).$$

Moreover, for the dual problem D *(resp.* D_h*), there exist optimal solutions; and any optimal solution* $\lambda^* \in \mathbb{R}^m \times \mathbb{R}^\ell_+$ *(resp.* $\lambda^*_h \in \mathbb{R}^m \times \mathbb{R}^\ell_+$*) satisfies* $\|\lambda^*\|_1 \le \frac{2M}{\varepsilon}$ *(resp.* $\|\lambda^*_h\|_1 \le \frac{2M}{\varepsilon}$*).*
(ii) *Let* $\gamma^* \in \Gamma(0)$*. Then* γ^* *is a (global) solution to the constrained problem* (2.7) *if and only if there exists* $\lambda^* \in \mathbb{R}^m \times \mathbb{R}^\ell_+$ *such that the following conditions hold true:*

$$\textit{[Complementarity]} \qquad \mathbb{E}^{\mathbb{P}^{\gamma^*}}\Big[\sum_{i=1}^{m+\ell} \lambda^*_i \Psi_i(X^{\gamma^*}_\cdot)\Big] = 0, \tag{3.17}$$

$$\textit{[Stationarity]} \qquad \mathbb{E}^{\mathbb{P}^{\gamma^*}}\big[\Upsilon(X^{\gamma^*}_\cdot, \lambda^*)\big] = \inf_{\gamma \in \Gamma} \mathbb{E}^{\mathbb{P}^{\gamma}}\big[\Upsilon(X^{\gamma}_\cdot, \lambda^*)\big]. \tag{3.18}$$

In this case, λ^* *is a solution to the dual problem (i.e.* $d(\lambda^*) = D(0)$*).*
(ii') *Let* $\gamma^*_h \in \Gamma_h(0)$*. Then* γ^*_h *is a (global) solution to the constrained problem* (2.12) *if and only if there exists* $\lambda^*_h \in \mathbb{R}^m \times \mathbb{R}^\ell_+$ *such that the following conditions hold true:*

$$\textit{[Complementarity]} \qquad \mathbb{E}^{\mathbb{P}^{\gamma^*_h}}\Big[\sum_{i=1}^{m+\ell} \lambda^*_{h,i} \Psi_i(\widehat{X}^{\gamma^*_h}_\cdot)\Big] = 0, \tag{3.19}$$

$$\textit{[Stationarity]} \qquad \mathbb{E}^{\mathbb{P}^{\gamma^*_h}}\big[\Upsilon(\widehat{X}^{\gamma^*_h}_\cdot, \lambda^*_h)\big] = \inf_{\gamma_h \in \Gamma_h} \mathbb{E}^{\mathbb{P}^{\gamma_h}}\big[\Upsilon(\widehat{X}^{\gamma_h}_\cdot, \lambda^*_h)\big]. \tag{3.20}$$

In this case, λ^*_h *is a solution to the discrete dual problem (i.e.* $d_h(\lambda^*_h) = D_h(0)$*).*

Remark 3.4. The estimation of the dual optimizer λ^* implies that it is enough to optimize λ in a compact set $B_{2M/\varepsilon}(0) := \{\lambda \in \mathbb{R}^m \times \mathbb{R}^\ell_+ \ : \|\lambda\|_1 \le 2M/\varepsilon\}$ for the dual problem (2.9).

Remark 3.5. Let γ^* be an optimal solution to the initial constrained control problem, then it is also a solution to the unconstrained control problem $d(\lambda^*)$ with an optimal dual optimizer λ^*. In the Markovian context, where μ, σ, Φ and Ψ_i depend only on the running process X_t in place of the path $X_{t\wedge\cdot}$, and under some regularity condition, one may deduce a necessary condition on

γ^* from the stochastic Pontryagin maximum principle of Peng [22]. This is exactly the extended stochastic Pontryagin maximum principle for the constrained optimal control problem, see e.g. Yong and Zhou [31], Bonnans and Silva [6], etc. The optimality conditions provided in Theorem 3.3 are in line with those of Pfeiffer in [24] in a non-linear Markovian setting.

We next provide an approximation result under Assumption 3.2, that is,

$$V_h(0) = D_h(0) \longrightarrow V(0) = D(0).$$

As the dual problems are equivalent to the optimization of $d(\lambda)$ or $d_h(\lambda)$ on a compact set $B_{2M/\varepsilon}(0)$, one can expect to have a uniform convergence rate between $d_h(\lambda)$ and $d(\lambda)$ on $B_{2M/\varepsilon}(0)$, which implies a convergence rate between $D_h(0)$ and $D(0)$. We formulate the convergence rate condition as follows.

Assumption 3.6. *There are constants $C > 0$ and $\rho > 0$ such that*

$$|d(\lambda) - d_h(\lambda)| \le Ch^\rho, \quad \textit{for all } \ \lambda \in B_{2M/\varepsilon}(0) \subset \mathbb{R}^m \times \mathbb{R}^\ell_+.$$

Remark 3.7. (i) Consider the Markovian context, where

$$(\mu, \sigma)(t, \omega, a) = (\mu_0, \sigma_0)(t, \omega_t, a)$$

for some $(\mu_0, \sigma_0)\colon [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}^n \times \mathbb{S}^d$ and $\Phi(\omega) = \Phi_0(\omega_T)$, $\Psi_i(\omega) = \Psi_{0,i}(\omega_T)$ for some functions $\Phi_0\colon \mathbb{R}^n \to \mathbb{R}$ and $\Psi_{0,i}\colon \mathbb{R}^n \to \mathbb{R}$. Assume that $\mu_0, \sigma_0, \Phi_0, \Psi_{0,i}$ are all bounded continuous, and $(\mu_0, \sigma_0)(t, x, a)$ is Lipschitz in x, $\frac{1}{2}$-Hölder in t uniformly on $[0,T] \times \mathbb{R}^n \times A$, and Φ_0, $\Psi_{0,i}$ are all Lipschitz in x, and further that for any $\delta > 0$, there exists a finite subset $A_\delta \subset A$ such that

$$\sup_{(t,x,a) \in [0,T] \times \mathbb{R}^n \times A} \inf_{a' \in A_\delta} \Big(\big\| (\mu_0, \sigma_0)(t, x, a) - (\mu_0, \sigma_0)(t, x, a') \big\| \Big) \ \le \ \delta.$$

Then under Assumption 3.2, the condition in Assumption 3.6 holds true for some constant $C > 0$ (independent of $\lambda \in B_{2M/\varepsilon}(0)$) and $\rho = \frac{1}{10}$. The above convergence rate result was first proved by Krylov's [18] based on a shaking coefficient argument, and then improved by Barles and Jakobsen [1].

(ii) In the non-Markovian context, assume that the sets $\{\mu(t, \omega, a) \ : a \in A\}$ and $\{\sigma\sigma^\top(t, \omega, a) \ : a \in A\}$ are uniformly bounded and independent of $\omega \in \Omega$, and Φ and Ψ_i are Lipschitz, then under Assumption 3.2, the condition in Assumption 3.6 holds true for some constant $C > 0$ and $\rho = \frac{1}{8}$. The above convergence rate result is obtained by Dolinsky [13] (see also Tan [29]) using a strong invariance principle argument.

(iii) In the non-Markovian context, assume that the volatility function σ is not controlled so that the optimal control problem in $d(\lambda)$ can be reformulated as a BSDE (backward stochastic differential equation). Convergence results for numerical schemes of BSDE lead to a rate $\rho = \frac{1}{2}$, see e.g. Bouchard and Touzi [9], Zhang [32], etc.

Theorem 3.8. *Let Assumptions 3.1 and 3.2 hold true. Then*

$$D_h(0) = V_h(0) \longrightarrow V(0) = D(0), \quad \textit{as} \ \ h \longrightarrow 0.$$

Suppose in addition that Assumption 3.6 holds true with some constants $C > 0$ and $\rho > 0$, then

$$|V_h(0) - V(0)| = |D_h(0) - D(0)| \ \le \ Ch^\rho. \tag{3.21}$$

3.3 Numerical algorithm to compute $D_h(0)$

It is not hard to verify that $\lambda \mapsto d_h(\lambda)$ is a concave function, therefore it is natural to compute

$$D_h(0) = \max_{\lambda \in B_{2M/\varepsilon}(0)} d_h(\lambda)$$

by the following gradient algorithm, which is an iterative method (w.r.t. k).

- Initialize the problem with $\lambda_0 = 0$.
- Given λ_k, solve the problem $d_h(\lambda_k)$, that is, find $\hat{\gamma}_k$ such that

$$\mathbb{E}^{\mathbb{P}^{\hat{\gamma}_k}}\big[\Phi\big(\widehat{X}_\cdot^{\hat{\gamma}_k}\big) + \lambda \cdot \Psi\big(\widehat{X}_\cdot^{\hat{\gamma}_k}\big)\big] = d_h(\lambda_k) = \inf_{\gamma_h \in \Gamma_h} \mathbb{E}^{\mathbb{P}^{\gamma_h}}\big[\Phi(\widehat{X}_\cdot^{\gamma_h}) + \lambda_h^* \cdot \Psi(\widehat{X}_\cdot^{\gamma_h})\big].$$

- With $(\lambda_k, \hat{\gamma}_k)$, compute λ_{k+1} by

$$\lambda_{k+1} := \Pi\big(\lambda_k + \theta_k \nabla d_h(\lambda_k)\big) \quad \text{with} \quad \nabla d_h(\lambda_k) := \mathbb{E}^{\mathbb{P}^{\hat{\gamma}_k}}\big[\Psi\big(\widehat{X}_\cdot^{\hat{\gamma}_k}\big)\big],$$

  where Π denote the projection operator from $\mathbb{R}^{m+\ell}$ on $B_{2M/\varepsilon}(0) \cap (\mathbb{R}^m \times \mathbb{R}^\ell_+)$.

We refer to Nesterov [21, Section 3.2.3] for a discussion on appropriate choices of stepsizes $\theta_k > 0$.

Choice of H_h and resolution of the (unconstrained) control problem $d_h(\lambda_k)$ In the purpose of approximating the initial continuous time constrained control problem, we need to choose good kernel functions H_h such that Assumption 3.2 holds true. There exist different choices of H_h, which correspond to different numerical schemes for the resolution of the optimal control problem (or nonlinear parabolic equation). To illustrate the idea, let us take the example of the finite difference scheme in the one-dimensional ($d = 1$) case. Following Kushner and Dupuis [19], one can interpret the finite difference scheme for the HJB equation as a controlled Markov chain system, where the controlled discrete time process has increments $(-\Delta x, 0, \Delta x)$ with different (controlled) probabilities between time $t_k := k\Delta t$ and $t_{k+1} := (k+1)\Delta t$. In the above definition, $h = (\Delta t, \Delta x)$ represents time-space steps parameters in the dicretization of $[0, T] \times \mathbb{R}$. In our context, the increment of the controlled process is given by $H_h(t_k, \widehat{\mathbf{x}}, a, U)$ given the past states $\mathbf{x} = (x_0, \cdots, x_k)$ and the control value a. Then H_h could be constructed explicitly such that, with r.v. $U \sim \mathcal{U}[0,1]$,

$$H_h(t_k, \widehat{\mathbf{x}}, a, U) = \begin{cases} \Delta x, & \text{with probability } p_+ = \mu(t_k, \widehat{\mathbf{x}}, a)\frac{\Delta t}{\Delta x} + \frac{1}{2}\sigma^2(t_k, \widehat{\mathbf{x}}, a)\frac{\Delta t}{\Delta x^2}, \\ -\Delta x, & \text{with probability } p_- = \frac{1}{2}\sigma^2(t_k, \widehat{\mathbf{x}}, a)\frac{\Delta t}{\Delta x^2}, \\ 0, & \text{with probability } 1 - p_+ - p_-. \end{cases}$$

In the above definition, one needs to choose parameters $(\Delta t, \Delta x)$ such that p_+, p_- and $1 - p_+ - p_-$ take value in $[0, 1]$, ensuring that $(p_+,\ p_-,\ 1 - p_+ - p_-)$ it is a probability function and $H_h(t_k, \widehat{\mathbf{x}}, a, U)$ is a random variable. Then by direct computation, one can check that H_h satisfies Assumption 3.2(ii).

Although Kushner and Dupuis [19] mainly discussed the finite difference scheme for optimal control problem in a Markovian framework, the same arguments still apply in the path-dependent framework, and on other (monotone) numerical schemes for optimal control problems, or equivalently to HJB equations (see e.g. Tan [29], Possamaï and Tan [25], Bonnans, Gianatti and Silva [4]). Let us also refer to Section 4 of Ren and Tan [26] for more examples on how to choose H_h to obtain other numerical schemes, including the semi-Lagrangian scheme, the probabilistic scheme of Fahim, Touzi and Warin [15], etc.

3.4 Further discussions

A strong formulation of the optimal control problem under constraints By fixing the probability space and the Brownian motion in the weak control term γ, one can obtain a strong formulation of the control problem. This can be achieved by just fixing the filtration $\mathbb{F}^\gamma$ as the Brownian filtration, as all other processes are assumed to be adapted to it.

Definition 3.9. (i) A weak control term

$$\gamma = (\Omega^\gamma, \mathcal{F}^\gamma, \mathbb{F}^\gamma = (\mathcal{F}_t^\gamma)_{t\in[0,T]}, \mathbb{P}^\gamma, X^\gamma, \alpha^\gamma, B^\gamma)$$

is called a strong control term if $\mathbb{F}^\gamma$ is the (augmented) filtration generated by B^γ.

(ii) A strong control term γ is called piecewise constant if α^γ is piecewise constant, i.e. $\alpha_s = \alpha_{t_i}$ for $s \in [t_i, t_{i+1})$ with some discret time grid $0 = t_0 < \cdots < t_N = T$.

Let us denote

$$\Gamma_S := \big\{\text{All strong control terms}\big\}, \quad \Gamma_{S,0} := \big\{\text{All piecewise strong control terms}\big\},$$

and

$$\Gamma_S(z) \ := \ \Gamma_S \cap \Gamma(z), \quad \text{for } z \in \mathbb{R}^m \times \mathbb{R}^\ell.$$

We can then introduce a strong formulation of the optimal control problem under constraints:

$$V_S(z) \ := \ \inf_{\gamma\in\Gamma_S(z)} \mathbb{E}^{\mathbb{P}^\gamma}\big[\Phi\big(X_\cdot^\gamma\big)\big].$$

As we will see in the technical proof part, one can approximate a weak control by strong control terms. However, the equality constraints $\mathbb{E}^{\mathbb{P}^\gamma}[\Psi_i(X_\cdot^\gamma)] = z_i$ may not be ensured in the approximation, and for this reason, we are not able to prove that $V_S(z) = V(z)$. In other words, with the presence of the equality constraints, $V_S(z)$ may lack some regularity in z. Nevertheless, we are able to prove that by considering the convex envelop or lower-semicontinuous envelop of V_S, we then obtain V. Let us denote by $V_S^{l.s.c.}$ the biggest lower-semicontinuous function dominated by V_S, and by V_S^{conv} the biggest convex function dominated by V_S.

Proposition 3.10. *Let Assumptions 3.1 and 3.2 hold true. Then,*

$$V(z) \ = \ V_S^{conv}(z) \ = \ V_S^{l.s.c.}(z), \quad \textit{for all } z \in B_{\varepsilon/2}(0),$$

where $B_{\varepsilon/2}(0) := \{z \in \mathbb{R}^{m+\ell} \ : \|z\|_1 \le \varepsilon/2\}$,

The proof will be completed in Section 5.5.

On Assumption 3.1 Item (i) in Assumption 3.1 is an integrability condition, which can be reasonably checked for concrete examples. The qualification condition in Assumption 3.1 (ii) turns to be more abstract. We provide below a more explicit equivalent formulation, which could be easier to check (at least numerically). Let

$$E := \{-1, 1\}^m \times \{1\}^\ell, \quad \text{and} \quad \Theta_1^+ := \Big\{(\theta_1, \cdots, \theta_{m+\ell}) \in \mathbb{R}_+^{m+\ell} \ : \sum_{i=1}^{m+\ell} \theta_i = 1\Big\},$$

which is a convex and compact subset of $\mathbb{R}_+^{m+\ell}$. Let $e \in E$, we denote

$$e\Psi(\cdot) := \big(e_1\Psi_1(\cdot), \cdots, e_{m+\ell}\Psi_{m+\ell}(\cdot)\big).$$

Proposition 3.11. *Assumption 3.1* (ii) *holds if and only if, for each* $e \in E$,

$$\max_{\theta\in\Theta_1^+}\inf_{\gamma\in\Gamma}\mathbb{E}\big[\theta\cdot e\Psi(X_\cdot^\gamma)\big] = \inf_{\gamma\in\Gamma}\max_{\theta\in\Theta_1^+}\mathbb{E}\big[\theta\cdot e\Psi(X_\cdot^\gamma)\big] \le -\varepsilon, \tag{3.22}$$

and

$$\max_{\theta\in\Theta_1^+}\inf_{\gamma_h\in\Gamma_h}\mathbb{E}\big[\theta\cdot e\Psi(\widehat{X}_\cdot^{\gamma_h})\big] = \inf_{\gamma_h\in\Gamma_h}\max_{\theta\in\Theta_1^+}\mathbb{E}\big[\theta\cdot e\Psi(\widehat{X}_\cdot^{\gamma_h})\big] \le -\varepsilon. \tag{3.23}$$

Remark 3.12. Notice that the optimization problems (3.22) and (3.23) can be computed by the same numerical algorithm as the one described in Section 3.3.

Proof. We will only provide the equivalence between (3.14) and (3.22). The equivalence between (3.15) and (3.23) follows by exactly the same arguments.

First, it is easy to see that the condition (3.14) is equivalent to that for all $e \in E$, there is some control $\gamma \in \Gamma$ such that

$$\mathbb{E}\big[e_i\Psi_i(X_\cdot^\gamma)\big] \le -\varepsilon, \quad \text{for all } i = 1, \cdots, m+\ell.$$

Next, we notice that

$$\max_{i=1,\cdots,m+\ell} c_i \;=\; \max_{\theta\in\Theta_1^+}\big(\theta\cdot c\big), \quad \text{for all } \; c = (c_1, \cdots, c_{m+\ell}) \in \mathbb{R}^{m+\ell},$$

then (3.14) is equivalent to

$$\inf_{\gamma\in\Gamma}\max_{\theta\in\Theta_1^+}\mathbb{E}\big[\theta\cdot e\Psi(X_\cdot^\gamma)\big] \le -\varepsilon, \quad \text{for each } e \in E.$$

To conclude, it is enough to prove the duality result in (3.22). Let us denote by $\mathcal{P} := \{\mathbb{P}^\gamma \circ (X^\gamma)^{-1} \; : \gamma \in \Gamma\}$, which is a convex set (see Lemma 5.7). Notice that Θ_1^+ is a convex compact subset of $\mathbb{R}^{m+\ell}$, then it follows by the classical minimax theorem that

$$\begin{aligned}\max_{\theta\in\Theta_1^+}\inf_{\gamma\in\Gamma}\mathbb{E}\big[\theta\cdot e\Psi(X_\cdot^\gamma)\big] &= \max_{\theta\in\Theta_1^+}\inf_{\mathbb{P}\in\mathcal{P}}\mathbb{E}^{\mathbb{P}}\big[\theta\cdot e\Psi(X_\cdot)\big] \\ &= \inf_{\mathbb{P}\in\mathcal{P}}\max_{\theta\in\Theta_1^+}\mathbb{E}^{\mathbb{P}}\big[\theta\cdot e\Psi(X_\cdot)\big] = \inf_{\gamma\in\Gamma}\max_{\theta\in\Theta_1^+}\mathbb{E}\big[\theta\cdot e\Psi(X_\cdot^\gamma)\big].\end{aligned}$$

□

4 Numerical examples

In this section we illustrate our theoretical results by two simple numerical examples with a linear quadratic structure, so that the reference value of the problem can be computed explicitly. Concretely, we set $A = \mathbb{R}$, $\mu(t,x,a) = a$, $\sigma(t,x,a) \equiv 1$, so that for all $\gamma \in \Gamma$, X^γ is given by

$$X_t^\gamma = x_0 + \int_0^t \alpha_s^\gamma ds + B_t^\gamma, \quad \text{where } \mathbb{E}^{\mathbb{P}^\gamma}\Big[\int_0^T \big(\alpha_s^\gamma\big)^2 ds\Big] < \infty. \tag{4.24}$$

Given $T > 0$ and $(a_T, b_T, c_T) \in \mathbb{R}^3$, we define the value function

$$V(x_0; T, a_T, b_T, c_T) \;:=\; \inf_{\gamma\in\Gamma}\mathbb{E}\Big[\int_0^T \frac{1}{2}(\alpha_t^\gamma)^2 dt + \frac{1}{2}a_T(X_T^\gamma)^2 + b_T X_T^\gamma + c_T\Big].$$

With the above linear quadratic structure, one can check that, if $1 + Ta_T > 0$, then

$$V(x_0; T, a_T, b_T, c_T) = \frac{1}{2} a_0 x_0^2 + b_0 x_0 + c_0,$$

where

$$a_0 = \frac{a_T}{1 + Ta_T}, \quad b_0 = \frac{b_T}{1 + Ta_T}, \quad c_0 = c_T - \frac{T b_T^2}{2(1 + Ta_T)} + \frac{1}{2} \ln(1 + Ta_T). \tag{4.25}$$

If $1 + Ta_T \leq 0$, then $V(T, a_T, b_T, c_T) = -\infty$.

Example 4.1. We consider the constrained optimization problem

$$\inf_{\gamma \in \Gamma} \mathbb{E}^{\mathbb{P}^\gamma} \Big[\int_0^1 \frac{1}{2} (\alpha_t^\gamma)^2 dt + (X_T^\gamma)^2 \Big], \quad \text{subject to: } \mathbb{E}^{\mathbb{P}^\gamma} \big[-X_T^\gamma + 1 \big] = 0,$$

where X^γ follows (4.24) with $x_0 = 0$. The dual criterion $d(\lambda)$ can be explicitly calculated with formula (4.25), applied with $a_1 = 2$, $b_1 = -\lambda$, and $c_1 = \lambda$. We obtain $d(\lambda) = -\frac{1}{6}\lambda^2 + \lambda + \frac{1}{2}\ln(3)$. The function d reaches its maximum $\frac{3}{2} + \frac{1}{2}\ln(3)$ at $\lambda^* = 3$.

Example 4.2. We next consider the problem

$$\inf_{\gamma \in \Gamma} \mathbb{E}^{\mathbb{P}^\gamma} \Big[\int_0^1 \frac{1}{2} (\alpha_t^\gamma)^2 dt + (X_T^\gamma)^2 \Big], \quad \text{subject to: } \mathbb{E}^{\mathbb{P}^\gamma} \big[(X_T^\gamma - 1)^2 \big] - \frac{1}{2} \leq 0,$$

where X^γ follows (4.24) with $x_0 = 0$. The dual criterion, obtained by applying (4.25) with $a_1 = 2 + 2\lambda$, $b_1 = -2\lambda$, $c_1 = \frac{1}{2}\lambda$, is given by

$$d(\lambda) = \begin{cases} -\frac{2\lambda^2}{3+2\lambda} + \frac{1}{2}\lambda + \frac{1}{2}\ln(3 + 2\lambda), & \text{if } \lambda \geq -3/2, \\ -\infty, & \text{otherwise.} \end{cases}$$

The function d reaches its maximum (approximately) 0.91 at $\lambda^* = -1 + \frac{1}{2}\sqrt{19} \approx 1.12$.

We implement the numerical procedure as described in Section 3.3, combined with the finite difference method for the discretization of the process. For $\Delta t = h > 0$, we set $\Delta x = \sqrt{\Delta t / 0.1}$. The results are presented on Figure 1 (with a logarithmic scale), for values of h ranging from 2×10^{-5} to 10^{-2}. The dual problem has been solved at a high precision, so that $|\nabla d_h(\lambda)| < 10^{-7}$. A rate of convergence equal to $1/2$ can be observed for $|V(0) - V_h(0)|$, which is consistent with the convergence results of numerical methods for BSDEs, see Remark 3.7. Interestingly, the approximate Lagrange multiplier λ_h converges at the same rate on these two examples.

5 Proofs

5.1 An abstract duality result

We consider in this subsection an abstract formulation of the optimal control problem with expectation constraints. This formulation will facilitate the presentation of the so called Lagrange relaxation approach, which is at the core of the proof of Theorem 3.3. Let us refer the reader to [16, Chapter XII] for a detailed presentation of this classical approach. For completeness, we first recall some basic definitions in convex analysis. For a convex function $f \colon \mathbb{R}^{m+\ell} \to \mathbb{R}$, the subdifferential $\partial f(x)$ is defined by

$$\partial f(x) = \big\{ \lambda \in \mathbb{R}^{m+\ell} \ : f(y) \geq f(x) + \langle \lambda, y - x \rangle, \ \forall y \in \mathbb{R}^{m+\ell} \big\}.$$

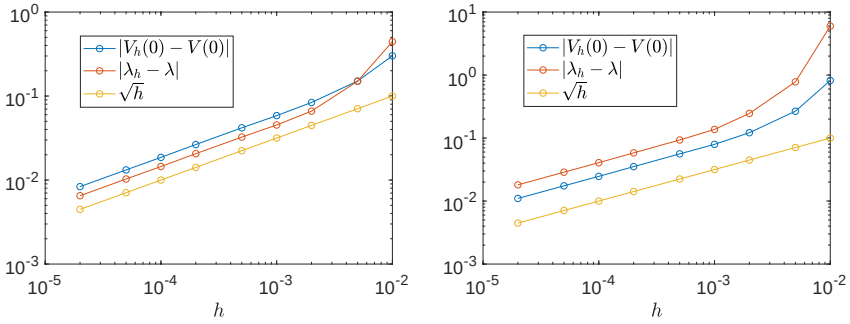

Figure 1: Approximation error for different discretization parameters for examples 1 (left) and 2 (right).

The superdifferential of a concave function is defined similarly, by changing the inequality sign in the above inequality. The conjugate $f^*\colon \mathbb{R}^{m+\ell} \to \mathbb{R}$ of the function f is defined by

$$f^*(\lambda) = \sup_{x \in \mathbb{R}^{m+\ell}} \Big(\langle \lambda, x \rangle - f(x) \Big).$$

All along this subsection, we consider an abstract measurable space $(\Omega, \mathcal{F})$, equipped with a convex subset $\mathcal{P}$ of $\mathcal{P}(\Omega)$, and some random variables ξ and $\zeta_1, ..., \zeta_{m+\ell}$. We make the following assumptions:

(A1) For all $\mathbb{P} \in \mathcal{P}$ and for all $i = 1, ..., m+\ell$, we have $\mathbb{E}^{\mathbb{P}}[|\xi|] < \infty$ and $\mathbb{E}^{\mathbb{P}}[|\zeta_i|] < \infty$;

(A2) $M := \sup_{\mathbb{P} \in \mathcal{P}} \left| \mathbb{E}^{\mathbb{P}}[\xi] \right| < \infty$;

(A3) There exists $\varepsilon > 0$ such that the following inclusion holds true:

$$B_\varepsilon(0) \subseteq \left\{ z \in \mathbb{R}^{m+\ell} \ : \exists \mathbb{P} \in \mathcal{P},\, \mathbb{E}^{\mathbb{P}}[\zeta] + z \in \{0_m\} \times \mathbb{R}^\ell_- \right\},$$

where $B_\varepsilon(0)$ denotes the ball of radius ε and center 0 for the supremum norm.

We aim at analyzing the following problem, referred to in this subsection as primal problem:

$$V(0) := \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^{\mathbb{P}}[\xi], \quad \text{subject to: } \begin{cases} \mathbb{E}^{\mathbb{P}}[\zeta_i] = 0, & \forall i = 1, ..., m, \\ \mathbb{E}^{\mathbb{P}}[\zeta_i] \leq 0, & \forall i = m+1, ..., m+\ell. \end{cases} \tag{5.26}$$

Observe that this is a convex problem, since $\mathcal{P}$ is assumed to be convex and the expectation is linear with respect to the involved probability measure. We next introduce the dual criterion $d\colon \mathbb{R}^{m+\ell} \to \mathbb{R}$, defined as follows:

$$d(\lambda) := \inf_{\mathbb{P} \in \mathcal{P}} \mathcal{L}(\mathbb{P}, \lambda), \quad \text{where } \mathcal{L}(\mathbb{P}, \lambda) := \mathbb{E}^{\mathbb{P}}[\xi] + \langle \lambda, \mathbb{E}^{\mathbb{P}}[\zeta] \rangle \text{ is the Lagrangian.}$$

Finally, the dual problem to (5.26) is given by

$$D(0) := \sup_{\lambda \in \mathbb{R}^{m+\ell}} d(\lambda), \quad \text{subject to: } \lambda \in \{0_m\} \times \mathbb{R}^\ell_+. \tag{5.27}$$

The main result of the subsection is the following proposition.

Proposition 5.1. *The primal problem* (5.26) *and the dual problem* (5.27) *have the same value. Moreover, the set of solutions to the dual problem is non-empty, coincide with* $\partial V(0)$*, and is bounded. More precisely, any dual optimizer* λ^* *satisfies*

$$\|\lambda\|_1 := \sum_{i=1}^{m+\ell} |\lambda_i| \leq \frac{2M}{\varepsilon}. \tag{5.28}$$

Proof. The key idea of the approach consists in parametrizing Problem (5.26) by introducing a variable z in the constraints. We introduce the function $V\colon \mathbb{R}^{m+\ell} \to \mathbb{R}$, defined as follows:

$$V(z) = \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\xi], \quad \text{subject to: } \mathbb{E}^{\mathbb{P}}[\zeta] + z \in \{0_m\} \times \mathbb{R}^{\ell}_{-}. \tag{5.29}$$

Step 1: The function V *is convex and finite in a neighborhood of 0.* Take $z^{(1)}, z^{(2)} \in \mathbb{R}^{m+\ell}$, and $\mathbb{P}_1, \mathbb{P}_2 \in \mathbb{P}$ such that $\mathbb{E}^{\mathbb{P}_i}[\zeta] + z^{(i)} \in \{0_m\} \times \mathbb{R}^{\ell}_{-}$, $i = 1, 2$. Let $\theta \in [0, 1]$. Let us set $z = \theta z^{(1)} + (1-\theta) z^{(2)}$ and $\mathbb{P} = \theta\mathbb{P}_1 + (1-\theta)\mathbb{P}_2$. By convexity of $\mathcal{P}$, we have that $\mathbb{P} \in \mathcal{P}$. Moreover,

$$\mathbb{E}^{\mathbb{P}}[\zeta] + z = \theta\big(\mathbb{E}^{\mathbb{P}_1}[\zeta] + z^{(1)}\big) + (1-\theta)\big(\mathbb{E}^{\mathbb{P}_2}[\zeta] + z^{(2)}\big) \in \{0_m\} \times \mathbb{R}^{\ell}_{-}.$$

Therefore,

$$V(z) \leq \mathbb{E}^{\mathbb{P}}[\xi] \leq \theta\mathbb{E}^{\mathbb{P}_1}[\xi] + (1-\theta)\mathbb{E}^{\mathbb{P}_2}[\xi].$$

Minimizing the right-hand side with respect to $\mathbb{P}_1$ and $\mathbb{P}_2$, we finally obtain that $V(z) \leq \theta V(z^{(1)}) + (1-\theta)V(z^{(2)})$, which concludes the proof of convexity. By Assumption (A2), the function V is bounded from below by $-M$. By Assumption (A3), we have that for all $z \in B_\varepsilon(0)$, the optimization problem associated with $V(z)$ is feasible, thus $V(z) < \infty$. This proves that V is finite in a neighbourhood of 0.

Step 2: Calculation of the conjugate function V^**.* For $\lambda \in \mathbb{R}^{m+\ell}$, we have

$$\begin{aligned} V^*(\lambda) &= \sup_{z\in\mathbb{R}^{m+\ell}} \big(\langle\lambda, z\rangle - V(z)\big) \\ &= \sup_{z\in\mathbb{R}^{m+\ell},\, \mathbb{P}\in\mathcal{P}} \langle\lambda, z\rangle - \mathbb{E}^{\mathbb{P}}[\xi], \quad \text{subject to: } \mathbb{E}^{\mathbb{P}}[\zeta] + z \in \{0_m\} \times \mathbb{R}^{\ell}_{-}. \end{aligned}$$

We make the changes of variable $z' = \mathbb{E}^{\mathbb{P}}[\zeta] + z$. Observing that

$$\langle\lambda, z\rangle - \mathbb{E}^{\mathbb{P}}[\xi] = \langle\lambda, z'\rangle - \mathcal{L}(\mathbb{P}, \lambda),$$

we deduce that

$$V^*(\lambda) = \sup_{z'\in\{0_m\}\times\mathbb{R}^{\ell}_{-}} \langle\lambda, z'\rangle - \inf_{\mathbb{P}\in\mathcal{P}} \mathcal{L}(\mathbb{P}, \lambda).$$

For all $\lambda \in \mathbb{R}^{m+\ell}$, we have

$$\sup_{z'\in\{0_m\}\times\mathbb{R}^{\ell}_{-}} \langle\lambda, z'\rangle = \begin{cases} 0, & \text{if } \lambda \in \mathbb{R}^m \times \mathbb{R}^{\ell}_{+}, \\ +\infty, & \text{otherwise.} \end{cases}$$

Therefore,

$$V^*(\lambda) = \begin{cases} -\inf_{\mathbb{P}\in\mathcal{P}} \mathcal{L}(\mathbb{P}, \lambda) = -d(\lambda), & \text{if } \lambda \in \mathbb{R}^m \times \mathbb{R}^{\ell}_{+}, \\ +\infty, & \text{otherwise.} \end{cases}$$

Step 3: Duality result. As a consequence of [5, Proposition 2.108], the mapping V is continuous at 0, since it is convex and finite in a neighborhood of 0. It follows with [5, Proposition 2.126] that $\partial V(0)$ is non-empty. It follows further with [5, Proposition 2.118] that $V(0) = V^{**}(0)$. Finally, the equivalence in [5, Relation 2.232] ensures that for all $\lambda \in \mathbb{R}^{m+\ell}$,

$$V(0) + V^*(\lambda) = \langle \lambda, 0 \rangle = 0 \Longleftrightarrow \lambda \in \partial V(0).$$

Therefore, we have

$$V(0) = V^{**}(0) = \sup_{\lambda \in \mathbb{R}^{m+\ell}} \langle 0, \lambda \rangle - V^*(\lambda) = \sup_{\lambda \in \mathbb{R}^m \times \mathbb{R}^\ell_+} d(\lambda),$$

proving that the primal and the dual problems have the same value. For all $\lambda \in \partial V(0)$, we have $V(0) = -V^*(\lambda)$, thus $V^*(\lambda)$ is finite, and necessarily, we have $\lambda \in \mathbb{R}_m \times \mathbb{R}^\ell_+$ and $V^*(\lambda) = -d(\lambda)$, which proves that λ is a solution to the dual problem. Conversely, if λ is a solution to the dual problem, then we have $V(0) = d(\lambda) = -V^*(\lambda)$ and therefore $\lambda \in \partial V(0)$.

Step 4: Boundedness of the set of dual solutions. Let $\lambda^* \in \partial V(0)$. Let z be defined by $z_i = \text{sign}(\lambda^*_i)\varepsilon$. By Assumptions (A2) and (A3), we have $V(z) \leq M$ and $V(0) \geq -M$. We have

$$M \geq V(z) \geq V(0) + \langle \lambda^*, z \rangle \geq -M + \varepsilon \sum_{i=1}^{m+\ell} |\lambda^*_i| = -M + \varepsilon \|\lambda^*\|_1.$$

Then it follows that (5.28) holds true. □

As consequence of Proposition 5.1, we obtain the following optimality conditions.

Corollary 5.2. *For all solutions $\mathbb{P}^*$ to the primal problem* (5.26) *and for all solutions λ^* to the dual problem* (5.27)*, one has*

$$\mathcal{L}(\mathbb{P}^*, \lambda^*) = \min_{\mathbb{P} \in \mathcal{P}} \mathcal{L}(\mathbb{P}, \lambda^*). \tag{5.30}$$

Moreover, for each $i = m+1, ..., m+\ell$, one has $\lambda^*_i = 0$ whenever $\mathbb{E}^{\mathbb{P}^*}[\zeta_i] < 0$, i.e. $\langle \lambda^*, \mathbb{E}^{\mathbb{P}^*}[\zeta] \rangle = 0$.

Proof. Since $\mathbb{E}^{\mathbb{P}^*}[\zeta] \in \{0_m\} \times \mathbb{R}^\ell_-$ and $\lambda^* \in \mathbb{R}^m \times \mathbb{R}^\ell_+$, we have

$$\mathcal{L}(\mathbb{P}^*, \lambda^*) = \mathbb{E}^{\mathbb{P}^*}[\xi] + \langle \lambda^*, \mathbb{E}^{\mathbb{P}^*}[\zeta] \rangle \leq \mathbb{E}^{\mathbb{P}^*}[\xi].$$

By Proposition 5.1, $\lambda^* \in \partial V(0)$. Therefore,

$$\mathcal{L}(\mathbb{P}^*, \lambda^*) \leq V(0) = -V^*(\lambda^*) = d(\lambda^*) = \inf_{\mathbb{P} \in \mathcal{P}} \mathcal{L}(\mathbb{P}, \lambda^*).$$

Of course, since $\mathbb{P}^* \in \mathcal{P}$, we also have that $\mathcal{L}(\mathbb{P}^*, \lambda^*) \geq \inf_{\mathbb{P} \in \mathcal{P}} \mathcal{L}(\mathbb{P}, \lambda^*)$. As a consequence, all the above inequalities are equalities. Equality (5.30) is proved, and we also have that $\langle \lambda^*, \mathbb{E}^{\mathbb{P}^*}[\zeta] \rangle = 0$, from which the complementarity condition easily follows. □

The necessary optimality conditions of Corollary 5.2 are also sufficient optimality conditions.

Lemma 5.3. *Let $\mathbb{P}^* \in \mathcal{P}$. Assume that there exists $\lambda^* \in \mathbb{R}^m \times \mathbb{R}^\ell_+$ such that $\langle \lambda^*, \mathbb{E}^{\mathbb{P}^*}[\zeta] \rangle = 0$ and such that*

$$\mathcal{L}(\mathbb{P}^*, \lambda^*) = \min_{\mathbb{P} \in \mathcal{P}} \mathcal{L}(\mathbb{P}, \lambda^*).$$

Then $\mathbb{P}^*$ is a solution to (5.26).

Proof. Let $\mathbb{P}$ be feasible for (5.26). Then the following inequalities and equality can be easily verified: $\mathbb{E}^{\mathbb{P}}[\xi] \geq \mathcal{L}(\mathbb{P}, \lambda^*) \geq \mathcal{L}(\mathbb{P}^*, \lambda^*) = \mathbb{E}^{\mathbb{P}^*}[\xi]$, which proves the lemma. □

We give now a property of the superdifferential of d, which in particular can be used in a numerical algorithm to solve the dual problem.

Lemma 5.4. *The map d is concave. Moreover, for all $\lambda \in \mathbb{R}^{m+\ell}$, for all $\mathbb{P} \in \mathcal{P}$ such that $d(\lambda) = \mathcal{L}(\mathbb{P}, \lambda)$, the vector $z := \mathbb{E}^{\mathbb{P}}[\zeta]$ lies in the superdifferential of d at λ.*

Proof. The Lagrangian $\mathcal{L}$ is affine with respect to λ, therefore concave. The mapping $d(\lambda)$ is expressed as an infimum of concave functions, thus it is concave. Let $\lambda \in \mathbb{R}^{m+\ell}$. Let $\mathbb{P}$ be such that $d(\lambda) = \mathcal{L}(\mathbb{P}, \lambda)$. For all $\lambda' \in \mathbb{R}^{m+\ell}$, we have

$$\mathcal{L}(\mathbb{P}, \lambda') = \mathcal{L}(\mathbb{P}, \lambda) + \langle \lambda' - \lambda, \mathbb{E}^{\mathbb{P}}[\zeta] \rangle = d(\lambda) + \langle \lambda' - \lambda, z \rangle.$$

Minimizing the left-hand side w.r.t. $\mathbb{P}$, we obtain that $d(\lambda') \leq d(\lambda) + \langle \lambda' - \lambda, z \rangle$, as was to be proved. □

We finish this subsection with a technical result concerning the Lipschitz continuity of the function V defined in (5.29).

Lemma 5.5. *The map V is Lipschitz continuous with modulus $\frac{4M}{\varepsilon}$ on $B_{\varepsilon/2}(0)$.*

Proof. Let z_1 and $z_2 \in B_{\varepsilon/2}(0)$ with $z_1 \neq z_2$. Take an arbitrary $\mathbb{P}_1 \in \mathcal{P}$ such that $\mathbb{E}^{\mathbb{P}_1}[\zeta] + z_1 \in \{0_m\} \times \mathbb{R}^{\ell}_{-}$ (whose existence is ensured by Assumption (A3)). Consider the half-line $\{z_1 + \theta(z_2 - z_1) \,|\, \theta \geq 0\}$. It has a unique intersection point z with the boundary of $B_{\varepsilon}(0)$. Let $\theta \geq 0$ be such that $z = z_1 + \theta(z_2 - z_1)$. Since $\|z\|_\infty = \varepsilon$ and $\|z_1\|_\infty \leq \varepsilon/2$, we have $\|z - z_1\|_\infty \geq \varepsilon/2$ and thus

$$\theta \;=\; \frac{\|z - z_1\|_\infty}{\|z_2 - z_1\|_\infty} \;\geq\; \frac{\varepsilon}{2\|z_2 - z_1\|_\infty}.$$

We also have $\theta > 1$, since the whole segment $\{z_1 + \theta(z_2 - z_1) \,:\, \theta \in [0,1]\}$ is included in $B_{\varepsilon/2}(0)$. By (A3), there exists $\mathbb{P} \in \mathcal{P}$ such that $\mathbb{E}^{\mathbb{P}}[\zeta] + z \in \{0_m\} \times \mathbb{R}^{\ell}_{-}$. Let us define $\mathbb{P}_2 = (1 - 1/\theta)\mathbb{P}_1 + (1/\theta)\mathbb{P}$. Since $1/\theta \leq 1$ and since $\mathcal{P}$ is convex, $\mathbb{P} \in \mathcal{P}$. Moreover, we have $z_2 = (1 - 1/\theta)z_1 + (1/\theta)z$, thus $\mathbb{E}^{\mathbb{P}_2}[\zeta] + z_2 \in \{0_m\} \times \mathbb{R}^{\ell}_{-}$. It follows that

$$V(z_2) \leq \mathbb{E}^{\mathbb{P}_2}[\xi] = \mathbb{E}^{\mathbb{P}_1}[\xi] + \frac{1}{\theta}\big(\mathbb{E}^{\mathbb{P}}[\xi] - \mathbb{E}^{\mathbb{P}_1}[\xi]\big).$$

Using (A2) and $1/\theta \leq 2\|z_2 - z_1\|_\infty/\varepsilon$, we finally obtain that

$$V(z_2) \leq V(z_1) + \frac{4M}{\varepsilon}\|z_2 - z_1\|_\infty,$$

which concludes the proof. □

5.2 Reformulation of the constrained control problem on the canonical space

In order to prove the duality result in Theorem 3.3, we will reformulate the continuous time and discrete time constrained control problems (2.7) and (2.12) in the framework of (5.26) in Section 5.1. Concretely, we will reformulate problems (2.7) and (2.12) as optimization problems over a space of probability measures on an appropriate canonical space. Moreover, the space of probability measures enjoys a good closeness property, which also plays an essential role to prove the approximation results in Theorem 3.8.

Recall that $\Omega := C([0,T],\mathbb{R}^n)$ denotes the canonical space of all $\mathbb{R}^n$–valued continuous paths on $[0,T]$, with canonical process $X = (X_t)_{0\le t\le T}$, and canonical filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$. Denote by $\mathbb{A}$ the collection of all Borel (positive) measures m on $[0,T]\times A$ whose marginal distribution on $[0,T]$ is the Lebesgue measure ds, i.e., $\mathbb{A} := \{m : m(ds,da) = m(s,da)ds\}$ for a measurable family of $(m(s,da))_{s\in[0,T]}$ of Borel probability measures on A. We also consider a subset $\mathbb{A}_0 \subset \mathbb{A}$, which consists of all measures $m(ds,da)$ such that $m(ds,da) = \delta_{\psi_s}(da)ds$ for some Borel measurable function $\psi\colon [0,T]\to A$. Let us denote by Λ the canonical element on $\mathbb{A}$ and denote

$$\Lambda_t(\phi) := \int_0^t \int_A \phi(s,a)\Lambda(ds,da)$$

for any bounded Borel measurable functions defined on $[0,T]\times A$. We then introduce a canonical filtration $\mathbb{F}^A = (\mathcal{F}^A_t)_{0\le t\le T}$ on $\mathbb{A}$ by

$$\mathcal{F}^A_t := \sigma\big\{\Lambda_s(\phi)\ : s\le t,\ \phi\in C_b([0,T]\times A)\big\}.$$

Next, we introduce an enlarged canonical space $\overline{\Omega} := \Omega\times\mathbb{A}$, which inherits the canonical processes (X,Λ), and is equipped with the enlarged canonical filtration $\overline{\mathbb{F}} = (\overline{\mathcal{F}}_t)_{0\le t\le T}$, defined by $\overline{\mathcal{F}}_t := \sigma\big(X_s,\Lambda_s(\phi)\ : s\le t,\ \phi\in C_b([0,T]\times U)\big)$. Denote also $\overline{\mathcal{F}} := \overline{\mathcal{F}}_T$, and by $\mathcal{P}(\overline{\Omega})$ the set of all probability measures on $(\overline{\Omega},\overline{\mathcal{F}})$, equipped with the weak convergence topology.

We now introduce a Wasserstein distance on a subspace of $\mathcal{P}(\overline{\Omega})$. Let $\mathcal{P}([0,T]\times A)$ denote the space of all probability measures m on $[0,T]\times A$, and $\mathcal{P}_2([0,T]\times A)$ the subspace of measures m such that $\int_0^T\int_A \big(\rho(a,a_0)\big)^2 m(ds,da) < \infty$, define the Wasserstein distance d_2 on $\mathcal{P}_2([0,T]\times A)$ by

$$d_2(m,m') := \Big(\inf_{\pi\in\Pi(m,m')}\int_0^T\int_A\int_0^T\int_A \big(\rho(a,a')^2 + |s-s'|^2\big)\pi(ds,da,ds',da')\Big)^{1/2},$$

where $\Pi(m,m')$ denotes the collection of joint distributions on $([0,T]\times A)\times([0,T]\times A)$ with marginal distribution m and m'. We can then extend the Wasserstein distance d_2 on $\mathbb{A}_2 := \{m\in\mathbb{A}\ : m/T\in\mathcal{P}_2([0,T]\times A)\}$ by

$$d_{\mathbb{A}_2}(m,m') := d_2(m/T, m'/T),\quad \text{for all } m,m'\in\mathbb{A}_2.$$

Now, let us denote $\mathcal{P}_2(\overline{\Omega})$ the set of all $\overline{\mathbb{P}}\in\mathcal{P}(\overline{\Omega})$, such that

$$\mathbb{E}^{\overline{\mathbb{P}}}\Big[\|X\|^2 + \int_0^T\int_A \rho(a_0,a)^2\Lambda(ds,da)\Big] < \infty. \tag{5.31}$$

It is easy to check that $\Lambda\in\mathbb{A}_2$, $\overline{\mathbb{P}}$–a.s. for all $\overline{\mathbb{P}}\in\mathcal{P}_2(\overline{\Omega})$. We then introduce the Wasserstein distance $\mathcal{W}_2$ on $\mathcal{P}_2(\overline{\Omega})$ by

$$\mathcal{W}_2(\overline{\mathbb{P}},\overline{\mathbb{P}}') := \Big(\inf_{\pi\in\Pi(\overline{\mathbb{P}},\overline{\mathbb{P}}')}\int_{\overline{\Omega}}\int_{\overline{\Omega}} \big(\|\omega-\omega'\|^2 + d^2_{\mathbb{A}_2}(m,m')\big)\pi\big(d(\omega,m),d(\omega',m')\big)\Big)^{1/2}. \tag{5.32}$$

Reformulation of the continuous time control problem Following El Karoui, Huu Nguyen and Jeanblanc [14], we can in fact reformulate the control problem on the canonical space by the martingale problem. Recall the definition of strong and weak control terms in Definitions 2.1 and 3.9. Given a function $\varphi\in C^2_b(\mathbb{R}^n)$, let us define an $\overline{\mathbb{F}}$-adapted continuous function $(C_t(\varphi))_{t\in[0,T]}$ by

$$C_t(\varphi) := \varphi(X_t) - \int_0^t\int_A \mathcal{L}\varphi(s,X_{s\wedge\cdot},a,X_s)\Lambda(ds,da),$$

with

$$\mathcal{L}\varphi(s,\omega,a,x) \;:=\; \mu(t,\omega_{t\wedge\cdot},a)\cdot D\varphi(x) + \frac{1}{2}\mathrm{Tr}\Big[\sigma\sigma^T(t,\omega_{t\wedge\cdot},a)D^2\varphi(x)\Big].$$

Definition 5.6. (i) A probability measure $\overline{\mathbb{P}}$ on $(\overline{\Omega},\overline{\mathcal{F}})$ is called a relaxed control rule if

$$\overline{\mathbb{P}}[X_0 = x_0] = 1, \quad \mathbb{E}^{\overline{\mathbb{P}}}\Big[\|X\|^2 + \int_0^T\int_A \rho(a_0,a)^2\Lambda(ds,da)\Big] < \infty, \tag{5.33}$$

and the process $(C_t(\varphi))_{t\in[0,T]}$ is a $(\overline{\mathbb{F}},\overline{\mathbb{P}})$-martingale for all functions $\varphi \in C_b^2(\mathbb{R}^n)$.

(ii) A probability measure $\overline{\mathbb{P}}$ is called a weak control rule if there exists a weak control term $\gamma\in\Gamma$ such that $\overline{\mathbb{P}} = \mathbb{P}^\gamma\circ\big(X^\gamma,\Lambda^\gamma\big)^{-1}$ with $\Lambda^\gamma(ds,da) := \delta_{\alpha_s^\gamma}(da)ds$.

(iii) A weak control rule $\overline{\mathbb{P}}$ is called a strong control rule if it is induced by a strong control term $\gamma\in\Gamma_S$, and it is called a piecewise strong control rule if $\gamma\in\Gamma_{S,0}$.

Let us denote

$$\begin{aligned}&\overline{\mathcal{P}}_R \text{ (resp. } \overline{\mathcal{P}}_W,\ \overline{\mathcal{P}}_S,\ \overline{\mathcal{P}}_{S,0})\\ &\quad:= \big\{\text{All relaxed (resp. weak, strong, piecewise strong) control rules}\big\},\end{aligned} \tag{5.34}$$

and, for $z\in\mathbb{R}^{m+\ell}$,

$$\begin{aligned}\overline{\mathcal{P}}_R(z) := \big\{\overline{\mathbb{P}}\in\overline{\mathcal{P}}_R \;&: \mathbb{E}^{\overline{\mathbb{P}}}[\Psi_i(X_\cdot)] = z_i,\ i=1,\cdots,m,\\ &\mathbb{E}^{\overline{\mathbb{P}}}[\Psi_{m+j}(X_\cdot)] \le z_{m+j},\ j=1,\cdots,\ell\big\},\end{aligned}$$

and

$$\overline{\mathcal{P}}_W(z) := \overline{\mathcal{P}}_W\cap\overline{\mathcal{P}}_R(z).$$

Consequently, one can redefine $V(z)$ in (2.7), and $d(\lambda)$ in (2.9) by

$$V(z) = \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}_W(z)}\mathbb{E}^{\overline{\mathbb{P}}}\big[\Phi(X_\cdot)\big], \text{ and } d(\lambda) = \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}_W}\mathbb{E}^{\overline{\mathbb{P}}}\big[\Phi(X_\cdot)+\lambda\cdot\Psi(X_\cdot)\big]. \tag{5.35}$$

Notice that the integrability condition (5.31) implies that $\overline{\mathcal{P}}_R\subset\mathcal{P}_2(\overline{\Omega})$.

Lemma 5.7. (i) *Both sets $\overline{\mathcal{P}}_W$ and $\overline{\mathcal{P}}_R$ are convex, and*

$$\overline{\mathcal{P}}_W \;=\; \big\{\overline{\mathbb{P}}\in\overline{\mathcal{P}}_R \;:\mathbb{P}[\Lambda\in\mathbb{A}_0]=1\big\}, \tag{5.36}$$

and $\overline{\mathcal{P}}_{S,0}\subset\overline{\mathcal{P}}_S\subset\overline{\mathcal{P}}_W\subset\overline{\mathcal{P}}_R$.

(ii) *Let Assumption 3.2 hold. Then $\overline{\mathcal{P}}_{S,0}$ is dense in $\overline{\mathcal{P}}_R$ under the Wasserstein $\mathcal{W}_2$ distance.*

Proof. (i) We first prove (5.36). Given a weak control term $\gamma\in\Gamma$, and let $\overline{\mathbb{P}} := \mathbb{P}^\gamma\circ(X^\gamma,\Lambda^\gamma)^{-1}$ with $\Lambda^\gamma := \delta_{\alpha_s^\gamma}(da)ds$, it is straightforward to check that $\overline{\mathbb{P}}\in\overline{\mathcal{P}}_R$ and satisfies $\overline{\mathbb{P}}[\Lambda\in\mathbb{A}_0] = \mathbb{P}^\gamma[\Lambda^\gamma\in\mathbb{A}_0] = 1$. On the other hand, let $\overline{\mathbb{P}}\in\overline{\mathcal{P}}_R$ such that $\overline{\mathbb{P}}[\Lambda\in\mathbb{A}_0]=1$, one can construct on a possible enlarged space of $\overline{\Omega}$, a $\overline{\mathbb{F}}$–predictable process α and a Brownian motion $W^{\overline{\mathbb{P}}}$ such that $\overline{\mathbb{P}}[\Lambda(da,ds) = \delta_{\alpha_s}(da)ds] = 1$ and

$$X_t = x_0 + \int_0^t\mu(s,X_s,\alpha_s)ds + \int_0^t\sigma(s,X_s,\alpha_s)dW_s^{\overline{\mathbb{P}}}, \;\; \overline{\mathbb{P}}\text{–a.s.}$$

It follows that the term $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{F}}, \overline{\mathbb{P}}, X, \alpha, W^{\overline{\mathbb{P}}})$ is a weak control term in Definition 2.6, and hence $\overline{\mathbb{P}} \in \overline{\mathcal{P}}_W$. This proves the equality in (5.36).

We next prove that $\overline{\mathcal{P}}_R$ is convex. Let $\overline{\mathbb{P}}_1, \overline{\mathbb{P}}_2 \in \overline{\mathcal{P}}_R$, and $\overline{\mathbb{P}} = \theta\overline{\mathbb{P}}_1 + (1-\theta)\overline{\mathbb{P}}_2$ for some $\theta \in [0,1]$. Then it is clear that $\overline{\mathbb{P}}$ satisfies (5.33) as $\overline{\mathbb{P}}_1$ and $\overline{\mathbb{P}}_2$. Further, as $(C_t(\varphi))_{t\in[0,T]}$ is a $\overline{\mathbb{F}}$–martingale under both $\overline{\mathbb{P}}_1$ and $\overline{\mathbb{P}}_2$, one has, for all $s \le t$ and $\overline{\mathcal{F}}_s$–measurable bounded r.v. ξ,

$$\begin{aligned}\mathbb{E}^{\overline{\mathbb{P}}}\big[\big(C_t(\varphi) - C_s(\varphi)\big)\xi\big] &= \theta\mathbb{E}^{\overline{\mathbb{P}}_1}\big[\big(C_t(\varphi) - C_s(\varphi)\big)\xi\big] + (1-\theta)\mathbb{E}^{\overline{\mathbb{P}}_2}\big[\big(C_t(\varphi) - C_s(\varphi)\big)\xi\big] \\ &= 0.\end{aligned}$$

Thus $(C_t(\varphi))_{t\in[0,T]}$ is also a $\overline{\mathbb{F}}$–martingale under $\overline{\mathbb{P}}$, and hence $\overline{\mathbb{P}} \in \overline{\mathcal{P}}_R$. Then $\overline{\mathcal{P}}_R$ is convex. Using (5.36) and the convexity of $\overline{\mathcal{P}}_R$, it follows that $\overline{\mathcal{P}}_W$ is also convex. Finally, the inclusion relation $\overline{\mathcal{P}}_{S,0} \subset \overline{\mathcal{P}}_S \subset \overline{\mathcal{P}}_W \subset \overline{\mathcal{P}}_R$ is trivial by their definitions and (5.36).

(ii) Finally, the approximation of a relaxed control rule by weak control rules is a classical result as illustrated by El Karoui, Huu Nguyen and Jeanblanc [14] under the weak convergence topology. For the density of $\overline{\mathcal{P}}_{S,0}$ in $\overline{\mathcal{P}}_R$ under $\mathcal{W}_2$, we can refer e.g. to Theorem 3.1 of Djete, Possamaï and Tan [12] for a proof with explicit construction. □

Reformulation of the discrete time control problem Notice that a discrete time process on grid $0 = t_0 < \cdots < t_N = T$ can be considered as a continuous time piecewise constant process on $[0,T]$. Concretely, let

$$\gamma = \big(\Omega^\gamma, \mathcal{F}^\gamma, \mathbb{F}^\gamma, \mathbb{P}^\gamma, \alpha^\gamma, X^\gamma, U_k^\gamma, k = 1, \cdots, N\big) \in \Gamma_h$$

be a weak discrete time control term, we (re-)define $(\widehat{X}^\gamma, \alpha^\gamma, \Lambda^\gamma)$ as processes on $[0,T]$ by

$$\widehat{X}_s^\gamma := (t_{k+1} - s)X_{t_k}^\gamma + (s - t_k)X_{t_{k+1}}^\gamma, \quad \alpha_s^\gamma := \alpha_{t_k}^\gamma, \quad \text{for all } s \in [t_k, t_{k+1}),$$

and $\Lambda^\gamma(ds, da) := \delta_{\alpha_s^\gamma}(da)ds$. Denote

$$\begin{aligned}\overline{\mathcal{P}}_W^h &:= \big\{\overline{\mathbb{P}}^h := \mathbb{P}^\gamma \circ \big(\widehat{X}^\gamma, \Lambda^\gamma\big)^{-1} \ : \gamma \in \Gamma_h\big\}, \\ \overline{\mathcal{P}}_W^h(z) &:= \big\{\overline{\mathbb{P}}^h := \mathbb{P}^\gamma \circ \big(\widehat{X}^\gamma, \Lambda^\gamma\big)^{-1} \ : \gamma \in \Gamma_h(z)\big\},\end{aligned}$$

so that V_h and d_h in (2.12) and (2.13) can be redefined by

$$V_h(z) = \sup_{\overline{\mathbb{P}}^h \in \overline{\mathcal{P}}_W^h(z)} \mathbb{E}^{\overline{\mathbb{P}}^h}\big[\Phi(X.)\big] \quad \text{and} \quad d_h(\lambda) = \sup_{\overline{\mathbb{P}}^h \in \overline{\mathcal{P}}_W^h} \mathbb{E}^{\overline{\mathbb{P}}^h}\big[\Phi(X.) + \lambda \cdot \Psi(X.)\big]. \quad (5.37)$$

Lemma 5.8. *The set $\overline{\mathcal{P}}_W^h$ is convex.*

Proof. Let us consider an enlarged canonical space $\widehat{\Omega} := \overline{\Omega} \times A^N \times [0,1]^N$ equipped with canonical element (X, Λ, α, U) with $\alpha = (\alpha_0, \cdots, \alpha_{N-1})$ and $U = (U_1, \cdots, U_N)$. The canonical filtration $\widehat{\mathbb{F}} = (\widehat{\mathcal{F}}_t)_{0\le t\le T}$ is defined by

$$\widehat{\mathcal{F}}_t := \sigma\big(X_s, \Lambda_s(\phi), \alpha_{t_i}, U_i \ : s \le t, \phi \in C_b([0,T] \times A), t_i \le t\big).$$

We define

$$\widehat{\mathcal{P}}_W^h := \big\{\mathbb{P}^\gamma \circ \big(X^\gamma, \Lambda^\gamma, \alpha^\gamma, U^\gamma\big)^{-1} \ : \gamma \in \Gamma_h\big\}, \ \text{ so that } \overline{\mathcal{P}}_W^h = \big\{\widehat{\mathbb{P}}|_{\overline{\Omega}} \ : \widehat{\mathbb{P}} \in \widehat{\mathcal{P}}_W^h\big\}.$$

Then to conclude the proof, it is enough to prove that $\widehat{\mathcal{P}}_W^h$ is convex, which implies immediately that $\overline{\mathcal{P}}_W^h$ is convex.

Let us re-define the process α on $[0,T]$ by $\alpha_s := \alpha_{t_k}$ for all $s \in [t_k, t_{k+1})$ for each $k = 0, \cdots, N-1$. We claim that $\widehat{\mathbb{P}} \in \widehat{\mathcal{P}}_W^h$ if and only if $\widehat{\mathbb{P}}[X_0 = x_0] = 1$, and for each $k = 1, \cdots, N$,

$$\begin{cases} X_{t_k} = X_{t_{k-1}} H_h(t_{k-1}, X_{t_{k-1}\wedge\cdot}, \alpha_{k-1}, U_k), \\ X_s = (t_{k+1} - s)X_{t_k} + (s - t_k)X_{t_{k+1}}, \quad s \in [t_k, t_{k+1}], & \widehat{\mathbb{P}}\text{–a.s.} \\ \Lambda(da, ds) = \delta_{\alpha_s}(da)ds, \text{ and } U_k \sim \mathcal{U}[0,1] \text{ is independent of } \widehat{\mathcal{F}}_{t_{k-1}}, \end{cases} \tag{5.38}$$

Indeed, for every $\gamma \in \Gamma_h$ and $\widehat{\mathbb{P}} = \mathbb{P}^\gamma \circ \left(X^\gamma, \Lambda^\gamma, \alpha^\gamma, U^\gamma\right)^{-1}$, it is straightforward to check that $\widehat{\mathbb{P}}$ satisfies $\widehat{\mathbb{P}}[X_0 = x_0] = 1$ and (5.38). Next, let $\widehat{\mathbb{P}}$ be a probability measure on $\widehat{\Omega}$ satisfying $\widehat{\mathbb{P}}[X_0 = x_0] = 1$ and (5.38), then it is easy to check that the term $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{F}}, \widehat{\mathbb{P}}, (X_{t_k})_{k=0,\cdots,N}, \alpha, U)$ consists of a discrete time weak control term, and hence $\widehat{\mathbb{P}} \in \widehat{\mathcal{P}}_W^h$.

We finally prove that $\widehat{\mathcal{P}}_W^h$ is convex. Let $\widehat{\mathbb{P}}_1, \widehat{\mathbb{P}}_2 \in \widehat{\mathcal{P}}_W^h$, and $\widehat{\mathbb{P}} := \theta\widehat{\mathbb{P}}_1 + (1-\theta)\widehat{\mathbb{P}}_2$ for some $\theta \in [0,1]$. Then $\widehat{\mathbb{P}}[X_0 = x_0] = \theta\widehat{\mathbb{P}}_1[X_0 = x_0] + (1-\theta)\widehat{\mathbb{P}}_2[X_0 = x_0] = 1$. Moreover, (5.38) holds true under $\widehat{\mathbb{P}}$, as it does under $\widehat{\mathbb{P}}_1$ and $\widehat{\mathbb{P}}_2$. Thus $\widehat{\mathbb{P}} \in \widehat{\mathcal{P}}_W^h$, which implies that $\widehat{\mathcal{P}}_W^h$ is convex, and so is $\overline{\mathcal{P}}_W^h$. □

5.3 Proof of the duality results in Theorem 3.3

Recall that $\overline{\mathcal{P}}_W$ and $\overline{\mathcal{P}}_W^h$ are convex by Lemmas 5.7 and 5.8, and (V, d) and (V_h, d_h) can be reformulated in (5.35) and (5.37) as in the framework of (5.26) in Section 5.1. Then Theorem 3.3 follows directly by Proposition 5.1, Corollary 5.2, and Lemma 5.3.

5.4 Proof of the approximation results in Theorem 3.8

To prove the general convergence result in Theorem 3.8, we will adapt the classical "weak convergence" arguments of Kushner and Dupuis [19] in our constrained context. In fact, we will work under the Wasserstein $\mathcal{W}_2$ distance. Note that Assumption 3.2 is in force in this subsection. Let us provide the proof in 4 steps.

Step 1 Recall that $\overline{\mathcal{P}}_R(z)$ is defined below (5.34), we will first prove that

$$V(0) = V_R(0), \quad \text{where } V_R(z) := \inf_{\overline{\mathbb{P}} \in \overline{\mathcal{P}}_R(z)} \mathbb{E}^{\overline{\mathbb{P}}}[\Phi(X_\cdot)]. \tag{5.39}$$

Lemma 5.9. *On $B_{\varepsilon/2}(0)$, the functions $V_R(\cdot)$, $V(\cdot)$, and $V_h(\cdot)$, for all $h > 0$, are Lipschitz with Lipschitz constant $\frac{4M}{\varepsilon}$.*

Proof. Notice that $\overline{\mathcal{P}}_R$ is convex, and $\overline{\mathcal{P}}_W$ is dense in $\overline{\mathcal{P}}_R$ under $\mathcal{W}_2$. Then it is easy to check that the set $\{\overline{\mathbb{P}}|_\Omega \ : \overline{\mathbb{P}} \in \overline{\mathcal{P}}_R\}$ satisfies all the preliminary conditions in Section 5.1, as $\{\overline{\mathbb{P}}|_\Omega \ : \overline{\mathbb{P}} \in \overline{\mathcal{P}}_W\}$ does. Then the Lipschitz property of V_R, V and V_h follows by Lemma 5.5. □

To prove the equality $V_R(0) = V(0)$ in (5.39), we first notice that $\overline{\mathcal{P}}_W(0) \subset \overline{\mathcal{P}}_R(0)$, and hence $V_R(0) \le V(0)$.

Next, let us fix $\delta > 0$, and $\overline{\mathbb{P}} \in \overline{\mathcal{P}}_R(0)$ be such that $\mathbb{E}^{\overline{\mathbb{P}}}[\Phi(X)] \ge V_R(0) - \delta$. By the density of $\overline{\mathcal{P}}_W$ in $\overline{\mathcal{P}}_R$ and the growth condition of Φ as well as Ψ_i in Assumption 3.2, there exists some $\overline{\mathbb{P}}' \in \overline{\mathcal{P}}_W$ such that

$$\left|\mathbb{E}^{\overline{\mathbb{P}}'}[\Phi(X)] - \mathbb{E}^{\overline{\mathbb{P}}}[\Phi(X)]\right| \le \delta, \quad \text{and } \left|\mathbb{E}^{\overline{\mathbb{P}}'}[\Psi_i(X)] - \mathbb{E}^{\overline{\mathbb{P}}}[\Psi_i(X)]\right| \le \delta, \ i = 1, \cdots, m+\ell.$$

This implies that for some $z \in B_\delta(0)$, one has $V(z) \le \mathbb{E}^{\overline{\mathbb{P}}'}[\Phi(X)] \le V_R(0) + 2\delta$. As $\delta > 0$ is arbitrary, and V_R and V are both Lipschitz on $B_{\varepsilon/2}(0)$, it follows that $V(0) \le V_R(0)$, and we hence conclude that $V_R(0) = V(0)$.

Step 2 We next prove that $\liminf_{h\to 0} V_h(0) \ge V(0)$. Let $(h_n)_{n\ge 1}$ be a sequence of positive real numbers such that $\lim_{n\to\infty} h_n = 0$ and $(\overline{\mathbb{P}}_n)_{n\ge 1}$ be a sequence such that $\overline{\mathbb{P}}_n \in \overline{\mathcal{P}}_W^{h_n}(0)$ for all $n \ge 1$, and

$$\liminf_{h\to 0} \; V_h(0) \;=\; \lim_{n\to\infty} \mathbb{E}^{\overline{\mathbb{P}}_n}\big[\Phi\big(X_\cdot\big)\big].$$

Lemma 5.10. *The sequence $(\overline{\mathbb{P}}_n)_{n\ge 1}$ is relatively compact under $\mathcal{W}_2$. Moreover, any limit point of $(\overline{\mathbb{P}}_n)_{n\ge 1}$ lies in $\overline{\mathcal{P}}_R(0)$.*

Proof. First, as A is compact, then $\mathbb{A}$ is also compact under the weak convergence topology. Further, for each $n \ge 1$, let us denote $t_k^n := kh_n$ for $k = 0, \cdots, N_n$, then for sequences $(\alpha_k^n, U_k^n)_{0\le k\le N_n - 1}$ (see (5.38))

$$X_{t_{k+1}^n} = X_{t_k^n} + H_{h_n}(t_k^n, X_{t_k^n\wedge\cdot}, \alpha_k^n, U_k^n), \quad \overline{\mathbb{P}}_n\text{–a.s.},$$

so that

$$\begin{aligned}\big|X_{t_{k+1}^n}\big|^2 \le 3x_0^2 + 3\Big|\sum_{i\le k} \mu(t_i^n, X_{t_i^n\wedge\cdot}, \alpha_i^n)h_n\Big|^2 \\ + 3\Big|\sum_{i\le k}\Big(H_{h_n}(t_i^n, X_{t_i^n\wedge\cdot}, \alpha_i^n, U_i^n) - \mu(t_i^n, X_{t_i^n\wedge\cdot}, \alpha_i^n)h_n\Big)\Big|^2.\end{aligned}$$

Let $S_k^n := \mathbb{E}^{\overline{\mathbb{P}}_n}\big[\sup_{0\le i\le k+1}|X_{t_i^n}|^2\big]$, using (2.4) and Assumption 3.2.(ii), it follows by direct computation that

$$S_{k+1}^n \;\le\; \sum_{i\le k} C(S_i^n + 1)h_n, \quad \text{for some constant } C \text{ independent of } n.$$

Then by the discrete time Gronwall lemma, one has

$$\sup_{n\ge 1} \mathbb{E}^{\overline{\mathbb{P}}_n}\Big[\sup_{0\le k\le N_n}\big|X_{t_k^n}\big|^2\Big] < \infty.$$

Consider the processes $\sum_{i\le k}\mu(t_i^n, X_{t_i^n\wedge\cdot}, \alpha_i^n)h_n$ and $\sum_{i\le k}\sigma\sigma^\top(t_i^n, X_{t_i^n\wedge\cdot}, \alpha_i^n)h_n$. It is easy to deduce that $(\overline{\mathbb{P}}_n|_\Omega)_{n\ge 1}$ is relatively compact under the weak convergence topology, by using Theorem 2.3 of [17]. Thus the sequence $(\overline{\mathbb{P}}_n)_{n\ge 1}$ is relatively compact under the weak convergence topology.

We can consider the 3rd moment to compute $\mathbb{E}^{\overline{\mathbb{P}}_n}\big[\big|\sum_{i\le k}\mu(t_i^n, X_{t_i^n\wedge\cdot}, \alpha_i^n)h_n\big|^3\big]$ and

$$\begin{aligned}&\mathbb{E}^{\overline{\mathbb{P}}_n}\Big[\sup_{k\le N_n}\Big|\sum_{i\le k}\Big(H_{h_n}(t_i^n, X_{t_i^n\wedge\cdot}, \alpha_i^n, U_i^n) - \mu(t_i^n, X_{t_i^n\wedge\cdot}, \alpha_i^n)h_n\Big)\Big|^3\Big] \\ &\le CT^{3/2}\mathbb{E}^{\overline{\mathbb{P}}_n}\Big[\Big(\frac{h_n}{T}\sum_{k\le N_n}\Big(H_{h_n}(t_k^n, X_{t_k^n\wedge\cdot}, \alpha_k^n, U_k^n) - \mu(t_i^n, X_{t_k^n\wedge\cdot}, \alpha_k^n)h_n\Big)^2\Big)^{3/2}\Big] \\ &\le CT^{3/2}\mathbb{E}^{\overline{\mathbb{P}}_n}\Big[\frac{h_n}{T}\sum_{k\le N_n}\Big|H_{h_n}(t_k^n, X_{t_k^n\wedge\cdot}, \alpha_k^n, U_k^n) - \mu(t_i^n, X_{t_k^n\wedge\cdot}, \alpha_k^n)h_n\Big|^3\Big],\end{aligned}$$

where the above two inequality follows by BDG inequality and Jensen's inequality. Then using again Assumption 3.2.(ii), one deduces that

$$\sup_{n\geq 1}\mathbb{E}^{\overline{\mathbb{P}}_n}\big[\|X\|^3\big]<\infty,$$

so that $(\overline{\mathbb{P}}_n)_{n\geq 1}$ is relatively compact w.r.t. the $\mathcal{W}_2$ distance (recall that A is compact).

Let $\overline{\mathbb{P}}$ be a limit of $(\overline{\mathbb{P}}_n)_{n\geq 1}$ w.r.t. the $\mathcal{W}_2$ distance. To prove that $\overline{\mathbb{P}}\in\overline{\mathcal{P}}_R$, it is enough to prove that, for any $\varphi\in C_b^2(\mathbb{R}^n)$, $0<s_1<\cdots s_k<s<t$, $f\in C_b(\mathbb{R}^{(n+p)\times k})$ and $\phi_j\in C_b([0,T]\times A)$, one has

$$\begin{aligned}&\mathbb{E}^{\overline{\mathbb{P}}}\Big[f(X_{s_i},\Lambda_{s_i}(\phi_j);\ 1\leq j\leq p,1\leq i\leq k)\big(C_t(\varphi)-C_s(\varphi)\big)\Big]\\&\quad=\lim_{n\to\infty}\mathbb{E}^{\overline{\mathbb{P}}_n}\Big[f(X_{s_i},\Lambda_{s_i}(\phi_j);\ 1\leq j\leq p,1\leq i\leq k)\big(C_t(\varphi)-C_s(\varphi)\big)\Big]=0.\end{aligned}$$

To prove the limit property, it is enough to notice that

$$\begin{aligned}&\mathbb{E}^{\overline{\mathbb{P}}_n}\Big[f(X_{s_i},\Lambda_{s_i}(\phi_j)\ ;1\leq j\leq p,1\leq i\leq k)\big(C_t(\varphi)-C_s(\varphi)\big)\Big]\\=\ &\mathbb{E}^{\mathbb{P}^{\gamma_{h_n}}}\Big[f\Big(\widehat{X}_{s_i}^{\gamma_{h_n}},\int_0^{s_i}\phi_j(\alpha_s^{\gamma_{h_n}})ds);j,i\Big)\\&\qquad\Big(\varphi\big(\widehat{X}_t^{\gamma_{h_n}}\big)-\varphi\big(\widehat{X}_s^{\gamma_{h_n}}\big)+\int_s^t\mathcal{L}\varphi\big(r,\widehat{X}_{r\wedge\cdot}^{\gamma_{h_n}},\alpha_r^{\gamma_{h_n}},\widehat{X}_r^{\gamma_{h_n}}\big)dr\Big)\Big].\end{aligned}$$

Applying Taylor formula on $\varphi\big(\widehat{X}_t^{\gamma_{h_n}}\big)-\varphi\big(\widehat{X}_s^{\gamma_{h_n}}\big)$, then using the dynamic of $X^{\gamma_{h_n}}$ in (2.11) and Assumption 3.2.(ii), it follows by direct computation that the limit equals to 0. Then $\overline{\mathbb{P}}\in\overline{\mathcal{P}}_R$.

Finally, $\overline{\mathbb{P}}$ is the limit of $(\overline{\mathbb{P}}_n)_{n\geq 1}$ w.r.t. the $\mathcal{W}_2$ distance and Ψ_i, $i=1,\cdots,m+\ell$ have all quadratic growth, thus $\mathbb{E}^{\overline{\mathbb{P}}_n}[\Psi_i(X_\cdot)]\to\mathbb{E}^{\overline{\mathbb{P}}}[\Psi_i(X)]$. As $\overline{\mathbb{P}}_n\in\overline{\mathcal{P}}_W^{h_n}(0)$, it follows that $\overline{\mathbb{P}}\in\overline{\mathcal{P}}_R(0)$. □

Let $(n_k)_{k\geq 1}$ be a subsequence such that $\overline{\mathbb{P}}_{n_k}\to\overline{\mathbb{P}}_*\in\overline{\mathcal{P}}_R(0)$, then

$$\liminf_{h\to 0}V_h(0)\ =\ \lim_{k\to\infty}\mathbb{E}^{\mathbb{P}_{n_k}}\big[\Phi\big(X_\cdot\big)\big]\ =\ \mathbb{E}^{\overline{\mathbb{P}}_*}\big[\Phi\big(X_\cdot\big)\big]\ \geq\ V(0).\tag{5.40}$$

Step 3 We prove here $\limsup_{h\to 0}V_h(0)\leq V(0)$.

Lemma 5.11. *Let $\overline{\mathbb{P}}\in\overline{\mathcal{P}}_R(0)$, then for any $\eta_0>0$, $\eta>0$ and $h_0>0$, there exist $h\leq h_0$, $z\in\mathbb{R}^m\times\mathbb{R}_+^\ell$ and a weak discrete time control γ_h such that $\|z\|_1\leq\eta$, $\gamma_h\in\Gamma_h(z)$ and*

$$\Big|\mathbb{E}^{\overline{\mathbb{P}}}\big[\Phi(X_\cdot)\big]-\mathbb{E}^{\mathbb{P}^{\gamma_h}}\big[\Phi\big(\widehat{X}_\cdot^{\gamma_h}\big)\big]\Big|\ \leq\ \eta_0.$$

Proof. Let $\overline{\mathbb{P}}\in\overline{\mathcal{P}}_R(0)$, $\eta_0>0$, $\eta>0$ and $h_0>0$. By Lemma 5.7, there is a piecewise constant strong control $\gamma^{*,1}\in\Gamma_{S,0}$ such that $\gamma^{*,1}\in\Gamma_S(z_1)$ for some $z_1\in\mathbb{R}^m\times\mathbb{R}_+^\ell$ satisfying $\|z_1\|_1\leq\eta/3$ and $\big|\mathbb{E}^{\overline{\mathbb{P}}}\big[\Phi(X_\cdot)\big]-\mathbb{E}^{\mathbb{P}^{\gamma^{*,1}}}\big[\Phi(X_\cdot^{\gamma^{*,1}})\big]\big|\leq\eta_0/3$.

Further, on the probability space $(\Omega^{\gamma^{*,1}},\mathcal{F}^{\gamma^{*,1}})$, one can approximate the piecewise constant control $\alpha^{\gamma^{*,1}}$ by piecewise constant control $\alpha^{*,2}$ such that

$$\gamma^{*,1}:=(\Omega^{\gamma^{*,1}},\mathcal{F}^{\gamma^{*,1}},\mathbb{F}^{\gamma^{*,1}},\mathbb{P}^{\gamma^{*,1}},X^{\gamma^{*,1}},B^{\gamma^{*,1}},\alpha^{\gamma^{*,2}})\in\mathcal{A}(z_2)$$

for some z_2 satisfying $\|z_2-z_1\|_1\leq\eta/3$, where each $k\geq 1$, $\alpha_{t_k}^{*,2}$ is a function of $(B_{t_i}^{\gamma^{*,1}}:i\leq k)$, which is almost surely continuous under the law of $(B_{t_i}^{\gamma^{*,1}}:i\leq k)$, and such that $\big|\mathbb{E}^{\mathbb{P}^{\gamma^{*,1}}}\big[\Phi(X_\cdot^{\gamma^{*,1}})\big]-\mathbb{E}^{\mathbb{P}^{\gamma^{*,1}}}\big[\Phi(X_\cdot^{\gamma^{*,2}})\big]\big|\leq\eta_0/3$.

Next, recall that $\alpha^{*,2}_{t_k}$ are a.s. continuous functions of $(B_{t_i} \ : i \le k)$ for all $k \ge 1$. One can consider a sequence of discrete time weak controls $(\gamma_n)_{n\ge1}$ such that $\gamma_n \in \Gamma_{h_n}$ for all $n \ge 1$, $h_n \to 0$, and moreover

$$\alpha^{\gamma_n}_{t_k} = \alpha^{*,2}_{t_k}(B^{\gamma_n}_{t_i} \ : i \le k),$$

where

$$B^{\gamma_n}_{t_i} := \sum_{j\le i-1} \sigma^{-1}(t_j, \widehat{X}^{\gamma_n}, \alpha^{\gamma_n}_{t_j})\Big(H_{h_n}((t_j, \widehat{X}^{\gamma_n}, \alpha^{\gamma_n}_{t_j}, U^{\gamma_n}_i) - \mu(t_j, \widehat{X}^{\gamma_n}, \alpha^{\gamma_n}_{t_j})h_n\Big).$$

Denote $\Lambda^{\gamma_n}(ds, da) := \delta_{\alpha^{\gamma_n}_s}(da)ds$, then by similar martingale problem arguments as in Lemma 5.10, one can check that

$$\overline{\mathbb{P}}^{\gamma_n} \circ \big(\widehat{X}^{\gamma_n}_\cdot, \widehat{B}^{\gamma_n}_\cdot, \Lambda^{\gamma_n}\big)^{-1} \ \to \ \mathbb{P}^* \circ \big(X^*_\cdot, B^*_\cdot, \delta_{\alpha^*_s}(da)da\big)^{-1},$$

for some probability space $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$, equipped with a Brownian motion B^* and X^* satisfies the SDE (2.6) with Brownian motion B^* and control process α^*, and moreover, α^* is a piecewise constant control process satisfying

$$\alpha^*_{t_k} = \alpha^{*,2}_{t_k}(B^*_{t_i} \ : i \le k).$$

Under the Lipschitz condition above (3.16), one has the uniqueness of the SDE (2.6), then it follows that $\mathbb{P}^{\gamma_n} \circ \big(\widehat{X}^{\gamma_n}_\cdot\big)^{-1} \to \mathbb{P}^* \circ (X^*)^{-1} = \mathbb{P}_0 \circ \big(X^{\alpha^{*,2}}_\cdot\big)^{-1}$ under $\mathcal{W}_2$, as $n \to \infty$. Notice that Φ and Ψ_i, $i = 1, \cdots, m+\ell$ are all of quadratic growth, then there is some discrete time weak control γ_h such that $\gamma_h \in \Gamma_h(z_3)$ for $h \le h_0$ and $\|z_3 - z_2\|_1 \le \eta/3$ and

$$\Big|\mathbb{E}^{\mathbb{P}_0}\big[\Phi(X^{\alpha^{*,2}}_\cdot)\big] - \mathbb{E}^{\mathbb{P}^{\gamma_h}}\big[\Phi\big(\widehat{X}^{\gamma_h}_\cdot\big)\big]\Big| \ \le \ \eta_0/3.$$

Finally, combining all the above estimates, it follows that γ_h is the required weak discrete time control. □

Let $\overline{\mathbb{P}}^* \in \overline{\mathcal{P}}_W(0)$ be an ε–optimal solution of the problem $V(0)$, i.e.

$$\mathbb{E}^{\overline{\mathbb{P}}^*}[\Phi(X_\cdot)] - \varepsilon \le V(0) = \inf_{\overline{\mathbb{P}} \in \overline{\mathcal{P}}_W(0)} \mathbb{E}^{\overline{\mathbb{P}}}[\Phi(X_\cdot)].$$

Let $\eta_0 > 0$ be an arbitrary constant, we use Lemma 5.11 to obtain a sequence $(\gamma_{h_n})_{n\ge1}$ and a sequence $(z_n)_{n\ge1}$ such that $\gamma_{h_n} \in \Gamma_{h_n}(z_n)$ for all $n \ge 1$, $\|z_n\|_1 \to 0$ and

$$\Big|\mathbb{E}^{\overline{\mathbb{P}}^*}\big[\Phi(X_\cdot)\big] - \mathbb{E}^{\mathbb{P}^{\gamma_{h_n}}}\big[\Phi\big(\widehat{X}^{\gamma_{h_n}}_\cdot\big)\big]\Big| \ \le \ \eta_0.$$

Using the Lipschitz property of $V_h(\cdot)$ in Lemma 5.9, it follows that

$$\begin{aligned}\limsup_{h\to0} V_h(0) \ &\le \ \limsup_{n\to\infty} \ \Big(V_h(z_n) + \frac{4M}{\varepsilon}\|z_n\|_1\Big) \\ &\le \ \limsup_{n\to\infty} \ \mathbb{E}^{\mathbb{P}^{\gamma_{h_n}}}\big[\Phi\big(\widehat{X}^{\gamma_{h_n}}_\cdot\big)\big] \ \le \ \mathbb{E}^{\overline{\mathbb{P}}^*}\big[\Phi(X_\cdot)\big] + \eta_0.\end{aligned}$$

Recall that $\overline{\mathbb{P}}^*$ is an ε–optimal solution of the problem $V(0)$ and $\eta_0 > 0$ is arbitrary, it follows that $\limsup_{h\to0} V_h(0) \le V(0)$. By (5.40), together with the duality results in Theorem 3.3, one can then conclude that

$$D_h(0) = V_h(0) \longrightarrow V(0) = D(0), \quad \text{as} \ \ h \longrightarrow 0.$$

Step 4 We finally prove the convergence rate result in (3.21). Notice that the dual optimizers λ^* for problem $D(0) = \sup_\lambda d(\lambda)$ and λ_h^* for $D_h(0) = \sup_{\lambda_h} d_h(\lambda)$ are both bounded by $\frac{2M}{\varepsilon}$ by Theorem 3.3, then under Assumption 3.6, one has

$$|D(0) - D_h(0)| \ \leq \ Ch^\rho.$$

Finally, (3.21) follows from $V_h(0) = D(0)$ and $V(0) = D(0)$ (Theorem 3.3).

5.5 Proof of Proposition 3.10

We first notice that $V \leq V_S$ as $\Gamma_S \subset \Gamma$. Further, V is convex on $\mathbb{R}^m \times \mathbb{R}^\ell$, and is continuous on $B_{\varepsilon/2}(0)$. Then

$$V(z) \ \leq \ V_S^{l.s.c.}(z) \text{ for } z \in B_{\varepsilon/2}(0), \quad \text{and} \quad V(z) \ \leq \ V_S^{conv}(z) \text{ for } z \in \mathbb{R}^m \times \mathbb{R}^\ell.$$

Let $z \in B_{\varepsilon/2}(0)$ so that $V(z) \in \mathbb{R}$. By similar arguments as in Lemma 5.11, for any $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(z)$, there is a sequence $(z_n, \overline{\mathbb{P}}_n)_{n\geq 1}$ such that $\overline{\mathbb{P}}_n \to \mathbb{P}$, $z_n \to z$ as $n \to \infty$, and $\overline{\mathbb{P}}_n \in \overline{\mathcal{P}}_S(z_n)$ for each $n \geq 1$. It follows that $V(z) \geq \lim_{n\to\infty} V_S(z_n)$ and hence

$$V(z) \ \geq \ V_S^{l.s.c.}(z), \text{ for } z \in B_{\varepsilon/2}(0).$$

Therefore, $V = V_S^{l.s.c.}$ on $B_{\varepsilon/2}(0)$. By the convexity and Lipschitz continuity of $V = V_S^{l.s.c.}$ on $B_{\varepsilon/2}(0)$, it follows that $V_S^{l.s.c.}$ is the convex envelop of V_S restricted on $B_{\varepsilon/2}(0)$, which implies that $V(z) = V_S^{l.s.c.}(z) \geq V_S^{conv}(z)$ for $z \in B_{\varepsilon/2}(0)$. Therefore, one has

$$V(z) = V_S^{l.s.c.}(z) = V_S^{conv}(z), \text{ for all } z \in B_{\varepsilon/2}(0).$$

References

[1] G. Barles and E. Jakobsen. Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations. *Mathematics of Computation*, 76(260):1861–1893, 2007.

[2] M. Beiglböck, A.M.G. Cox, and M. Huesmann. Optimal transport and Skorokhod embedding. *Inventiones mathematicae*, 208(2):327–400, 2017.

[3] O. Bokanowski, A. Picarelli, and H. Zidani. State-constrained stochastic optimal control problems via reachability approach. *SIAM Journal on Control and Optimization*, 54(5):2568–2593, 2016.

[4] J.F. Bonnans, J. Gianatti, and F.J. Silva. On the time discretization of stochastic optimal control problems: the dynamic programming approach. *ESAIM: Control, Optimisation and Calculus of Variations*, 25:63, 2019.

[5] J.F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer Science & Business Media, 2013.

[6] 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, 2012.

[7] B. Bouchard, R. Elie, and C. Imbert. Optimal control under stochastic target constraints. *SIAM Journal on Control and Optimization*, 48(5):3501–3531, 2010.

[8] B. Bouchard and M. Nutz. Weak dynamic programming for generalized state constraints. *SIAM Journal on Control and Optimization*, 50(6):3344–3373, 2012.

[9] B. Bouchard and N. Touzi. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Processes and their applications*, 111(2):175–206, 2004.

[10] Y.-L. Chow, X. Yu, and C. Zhou. On dynamic programming principle for stochastic control under expectation constraints. *Journal of Optimization Theory and Applications*, pages 1–16, 2020.

[11] L. Denis and C. Martini. A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. *The Annals of Applied Probability*, 16(2):827–852, 2006.

[12] F.M. Djete, D. Possamaï, and X. Tan. McKean-Vlasov optimal control: limit theory and equivalence between different formulations. *arXiv preprint arXiv:2001.00925*, 2020.

[13] Y. Dolinsky. Numerical schemes for G-expectations. *Electronic Journal of Probability*, 17, 2012.

[14] N. El Karoui, D. Nguyen Huu, and M. Jeanblanc-Picqué. Compactification methods in the control of degenerate diffusions: existence of an optimal control. *Stochastics: an international journal of probability and stochastic processes*, 20(3):169–219, 1987.

[15] A. Fahim, N. Touzi, and X. Warin. A probabilistic numerical method for fully nonlinear parabolic PDEs. *The Annals of Applied Probability*, 21(4):1322–1364, 2011.

[16] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms I: Fundamentals*, volume 305. Springer science & business media, 2013.

[17] J. Jacod, J. Memin, and M. Metivier. On tightness and stopping times. *Stochastic Processes and their Applications*, 14(2):109–146, 1983.

[18] N.V. Krylov. On the rate of convergence of finite-difference approximations for Bellmans equations with variable coefficients. *Probability Theory & Related Fields*, 117(1), 2000.

[19] H. Kushner and P.G. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24. Springer Science & Business Media, 2013.

[20] T. Mikami and M. Thieullen. Duality theorem for the stochastic optimal control problem. *Stochastic processes and their applications*, 116(12):1815–1835, 2006.

[21] Y. Nesterov. *Introductory lectures on convex optimization: A basic course*, volume 87. Springer Science & Business Media, 2013.

[22] S. Peng. A general stochastic maximum principle for optimal control problems. *SIAM Journal on control and optimization*, 28(4):966–979, 1990.

[23] L. Pfeiffer. Two approaches to stochastic optimal control problems with a final-time expectation constraint. *Applied Mathematics & Optimization*, 77(2):377–404, 2018.

[24] L. Pfeiffer. Optimality conditions in variational form for non-linear constrained stochastic control problems. *Mathematical Control & Related Fields*, 2019, online first.

[25] D. Possamaï and X. Tan. Weak approximation of second-order BSDEs. *The Annals of Applied Probability*, 25(5):2535–2562, 2015.

[26] Z. Ren and X. Tan. On the convergence of monotone schemes for path-dependent PDEs. *Stochastic Processes and their Applications*, 127(6):1738–1762, 2017.

[27] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2014.

[28] H.M. Soner and N. Touzi. Stochastic target problems, dynamic programming, and viscosity solutions. *SIAM Journal on Control and Optimization*, 41(2):404–424, 2002.

[29] X. Tan. Discrete-time probabilistic approximation of path-dependent stochastic control problems. *The Annals of Applied Probability*, 24(5):1803–1834, 2014.

[30] X. Tan and N. Touzi. Optimal transportation under controlled stochastic dynamics. *The annals of probability*, 41(5):3201–3240, 2013.

[31] J. Yong and X.Y. Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Science & Business Media, 1999.

[32] J. Zhang. A numerical scheme for BSDEs. *The annals of applied probability*, 14(1):459–488, 2004.