A level-set approach to the control of state-constrained McKean-Vlasov equations: application to renewable energy storage and portfolio selection*

Maximilien GERMAIN[†] Huyên PHAM[‡] Xavier WARIN[§]

December 21, 2021

Abstract

We consider the control of McKean-Vlasov dynamics (or mean-field control) with probabilistic state constraints. We rely on a level-set approach which provides a representation of the constrained problem in terms of an unconstrained one with exact penalization and running maximum or integral cost. The method is then extended to the common noise setting. Our work extends (Bokanowski, Picarelli, and Zidani, SIAM J. Control Optim. 54.5 (2016), pp. 2568–2593) and (Bokanowski, Picarelli, and Zidani, Appl. Math. Optim. 71 (2015), pp. 125–163) to a mean-field setting.

The reformulation as an unconstrained problem is particularly suitable for the numerical resolution of the problem, that is achieved from an extension of a machine learning algorithm from (Carmona, Laurière, arXiv:1908.01613 to appear in Ann. Appl. Prob., 2019). A first application concerns the storage of renewable electricity in the presence of mean-field price impact and another one focuses on a mean-variance portfolio selection problem with probabilistic constraints on the wealth. We also illustrate our approach for a direct numerical resolution of the primal Markowitz continuous-time problem without relying on duality.

Keywords: mean-field control, state constraints, neural networks.

AMS subject classification: 49N80, 49M99, 68T07, 93E20.

1 Introduction

The control of McKean-Vlasov dynamics, also known as mean-field control problem, has attracted a lot of interest over the last years since the emergence of the mean-field game theory. There is now an important literature on this topic addressing on one hand the theoretical aspects either by dynamic programming approach (see [34, 39, 38, 21]), or by maximum principle (see [14]), and on the other hand the numerous applications in economics and finance, and we refer to the two-volume monographs [15, 16] for an exhaustive and detailed treatment of this area.

In this paper, we aim to study control of McKean-Vlasov dynamics under the additional presence of state constraints in law. The consideration of probabilistic constraints (usually in expectation or in target form) for standard stochastic control has many practical applications, notably in finance with quantile and CVaR type constraints, and is the subject of many papers, we refer to [41, 11, 30, 19, 37, 4] for an overview.

*This work was supported by FiME (Finance for Energy Market Research Centre) and the "Finance et Développement Durable - Approches Quantitatives" EDF - CACIB Chair. We thank Marianne Akian, Olivier Bokanowski, and Nadia Oudjane for useful comments.

[†]EDF R&D, Université de Paris and Sorbonne Université, CNRS, Laboratoire de Probabilités, Statistique et Modélisation, F-75013 Paris, France mgermain at lpsm.paris

[‡]Université de Paris and Sorbonne Université, CNRS, Laboratoire de Probabilités, Statistique et Modélisation, F-75013 Paris, France, CREST-ENSAE & FiME pham at lpsm.paris

[§]EDF R&D & FiME xavier.warin at edf.fr

There exists some recent works dealing with mean-field control under some specific law state constraints. For example, the paper [18] solves mean-field control with delay and smooth expectation terminal constraint (and without dependence with respect to the law of the control). In the case of mean field games, state constraints are considered by [12, 13, 28, 33, 3]. In these cited works the state belongs to a compact set, which corresponds to a particular case of our constraints in distribution. Related literature includes the recent work [10] which studies a mean-field target problem where the aim is to find the initial laws of a controlled McKean-Vlasov process satisfying a law constraint, but only at terminal time. The paper [23] also studies these terminal constraint in law for the control of a standard diffusion process. Next, it has been extended in [24] to a running law constraint for the control of a standard diffusion process with McKean-Vlasov type cost through the control of a Fokker-Planck equation. Several works also consider directly the optimal control of Fokker-Planck equations in the Wasserstein space with terminal or running constraints, such as [8, 9] through Pontryagin principle, in the deterministic case without diffusion.

In this paper, we consider general running (at discrete or continuous time) and terminal constraints in law, and extend the level-set approach [6, 7] (see also [2] in the deterministic case) to our mean-field setting. This enables us to reformulate the constrained McKean-Vlasov control problem into an unconstrained mean-field control problem with an auxiliary state variable, and a running path-dependent supremum cost or alternatively a non path-dependent integral cost over the constrained functions. Such equivalent representations of the control problem with exact penalization turns out to be quite useful for an efficient numerical resolution of the original constrained mean-field control problem. We shall actually adapt the machine learning algorithm in [17] for solving two applications in renewable energy storage and in portfolio selection.

The outline of the paper is organized as follows. Section 2 develops the level-set approach in our constrained mean-field setting with supremum term. We present in Section 3 the alternative level-set formulation with integral term, and discuss when the optimization over open-loop controls yields the same value than the optimization over closed-loop controls. This will be useful for numerical purpose in the approximation of optimal controls. The method is then extended in Section 4 to the common noise setting. Finally, we present in Section 5 the applications and numerical tests.

2 Mean-field control with state constraints

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which is defined a d-dimensional Brownian motion W with associated filtration $\mathbb{F} = (\mathcal{F}_t)_t$ augmented with $\mathbb{P}$-null sets. We assume that $\mathcal{F}_0$ is "rich enough" in the sense that any probability measure μ on $\mathbb{R}^d$ can be represented as the distribution law of some $\mathcal{F}_0$-measurable random variable. This is satisfied whenever the probability space $(\Omega, \mathcal{F}_0, \mathbb{P})$ is atomless.

We consider the following cost and dynamics:

$$J(X_0, \alpha) = \mathbb{E}\Big[\int_0^T f\big(s, X_s^\alpha, \alpha_s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}\big) \ \mathrm{d}s + g\big(X_T^\alpha, \mathbb{P}_{X_T^\alpha}\big)\Big] \tag{2.1}$$

$$X_t^\alpha = X_0 + \int_0^t b\big(s, X_s^\alpha, \alpha_s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}\big) \ \mathrm{d}s + \int_0^t \sigma\big(s, X_s^\alpha, \alpha_t, \mathbb{P}_{(X_s^\alpha, \alpha_s)}\big) \ \mathrm{d}W_s, \tag{2.2}$$

where $\mathbb{P}_{(X_s^\alpha, \alpha_s)}$ is the joint law of (X_s^α, α_s) under $\mathbb{P}$ and X_0 is a given random variable in $L^2(\mathcal{F}_0, \mathbb{R}^d)$. The control α belongs to a set $\mathcal{A}$ of $\mathbb{F}$-progressively measurable processes with values in a set $A \subseteq \mathbb{R}^q$. The coefficients b and σ are measurable functions from $[0, T] \times \mathbb{R}^d \times A \times \mathcal{P}_2(\mathbb{R}^d \times A)$ into $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$, where $\mathcal{P}_2(E)$ is the set of square integrable probability measures on the metric space E, equipped with the 2-Wasserstein distance $\mathcal{W}_2$. We make some standard Lipschitz conditions on b, σ in order to ensure that equation (2.2) is well-defined and admits a unique strong solution, which is square-integrable. The function f is a real-valued measurable function on $[0, T] \times \mathbb{R}^d \times A \times \mathcal{P}_2(\mathbb{R}^d \times A)$, while g is a measurable function on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, and we assume that f and g satisfy some linear growth condition which ensures that the functional in (2.1) is well-defined.

Furthermore, the law of the controlled McKean-Vlasov process X is constrained to verify

$$\Psi(t, \mathbb{P}_{X_t^\alpha}) \leq \ 0, \quad 0 \leq t \leq T, \tag{2.3}$$

where $\Psi = (\Psi^l)_{1\le l\le k}$ is a given function from $[0,T] \times \mathcal{P}_2(\mathbb{R}^d)$ into $\mathbb{R}^k$. Here, the multi-dimensional constraint $\Psi(t,\mu) \le 0$ has to be understood componentwise, i.e., $\Psi^l(t,\mu) \le 0$, $l = 1, \cdots, k$. The problem of interest is therefore

$$V := \inf_{\alpha\in\mathcal{A}} \big\{ J(X_0,\alpha) : \Psi(t, \mathbb{P}_{X_t^\alpha}) \le 0, \ \forall\, t \in [0,T] \big\}.$$

By convention the infimum of the empty set is $+\infty$. When needed, we will sometimes use the notation V^Ψ to emphasize the dependence of the value function on Ψ. Clearly, $\Psi \le \Psi'$ (meaning that for each component $\Psi^l \le \Psi'^l$, $l = 1 \cdots k$) implies $V^\Psi \le V^{\Psi'}$.

Remark 2.1. *This very general type of constraints includes for instance:*

- *Controlled McKean-Vlasov process X constrained to stay inside a non-empty closed set $\mathcal{K}_t \subseteq \mathbb{R}^d$ with probability larger than a threshold $p_t \in [0,1]$, namely*

$$\mathbb{P}(X_t^\alpha \in \mathcal{K}_t) \ge p_t, \ \forall\, t \in [0,T],$$

  with $\Psi : (t,\mu) \mapsto p_t - \mu(\mathcal{K}_t)$. With $p_t = 1$, $\forall t \in [0,T]$ it yields almost sure constraints.

- *Almost sure contraints on the state, $X_t^\alpha \in \mathcal{K}_t$, $\forall t \in [0,T]$ $\mathbb{P}$ a.s., with*

$$\Psi : (t,\mu) \mapsto \int_{\mathbb{R}^d} d_{\mathcal{K}_t}(x)\ \mu(\mathrm{d}x),$$

  where $d_{\mathcal{K}_t}$ is the distance function to the non-empty closed set $\mathcal{K}_t$.

- *The case of a Wasserstein ball constraint around a benchmark law η_t in the form $\mathcal{W}_2(\mathbb{P}_{X_t^\alpha}, \eta_t) \le \delta_t$ with*

$$\Psi : (t,\mu) \mapsto \mathcal{W}_2(\mu, \eta_t) - \delta_t.$$

  This is the constraint considered in [36] at terminal time.

- *A terminal constraint in law $\varphi(\mathbb{P}_{X_T^\alpha}) \le 0$ as in [23] with*

$$\Psi : (t,\mu) \mapsto \varphi(\mu)\mathbb{1}_{t=T}.$$

- *Terminal constraint in law $\mathbb{P}_{X_T^\alpha} \in \mathbb{K} \subset \mathcal{P}_2(\mathbb{R}^d)$ as in [10] with*

$$\Psi : (t,\mu) \mapsto (1 - \mathbb{1}_{\mu\in\mathbb{K}})\mathbb{1}_{t=T}.$$

- *The case of discrete time constraints $\phi(t_i, \mathbb{P}_{X_{t_i}^\alpha}) \le 0$ for $t_1 < \cdots < t_k$ with*

$$\Psi : (t,\mu) \mapsto \phi(t,\mu)\mathbb{1}_{t\in\{t_1,\cdots,t_k\}}.$$

Even though this problem seems much more involved than the standard stochastic control problem with state constraints investigated in [7], thanks to an adequate reformulation, it turns out that we can adapt the main ideas from this paper to our framework and construct similarly an unconstrained auxiliary problem (in infinite dimension).

2.1 A target problem and an associated control problem

Given $z \in \mathbb{R}$, and $\alpha \in \mathcal{A}$, define a new state variable

$$Z_t^{z,\alpha} := z - \mathbb{E}\Big[\int_0^t f\big(s, X_s^\alpha, \alpha_s, \mathbb{P}_{(X_s^\alpha,\alpha_s)}\big)\ \mathrm{d}s\Big] = z - \int_0^t \widehat{f}\big(s, \mathbb{P}_{(X_s^\alpha,\alpha_s)}\big)\ \mathrm{d}s, \quad 0 \le t \le T, \tag{2.4}$$

where $\widehat{f}$ is the function defined on $[0,T] \times \mathcal{P}_2(\mathbb{R}^d \times A)$ by $\widehat{f}(t,\nu) = \int_{\mathbb{R}^d\times A} f\big(t,x,a,\nu\big)\ \nu(\mathrm{d}x,\mathrm{d}a)$. We also denote by $\widehat{g}$ the function defined on $\mathcal{P}_2(\mathbb{R}^d)$ by $\widehat{g}(\mu) = \int_{\mathbb{R}^d} g(x,\mu)\mu(\mathrm{d}x)$.

Lemma 2.2. *The value function admits the* **deterministic target problem** *representation*

$$V = \inf\{z \in \mathbb{R} \mid \exists\, \alpha \in \mathcal{A} \text{ s.t. } \widehat{g}(\mathbb{P}_{X_T^\alpha}) \leq Z_T^{z,\alpha},\ \Psi(s, \mathbb{P}_{X_s^\alpha}) \leq 0,\ \forall\, s \in [0,T]\}.$$

Proof. We first observe from the definition of V in (2.3) that it can be rewritten as

$$V = \inf\{z \in \mathbb{R} \mid \exists\, \alpha \in \mathcal{A} \text{ s.t. } J(X_0, \alpha) \leq z,\ \Psi(t, \mathbb{P}_{X_t^\alpha}) \leq 0,\ \forall\, t \in [0,T]\}.$$

Next, by noting that the cost functional is written as

$$J(X_0, \alpha) = \int_0^T \widehat{f}\big(t, \mathbb{P}_{(X_t^\alpha, \alpha_t)}\big)\ \mathrm{d}t + \widehat{g}(\mathbb{P}_{X_T^\alpha}),$$

the result then follows immediately by the definition of $Z^{z,\alpha}$ in (2.4). □

We want to link this representation to the **zero-level set** of the solution of an auxiliary unconstrained control problem. Define the **auxiliary unconstrained deterministic** control problem:

$$\mathcal{Y}^\Psi : z \in \mathbb{R} \mapsto \inf_{\alpha \in \mathcal{A}} \Big[\{\widehat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z,\alpha}\}_+ + \sum_{l=1}^{k} \sup_{s \in [0,T]} \{\Psi^l(s, \mathbb{P}_{X_s^\alpha})\}_+\Big], \tag{2.5}$$

with the notation $\{x\}_+ = \max(x, 0)$ for the positive part. We see that $\mathcal{Y}^\Psi(z) \geq 0$.

By classical estimates on McKean-Vlasov equations we can obtain continuity and growth conditions on $\mathcal{Y}^\Psi$. The proof of Proposition 2.3 is given in Section 2.3.

Proposition 2.3. $\mathcal{Y}^\Psi$ *verifies*

1. $\mathcal{Y}^\Psi$ *is 1-Lipschitz. For any* $z, z' \in \mathbb{R}$

$$|\mathcal{Y}^\Psi(z) - \mathcal{Y}^\Psi(z')| \leq |z - z'|.$$

2. $\mathcal{Y}^\Psi$ *is non-increasing. Thus if* $\mathcal{Y}^\Psi(z_0) = 0$ *then* $\mathcal{Y}^\Psi(z) = 0$ *for all* $z \geq z_0$.

Define the infimum of the zero level-set

$$\mathcal{Z}^\Psi := \inf\{z \in \mathbb{R} \mid \mathcal{Y}^\Psi(z) = 0\}. \tag{2.6}$$

We prove a first result linking the auxiliary control problem with the original constrained problem. Solving this easier problem provides bounds on the value function, by making the constraint function vary.

Theorem 2.4.

1. *If for some* $z \in \mathbb{R}$ $\exists\, \alpha \in \mathcal{A}$ s.t. $\widehat{g}(\mathbb{P}_{X_T^\alpha}) \leq Z_T^{z,\alpha},\ \Psi(s, \mathbb{P}_{X_s^\alpha}) \leq 0,\ \forall\, s \in [0,T]$ *then* $\mathcal{Y}^\Psi(z) = 0$.
2. *If* V^Ψ *is finite then* $\mathcal{Y}^\Psi(V^\Psi) = 0$. *Thus* $\mathcal{Z}^\Psi \leq V^\Psi$.
3. *Define* $1_k = (1, \cdots, 1) \in \mathbb{R}^k$. *We have the upper bound*

$$V^\Psi \leq \inf_{\varepsilon > 0} \mathcal{Z}^{\Psi + \varepsilon 1_k}.$$

To sum up, when $V^\Psi < +\infty$, Theorem 2.4 provides the bounds

$$\mathcal{Z}^\Psi \leq V^\Psi \leq \inf_{\varepsilon > 0} \mathcal{Z}^{\Psi + \varepsilon 1_k}.$$

The proof of Theorem 2.4 is given in Section 2.3.

Remark 2.5. *In the easier case where optimal controls exist for the auxiliary problem, as assumed in [7], and when* Ψ *is continuous, similar arguments as in [7] (and Section 4) directly prove that* $\mathcal{Z}^\Psi = V^\Psi$ *and that an optimal control* α^* *associated to the auxiliary problem* $\mathcal{Y}^\Psi(V)$ *is optimal for the original problem. However some difficulties arise when trying to remove this assumption.*

Remark 2.6. *If there exists $\varepsilon_0 > 0$ such that $V^{\Psi+\varepsilon_0 1_k} < \infty$ then $\mathcal{Z}^{\Psi+\varepsilon_0 1_k} \leq V^{\Psi+\varepsilon_0 1_k} < \infty$ by Theorem 2.4. Thus the right-hand side of the previous inequality is finite.*

On the other hand, if we consider for instance a one-dimensional terminal constraint in law $\varphi(\mathbb{P}_{X_T^\alpha}) \leq 0$, it is represented with

$$\Psi : (t, \mu) \mapsto \varphi(\mu)\mathbb{1}_{t=T},$$

and we see that the constraint $\Psi(t, \mu) + \varepsilon \leq 0$ would never be verified for any $t < T$ and any $\varepsilon > 0$, hence $V^{\Psi+\varepsilon} = \infty$.

In view of the above example in Remark 2.6, we introduce a modified constraint function in order to deal with discrete time constraints, and also with a.s. constraints. Given a constraint function $\Psi(t, \mu)$, we define

$$\overline{\Psi}_\kappa(t, \mu) := \Psi(t, \mu) - \kappa \sum_{l=1}^{k} \mathbb{1}_{\{\Psi^l(t,\mu)\leq 0\}} e_l, \tag{2.7}$$

with $\kappa > 0$ and e_l the l-th component of the canonical basis of $\mathbb{R}^k$. Then it is immediate to see that

$$V^\Psi = V^{\overline{\Psi}_\kappa}, \ \mathcal{Y}^\Psi = \mathcal{Y}^{\overline{\Psi}_\kappa}, \ \mathcal{Z}^\Psi = \mathcal{Z}^{\overline{\Psi}_\kappa}.$$

Remark 2.7. *Notice that by taking $\varepsilon_0 < \kappa$, and assuming that $V^\Psi < \infty$, we have $\mathcal{Z}^{\overline{\Psi}_\kappa+\varepsilon_0 1_k} < \infty$. Indeed, by applying Theorem 2.4 to $\overline{\Psi}_\kappa$, we have $\mathcal{Z}^{\overline{\Psi}_\kappa+\varepsilon_0 1_k} \leq V^{\overline{\Psi}_\kappa+\varepsilon_0 1_k}$. Moreover, by observing that an admissible control for the original problem V^Ψ is also admissible for the auxiliary problem with constraint function $\overline{\Psi}_\kappa + \varepsilon_0 1_k$, by definition of $\overline{\Psi}_\kappa$, this implies that $V^{\overline{\Psi}_\kappa+\varepsilon_0 1_k} < \infty$.*

2.2 Representation of the value function

Now we prove under some assumptions on the constraints the continuity property $\mathcal{Z}^{\overline{\Psi}_\kappa} = \inf_{\varepsilon>0} \mathcal{Z}^{\overline{\Psi}_\kappa+\varepsilon 1_k}$ in order to obtain a characterization of the original value function V^Ψ. The result relies on convexity arguments.

Lemma 2.8. *$(z, \varepsilon) \in \mathbb{R} \times \mathbb{R} \mapsto \mathcal{Y}^{\Psi+\varepsilon 1_k}(z)$ is jointly convex.*

The proof of Lemma 2.8 is given in Section 2.3.

Proposition 2.9. *$\mathcal{Y}^\Psi$ being convex, positive and non-increasing, if $\mathcal{Z}^\Psi < \infty$ then $\mathcal{Y}^\Psi$ is decreasing on $(-\infty, \mathcal{Z}^\Psi]$ then $\mathcal{Y}^\Psi(z) = 0$ on $[\mathcal{Z}^\Psi, \infty)$.*

Proof. By contradiction, if $\mathcal{Y}^\Psi(a) = \mathcal{Y}^\Psi(b) > 0$ with $a < b$ then by monotonicity $\mathcal{Y}^\Psi([a, b]) = \{\mathcal{Y}^\Psi(a)\}$ and $0 \in \partial\mathcal{Y}^\Psi(a)$ thus $\mathcal{Y}^\Psi(x) \geq \mathcal{Y}^\Psi(a) > 0$ $\forall x \in \mathbb{R}$ which is not the case because $\mathcal{Z}^\Psi < \infty$. As a consequence, $\mathcal{Y}^\Psi$ is decreasing. Then by continuity of $\mathcal{Y}^\Psi$ and definition of $\mathcal{Z}^\Psi$ we obtain $\mathcal{Y}^\Psi(\mathcal{Z}^\Psi) = 0$. □

Theorem 2.10. *Assume that $V^\Psi < \infty$. Then we have the representation*

$$\mathcal{Z}^\Psi = V^\Psi.$$

Moreover ε-optimal controls α^ε for the auxiliary problem $\mathcal{Y}^\Psi(V^\Psi)$ are ε-admissible ε-optimal controls for the original problem in the sense that

$$J(X_0, \alpha^\varepsilon) \leq V^\Psi + \varepsilon, \ \sup_{0\leq s\leq T} \Psi(s, \mathbb{P}_{X_s^{\alpha^\varepsilon}}) \leq \varepsilon.$$

Proof of Theorem 2.10. We prove the continuity of $\mathcal{Z}^{\overline{\Psi}_\kappa}$ along the curve $\mathcal{Z}^{\overline{\Psi}_\kappa+\varepsilon 1_k}$ for $\varepsilon \in \mathbb{R}$ where $\overline{\Psi}_\kappa$ is defined in (2.7).

Let $\kappa > 0$ and $\varepsilon_0 < \kappa$. By Remark 2.7, we know that $\mathcal{Z}^{\overline{\Psi}_\kappa+\varepsilon_0 1_k} < \infty$. We consider the optimization problem

$$\Phi : \varepsilon \in \mathbb{R} \mapsto \inf_z \ z + \chi^\Xi(\varepsilon, z),$$

where χ^Ξ is the indicator function of the non-empty admissible set $\Xi = \{(\varepsilon, z) \in \mathbb{R}^2 \mid \mathcal{Y}^{\overline{\Psi}_\kappa + \varepsilon 1_k}(z) = 0\} = \{(\varepsilon, z) \in \mathbb{R}^2 \mid \mathcal{Y}^{\overline{\Psi}_\kappa + \varepsilon 1_k}(z) \leq 0\}$, namely

$$\chi^\Xi(\varepsilon, z) = \begin{cases} 0 & \text{if } (\varepsilon, z) \in \Xi \\ +\infty & \text{otherwise.} \end{cases}$$

Note that $\mathcal{Z}^{\overline{\Psi}_\kappa + \varepsilon 1_k} = \inf\{z \mid \mathcal{Y}^{\overline{\Psi}_\kappa + \varepsilon 1_k}(z) = 0\} = \Phi(\varepsilon)$. By Proposition 2.8, $\mathcal{Y}^{\overline{\Psi}_\kappa + \varepsilon 1_k}(z)$ is jointly convex thus Ξ is convex. Hence, $(\varepsilon, z) \mapsto z + \chi^\Xi(\varepsilon, z)$ is jointly convex. Now $\Phi(\varepsilon)$ is convex as the marginal of a jointly convex function. As a consequence, $\varepsilon \in \mathbb{R} \mapsto \Phi(\varepsilon)$ is continuous in zero by noticing that $\varepsilon \in (-\infty, \varepsilon_0) \mapsto \Phi(\varepsilon) \leq \mathcal{Z}^{\overline{\Psi}_\kappa + \varepsilon_0 1_k} < +\infty$ and applying Lemma 2.1 from [27]. As a consequence $\mathcal{Z}^{\overline{\Psi}_\kappa} = \inf_{\varepsilon > 0} \mathcal{Z}^{\overline{\Psi}_\kappa + \varepsilon 1_k}$. Therefore by Theorem 2.4 applied to $\overline{\Psi}_\kappa$, we obtain $\mathcal{Z}^{\overline{\Psi}_\kappa} = V^{\overline{\Psi}_\kappa}$. Then recalling that $\mathcal{Z}^\Psi = \mathcal{Z}^{\overline{\Psi}_\kappa}$, $V^\Psi = V^{\overline{\Psi}_\kappa}$, the result follows.
Concerning the controls, take $\varepsilon > 0$, and consider an ε-optimal control $\alpha^\varepsilon \in \mathcal{A}$ such that

$$\{\widehat{g}(\mathbb{P}_T^{\alpha^\varepsilon}) - Z_T^{\mathcal{Z}^\Psi, \alpha^\varepsilon}\}_+ + \sum_{l=1}^k \sup_{s \in [0,T]} \{\Psi^l(s, \mathbb{P}_{X_s^{\alpha^\varepsilon}})\}_+ \leq \varepsilon.$$

The two terms on the l.h.s. being non-negative, they both are smaller than ε and thus

$$\widehat{g}(\mathbb{P}_T^{\alpha^\varepsilon}) \leq Z_T^{\mathcal{Z}^\Psi, \alpha^\varepsilon} + \varepsilon, \text{ and } \Psi^l(s, \mathbb{P}_{X_s^{\alpha^\varepsilon}}) \leq \varepsilon, \ \forall\, s \in [0, T], \ \forall l = 1, \cdots, k.$$

Hence

$$J(X_0, \alpha^\varepsilon) \leq \mathcal{Z}^\Psi + \varepsilon = V^\Psi + \varepsilon$$

and

$$\Psi(s, \mathbb{P}_{X_s^{\alpha^\varepsilon}}) \leq \varepsilon, \ \forall\, s \in [0, T].$$

□

2.3 Proofs

Proof of Proposition 2.3. 1) By the inequalities $|\inf_u A(u) - \inf_u B(u)| \leq \sup_u |A(u) - B(u)|$, $|\sup_u A(u) - \sup_u B(u)| \leq \sup_u |A(u) - B(u)|$we obtain for any $z, z' \in \mathbb{R}$

$$\begin{aligned}
&|\mathcal{Y}^\Psi(z) - \mathcal{Y}^\Psi(z')| \\
&= |\inf_{\alpha \in \mathcal{A}} \Big[\{\widehat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z,\alpha}\}_+ + \sum_{l=1}^k \sup_{s \in [t,T]} \{\Psi^l(s, \mathbb{P}_{X_s^\alpha})\}_+\Big] \\
&- \inf_{\alpha \in \mathcal{A}} \Big[\{\widehat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z',\alpha}\}_+ + \sum_{l=1}^k \sup_{s \in [t,T]} \{\Psi^l(s, \mathbb{P}_{X_s^\alpha})\}_+\Big]| \\
&\leq \sup_{\alpha \in \mathcal{A}} |\{\widehat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z,\alpha}\}_+ - \{\widehat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z',\alpha}\}_+ + \sum_{l=1}^k \sup_{s \in [t,T]} \{\Psi^l(s, \mathbb{P}_{X_s^\alpha})\}_+ - \sum_{l=1}^k \sup_{s \in [t,T]} \{\Psi^l(s, \mathbb{P}_{X_s^\alpha})\}_+| \\
&\leq \sup_{\alpha \in \mathcal{A}} |Z_T^{z,\alpha} - Z_T^{z',\alpha}| = |z - z'|,
\end{aligned}$$

by 1-Lipschitz continuity of $x \mapsto \{x\}_+$.

2) Denote by

$$L^\Psi(z, \alpha) = \{\widehat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z,\alpha}\}_+ + \sum_{l=1}^k \sup_{s \in [0,T]} \{\Psi^l(s, \mathbb{P}_{X_s^\alpha})\}_+,$$

so that $\mathcal{Y}^\Psi(z) = \inf_{\alpha \in \mathcal{A}} L^\Psi(z, \alpha)$. Then, it is clear that

$$z \leq z' \implies L^{\Psi}(z', \alpha) \leq L^{\Psi}(z, \alpha)$$

hence by minimizing, the same monotonicity property holds also for the value function

$$z \leq z' \implies \mathcal{Y}^{\Psi}(z') \leq \mathcal{Y}^{\Psi}(z).$$

□

Proof of Theorem 2.4. 1) $\exists\, \alpha \in \mathcal{A},\ \widehat{g}(\mathbb{P}_{X_T^\alpha}) \leq Z_T^{z,\alpha}$ and $\Psi(s, \mathbb{P}_{X_s^\alpha}) \leq 0,\ \forall\ s \in [0,T]$. Therefore

$$\{\widehat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z,\alpha}\}_+ + \sum_{l=1}^{k} \sup_{s\in[0,T]} \{\Psi^l(s, \mathbb{P}_{X_s^\alpha})\}_+ = 0$$

and by non-negativity of $\mathcal{Y}$ we obtain $\mathcal{Y}^{\Psi}(z) = 0$

2) By continuity of $\mathcal{Y}$ (Proposition 2.3) and 1), we obtain $\mathcal{Y}^{\Psi}(V^{\Psi}) = 0$ by taking admissible ε-optimal controls for the original problem and taking the limit $\varepsilon \to 0$. By definition of $\mathcal{Z}^{\Psi}$ the property is established.

3) We assume that exists $\varepsilon_0 > 0$ such that $\mathcal{Z}^{\Psi+\varepsilon_0 1_k} < +\infty$. If it is not the case then $\inf_{\varepsilon>0} \mathcal{Z}^{\Psi+\varepsilon 1_k} = +\infty$ and the inequality is verified. Let $0 < \varepsilon < \varepsilon_0$ satisfying $\mathcal{Z}^{\Psi+\varepsilon 1_k} < \infty$. By continuity of $\mathcal{Y}$ in the z variable (Proposition 2.3), $\mathcal{Y}^{\Psi+\varepsilon 1_k}(\mathcal{Z}^{\Psi+\varepsilon 1_k}) = 0$. Then by definition of $\mathcal{Y}^{\Psi+\varepsilon 1_k}$, for $0 < \varepsilon' \leq \varepsilon, \exists\, \alpha^{\varepsilon'} \in \mathcal{A}$ such that

$$\{\widehat{g}(\mathbb{P}_T^{\alpha^{\varepsilon'}}) - Z_T^{\mathcal{Z}^{\Psi+\varepsilon 1_k}, \alpha^{\varepsilon'}}\}_+ + \sum_{l=1}^{k} \sup_{s\in[0,T]} \{\Psi^l(s, \mathbb{P}_{X_s^\alpha}) + \varepsilon\}_+ \leq \varepsilon'.$$

The two terms on the l.h.s. being non-negative, they both are smaller than ε' and thus

$$\widehat{g}(\mathbb{P}_T^{\alpha^{\varepsilon'}}) \leq Z_T^{\mathcal{Z}^{\Psi+\varepsilon 1_k}, \alpha^{\varepsilon'}} + \varepsilon', \text{ and } \Psi^l(s, \mathbb{P}_{X_s^\alpha}) \leq \varepsilon' - \varepsilon \leq 0,\ \forall\ s \in [0,T],\ \forall l = 1, \cdots, k.$$

Hence

$$J(\alpha^{\varepsilon'}) \leq \mathcal{Z}^{\Psi+\varepsilon 1_k} + \varepsilon'$$

and

$$\Psi(s, \mathbb{P}_{X_s^\alpha}) \leq 0,\ \forall\ s \in [0,T].$$

Therefore by arbitrariness of ε' verifying $0 < \varepsilon' < \varepsilon$ we conclude that $V^{\Psi} \leq \mathcal{Z}^{\Psi+\varepsilon 1_k}$. By arbitrariness of ε verifying $0 < \varepsilon < \varepsilon_0$ it follows

$$V^{\Psi} \leq \inf_{\varepsilon\in(0,\varepsilon_0)} \mathcal{Z}^{\Psi+\varepsilon 1_k} = \inf_{\varepsilon>0} \mathcal{Z}^{\Psi+\varepsilon 1_k},$$

where the last equality comes from the non-increasing property of $\mathcal{Z}^{\Psi+\varepsilon 1_k}$ w.r.t. ε. □

Proof of Lemma 2.8. 1. Let $0 \leq \lambda \leq 1$. Then for $z, z', \varepsilon, \varepsilon' \in \mathbb{R}$

$$\begin{aligned}
&L^{\Psi+\lambda\varepsilon 1_k+(1-\lambda)\varepsilon' 1_k}(\lambda z + (1-\lambda)z', \alpha) \\
&= \{\lambda\Big(\widehat{g}(\mathbb{P}_{X_T^\alpha}) + \int_0^T \widehat{f}\big(s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}\big)\ \mathrm{d}s - z\Big) + (1-\lambda)\Big(\widehat{g}(\mathbb{P}_{X_T^\alpha}) + \int_0^T \widehat{f}\big(s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}\big)\ \mathrm{d}s - z'\Big)\}_+ \\
&+ \sum_{l=1}^{k} \sup_{s\in[0,T]} \{\lambda\Psi^l(s, \mathbb{P}_{X_s^\alpha}) + \lambda\varepsilon + (1-\lambda)\Psi^l(s, \mathbb{P}_{X_s^\alpha}) + (1-\lambda)\varepsilon'\}_+ \\
&\leq \lambda\{\widehat{g}(\mathbb{P}_{X_T^\alpha}) + \int_0^T \widehat{f}\big(s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}\big)\ \mathrm{d}s - z\}_+ + (1-\lambda)\{\widehat{g}(\mathbb{P}_{X_T^\alpha}) + \int_0^T \widehat{f}\big(s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}\big)\ \mathrm{d}s - z'\}_+ \\
&+ \sum_{l=1}^{k} \sup_{s\in[0,T]} \lambda\{\Psi^l(s, \mathbb{P}_{X_s^\alpha}) + \varepsilon\}_+ + (1-\lambda)\{\Psi^l(s, \mathbb{P}_{X_s^\alpha}) + \varepsilon'\}_+ \\
&\leq \lambda L^{\Psi+\varepsilon 1_k}(z, \alpha) + (1-\lambda) L^{\Psi'+\varepsilon' 1_k}(z', \alpha)
\end{aligned}$$

by convexity of $x \mapsto \{x\}_+$. By minimizing over the controls, the result follows. □

2.4 Potential extension towards dynamic programming

If one wants to use dynamic programming in order to solve the auxiliary control problem, it requires to write it down under a Markovian dynamic formulation. Define

$$X_s^{t,\xi,\alpha} = \xi + \int_t^s b\big(u, X_u^{t,\xi,\alpha}, \alpha_u, \mathbb{P}_{(X_u^{t,\xi,\alpha},\alpha_u)}\big)\ \mathrm{d}u + \int_t^s \sigma\big(u, X_u^{t,\xi,\alpha}, \alpha_u, \mathbb{P}_{(X_u^{t,\xi,\alpha},\alpha_u)}\big)\ \mathrm{d}W_u,$$

for $t \in [0,T]$, and $\xi \in L^2(\mathcal{F}_t, \mathbb{R}^d)$, and notice that we have the flow property

$$X_r^{t,\xi,\alpha} \ = \ X_r^{s,X_s^{t,\xi,\alpha},\alpha}, \quad \mathbb{P}_{X_r^{t,\xi,\alpha}} \ = \ \mathbb{P}_{X_r^{s,X_s^{t,\xi,\alpha},\alpha}}, \ \forall\ 0 \le s \le r \le T,$$

coming from existence and pathwise uniqueness in (2.2). We thus consider the cost function

$$J(t,\xi,\alpha) := \mathbb{E}\Big[\int_t^T f\big(s, X_s^{t,\xi,\alpha}, \alpha_s, \mathbb{P}_{(X_s^{t,\xi,\alpha},\alpha_s)}\big)\ \mathrm{d}s + g\big(X_T^{t,\xi,\alpha}, \mathbb{P}_{X_T^{t,\xi,\alpha}}\big)\Big],$$

and the value function

$$V(t,\xi) := \inf_{\alpha\in\mathcal{A}}\{J(t,\xi,\alpha) \mid \Psi(s, \mathbb{P}_{X_s^{t,\xi,\alpha}}) \le 0,\ \forall\ s \in [t,T]\}.$$

Then we introduce the auxiliary state variable

$$Z_r^{t,\xi,z,\alpha} := z - \mathbb{E}\Big[\int_t^r f\big(s, X_s^{t,\xi,\alpha}, \alpha_s, \mathbb{P}_{(X_s^{t,\xi,\alpha},\alpha_s)}\big)\ \mathrm{d}s\Big] = z - \int_t^r \widehat{f}\big(s, \mathbb{P}_{(X_s^{t,\xi,\alpha},\alpha_s)}\big)\ \mathrm{d}s, \quad t \le r \le T,$$

and the auxiliary value function is given by

$$\begin{aligned}\mathcal{Y}^\Psi(t,\xi,z) &= \inf_{\alpha\in\mathcal{A}}\Big[\{\widehat{g}(\mathbb{P}_{X_T^{t,\xi,\alpha}}) - Z^\alpha_{t,\xi,z}(T)\}_+ + \sup_{s\in[t,u]}\{\Psi^l(s,\mathbb{P}_{X_s^{t,\xi,\alpha}})\}_+\Big)\Big] \\ &=: \inf_{\alpha\in\mathcal{A}} L^\Psi(t,\xi,z,\alpha). \end{aligned} \tag{2.8}$$

We can treat the non-Markovian formulation of this problem by introducing as in [6] an additional state variable $Y_u^{t,\xi,\alpha,m} = \Big(\sum_{l=1}^k \sup_{s\in[t,u]}\{\Psi^l(s,\mathbb{P}_{X_s^{t,\xi,\alpha}})\}_+\Big) \vee m \ge 0$ for $u \ge t$ with $m \in \mathbb{R}$ and the value function

$$\overline{\mathcal{Y}}^\Psi(t,\xi,z,m) = \inf_{\alpha\in\mathcal{A}}\Big[\{\widehat{g}(\mathbb{P}_{X_T^{t,\xi,\alpha}}) - Z^\alpha_{t,\xi,z}(T)\}_+ + Y_T^{t,\xi,\alpha,m}\Big] =: \inf_{\alpha\in\mathcal{A}} \overline{L}^\Psi(t,\xi,z,m,\alpha).$$

The two problems are related by

$$\mathcal{Y}^\Psi(t,\xi,z) = \overline{\mathcal{Y}}^\Psi(t,\xi,z,\sum_{l=1}^k\{\Psi^l(t,\mathbb{P}_{X_t^{t,\xi,\alpha}})\}_+).$$

With this formulation, the problem (2.8) becomes a Mayer-type Markovian optimal control problem in the augmented state space $[0,T] \times L^2(\mathcal{F}_0,\mathbb{R}^d) \times \mathbb{R} \times \mathbb{R}$. As mentioned in [6], this procedure is used for instance for hedging lookback options in finance, see e.g. [29]. Now the infimum of the zero level-set is given by

$$\mathcal{Z}^\Psi(t,\xi) := \inf\{z \in \mathbb{R} \mid \overline{\mathcal{Y}}^\Psi(t,\xi,z,0) = 0\}.$$

Indeed note that $\overline{\mathcal{Y}}^\Psi(t,\xi,z,m) = 0 \iff m \le 0$ and $\overline{\mathcal{Y}}^\Psi(t,\xi,z,0) = 0$.

The Lipschitz and convexity properties of the value function are proven exactly as in Proposition 2.3 but we detail here the continuity in space and in the running maximum variable m.

Assumption 2.11. *Ψ, f, g, b, σ are Lipschitz continuous uniformly with respect to to other variables. Namely exists $[\Psi], [f], [g], [b], [\sigma], L > 0$ and locally bounded functions $h, \ell, \mathfrak{L} : [0, +\infty) \mapsto [0, +\infty)$ such that for any $t \in [0, T],\ x, x' \in \mathbb{R}^d, \mu \in \mathcal{P}_2(\mathbb{R}^d), \nu, \nu' \in \mathcal{P}_2(\mathbb{R}^d \times A),\ a \in A$*

$$\begin{aligned}
|\Psi(t,\mu) - \Psi(t,\mu')| &\le [\Psi]\mathcal{W}_2(\mu,\mu') \\
|f(t,x,a,\nu) - f(t,x',a,\nu')| &\le [f](|x-x'| + \mathcal{W}_2(\nu,\nu')) \\
|g(x,\mu) - g(x,\mu')| &\le [g](|x-x'| + \mathcal{W}_2(\mu,\mu')) \\
|b(t,x,a,\nu) - b(t,x',a,\nu')| &\le [b](|x-x'| + \mathcal{W}_2(\nu,\nu')) \\
|\sigma(t,x,a,\nu) - \sigma(t,x',a,\nu')| &\le [\sigma](|x-x'| + \mathcal{W}_2(\nu,\nu')) \\
|b(t,0,a,\delta_0 \otimes \mu)| + |\sigma(t,0,a,\delta_0 \otimes \mu)| + |f(t,0,a,\delta_0 \otimes \mu)| &\le L \\
|f(t,x,a,\nu)| &\le h(\|\nu\|_2)(1+|x|^2) \\
|g(x,\mu)| &\le \ell(\|\mu\|_2)(1+|x|^2) \\
|\Psi(t,\mu)| &\le \mathfrak{L}(\|\mu\|_2).
\end{aligned}$$

Proposition 2.12. *Under Assumption 2.11 $\overline{\mathcal{Y}}^{\Psi}$ is Lipschitz continuous: there exists $C > 0$ such that for any $t \in [0, T],\ \xi, \xi' \in L^2(\mathcal{F}_t, \mathbb{R}^d), m, m' \in \mathbb{R}$*

$$|\overline{\mathcal{Y}}^{\Psi}(t,\xi,z,m) - \overline{\mathcal{Y}}^{\Psi}(t,\xi',z',m')| \le |z-z'| + |m-m'| + C\sqrt{\mathbb{E}|\xi-\xi'|^2}.$$

Proof of Proposition 2.12. By the inequalities $|\inf_u A(u) - \inf_u B(u)| \le \sup_u |A(u) - B(u)|$, $|\sup_u A(u) - \sup_u B(u)| \le \sup_u |A(u) - B(u)|$, and $|a \vee b - c \vee d| \le |a-c| \vee |b-d| \le |a-c| + |b-d|$ we obtain for any $\xi, \xi' \in L^2(\mathcal{F}_t, \mathbb{R}^d)$ (if Ψ is not continuous consider $\xi = \xi'$)

$$\begin{aligned}
&|\overline{\mathcal{Y}}^{\Psi}(t,\xi,z,m) - \overline{\mathcal{Y}}^{\Psi}(t,\xi',z',m')| \\
&\le \sup_{\alpha\in\mathcal{A}} |\{\widehat{g}(\mathbb{P}_{X_T^{t,\xi,\alpha}}) - Z_T^{t,\xi,z,\alpha}\}_+ - \{\widehat{g}(\mathbb{P}_{X_T^{t,\xi',\alpha}}) - Z_T^{t,\xi',z',\alpha}\}_+ + \\
&\sup_{s\in[t,T]} \{\Psi(s,\mathbb{P}_{X_s^{t,\xi,\alpha}})\}_+ \vee m - \sup_{s\in[t,T]} \{\Psi(s,\mathbb{P}_{X_s^{t,\xi',\alpha}})\}_+ \vee m'| \\
&\le \sup_{\alpha\in\mathcal{A}} (|\widehat{g}(\mathbb{P}_{X_T^{t,\xi,\alpha}}) - \widehat{g}(\mathbb{P}_{X_T^{t,\xi',\alpha}})| + |Z_T^{t,\xi,z,\alpha} - Z_T^{t,\xi',z',\alpha}| + |\sup_{s\in[t,T]} \{\Psi(s,\mathbb{P}_{X_s^{t,\xi,\alpha}})\}_+ - \sup_{s\in[t,T]} \{\Psi(s,\mathbb{P}_{X_s^{t,\xi',\alpha}})\}_+| \\
&+ |m-m'|) \\
&\le \sup_{\alpha\in\mathcal{A}} \Big(|\mathbb{E}[g(X_T^{t,\xi,\alpha}, \mathbb{P}_{X_T^{t,\xi,\alpha}}) - g(X_T^{t,\xi',\alpha}, \mathbb{P}_{X_T^{t,\xi',\alpha}})]| + |z-z'| \\
&+ \Big|\mathbb{E}\Big[\int_t^T f\big(s, X_s^{t,\xi,\alpha}, \alpha_s, \mathbb{P}_{(X_s^{t,\xi,\alpha},\alpha_s)}\big)\, \mathrm{d}s - \int_t^T f\big(s, X_s^{t,\xi',\alpha}, \alpha_s, \mathbb{P}_{(X_s^{t,\xi',\alpha},\alpha_s)}\big)\, \mathrm{d}s\Big]\Big|\Big) \\
&+ \sup_{\alpha\in\mathcal{A}} \sup_{s\in[t,T]} |\{\Psi(s,\mathbb{P}_{X_s^{t,\xi,\alpha}})\}_+ - \{\Psi(s,\mathbb{P}_{X_s^{t,\xi',\alpha}})\}_+| + |m-m'| \\
&\le [\widehat{g}] \sup_{\alpha\in\mathcal{A}} (\mathbb{E}|X_T^{t,\xi,\alpha} - X_T^{t,\xi',\alpha}| + \mathcal{W}_2(\mathbb{P}_{X_T^{t,\xi,\alpha}}, \mathbb{P}_{X_T^{t,\xi',\alpha}})) + |z-z'| + |m-m'| + [\Psi] \sup_{\alpha\in\mathcal{A}} \sup_{s\in[t,T]} \mathcal{W}_2(\mathbb{P}_{X_s^{t,\xi,\alpha}}, \mathbb{P}_{X_s^{t,\xi',\alpha}}) \\
&+ T[f] \sup_{\alpha\in\mathcal{A}} \{\mathbb{E}[\sup_{s\in[t,T]} |X_s^{t,\xi,\alpha} - X_s^{t,\xi',\alpha}|] + \sup_{s\in[t,T]} \mathcal{W}_2(\mathbb{P}_{X_s^{t,\xi,\alpha}}, \mathbb{P}_{X_s^{t,\xi',\alpha}})\},
\end{aligned}$$

by Lipschitz continuity of $\Psi, x \mapsto \{x\}_+$. We recall the estimates

$$\begin{aligned}
&\sup_{s\in[t,T]} \mathcal{W}_2(\mathbb{P}_{X_s^{t,\xi,\alpha}}, \mathbb{P}_{X_s^{t,\xi',\alpha}}) = \sqrt{\sup_{s\in[t,T]} \mathcal{W}_2(\mathbb{P}_{X_s^{t,\xi,\alpha}}, \mathbb{P}_{X_s^{t,\xi',\alpha}})^2} \le C\sqrt{\mathbb{E}|\xi-\xi'|^2} \\
&\mathbb{E}[\sup_{s\in[t,T]} |X_s^{t,\xi,\alpha} - X_s^{t,\xi',\alpha}|] \le C\mathbb{E}|\xi-\xi'| \le C\sqrt{\mathbb{E}|\xi-\xi'|^2},
\end{aligned}$$

obtained by standard arguments (see e.g. the proof of Proposition 3.3 in [21]). Then the result follows. □

Proposition 2.13 (Law invariance properties)**.** *Under Assumption 2.11, we have law invariance of* $\overline{\mathcal{Y}}^{\Psi}$ *and* $\mathcal{Z}^{\Psi}$*, namely if* ξ, η *are* $\mathcal{F}_t$*-adapted square integrable with the same law, for any* $(t, z, m) \in [0,T] \times \mathbb{R} \times \mathbb{R}$

$$\begin{aligned}\overline{\mathcal{Y}}^{\Psi}(t, \xi, z, m) &= \overline{\mathcal{Y}}^{\Psi}(t, \eta, z, m) \\ \mathcal{Z}^{\Psi}(t, \xi) &= \mathcal{Z}^{\Psi}(t, \eta).\end{aligned}$$

Therefore we can define the lifted functions y^{Ψ}, z^{Ψ} *on* $[0,T] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}$ *(respectively* $[0,T] \times \mathcal{P}_2(\mathbb{R}^d)$ *by* $y^{\Psi}(t, \mathbb{P}_{\xi}, z, m) := \overline{\mathcal{Y}}^{\Psi}(t, \xi, z, m)$ *and* $z^{\Psi}(t, \mathbb{P}_{\xi}, z, m) := \mathcal{Z}^{\Psi}(t, \xi, z, m)$*.*

Proof. Apply the same arguments as in Theorem 3.5. from [20] to the unconstrained Markovian value function $\mathcal{Y}^{\Psi}$ on the extended state space. In particular use the continuity of $\mathcal{Y}^{\Psi}$ from Proposition 2.3 and notice for a given control α that in Step 1 of Theorem 3.5. from [20] the equality in law

$$\begin{aligned}&((X_s^{t,\xi,\alpha})_{s\in[t,T]}, (Z_s^{t,\xi,z,\alpha})_{s\in[t,T]}, (Y_u^{t,\xi,\alpha,m})_{s\in[t,T]}, (\alpha_s)_{s\in[t,T]}) \\ \overset{\mathcal{L}}{=}\ &((X_s^{t,\eta,\beta})_{s\in[t,T]}, (Z^{\beta,t,\eta,z})_{s\in[t,T]}, (Y_u^{t,\eta,\beta,m})_{s\in[t,T]}, (\beta_s)_{s\in[t,T]}),\end{aligned}$$

holds true with a_s defined in Lemma B.2. from [20] (verifying in particular the equality in law $\alpha_s = a_s(\xi, U_{\xi})$) and $\beta_s = a_s(\eta, U_{\eta})$ where U_{η} (respectively U_{ξ}) is a $\mathcal{F}_t$-adapted uniform random variable on $[0,1]$ independent of η (respectively ξ). Then use the definition (2.6) to obtain the same law invariance property for $\mathcal{Z}^{\Psi}$ too. □

Theorem 2.4 and Theorem 2.10 are still valid in the the dynamic case, by applying the exact same arguments. More precisely for any $(t, \xi) \in [0,T] \times L^2(\mathcal{F}_t, \mathbb{R}^d)$, if $V^{\Psi}(t, \xi) < \infty$ then

$$\mathcal{Z}^{\Psi}(t, \xi) \le V^{\Psi}(t, \xi) \le \inf_{\varepsilon > 0} \mathcal{Z}^{\Psi + \varepsilon 1_k}(t, \xi).$$

Similarly, arguments like in Theorem 2.10 prove that

$$\mathcal{Z}^{\Psi}(t, \xi) = V^{\Psi}(t, \xi),$$

if $V^{\Psi}(t, \xi) < \infty$.

If the value function is law invariant (see Proposition 2.13) and Theorem 2.10 holds true, we expect y to be formally (by combining arguments from [6, 21]) characterized by a Master Bellman equation in Wassertein space with oblique derivative boundary conditions.

3 An alternative auxiliary problem

We study the constrained McKean-Vlasov control problem

$$V := \inf_{\alpha \in \mathcal{A}} \big\{ J(X_0, \alpha) : \Psi(t, \mathbb{P}_{X_t^{\alpha}}) \le 0,\ \forall\ t \in [0,T],\ \varphi(\mathbb{P}_{X_T^{\alpha}}) \le 0 \big\},$$

where we now assume that the running constraint Ψ is continuous (hence, no discrete time constraints, see Remark 4.4), and with a terminal constraint function φ. We now consider an alternative auxiliary control problem as in [7]:

$$w(z) := \inf_{\alpha \in \mathcal{A}} \Big[\{\widehat{g}(\mathbb{P}_{X_T^{\alpha}}) - Z_T^{z,\alpha}\}_+ + \sum_{l=1}^{k} \int_0^T \{\Psi^l(s, \mathbb{P}_{X_s^{\alpha}})\}_+ \,\mathrm{d}s + \{\varphi(\mathbb{P}_{X_T^{\alpha}})\}_+ \Big]. \tag{3.1}$$

Compared to the control problem (2.5) of the previous section, the penalization term of the constrained function Ψ is in integral form instead of a supremum form. It follows that this problem is not path-dependent, and we shall show that it also provides a similar representation of the value function by its zero level set:

$$V = \inf\{z \in \mathbb{R} : w(z) = 0\},$$

but under the additional assumption that optimal controls do exist. Actually, we prove this result in the more general case with common noise in the next section.

The mean-field control problem (3.1) is Markovian with respect to the state variables $(X_t^\alpha, \mathbb{P}_{X_t^\alpha}, Z_t^{z,\alpha})$, and it is known from [20] that the infimum over open-loop controls α in $\mathcal{A}$ can be taken equivalently over randomized feedback policies, i.e. controls α in the form: $\alpha_t = \mathfrak{a}(t, X_t^\alpha, \mathbb{P}_{X_t^\alpha}, Z_t^{z,\alpha}, U)$, for some deterministic function $\mathfrak{a}$ from $[0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R} \times [0,1]$ into A, where U is an $\mathcal{F}_0$-measurable uniform random variable on $[0,1]$.

Let us now discuss conditions under which the infimum in (3.1) can be taken equivalently over (deterministic) feedback policies, i.e. for controls α in the form: $\alpha_t = \mathfrak{a}(t, X_t^\alpha, \mathbb{P}_{X_t^\alpha}, Z_t^{z,\alpha})$, for some deterministic function $\mathfrak{a}$ from $[0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}$ into A. This will be helpful for numerical purpose in Section 5. We assume on top of Assumption 2.11 that the running cost f, the drift b and the volatility coefficient σ do not depend on the law of the control process. We also assume that the running cost $f = f(t,x,\mu)$ does not depend on the control argument. The terminal constraint function φ should also verify the same assumptions as the terminal cost function g, namely Lipschitz continuity and local boundedness (see Assumption 2.11).

In this case, the corresponding dynamic auxiliary problem of (3.1) is written as

$$\mathrm{w}(t,\mu,z) = \inf_{\alpha \in \mathcal{A}} \Big[\{\widehat{g}(\mathbb{P}_{X_T^{t,\xi,\alpha}}) - Z_T^{t,\xi,z,\alpha}\}_+ + \sum_{l=1}^k \int_0^T \{\Psi^l(s, \mathbb{P}_{X_s^{t,\xi,\alpha}})\}_+ \, \mathrm{d}s + \{\varphi(\mathbb{P}_{X_T^{t,\xi,\alpha}})\}_+ \Big] \tag{3.2}$$

$$X_r^{t,\xi,\alpha} = \xi + \int_t^r b\big(s, X_s^{t,\xi,\alpha}, \alpha_s, \mathbb{P}_{X_s^{t,\xi,\alpha}}\big) \, \mathrm{d}s + \int_t^r \sigma\big(s, X_s^{t,\xi,\alpha}, \alpha_t, \mathbb{P}_{X_s^{t,\xi,\alpha}}\big) \, \mathrm{d}W_s, \quad \xi \sim \mu,$$

$$Z_r^{t,\xi,z,\alpha} = z - \int_t^r \bar{f}\big(s, \mathbb{P}_{X_s^{t,\xi,\alpha}}\big) \, \mathrm{d}s, \quad r \geq t,$$

where $\bar{f}$ is the function defined on $[0,T] \times \mathcal{P}_2(\mathbb{R}^d)$ by $\bar{f}(t,\mu) = \int_{\mathbb{R}^d} f\big(t,x,\mu\big) \, \mu(\mathrm{d}x)$. Note that we have applied Theorem 3.5 from [20] to obtain the law invariance of the auxiliary value function which can be written as a function of the measure μ. From Theorem 3.5, Proposition 5.6. 2), and equation (5.17) in [20] (see also Remark 5.2. from [21] and Section 6 in [38]) we see that the Bellman equation for problem (3.2) is:

$$\begin{cases} \partial_t \mathrm{w}(t,\mu,z) + \mathbb{E}[\inf_{a \in A}\{b\big(t,\xi,a,\mu\big)\partial_\mu \mathrm{w}(t,\mu,z)(\xi) - \bar{f}(t,\mu)\partial_z \mathrm{w}(t,\mu,z) \\ \quad + \frac{1}{2}\mathrm{Tr}(\sigma\sigma^\top\big(t,\xi,a,\mu\big)\partial_x\partial_\mu \mathrm{w}(t,\mu,z)(\xi))\}] + \sum_{l=1}^k \{\Psi^l(t,\mu)\}_+ = 0 \text{ for } (t,\mu,z) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R} \\ \mathrm{w}(T,\mu,z) = \{\widehat{g}(\mu) - z\}_+ + \{\varphi(\mu)\}_+ \text{ for } (\mu,z) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}\,. \end{cases} \tag{3.3}$$

By assuming that w is a smooth solution to this Bellman equation, and when the infimum in

$$\inf_{a \in A}\{b\big(t,x,a,\mu\big)\partial_\mu \mathrm{w}(t,\mu,z)(x) - \bar{f}(t,\mu)\partial_z \mathrm{w}(t,\mu,z) + \frac{1}{2}\mathrm{Tr}(\sigma\sigma^\top\big(t,x,a,\mu\big)\partial_x\partial_\mu \mathrm{w}(t,\mu,z)(x))\}$$

is attained for some measurable function $\hat{\mathfrak{a}}(t,x,\mu,z)$ on $[0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}$, we get an optimal control for (3.1) given in feedback form by $\alpha_t^* = \hat{\mathfrak{a}}(t, X_t^{\alpha^*}, \mathbb{P}_{X_t^{\alpha^*}}, Z_t^{z,\alpha^*})$, $0 \leq t \leq T$, which shows that one can restrict in (3.1) to deterministic feedback policies.

4 Extension to the common noise setting

We briefly discuss how the state constraints can be extended to mean-field control problems with common noise. In this case, in contrast with the previous section, we need to assume the existence of optimal control for the auxiliary unconstrained problem. It is similar to the assumption made by [7]. Let W^0 be a p-dimensional Brownian motion independent of W, and denote by $\mathbb{F}^0 = (\mathcal{F}_t^0)_t$ the filtration generated by W^0. We consider the following cost and dynamics:

$$J(\alpha) = \mathbb{E}\Big[\int_0^T f\big(t, X_t^\alpha, \alpha_t, \mathbb{P}^{W^0}_{(X_t^\alpha,\alpha_t)}\big) \, \mathrm{d}t + g\big(X_T^\alpha, \mathbb{P}^{W^0}_{X_T}\big)\Big]$$

$$\mathrm{d}X_t^\alpha = b\big(t, X_t^\alpha, \alpha_t, \mathbb{P}^{W^0}_{(X_t^\alpha,\alpha_t)}\big) \, \mathrm{d}t + \sigma\big(t, X_t^\alpha, \alpha_t, \mathbb{P}^{W^0}_{(X_t^\alpha,\alpha_t)}\big) \, \mathrm{d}W_t + \sigma^0\big(t, X_t^\alpha, \alpha_t, \mathbb{P}^{W^0}_{(X_t^\alpha,\alpha_t)}\big) \, \mathrm{d}W_t^0.$$

where $\mathbb{P}^{W^0}_{(X_t^\alpha,\alpha_t)}$ is the joint conditional law of (X_t^α,α_t) given W^0. The control process α belongs to a set $\mathcal{A}$ of $\mathbb{F}$-progressively measurable processes with values in a set $A\subset\mathbb{R}^q$.

The controlled McKean-Vlasov process X is constrained to verify $\Psi(t,\mathbb{P}^{W_0}_{X_t^\alpha})\le 0$ and $\varphi(\mathbb{P}^{W_0}_{X_T^\alpha})\le 0$. The proofs still follow the arguments from [7] but are slightly more involved than in Section 2 due to the additional noise appearing in the conditional law with respect to the common noise. We refer to [39, 26] for the dynamic programming approach to these problems. The problem of interest is

$$V^0=\inf_{\alpha\in\mathcal{A}}\{J(\alpha)\mid\Psi(t,\mathbb{P}^{W_0}_{X_t^\alpha})\le 0,\ \forall\, t\in[0,T],\ \varphi(\mathbb{P}^{W_0}_{X_T^\alpha})\le 0\ \}.$$

4.1 Representation by a stochastic target problem and an associated control problem

Given $z\in\mathbb{R}$, $\alpha\in\mathcal{A}$, and $\beta\in L^2(\mathbb{F}^0,\mathbb{R}^p)$, the set of $\mathbb{R}^p$-valued $\mathbb{F}^0$-adapted processes β s.t. $\mathbb{E}[\int_0^T|\beta_t|^2\mathrm{d}t]<\infty$, define

$$Z_t^{z,\alpha,\beta}:=z-\int_0^t\widehat{f}(s,\mathbb{P}^{W^0}_{(X_s^\alpha,\alpha_s)})\,\mathrm{d}s+\int_0^t\beta_s\,\mathrm{d}W_s^0,\quad 0\le t\le T. \tag{4.1}$$

Lemma 4.1. *The value function admits the* ***stochastic target problem*** *representation*

$$\begin{aligned}V^0=\inf\{z\in\mathbb{R}\mid\exists\,(\alpha,\beta)\in\mathcal{A}\times L^2(\mathbb{F}^0,\mathbb{R}^p)\text{ s.t. }&\widehat{g}(\mathbb{P}^{W^0}_{X_T^\alpha})\le Z_T^{z,\alpha,\beta},\\ &\Psi(t,\mathbb{P}^{W^0}_{X_t^\alpha})\le 0,\ \forall\,t\in[0,T],\ \varphi(\mathbb{P}^{W^0}_{X_T^\alpha})\le 0,\ \mathbb{P}\text{ a.s.}\}.\end{aligned}$$

Lemma 4.1 is proven in Section 4.2.

Define the **auxiliary unconstrained** control problem

$$\mathcal{U}(z):=\inf_{(\alpha,\beta)\in\mathcal{A}\times L^2(\mathbb{F}^0,\mathbb{R}^p)}\mathbb{E}\Big[\{\widehat{g}(\mathbb{P}^{W^0}_{X_T^\alpha})-Z_T^{z,\alpha,\beta}\}_++\sum_{l=1}^k\int_0^T\{\Psi^l(s,\mathbb{P}^{W^0}_{X_s^\alpha})\}_+\,\mathrm{d}s+\{\varphi(\mathbb{P}^{W^0}_{X_T^\alpha})\}_+\Big] \tag{4.2}$$

for $z\in\mathbb{R}$. We notice that $\mathcal{U}(z)\ge 0$.

Proposition 4.2. *$\mathcal{U}$ is 1-Lipschitz. For any $z,z'\in\mathbb{R}$*

$$|\mathcal{U}(z)-\mathcal{U}(z')|\le|z-z'|.$$

Proposition (4.2) is proven exactly as (2.3).

Assumption 4.3. *Problem* (4.2) *admits an optimal control for any $z\in\mathbb{R}$ and the constraint function $(t,\mu)\in[0,T]\times\mathcal{P}_2(\mathbb{R}^d)\mapsto\Psi(t,\mu)$ is continuous.*

Remark 4.4. *Please note that the integral penalization in* (4.2) *does not allow to consider discrete times constraints (except at terminal time) because the contribution to the integral would be null and the constraint function Ψ would be discontinuous in time. We could consider discrete time constraints in the objective of the auxiliary problem by adding a sum of functions of $\mathbb{P}^{W^0}_{X_{t_i}^\alpha}$ for some $(t_i)_i\in[0,T]$ but it would lose its standard Bolza form.*

Define $\mathcal{Z}=\inf\{z\in\mathbb{R}\mid\mathcal{U}(z)=0\}$.

Theorem 4.5.
1. *If $\exists\,(\alpha,\beta)\in\mathcal{A}\times L^2(\mathbb{F}^0,\mathbb{R}^p)$, $\widehat{g}(\mathbb{P}^{W^0}_{X_T^\alpha})\le Z_T^{z,\alpha,\beta}$, $\Psi(s,\mathbb{P}^{W^0}_{X_s^\alpha})\le 0,\ \forall\,s\in[0,T]$, and $\varphi(\mathbb{P}^{W^0}_{X_T^\alpha})\le 0$, $\mathbb{P}$ a.s. then $\mathcal{U}(z)=0$. Hence $\mathcal{Z}\le V^0$.*
2. *The value function verifies $V^0\le\mathcal{Z}$ thus $V^0=\mathcal{Z}$. Moreover optimal controls for the problem $\mathcal{U}(\mathcal{Z})=0$ are optimal for the original problem.*

Theorem 4.5 is proven in Section 4.2.

4.2 Proofs in the common noise framework

Proof of Lemma 4.1. We first observe that

$$V^0 = \inf\{z \in \mathbb{R} \mid \exists\, \alpha \in \mathcal{A} \text{ s.t. } \mathbb{E}\Big[\int_0^T \widehat{f}(s, \mathbb{P}^{W^0}_{(X^\alpha_s, \alpha_s)})\, \mathrm{d}s + \widehat{g}(\mathbb{P}^{W^0}_{X^\alpha_T})\Big] \le z,$$
$$\Psi(s, \mathbb{P}^{W^0}_{X^\alpha_s}) \le 0,\ \forall\, s \in [0,T],\ \mathbb{P}^0 \text{ a.s.}\}.$$

We need to prove that for $z \in \mathbb{R}$

$$\exists\, (\alpha,\beta) \in \mathcal{A} \times L^2(\mathbb{F}^0, \mathbb{R}^p) \text{ s.t. } \widehat{g}(\mathbb{P}^{W^0}_{X^\alpha_T}) \le Z^{z,\alpha,\beta}_T,\ \Psi(s, \mathbb{P}^{W^0}_{X^\alpha_s}) \le 0,\ \forall\, s \in [0,T],\ \varphi(\mathbb{P}^{W^0}_{X^\alpha_T}) \le 0,\ \mathbb{P}^0 \text{ a.s.}, \tag{4.3}$$

and

$$\exists\, \alpha \in \mathcal{A} \text{ s.t. } \mathbb{E}\Big[\int_0^T \widehat{f}(s, \mathbb{P}^{W^0}_{(X^\alpha_s, \alpha_s)})\, \mathrm{d}s + \widehat{g}(\mathbb{P}^{W^0}_{X^\alpha_T})\Big] \le z,\ \Psi(s, \mathbb{P}^{W^0}_{X^\alpha_s}) \le 0,\ \forall\, s \in [0,T],\ \varphi(\mathbb{P}^{W^0}_{X^\alpha_T}) \le 0,\ \mathbb{P}^0 \text{ a.s.}, \tag{4.4}$$

are equivalent. It is immediate to see that (4.3) $\implies$ (4.4) by taking the expectation and noticing that the Itô integral is a true martingale. Conversely, assuming (4.4), the martingale representation theorem provides a process $\widehat{\beta}$ such that

$$\begin{aligned} z &\ge \mathbb{E}\Big[\int_0^T \widehat{f}(s, \mathbb{P}^{W^0}_{(X^\alpha_s, \alpha_s)})\, \mathrm{d}s + \widehat{g}(\mathbb{P}^{W^0}_{X^\alpha_T})\Big] \\ &= \int_0^T \widehat{f}(s, \mathbb{P}^{W^0}_{(X^\alpha_s, \alpha_s)})\, \mathrm{d}s + \widehat{g}(\mathbb{P}^{W^0}_{X^\alpha_T}) - \int_0^T \widehat{\beta}_s\, \mathrm{d}W^0_s. \end{aligned}$$

Thus by (4.1)

$$Z^{z,\alpha,\widehat{\beta}}_T \ge \widehat{g}(\mathbb{P}^{W^0}_{X^\alpha_T}),\ \mathbb{P}^0 \text{ a.s.},$$

and we see that (4.4) $\implies$ (4.3). Then the result follows. □

Proof of Theorem 4.5. 1) $\exists\, (\alpha,\beta) \in \mathcal{A} \times L^2(\mathbb{F}^0, \mathbb{R}^p)$, $\widehat{g}(\mathbb{P}^{W^0}_{X^\alpha_T}) \le Z^{z,\alpha,\beta}_T$, $\Psi(s, \mathbb{P}^{W^0}_{X^\alpha_s}) \le 0$, $\forall\, s \in [0,T]$ and $\varphi(\mathbb{P}^{W^0}_{X^\alpha_T}) \le 0$, $\mathbb{P}^0$ a.s.. Therefore

$$\{\widehat{g}(\mathbb{P}^{W^0}_{X^\alpha_T}) - Z^{z,\alpha,\beta}_T\}_+ + \sum_{l=1}^k \int_0^T \{\Psi^l(s, \mathbb{P}^{W^0}_{X^\alpha_s})\}_+\, \mathrm{d}s + \{\varphi(\mathbb{P}^{W^0}_{X^\alpha_T})\}_+ = 0,\ \mathbb{P}^0 \text{ a.s.}$$

and by non-negativity of $\mathcal{U}$ we obtain $\mathcal{U}(z) = 0$. Then with optimal controls α^*, β^* we obtain $\mathcal{U}(V^0) = 0$. By definition of $\mathcal{Z}$ the property is established.

2) By 1) and the continuity given by Proposition 4.2, we obtain $\mathcal{U}(\mathcal{Z}) = 0$. Then by Assumption 4.3 $\exists\, (\alpha,\beta) \in \mathcal{A} \times L^2(\mathbb{F}^0, \mathbb{R}^p)$ such that

$$\mathbb{E}^0\Big[\{\widehat{g}(\mathbb{P}^{W^0}_{X^\alpha_T}) - Z^{\mathcal{Z},\alpha,\beta}_T\}_+ + \sum_{l=1}^k \int_0^T \{\Psi(s, \mathbb{P}^{W^0}_{X^\alpha_s})\}_+\, \mathrm{d}s + \{\varphi(\mathbb{P}^{W^0}_{X^\alpha_T})\}_+\Big] = 0.$$

The three terms on the l.h.s. being non-negative, they are in fact null $\mathbb{P}$ a.s. Thus

$$(\mathbb{P}^{W^0}_{X^\alpha_T}, Z^{\mathcal{Z},\alpha,\beta}_T) \in \mathrm{Epi}(\widehat{g}),\ \Psi(s, \mathbb{P}^{W^0}_{X^\alpha_s}) \le 0\ \forall\, s \in [0,T], \text{ and } \varphi(\mathbb{P}^{W^0}_{X^\alpha_T}) \le 0 \quad \mathbb{P} \text{ a.s.}$$

by continuity of Ψ and of $s \mapsto \mathbb{P}^{W^0}_{X^\alpha_s}$, which means $V^0 \le \mathcal{Z}$. By 1) it yields $V^0 = \mathcal{Z}$. As a consequence the previous proof provides an optimal control α for the original problem. □

5 Applications and numerical tests

We design several machine learning methods to solve this problem. We discretize the problem in time, parametrize the control by a neural network and directly minimize the cost. When the constraints are almost sure, we can sometimes enforce them by choosing an appropriate neural network architecture, for instance in storage problems. A more adaptive alternative is to solve the unconstrained auxiliary problem. We propose an extension of the first algorithm from [17] to achieve this task. Thus we obtain a machine learning method able to solve state constrained mean field control problems.

5.1 Algorithms

We solve the auxiliary problem in the simpler case without common noise with a first algorithm. We fix a relevant line segment K of $\mathbb{R}$ on which we are going to explore the potential values of the problem. For instance we know that the value is greater than the value of the unconstrained problem V therefore it is useless to compute the auxiliary value function for $z \leq V$. We discretize the problem in time on the grid $t_k := \frac{kT}{N}$. We call $\Delta t := \frac{kT}{N}$ and the Brownian increment $\Delta W_i := W_{t_{i+1}} - W_{t_i}$. For $j = 1, \cdots, N$, ΔW_i^j (respectively X_0^j) correspond to samples from N independent Brownian motions W^j (respectively from N independent random variables with law μ_0). For training we discretize K by using N points. We choose ε as a small parameter, typically of order 10^{-6}. We refer to [5] for results on the numerical approximation of level sets with a given threshold in the context of constrained deterministic optimal control. We propose the following extension of the Method 1 from [17]. It is tested in Subsection 5.2. It can indeed also be used to solve unconstrained problem.

Remark 5.1. *We point out that adding an additional parameter $\Lambda > 0$ in front of the constraint function does not modify the representation results. In that case we solve the following auxiliary problem*

$$\mathcal{Y}_\Lambda^\Psi := z \in \mathbb{R} \mapsto \inf_{\alpha \in \mathcal{A}} \Big[\{\widehat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z,\alpha}\}_+ + \Lambda \sum_{l=1}^k \int_0^T \{\Psi^l(s, \mathbb{P}_{X_s^\alpha})\}_+ \, \mathrm{d}s + \Lambda \varphi(\mathbb{P}_{X_T^\alpha}) \Big].$$

We discretize the problem in time and use a neural network by time step, since a single network taking time as input is usually not sufficient enough for complex problems, as shown in [43]. In view of the discussion about closed-loop controls in Section 3, the neural network representing the control at each time step takes as inputs the current states X and $Z_i^{z,\alpha}$ where z is taken on a discretization of K. The method is described in Algorithm 1 with an example in Section 5.2. Solving (3.3) with the approach of [31] would provide another numerical method for mean-field control with state constraints. The extension to the common noise case is given in Algorithm 2 where the neural network for the control at each time step t_i takes in addition as input the current value of the common noise $W_{t_i}^0$. Notice that in general, the control may depend on the past values of the common noise, which could be taken into account in the neural network by taking as inputs the past increments of the common noise $\Delta W_0^0, \dots, \Delta W_{i-1}^0$, where $\Delta W_i^0 = W_{t_{i+1}}^0 - W_{t_i}^0$. The neural network for the auxiliary control β at each time t_i takes as inputs the current state $Z_i^{z,\alpha}$ and the current value of the common noise. An illustration is given in Section 5.3.

Algorithm 1: Algorithm to solve mean-field control with probabilistic constraints

For a discretization $z_1 < \cdots < z_M$ of K, minimize over neural networks $(\alpha_i)_{i\in 0,\cdots,N_T-1}$: $\mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^q$ the loss function

$$\sum_{m=1}^{M} w_\Lambda(z_m)$$

with w_Λ defined by

$$w_\Lambda(z) := \mathbb{E}\Big[\{\frac{1}{N}\sum_{l=1}^{N} g\Big(X^l_{N_T}, \frac{1}{N}\sum_{j=1}^{N}\delta_{X^j_{N_T}}\Big) - Z^{z,\alpha}_{N_T}\}_+ + \Lambda\sum_{m=1}^{k}\sum_{i=1}^{N_T}\{\Psi^m\Big(t_i, \frac{1}{N}\sum_{j=1}^{N}\delta_{X^j_i}\Big)\}_+ + \Lambda\{\varphi\Big(\frac{1}{N}\sum_{j=1}^{N}\delta_{X^j_{N_T}}\Big)\}_+\Big].$$

/* Auxiliary problem */

and for $i = 0, \cdots, N_T - 1$, $j = 1, \cdots, N$

$$X^j_{i+1} = X^j_i + b\big(t_i, X^j_i, \alpha_i(X^j_i, Z^{z,\alpha}_i), \overline{\mu}_i\big)\Delta t + \sigma\big(t_i, X^j_i, \alpha_i(X^j_i, Z^{z,\alpha}_i), \overline{\mu}_i\big)\Delta W^j_i, \quad X^j_0 \sim \mu_0$$

$$Z^{z,\alpha}_{i+1} = Z^{z,\alpha}_i - \frac{1}{N}\sum_{l=1}^{N} f\big(t_i, X^l_i, \alpha_i(X^l_i, Z^{z,\alpha}_i), \overline{\mu}_i\big)\ \Delta t, \quad Z^{z,\alpha}_0 = z$$

$$\overline{\mu}_i = \frac{1}{N}\sum_{j=1}^{N}\delta_{(X^j_i, \alpha_i(X^j_i, Z^{z,\alpha}_i))}$$

/* Particle approximations */

Define α^* as the solution to this minimization problem.
Then, compute $V_0 = \inf\{z_i,\ i \in [\![1, M]\!] \mid w_\Lambda(z_i) \leq \varepsilon\}$ with $\alpha = \alpha^*$ in the dynamics.

/* Recovering the cost of the original problem */

Return the value V_0 and the optimal controls $\hat{\alpha}_i : x \mapsto \alpha^*_i(x, Z^{V_0,\alpha^*}_i)$ for $i = 0, \cdots, N_T - 1$.

/* Recovering the control of the original problem */

5.2 Mean-variance problem with state constraints

We consider the celebrated Markowitz portfolio selection problem where an investor can invest at any time t an amount α_t in a risky asset (assumed for simplicity to follow a Black-Scholes model with constant rate of return r and volatility $\sigma > 0$), hence generating a wealth process $X = X^\alpha$ with dynamics

$$\mathrm{d}X_t = \alpha_t r\ \mathrm{d}t + \alpha_t\sigma\ \mathrm{d}W_t, \quad 0 \leq t \leq T,\ X_0 = x_0 \in \mathbb{R}\,.$$

The goal is then to minimize over portfolio control α the mean-variance criterion :

$$\inf_\alpha\ J(\alpha) = \lambda \mathrm{Var}(X^\alpha_T) - \mathbb{E}[X^\alpha_T] \tag{5.1}$$

where $\lambda > 0$ is a parameter related to the risk aversion of the investor. We will add to this standard problem a conditional expectation constraint in the form

$$E[X^\alpha_t \mid X^\alpha_t \leq \theta] \geq \delta,\ \text{if}\ \mathbb{P}(X^\alpha_t \leq \theta) > 0,$$

with $\delta < \theta$, which can be reformulated as

$$0 \geq (\delta - E[X^\alpha_t \mid X^\alpha_t \leq \theta])\mathbb{P}(X^\alpha_t \leq \theta).$$

Algorithm 2: Algorithm to solve mean-field control with probabilistic constraints and common noise

For a discretization $z_1 < \cdots < z_M$ of K, minimize over neural networks $(\alpha_i)_{i\in 0,\cdots,N_T-1}$: $\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^p \to \mathbb{R}^q$ and $(\beta_i)_{i\in 0,\cdots,N_T-1}$: $\mathbb{R} \times \mathbb{R}^p \to \mathbb{R}^p$ the loss function

$$\sum_{m=1}^{M} w_\Lambda(z_m)$$

with w_Λ defined by

$$w_\Lambda(z) := \mathbb{E}\Big[\{\frac{1}{N}\sum_{l=1}^{N} g\Big(X^l_{N_T}, \frac{1}{N}\sum_{j=1}^{N}\delta_{X^j_{N_T}}\Big) - Z^{z,\alpha,\beta}_{N_T}\}_+ + \Lambda \sum_{m=1}^{k}\sum_{i=1}^{N_T}\{\Psi^m\Big(t_i, \frac{1}{N}\sum_{j=1}^{N}\delta_{X^j_i}\Big)\}_+ \ \Delta t$$
$$+ \Lambda\{\varphi\Big(\frac{1}{N}\sum_{j=1}^{N}\delta_{X^j_{N_T}}\Big)\}_+\Big].$$

`/* Auxiliary problem */`

and for $i = 0, \cdots, N_T - 1$, $j = 1, \cdots, N$

$$X^j_{i+1} = X^j_i + b\big(t_i, X^j_i, \alpha_i(X^j_i, Z^{z,\alpha,\beta}_i, W^0_{t_i}), \overline{\mu}_i\big)\Delta t + \sigma\big(t_i, X^j_i, \alpha_i(X^j_i, Z^{z,\alpha,\beta}_i, W^0_{t_i}), \overline{\mu}_i\big)\Delta W^j_i$$
$$+ \sigma_0\big(t_i, X^j_i, \alpha_i(X^j_i, Z^{z,\alpha,\beta}_i, W^0_{t_i}), \overline{\mu}_i\big)\Delta W^0_i, \quad X^j_0 \sim \mu_0$$
$$Z^{z,\alpha,\beta}_{i+1} = Z^{z,\alpha,\beta}_i - \frac{1}{N}\sum_{l=1}^{N} f\big(t_i, X^l_i, \alpha_i(X^l_i, Z^{z,\alpha,\beta}_i, W^0_{t_i}), \overline{\mu}_i\big)\ \Delta t + \beta_i(Z^{z,\alpha,\beta}_i, W^0_{t_i})\ \Delta W^0_i, \quad Z^{z,\alpha,\beta}_0 = z$$

$$\overline{\mu}_i = \frac{1}{N}\sum_{j=1}^{N}\delta_{(X^j_i, \alpha_i(X^j_i, Z^{z,\alpha,\beta}_i, W^0_{t_i}))}$$

`/* Particle approximations */`

Define (α^*, β^*) as the solution to this minimization problem.

Then, compute $V_0 = \inf\{z_i,\ i \in [\![1, M]\!] \mid w_\Lambda(z_i) \le \varepsilon\}$ with $\alpha = \alpha^*$ and $\beta = \beta^*$ in the dynamics.

`/* Recovering the cost of the original problem */`

Return the value V_0 and the optimal controls $\hat{\alpha}_i : x \mapsto \alpha^*_i(x, Z^{V_0,\alpha^*,\beta^*}_i, W^0_{t_i})$ for $i = 0, \cdots, N_T - 1$.

`/* Recovering the control of the original problem */`

The auxiliary deterministic unconstrained control problem is therefore

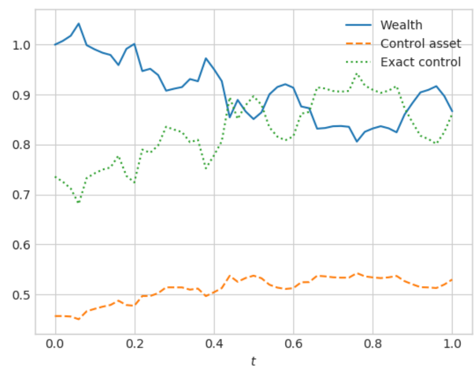
$$\mathcal{Y}_\Lambda(z) := \inf_{\alpha \in \mathcal{A}} \left[\{\lambda \text{Var}(X_T^\alpha) - \mathbb{E}[X_T^\alpha] - z\}_+ + \Lambda \int_0^T \{(\delta - E[X_s^\alpha \mid X_s^\alpha \leq \theta])\mathbb{P}(X_s^\alpha \leq \theta)\}_+ \, \mathrm{d}s \right]$$

with the dynamics $\mathrm{d}X_s^\alpha = \alpha_s r \, \mathrm{d}s + \alpha_s \sigma \, \mathrm{d}W_s$, which corresponds to the constraint function $\Psi(t, \mu) \mapsto (\delta - E_\mu[\xi \mid \xi \leq \theta])\mu((-\infty, \theta])$. We have the representation $J(\alpha^*) = \mathcal{Z} = \inf\{z \in \mathbb{R} \mid \mathcal{Y}_\Lambda(z) = 0\}$. Indeed we see that the null control is admissible with the modified constraint $E[X_t^\alpha \mid X_t^\alpha \leq \theta]\mathbb{P}(X_t^\alpha \leq \theta) = 0 \geq (\delta + \varepsilon)\mathbb{P}(X_t^\alpha \leq \theta) = 0, \ \forall t \in [0, T]$ for any $0 < \varepsilon < \theta - \delta$ because $x_0 \geq \theta$ hence $\mathbb{P}(X_t^\alpha \leq \theta) = 0$ so we can apply Theorem 2.10. For practical application, other constraints could be considered like almost sure constraints on the portfolio weights as in [42]. Instead of the dualization method used by [35], constraints on the law of the tracking error with respect to a reference portfolio could be enforced.

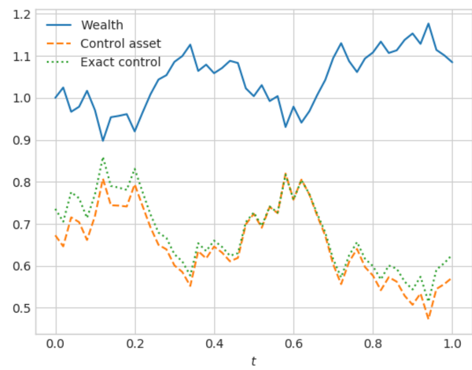
For numerical tests we take $r = 0.15, \ \sigma = 0.35, \ \lambda = 1$. We choose $x_0 = 1, \theta = 0.9, \ \delta = 0.8$ and solve

$$\begin{aligned} &\inf_\alpha J(\alpha) = \lambda \text{Var}(X_T^\alpha) - \mathbb{E}[X_T^\alpha] \\ &\quad \mathrm{d}X_t = \alpha_t r \, \mathrm{d}t + \alpha_t \sigma \, \mathrm{d}W_t, \\ &\quad (0.8 - E[X_t^\alpha \mid X_t^\alpha \leq 0.9])\mathbb{P}(X_t^\alpha \leq 0.9) \leq 0, \ \forall t \in [0, T]. \end{aligned} \tag{5.2}$$

We compare the controls from Algorithm 1 with the exact optimal ones in the unconstrained case for which we have an analytical value. We also solve without constraints for comparison and plot the final time histograms. We solve the unconstrained case with algorithm 1 and the one from [17] for comparison. We take 50 time steps for the time discretization and a batch size of 20000. We use an feedforward architecture with two hidden layers of 15 neurons. We perform 15000 gradient descent iterations thanks to the Tensorflow library. The true value $v = J(\alpha^*)$ is -1.05041 without constraints. We also have the upper bound $-1.$ for the value in the constrained case, corresponding to the identically null control and wealth process $X_t = 1 \ \forall t \in [0, T]$. With constraint we choose $K = [-1.047, -1.041]$, without constraint we take $K = [-1.07, -1.03]$, discretized by regular grids with 25 points.

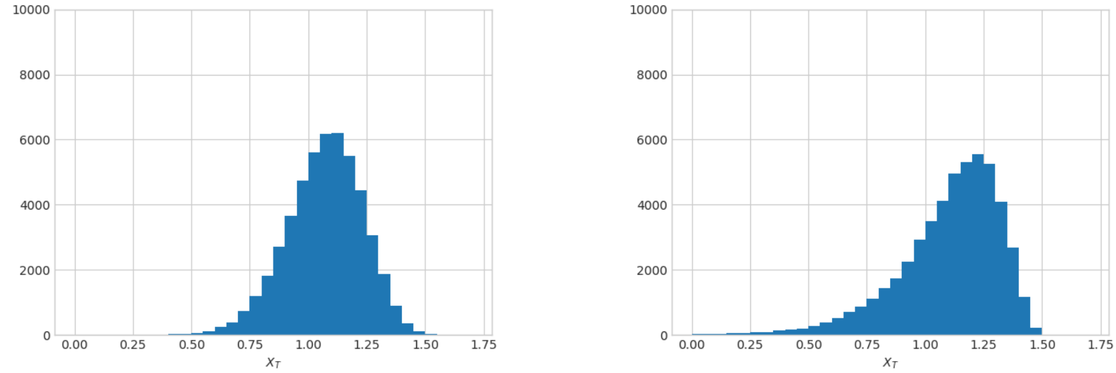


Problem (5.2)



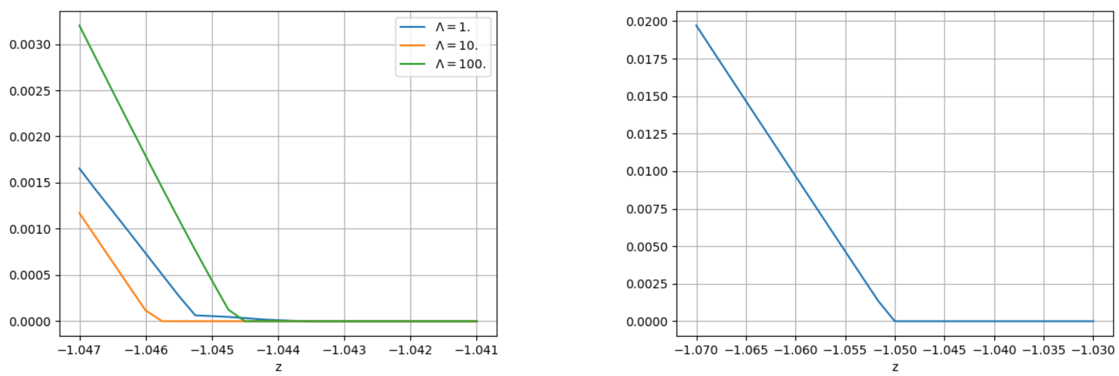
No constraint, problem (5.1)

Figure 1: Sample path of the controlled process X_t^α, with the analytical optimal control (for the unconstrained case) and the computed control. On the left figure we don't have the true control but plot the unconstrained one for comparison. Here $\Lambda = 100$



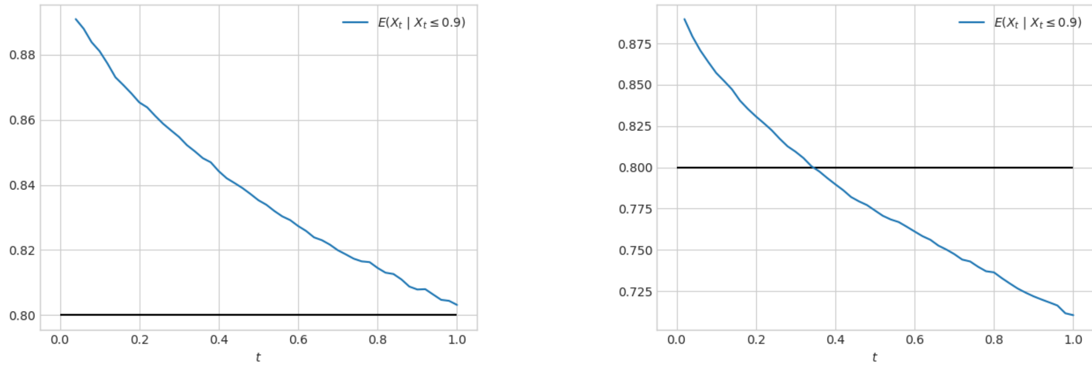
Problem (5.2) | No constraint, problem (5.1)

Figure 2: Histogram of X_T^α for 50000 samples. Here $\Lambda = 100$.



Problem (5.2) | No constraint, problem (5.1)

Figure 3: Auxiliary value function $\mathcal{Y}_\Lambda(z)$ for several values of Λ in the constrained case, auxiliary value function $\mathcal{Y}(z)$ in the unconstrained case



Problem (5.2) | No constraint, problem (5.1)

Figure 4: Conditional expectation $E[X_t^\alpha \mid X_t^\alpha \leq 0.9]$ estimated with 50000 samples. The black line corresponds to $\delta = 0.8$. Here $\Lambda = 100$

In Figure 2 we observe the shift of the distribution of the final wealth thanks to the constraint (on the left) with less probable large losses but also less probable large gains. Indeed Figure 4 confirms that the conditional expectation constraint is verified when we solve the corresponding problem through our level set approach. We see in Figure 3 that the more Λ is large the more the auxiliary value function becomes affine before reaching zero. Additional results are presented in Table 1.

Our method can also handle directly the primal of the mean-variance problem, that is to maximize over portfolio control α the expected terminal wealth under a terminal variance constraint:

$$
\begin{aligned}
\inf_{\alpha} \; & \bar{J}(\alpha) = -\mathbb{E}[X_T^\alpha] \\
& \mathrm{d}X_t = \; \alpha_t r \; \mathrm{d}t + \alpha_t \sigma \; \mathrm{d}W_t, \\
& \mathrm{Var}(X_T^\alpha) \leq \vartheta.
\end{aligned}
\tag{5.3}
$$

which give the same optimal control as Problem (5.1) under the correspondence $\lambda = \sqrt{\frac{\exp(\sigma^{-2} r^2 T) - 1}{4\vartheta}}$. This problem allows us to consider a constrained problem with an analytical solution. In this case $\mathrm{Var}(X_T^{\alpha^*}) = \vartheta$ thus $J(\alpha^*) = \lambda \mathrm{Var}(X_T^{\alpha^*}) + \bar{J}(\alpha^*) = \lambda\vartheta + \bar{J}(\alpha^*)$. See Remark 2.5 from [25]. For comparison with Problem (5.1) we thus report $\lambda\vartheta + \bar{J}(\alpha^*)$ for Problem (5.3) in Table 1 and choose $\vartheta = \frac{\exp(\sigma^{-2} r^2 T) - 1}{4\lambda^2} = 0.0504$. In this case the auxiliary deterministic unconstrained control problem is now

$$
\begin{aligned}
\mathcal{U}_\Lambda(z) = \inf_{\alpha \in \mathcal{A}} & \Big[\{ -\mathbb{E}[X_T^\alpha] - z \}_+ + \Lambda \{ \mathrm{Var}(X_T^\alpha) - \vartheta \}_+ \Big] \\
& \mathrm{d}X_t = \; \alpha_t r \; \mathrm{d}t + \alpha_t \sigma \; \mathrm{d}W_t,
\end{aligned}
$$

which corresponds to the constraint function $\Psi(t, \mu) \mapsto (\mathrm{Var}(\mu) - \vartheta)_+ \mathbb{1}_{t=T}$ and the modified constraint function $\overline{\Psi}_\eta(t, \mu) \mapsto (\mathrm{Var}(\mu) - \vartheta)_+ \mathbb{1}_{t=T} - \eta \mathbb{1}_{t<T}$ (see Remark 2.6). Theorem 2.10 still applies as far as the null control is admissible with the modified constraint $(\mathrm{Var}(\mu) - \vartheta)_+ \mathbb{1}_{t=T} + \varepsilon - \eta \mathbb{1}_{t<T} \leq 0$ for any $0 < \varepsilon < \eta$ and any $t \in [0, T]$.

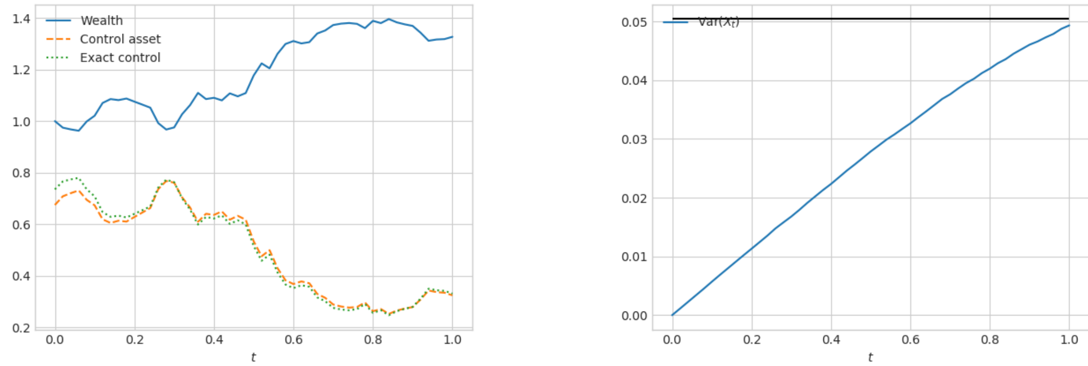


Figure 5: Sample trajectory of the controlled process X_t^α and the control for problem (5.3) (left). Variance $\mathrm{Var}(X_t)$ estimated with 50000 samples for problem (5.3) (right) with $\Lambda = 10$

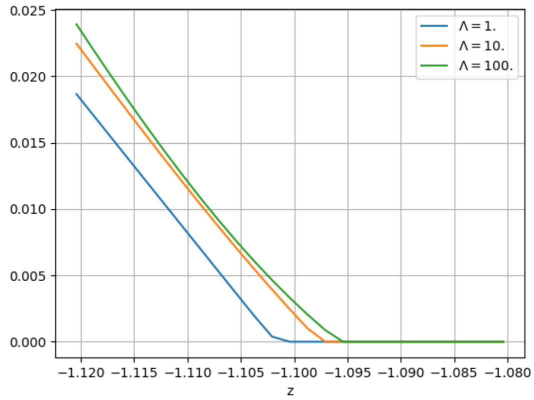


Figure 6: Auxiliary value function $\mathcal{U}_\Lambda(z)$ for several values of Λ

Figure 5 shows that we recover the optimal control for the problem and that the terminal variance constraint is satisfied. We see in Figure 6 that similarly as in Figure 3, for large values of Λ the auxiliary value function is affine before reaching zero. In this case the exact solution is -1.10 which is very close to the point in which the affine part reaches zero.

Problem	Average	Std	True val.	Error	$\mathbb{E}[X_T^{\alpha^*}]$	True $\mathbb{E}[X_T^{\alpha^*}]$	$\mathrm{Var}(X_T^{\alpha^*})$	True $\mathrm{Var}(X_T^{\alpha^*})$
(5.2)	-1.045	0.0005	?	?	1.07	?	0.027	?
(5.3)	-1.048	0.0017	-1.050	0.22	1.10	1.10	0.049	0.050
(5.1)	-1.050	0.0009	-1.050	0.07	1.10	1.10	0.050	0.050
(5.1) [17]	-1.052	0.0022	-1.050	0.13	1.10	1.10	0.053	0.050

Table 1: Estimate of the solution with maturity $T = 1$. Average and standard deviation observed over 10 independent runs are reported, with the relative error (in %). We also report the terminal expectation and variance of the approximated optimally controlled process for a single run. '?' means that we don't have a reference value. For problem (5.1) we take $\Lambda = 10$ and for problem (5.2) we choose $\Lambda = 100$

In Table 1 we observe that our method gives a small variance for the results over several runs. In the case where an analytical solution is known, the value of the control problem is computed accurately with less than 0.5% of relative error. The expectation and variance of the terminal value of the optimally controlled process are also very close to their theoretical values. In the case of a conditional expectation constraint, even though we don't have an exact solution we notice that the value is close to the unconstrained value hence since our solution is admissible, we expect to be near optimality. On the unconstrained problem (5.1) our scheme and the one from [17] give similar results.

5.3 Optimal storage of wind-generated electricity

We consider N wind turbines with N associated batteries. Define the productions P_t^i, storage levels X_t^i, storage injection α_t^i for which we provide a typical range[1]. We consider the following constraints

$$\begin{cases} 0 \leq X_t^i \leq X_{\max} \longrightarrow \text{limited storage capacity (1kWh} - \text{10 MWh)} \\ \underline{\alpha} \leq \alpha_t^i \leq \overline{\alpha} \longrightarrow \text{limited injection/withdrawal capacity (10 kW} - \text{10MW)} \end{cases}$$

with $X_{\max} \geq 0$, $\underline{\alpha} \leq 0 \leq \overline{\alpha}$. Define the spot price of electricity S_t without wind power, $\widetilde{S}_t$ the price with wind production. Selling a quantity $P_t^i - \alpha_t^i$ on the market, producer i obtains a revenue $\widetilde{S}_t(P_t^i - \alpha_t^i)$ (if $P_t^i - \alpha_t^i < 0$ the producer is buying from the market) where the market price is affected by linear price impact

$$\widetilde{S}_t = S_t - \frac{\Theta(N)}{N} \sum_{i=1}^{N} (P_t^i - \alpha_t^i),$$

modeling the impact of intermittent renewable production on the market. Θ is positive, non-decreasing and bounded. We call $\Theta_\infty = \lim_{N\to\infty} \Theta(N) < \infty$. We consider $N+2$ independent Brownian motions $W_t^0, B_t^0, W_t^1, \cdots, W_t^N$ and the following dynamics for the producers $i = 1, \cdots, N$ state processes

$$\begin{cases} \mathrm{d}X_t^i = \alpha_t^i \ \mathrm{d}t \\ \mathrm{d}P_t^i = \iota(\phi P_{\max} - P_t^i) \ \mathrm{d}t + \sigma_p (P_t \wedge \{P_{\max} - P_t^i\})_+ (\rho \ \mathrm{d}W_t^0 + \sqrt{1-\rho^2} \ \mathrm{d}W_t^i) \\ \mathrm{d}F(t,T) = F(t,T)\sigma_f e^{-a(T-t)} \mathrm{d}B_t^0 \\ S_t = F(t,t). \end{cases}$$

In the production dynamics, the common noises W_t^0, B_t^0 corresponds to the global weather and the market price randomness whereas the idiosyncratic noises W_t^i for $i > 1$ model the local weather, independent from one wind turbine to another. We call $\mathbb{F}^0$ the filtration generated by W_t^0, B_t^0. The productions P_t^i are bounded processes and the price S_t is positive. Of course the modified price $\widetilde{S}_t$ in the presence of renewable producers can become negative, as empirically observed in some overproduction events. However it stays bounded by below in our model. Producer i gain function to maximize is

$$J^i(\alpha_1, \cdots, \alpha_N) = \mathbb{E}\Big[\int_0^T \{S_t(P_t^i - \alpha_t^i) - \frac{\Theta(N)}{N}(P_t^i - \alpha_t^i) \sum_{j=1}^{N} (P_t^j - \alpha_t^j)\} \ \mathrm{d}t\Big].$$

The related mean field control problem for a central planner is therefore

$$\begin{aligned} &- \inf_{\alpha \in \mathcal{A}} \mathbb{E}\Big[\int_0^T \{-S_t(P_t - \alpha_t) + \Theta_\infty (P_t - \alpha_t) \mathbb{E}[P_t - \alpha_t | \mathbb{F}^0]\} \ \mathrm{d}t\Big] \\ &\qquad \mathrm{d}X_t = \alpha_t \ \mathrm{d}t \\ &\qquad \mathrm{d}P_t = \iota(\phi P_{\max} - P_t) \ \mathrm{d}t + \sigma_p (P_t \wedge \{P_{\max} - P_t\})_+ (\rho \ \mathrm{d}W_t^0 + \sqrt{1-\rho^2} \ \mathrm{d}W_t^1) \\ &\qquad \mathrm{d}F(t,T) = F(t,T)\sigma_f e^{-a(T-t)} \mathrm{d}B_t^0 \\ &\qquad S_t = F(t,t) \\ &\qquad 0 \leq X_t \leq X_{\max} \ \mathbb{P} \text{ a.s.} \end{aligned}$$

Here the state is $(X_t, P_t, S_t) \in \mathbb{R}^3$ hence the distribution of the state lives in $\mathcal{P}_2(\mathbb{R}^3)$. The set $\mathcal{A}$ corresponds to progressively measurable controls with values in the compact set $[\underline{\alpha}, \overline{\alpha}]$. A similar problem is solved by [1] without any storage constraints by Pontryagin principle. With constraints but without

[1] https://css.umich.edu/factsheets/us-grid-energy-storage-factsheet

mean-field interaction, a close problem is solved by [40]. For instance $X_{\max} = 0$ corresponds to the much simpler problem without storage nor control of the valuation of a wind power park. See also [22, 43]. To represent the almost sure constraint $0 \leq X_t \leq X_{\max}$ we choose as constrained function

$$\Psi : \mu \in \mathcal{P}_2(\mathbb{R}^3) \mapsto \int_{\mathbb{R}} \{(-x)_+^2 + (x - X_{\max})_+^2\} \ \mu_1(\mathrm{d}x),$$

where μ_1 is the first marginal law of the measure μ.

The auxiliary unconstrained control problem is therefore

$$\begin{aligned} w(z) := -\inf_{\alpha,\beta^{0,1},\beta^{0,2} \in \mathcal{A}\times L^2 \times L^2} \mathbb{E}\Big[\{\int_0^T \mathbb{E}[-S_t(P_t-\alpha_t) + \Theta_\infty(P_t-\alpha_t)\mathbb{E}[P_t-\alpha_t|\mathbb{F}^0]|\mathbb{F}^0] \ \mathrm{d}t - z \\ - \int_0^T \beta_s^{0,1} \ \mathrm{d}W_s^0 - \int_0^T \beta_s^{0,2} \ \mathrm{d}B_s^0\}_+ + \frac{1}{\epsilon}\int_0^T \mathbb{E}[(-X_s)_+^2 + (X_s - X_{\max})_+^2] \ \mathrm{d}s\Big] \end{aligned} \tag{5.4}$$

$$\begin{aligned} &\mathrm{d}X_t = \alpha_t \ \mathrm{d}t \\ &\mathrm{d}P_t = \iota(\phi P_{\max} - P_t) \ \mathrm{d}t + \sigma_p(P_t \wedge \{P_{\max} - P_t\})_+ \ (\rho \ \mathrm{d}W_t^0 + \sqrt{1-\rho^2} \ \mathrm{d}W_t^1) \\ &\mathrm{d}F(t,T) = F(t,T)\sigma_f e^{-a(T-t)}\mathrm{d}B_t^0 \\ &S_t = F(t,t) \end{aligned}$$

where ϵ is a small term used to force the a.s. constraints.

We consider the standard stochastic control benchmark with only common noise for the production ($\rho = 1$). It corresponds to a single very large wind farm where all wind turbines produce the same power. The problem degenerates as

$$\begin{aligned} &-\inf_{\alpha \in \mathcal{A}} \mathbb{E}\Big[\int_0^T \{(-S_t + \Theta_\infty(P_t - \alpha_t))(P - \alpha_t)\} \ \mathrm{d}t\Big] \\ &\mathrm{d}X_t = \alpha_t \ \mathrm{d}t \\ &\mathrm{d}P_t = \iota(\phi P_{\max} - P_t) \ \mathrm{d}t + \sigma_p(P_t \wedge \{P_{\max} - P_t\})_+ \ \mathrm{d}W_t^0 \\ &\mathrm{d}F(t,T) = F(t,T)\sigma_f e^{-a(T-t)}\mathrm{d}B_t^0 \\ &S_t = F(t,t) \\ &0 \leq X_t \leq X_{\max} \ \mathbb{P} \text{ a.s.} \end{aligned} \tag{5.5}$$

and equation (5.4) gives

$$\begin{aligned} w(z) := -\inf_{\alpha,\beta \in \mathcal{A}\times L^2} \mathbb{E}\Big[((Y^{\alpha,\beta} - z)_+ + \frac{1}{\epsilon}\int_0^T \mathbb{E}[(-X_s)_+^2 + (X_s - X_{\max})_+^2] \ \mathrm{d}s\Big] \\ \mathrm{d}X_t = \alpha_t \ \mathrm{d}t \\ \mathrm{d}P_t = \iota(\phi P_{\max} - P_t) \ \mathrm{d}t + \sigma_p(P_t \wedge \{P_{\max} - P_t\})_+ \ \mathrm{d}W_t^0 \\ \mathrm{d}F(t,T) = F(t,T)\sigma_f e^{-a(T-t)}\mathrm{d}B_t^0 \\ S_t = F(t,t) \end{aligned}$$

where

$$Y^{\alpha,\beta} = \int_0^T (-S_t + \Theta_\infty(P_t - \alpha_t))(P_t - \alpha_t) \ \mathrm{d}t - \int_0^T \beta_s^{0,1} \ \mathrm{d}W_s^0 - \int_0^T \beta_s^{0,2} \ \mathrm{d}B_s^0. \tag{5.6}$$

The solution of our optimization problem is then $z^* = \sup\{z \mid \hat{w}(z) = 0\}$ where $\hat{w}(z) := -w(z)$. Remark now that (5.6) will be estimated discretizing the integral $\int_0^T \beta_s^{0,1} \ \mathrm{d}W_s^0$ and $\int_0^T \beta_s^{0,2} \ \mathrm{d}B_s^0$ using an Euler scheme for the underlying processes and therefore $\hat{w}(z)$ will be above 0 except for low values of z due to

the variance of the $Y^{\alpha*,\beta}$ estimator that cannot be reduced to 0.
In order to reduce the variance of $Y^{\alpha,\beta}$, we propose to modify $Y^{\alpha,\beta}$ as follows :

$$Y^{\alpha,\beta} = \int_0^T (-S_t + \Theta_\infty(P_t - \alpha_t))(P_t - \alpha_t)\ \mathrm{d}t - \int_0^T (-S_t + \Theta_\infty(P_t - \hat{\alpha}_t))(P_t - \hat{\alpha}_t)\ \mathrm{d}t + \mathbb{E}[\int_0^T (-S_t + \Theta_\infty(P - \hat{\alpha}_t))(P_t - \hat{\alpha}_t)\ \mathrm{d}t] - \int_0^T \beta_s^{0,1}\ \mathrm{d}W_s^0 - \int_0^T \beta_s^{0,2}\ \mathrm{d}B_s^0$$

where $\hat{\alpha}_t$ is the rough estimation of the optimal **deterministic** command maximizing the gain.
We take $T = 40$, $N_T = 40$ time steps, $X_{\max} = 1$, $X_0 = 0.5$, $P_0 = 0.12$, $F(0,t) = 30 + 5\cos(\frac{2\pi t}{N}) + \cos(\frac{2\pi t}{7})$, $\sigma_f = 0.3$, $a = 0.16$, $\iota = 0.2$, $\sigma_p = 0.2$, $\phi = 0.3$, $P_{\max} = 0.2$, $-0.2 \leq \alpha \leq 0.2$, $\Theta(N) = 10$.
The network depends on P_t, S_t, X_t and z where z takes some deterministic values on a grid with the same spacing. The global curve is therefore approximated by a single run.
The grid is taken from 107 to 127 with a spacing of 0.5. The neural networks have two hidden layers with 14 neurons on each layer. We take a ϵ parameter equal to 10^{-4}. The number of gradient iterations is set to 50000 with a learning rate equal to 2×10^{-3} and every 100 iterations a more accurate estimate of $\hat{w}$ is calculated.
We give the $\hat{w}$ function on figure 7.

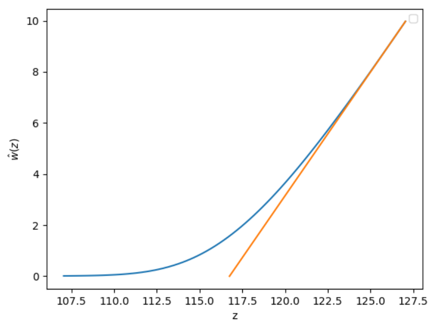


Figure 7: $\hat{w}$ function value for the storage problem

Using Dynamic Programming with the StOpt library [32], we get an optimal value equal to 117.28 while a direct optimization of (5.5) using some neural networks as in [43], [17] we get a value of 117.11. Encouraged by Remark 5.1, Figure 3, Figure 6 and the related comments, we empirically estimate the value function by the point where the linear part of the auxiliary function reaches zero when $\Lambda = \frac{1}{\varepsilon}$ is sufficiently large. The estimated value is 116.75, close to the reference solutions. On figure 8, we compare trajectories obtained by Dynamic Programming and by the Level Set approach : they are accurately calculated.

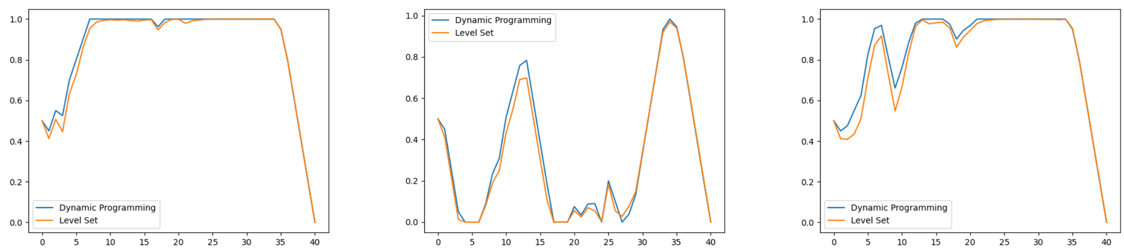


Figure 8: Storage trajectories with the Level Set and Dynamic Programming method.

References

[1] Clémence Alasseur, Imen Ben Taher, and Anis Matoussi. "An Extended Mean Field Game for Storage in Smart Grids". In: *Journal of Optimization Theory and Applications* 184 (2020), 644–670.

[2] Albert Altarovici, Olivier Bokanowski, and Hasnaa Zidani. "A general Hamilton-Jacobi framework for non-linear state-constrained control problems". In: *ESAIM: COCV* 19.2 (2013), pp. 337–357.

[3] Saeed Sadeghi Arjmand and Guilherme Mazanti. "Nonsmooth mean field games with state constraints". In: *arXiv:2110.15713* (2021).

[4] Alessandro Balata, Michael Ludkovski, Aditya Maheshwari, and Jan Palczewski. "Statistical learning for probability-constrained stochastic optimal control". In: *European Journal of Operational Research* 290.2 (2021), pp. 640–656.

[5] Olivier Bokanowski, Nidhal Gammoudi, and Hasnaa Zidani. "Optimistic Planning Algorithms For State-Constrained Optimal Control Problems". preprint. July 2021.

[6] Olivier Bokanowski, Athena Picarelli, and Hasnaa Zidani. "Dynamic Programming and Error Estimates for Stochastic Control Problems with Maximum Cost". In: *Appl Math Optim* 71 (2015), pp. 125–163.

[7] Olivier Bokanowski, Athena Picarelli, and Hasnaa Zidani. "State-Constrained Stochastic Optimal Control Problems via Reachability Approach". In: *SIAM Journal on Control and Optimization* 54.5 (2016), pp. 2568–2593.

[8] Benoît Bonnet. "A Pontryagin Maximum Principle in Wasserstein spaces for constrained optimal control problems". In: *ESAIM: COCV* 25 (2019), p. 52.

[9] Benoît Bonnet and Hélène Frankowska. "Necessary Optimality Conditions for Optimal Control Problems in Wasserstein Spaces". In: *Appl Math Optim* (2021).

[10] Bruno Bouchard, Boualem Djehiche, and Idris Kharroubi. "Quenched Mass Transport of Particles Toward a Target". In: *Journal of Optimization Theory and Applications* 186.2 (2020), pp. 345–374.

[11] Bruno Bouchard, Romuald Elie, and Cyril Imbert. "Optimal Control under Stochastic Target Constraints". In: *SIAM Journal on Control and Optimization* 48.5 (2010), pp. 3501–3531.

[12] Piermarco Cannarsa and Rossana Capuani. "Existence and Uniqueness for Mean Field Games with State Constraints". In: *PDE Models for Multi-Agent Phenomena*. Ed. by Cardaliaguet Pierre, Porretta A., and Salvarani F. Vol. 28. Springer INdAM Series. Springer, Cham, 2018.

[13] Piermarco Cannarsa, Rossana Capuani, and Pierre Cardaliaguet. "Mean Field Games with state constraints: from mild to pointwise solutions of the PDE system". In: *Calculus of Variations and Partial Differential Equations* 60 (2021).

[14] René Carmona and François Delarue. "Forward–backward stochastic differential equations and controlled McKean–Vlasov dynamics". In: *The Annals of Probability* 43.5 (2015), pp. 2647 –2700.

[15] René Carmona and François Delarue. *Probabilistic Theory of Mean Field Games: vol. I, Mean Field FBSDEs, Control, and Games*. Springer, 2018.

[16] René Carmona and François Delarue. *Probabilistic Theory of Mean Field Games: vol. II, Mean Field game with common noise and Master equations*. Springer, 2018.

[17] René Carmona and Mathieu Laurière. "Convergence analysis of machine learning algorithms for the numerical solution of mean-field control and games: II The finite horizon case". In: *arXiv:1908.01613, to appear in The Annals of Applied Probability* (2019).

[18] Li Chen and Jiandong Wang. "Maximum principle for delayed stochastic mean-field control problem with state constraint". In: *Advances in Difference Equations* 348 (2019).

[19] Yuk-Loong Chow, Xiang Yu, and Chao Zhou. "On Dynamic Programming Principle for Stochastic Control under Expectation Constraints". In: *Journal of Optimization Theory and Applications* 185 (2020), 803–818.

[20] Andrea Cosso, Fausto Gozzi, Idris Kharroubi, Huyên Pham, and Mauro Rosestolato. "Optimal control of path-dependent McKean-Vlasov SDEs in infinite dimension". In: *arXiv:2012.14772* (2020).

[21] Andrea Cosso and Huyên Pham. "Zero-sum stochastic differential games of generalized McKean–Vlasov type". In: *Journal de Mathématiques Pures et Appliquées* 129 (2019), pp. 180 –212.

[22] Nicolas Curin et al. "A deep learning model for gas storage optimization". In: *arXiv:2102.01980* (2021).

[23] Samuel Daudin. "Optimal Control of Diffusion Processes with Terminal Constraint in Law". In: *arXiv:2012.10707* (2020).

[24] Samuel Daudin. "Optimal control of the Fokker-Planck equation under state constraints in the Wasserstein space". In: *arXiv:2109.14978* (2021).

[25] Carmine De Franco, Johann Nicolle, and Huyên Pham. "Bayesian learning for the Markowitz portfolio selection problem". In: *International Journal of Theoretical and Applied Finance* 22.07 (2019), p. 1950037.

[26] Mao Fabrice Djete, Dylan Possamaï, and Xiaolu Tan. "McKean-Vlasov optimal control: the dynamic programming principle". In: *arXiv:1907.08860* (2020).

[27] Ivar Ekeland and Roger Témam. *Convex Analysis and Variational Problems*. Society for Industrial and Applied Mathematics, 1999.

[28] Guanxing Fu and Ulrich Horst. "Mean-Field Leader-Follower Games with Terminal State Constraint". In: *SIAM Journal on Control and Optimization* 58.4 (2020), pp. 2078–2113.

[29] Alfred Galichon, Pierre Henry-Labordère, and Nizar Touzi. "A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options". In: *The Annals of Applied Probability* 24.1 (2014), pp. 312 –336.

[30] Abebe Geletu, Michael Klöppel, Hui Zhang, and Pu Li. "Advances and applications of chance-constrained approaches to systems optimisation under uncertainty". In: *International Journal of Systems Science* 44.7 (2013), pp. 1209–1232.

[31] M. Germain, M. Laurière, H. Pham, and X. Warin. "DeepSets and derivative networks for solving symmetric PDEs". In: *arXiv: 2103.00838* (2021).

[32] Hugo Gevret et al. *STochastic OPTimization library in C++*. 2018. URL: https://hal.archives-ouvertes.fr/hal-01361291.

[33] Jameson Graber and Sergio Mayorga. "A note on mean field games of controls with state constraints: existence of mild solutions". In: *arXiv:2109.11655* (2021).

[34] Mathieu Laurière and Olivier Pironneau. "Dynamic programming for mean-field type control". In: *Comptes Rendus Mathematique* 352.9 (2014), pp. 707–713.

[35] William Lefebvre, Grégoire Loeper, and Huyên Pham. "Mean-Variance Portfolio Selection with Tracking Error Penalization". In: *Mathematics* 8.11 (2020).

[36] Silvana M. Pesenti and Sebastian Jaimungal. "Portfolio Optimisation within a Wasserstein Ball". In: *Available at SSRN: https://ssrn.com/abstract=3744994* (2021).

[37] Laurent Pfeiffer, Xiaolu Tan, and Yu-Long Zhou. "Duality and Approximation of Stochastic Optimal Control Problems under Expectation Constraints". In: *SIAM Journal on Control and Optimization* 59.5 (2021), pp. 3231–3260.

[38] Huyên Pham and Xiaoli Wei. "Bellman equation and viscosity solutions for mean-field stochastic control problem". In: *ESAIM: COCV* 24.1 (2018), pp. 437–461.

[39] Huyên Pham and Xiaoli Wei. "Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics". In: *SIAM Journal on Control and Optimization* 55.2 (2017), pp. 1069–1101.

[40] Athena Picarelli and Tiziano Vargiolu. "Optimal management of pumped hydroelectric production with state constrained optimal control". In: *Journal of Economic Dynamics and Control* (2020), p. 103940.

[41] H. Mete Soner and Nizar Touzi. "Stochastic Target Problems, Dynamic Programming, and Viscosity Solutions". In: *SIAM Journal on Control and Optimization* 41.2 (2002), pp. 404–424.

[42] Xavier Warin. "Deep learning for efficient frontier calculation in finance". In: *arXiv:2101.02044* (2021).

[43] Xavier Warin. "Reservoir optimization and Machine Learning methods". In: *arXiv:2106.08097* (2021).