

1 **A LEVEL-SET APPROACH TO THE CONTROL OF**
2 **STATE-CONSTRAINED MCKEAN-VLASOV EQUATIONS:**
3 **APPLICATION TO RENEWABLE ENERGY STORAGE AND**
4 **PORTFOLIO SELECTION**

MAXIMILIEN GERMAIN

EDF R&D, Université Paris Cité and Sorbonne Université,
CNRS, Laboratoire de Probabilités, Statistique et Modélisation
F-75013 Paris, France

HUYÊN PHAM *

Université Paris Cité and Sorbonne Université, CNRS,
Laboratoire de Probabilités, Statistique et Modélisation, CREST-ENSAE & FiME
F-75013 Paris, France

XAVIER WARIN

EDF R&D & FiME
91120 Palaiseau, France

(Communicated by the associate editor name)

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.

2020 *Mathematics Subject Classification.* Primary: 49N80, 49M99 ; Secondary: 68T07, 93E20.

Key words and phrases. Mean-field control, state constraints, neural networks, McKean-Vlasov dynamics, probabilistic constraints.

This work was supported by FiME (Finance for Energy Market Research Centre) and the “Finance et Développement Durable - Approches Quantitatives” EDF - CACIB Chair.

*Corresponding author: Huyên Pham.

1 **1. Introduction.** The control of McKean-Vlasov dynamics, also known as mean-
 2 field control problem, has attracted a lot of interest over the last years since the
 3 emergence of the mean-field game theory. There is now an important literature
 4 on this topic addressing on one hand the theoretical aspects either by dynamic
 5 programming approach (see [32, 37, 38, 21]), or by maximum principle (see [14]),
 6 and on the other hand the numerous applications in economics and finance, and
 7 we refer to the two-volume monographs [15, 16] for an exhaustive and detailed
 8 treatment of this area.

9 In this paper, we aim to study control of McKean-Vlasov dynamics under the
 10 additional presence of state constraints in law. The consideration of probabilistic
 11 constraints (usually in expectation or in target form) for standard stochastic control
 12 has many practical applications, notably in finance with quantile and CVaR type
 13 constraints, and is the subject of many papers, we refer to [41, 11, 28, 19, 36, 4] for
 14 an overview.

15 There exists some recent works dealing with mean-field control under some spe-
 16 cific law state constraints. For example, the paper [18] solves mean-field control
 17 with delay and smooth expectation terminal constraint (and without dependence
 18 with respect to the law of the control). In the case of mean field games, state
 19 constraints are considered by [12, 13, 26, 30, 3]. In these cited works the state
 20 belongs to a compact set, which corresponds to a particular case of our constraints
 21 in distribution. Related literature includes the recent work [10] which studies a
 22 mean-field target problem where the aim is to find the initial laws of a controlled
 23 McKean-Vlasov process satisfying a law constraint, but only at terminal time. The
 24 paper [23] also studies these terminal constraint in law for the control of a standard
 25 diffusion process. Next, it has been extended in [24] to a running law constraint for
 26 the control of a standard diffusion process with McKean-Vlasov type cost through
 27 the control of a Fokker-Planck equation. Several works also consider directly the op-
 28 timal control of Fokker-Planck equations in the Wasserstein space with terminal or
 29 running constraints, such as [8, 9] through Pontryagin principle, in the deterministic
 30 case without diffusion.

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

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

1 **2. Mean-field control with state constraints.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability
 2 space on which is defined a d -dimensional Brownian motion W with associated fil-
 3 tration $\mathbb{F} = (\mathcal{F}_t)_t$ augmented with \mathbb{P} -null sets. We assume that \mathcal{F}_0 is “rich enough”
 4 in the sense that any probability measure μ on \mathbb{R}^d can be represented as the distri-
 5 bution law of some \mathcal{F}_0 -measurable random variable. This is satisfied whenever the
 6 probability space $(\Omega, \mathcal{F}_0, \mathbb{P})$ is atomless, see [15], p.352.

We consider the following cost and dynamics:

$$J(X_0, \alpha) = \mathbb{E} \left[\int_0^T f(s, X_s^\alpha, \alpha_s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}) ds + g(X_T^\alpha, \mathbb{P}_{X_T^\alpha}) \right] \quad (1)$$

$$X_t^\alpha = X_0 + \int_0^t b(s, X_s^\alpha, \alpha_s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}) ds + \int_0^t \sigma(s, X_s^\alpha, \alpha_s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}) dW_s,$$

7 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
 8 in $L^2(\mathcal{F}_0, \mathbb{R}^d)$. The control α belongs to a set \mathcal{A} of \mathbb{F} -progressively measurable
 9 processes with values in a set $A \subseteq \mathbb{R}^q$. The coefficients b and σ are measurable
 10 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
 11 of square integrable probability measures on the metric space E , equipped with the
 12 2-Wasserstein distance \mathcal{W}_2 . We make some standard Lipschitz conditions on b, σ in
 13 order to ensure that equation (1) is well-defined and admits a unique strong solution,
 14 which is square-integrable. The function f is a real-valued measurable function on
 15 $[0, T] \times \mathbb{R}^d \times A \times \mathcal{P}_2(\mathbb{R}^d \times A)$, while g is a measurable function on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, and
 16 we assume that f and g satisfy some linear growth condition which ensures that
 17 the functional in (1) is well-defined.

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

$$\Psi(t, \mathbb{P}_{X_t^\alpha}) \leq 0, \quad 0 \leq t \leq T, \quad (2)$$

where Ψ_l , $l = 1, \dots, k$, are given functions from $[0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ into \mathbb{R} , and $\Psi :=$
 $\max_{1 \leq l \leq k} \Psi_l$. In other words, the constraint $\Psi(t, \mu) \leq 0$ means that $\Psi^l(t, \mu) \leq 0$, l
 $= 1, \dots, k$. The problem of interest is therefore

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

18 By convention the infimum of the empty set is $+\infty$. When needed, we will sometimes
 19 use the notation V^Ψ to emphasize the dependence of the value function on Ψ .
 20 Clearly, $\Psi \leq \Psi'$ implies $V^\Psi \leq V^{\Psi'}$.

21 **Remark 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

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

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

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

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

24 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 $\mathcal{W}_2(\mathbb{P}_{X_t^\alpha}, \eta_t) \leq \delta_t$ with

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

1 This is the constraint considered in [35] at terminal time.

- A terminal constraint in law $\varphi(\mathbb{P}_{X_T^\alpha}) \leq 0$ 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 $\phi(t_i, \mathbb{P}_{X_{t_i}^\alpha}) \leq 0$ for $t_1 < \dots < t_k$ with

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

2 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

$$\begin{aligned} Z_t^{z, \alpha} &:= z - \mathbb{E} \left[\int_0^t f(s, X_s^\alpha, \alpha_s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}) ds \right] \\ &= z - \int_0^t \widehat{f}(s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}) ds, \quad 0 \leq t \leq T, \end{aligned} \quad (3)$$

7 where \widehat{f} is the function defined on $[0, T] \times \mathcal{P}_2(\mathbb{R}^d \times \mathcal{A})$ by $\widehat{f}(t, \nu)$
8 $= \int_{\mathbb{R}^d \times \mathcal{A}} f(t, x, a, \nu) \nu(dx, da)$. We also denote by \widehat{g} the function defined on $\mathcal{P}_2(\mathbb{R}^d)$
9 by $\widehat{g}(\mu) = \int_{\mathbb{R}^d} g(x, \mu) \mu(dx)$.

Lemma 2.1. *The value function admits the **deterministic target problem** representation*

$$V = \inf \{ z \in \mathbb{R} \mid \exists \alpha \in \mathcal{A} \text{ s.t. } \widehat{g}(\mathbb{P}_{X_T^\alpha}) \leq Z_T^{z, \alpha}, \Psi(s, \mathbb{P}_{X_s^\alpha}) \leq 0, \forall s \in [0, T] \}. \quad (4)$$

Proof. We first observe from the definition of V in (2) 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

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

10 the result then follows immediately by the definition of $Z^{z, \alpha}$ in (3). \square

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

$$\mathcal{Y}^\Psi : z \in \mathbb{R} \mapsto \inf_{\alpha \in \mathcal{A}} \left[\widehat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z, \alpha} \right]_+ + \sup_{s \in [0, T]} \{ \Psi(s, \mathbb{P}_{X_s^\alpha}) \}_+, \quad (5)$$

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

1 By classical estimates on McKean-Vlasov equations we can obtain continuity and
2 growth conditions on \mathcal{Y}^Ψ . The proof of Proposition 1 is given in Section 2.3.

3 **Proposition 1.** \mathcal{Y}^Ψ verifies

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

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

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

5 Define the infimum of the zero level-set

$$\mathcal{Z}^\Psi := \inf\{z \in \mathbb{R} \mid \mathcal{Y}^\Psi(z) = 0\}. \quad (6)$$

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

9 **Theorem 2.2.** 1. If for some $z \in \mathbb{R} \exists \alpha \in \mathcal{A}$ s.t. $\widehat{g}(\mathbb{P}_{X_T^\alpha}) \leq Z_T^{z, \alpha}$, $\Psi(s, \mathbb{P}_{X_s^\alpha}) \leq$
10 $0, \forall s \in [0, T]$ then $\mathcal{Y}^\Psi(z) = 0$.
11 2. If V^Ψ is finite then $\mathcal{Y}^\Psi(V^\Psi) = 0$. Thus $\mathcal{Z}^\Psi \leq V^\Psi$.
3. We have the upper bound

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

12 To sum up, when $-\infty < V^\Psi < +\infty$, Theorem 2.2 provides the bounds

$$\mathcal{Z}^\Psi \leq V^\Psi \leq \inf_{\varepsilon > 0} \mathcal{Z}^{\Psi + \varepsilon}. \quad (7)$$

13 The proof of Theorem 2.2 is given in Section 2.3.

14 **Remark 2.** In the easier case where optimal controls exist for the auxiliary problem,
15 as assumed in [7], similar arguments as in [7] (and Section 4) directly prove that
16 $\mathcal{Z}^\Psi = V^\Psi$ and that an optimal control α^* associated to the auxiliary problem
17 $\mathcal{Y}^\Psi(V)$ is optimal for the original problem. However some difficulties arise when
18 trying to remove this assumption.

19 **Remark 3.** If there exists $\varepsilon_0 > 0$ such that $V^{\Psi + \varepsilon_0} < +\infty$ then $\mathcal{Z}^{\Psi + \varepsilon_0} \leq V^{\Psi + \varepsilon_0} <$
20 $+\infty$ by Theorem 2.2. Thus the right-hand side of (7) is finite.

On the other hand, we need to be careful about the form of the constraint function
if we want to use (7). There are cases in which one choice of Ψ gives a infinite right-
hand side in this equation but another representation of the constraint gives an finite
right-hand side. Let us consider for instance a one-dimensional terminal constraint
in law $\varphi(\mathbb{P}_{X_T^\alpha}) \leq 0$, represented by

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

21 We see that the constraint $\Psi(t, \mu) + \varepsilon \leq 0$ would never be verified for any $t < T$
22 and any $\varepsilon > 0$, hence $V^{\Psi + \varepsilon} = +\infty$ and $\mathcal{Z}^{\Psi + \varepsilon} = +\infty$. In that case there would be
23 a gap in (7) and one wouldn't be able to conclude that $V^\Psi = \mathcal{Z}^\Psi$.

In view of the above example in Remark 3, we introduce a modified constraint
function in order to deal with discrete time constraints, and also with a.s. con-
straints. Given a constraint function $\Psi(t, \mu)$, we define

$$\overline{\Psi}_\kappa(t, \mu) := \Psi(t, \mu) - \kappa \mathbb{1}_{\{\Psi(t, \mu) \leq 0\}}, \quad (8)$$

1 with $\kappa > 0$. By observing that $\bar{\Psi}_\kappa(t, \mu) \leq 0 \Leftrightarrow \Psi(t, \mu) \leq 0$, it follows that

$$V^\Psi = V^{\bar{\Psi}_\kappa}, \mathcal{Y}^\Psi = \mathcal{Y}^{\bar{\Psi}_\kappa}, \mathcal{Z}^\Psi = \mathcal{Z}^{\bar{\Psi}_\kappa}. \quad (9)$$

2 This auxiliary function is only used for the theoretical results in the next section
3 but not for the numerics.

4 **Remark 4.** Notice that by taking $\varepsilon_0 < \kappa$, and assuming that $V^\Psi < \infty$, we have
5 $\mathcal{Z}^{\bar{\Psi}_\kappa + \varepsilon_0} < \infty$. Indeed, by applying Theorem 2.2 to $\bar{\Psi}_\kappa$, we have $\mathcal{Z}^{\bar{\Psi}_\kappa + \varepsilon_0} \leq V^{\bar{\Psi}_\kappa + \varepsilon_0}$.
6 Moreover, by observing that an admissible control for the original problem V^Ψ
7 is also admissible for the auxiliary problem with constraint function $\bar{\Psi}_\kappa + \varepsilon_0$, by
8 definition of $\bar{\Psi}_\kappa$, this implies that $V^{\bar{\Psi}_\kappa + \varepsilon_0} < \infty$.

9 **2.2. Representation of the value function.** Now we prove under some assump-
10 tions on the constraints the continuity property $\mathcal{Z}^{\bar{\Psi}_\kappa} = \inf_{\varepsilon > 0} \mathcal{Z}^{\bar{\Psi}_\kappa + \varepsilon}$ in order to
11 obtain a characterization of the original value function V^Ψ . The result relies on
12 convexity arguments, and holds true within the linear-convex model assumption:

Assumption 1 (Lin-Conv). *The coefficients b valued in \mathbb{R}^d , $\sigma = (\sigma_j)_{1 \leq j \leq d}$ valued
in $\mathbb{R}^{d \times d}$ of the controlled mean-field equation are in the linear form:*

$$b(t, x, a, \nu) = \beta(t) + B(t)x + C(t)a + \bar{B}(t) \int x\nu(dx, da) + \bar{C}(t) \int a\nu(dx, da)$$

$$\sigma_j(t, x, a, \nu) = \gamma_j(t) + D_j(t)x + F_j(t)a + \bar{D}_j(t) \int x\nu(dx, da) + \bar{F}_j(t) \int a\nu(dx, da),$$

13 for $(t, x, a, \nu) \in [0, T] \times \mathbb{R}^d \times A \times \mathcal{P}_2(\mathbb{R}^d \times A)$, with A convex set of \mathbb{R}^q , and with
14 bounded measurable function $\beta, \gamma_j, B, D_j, C, F_j, \bar{B}, \bar{D}_j, \bar{C}$, and \bar{F}_j , $j = 1, \dots, d$,
15 on $[0, T]$, valued respectively on $\mathbb{R}^d, \mathbb{R}^d, \mathbb{R}^{d \times d}, \mathbb{R}^{d \times d}, \mathbb{R}^{d \times q}, \mathbb{R}^{d \times q}, \mathbb{R}^{d \times d}, \mathbb{R}^{d \times d}$,
16 $\mathbb{R}^{d \times q}, \mathbb{R}^{d \times q}$.

The cost functions f, g and the constraint functions Ψ_l , $l = 1, \dots, k$, are in the
form:

$$f(t, x, a, \nu) = \tilde{f}(t, x, a, \int x\nu(dx, da), \int a\nu(dx, da)) \quad g(x, \mu) = \tilde{g}(x, \int x\mu(dx)),$$

$$\Psi_l(t, \mu) = \tilde{\Psi}_l(t, \int x\mu(dx)).$$

17 for $(t, x, a, \mu, \nu) \in [0, T] \times \mathbb{R}^d \times A \times \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d \times A)$, where $\tilde{f}(t, \cdot)$ is convex on
18 $\mathbb{R}^d \times \mathbb{R}^q \times \mathbb{R}^d \times \mathbb{R}^q$, and $\tilde{g}, \tilde{\Psi}_l(t, \cdot)$ are convex on $\mathbb{R}^d \times \mathbb{R}^d$.

19 **Lemma 2.3.** *Under Assumption 1, the function $(z, \varepsilon) \in \mathbb{R} \times \mathbb{R} \mapsto \mathcal{Y}^{\Psi + \varepsilon}(z)$ is
20 convex.*

21 The proof of Lemma 2.3 is given in Section 2.3.

22 **Proposition 2.** *Let Assumption 1 hold. Then, \mathcal{Y}^Ψ being convex, positive and non-
23 increasing, if $\mathcal{Z}^\Psi < \infty$ then \mathcal{Y}^Ψ is decreasing on $(-\infty, \mathcal{Z}^\Psi]$ and $\mathcal{Y}^\Psi(z) = 0$ on
24 $[\mathcal{Z}^\Psi, \infty)$.*

25 *Proof.* By contradiction, if $\mathcal{Y}^\Psi(a) = \mathcal{Y}^\Psi(b) > 0$ with $a < b$ then by monotonicity
26 $\mathcal{Y}^\Psi([a, b]) = \{\mathcal{Y}^\Psi(a)\}$ 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
27 is not the case because $\mathcal{Z}^\Psi < \infty$. As a consequence, \mathcal{Y}^Ψ is decreasing. Then by
28 continuity of \mathcal{Y}^Ψ and definition of \mathcal{Z}^Ψ we obtain $\mathcal{Y}^\Psi(\mathcal{Z}^\Psi) = 0$. \square

Theorem 2.4. *Under Assumption 1, assume that $-\infty < V^\Psi < \infty$. 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

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

- 1 *Proof of Theorem 2.4.* We prove the continuity of $\mathcal{Z}^{\bar{\Psi}_\kappa}$ along the curve $\mathcal{Z}^{\bar{\Psi}_\kappa + \varepsilon}$ for $\varepsilon \in \mathbb{R}$ where $\bar{\Psi}_\kappa$ is defined in (8).

Let $\kappa > 0$ and $\varepsilon_0 < \kappa$. By Remark 4, we know that $\mathcal{Z}^{\bar{\Psi}_\kappa + \varepsilon_0} < \infty$. We consider the function

$$\Phi(\varepsilon) := \mathcal{Z}^{\bar{\Psi}_\kappa + \varepsilon} = \inf\{z \in \mathbb{R} : \mathcal{Y}^{\bar{\Psi}_\kappa + \varepsilon}(z) \leq 0\} < \infty, \quad \varepsilon < \varepsilon_0,$$

and observe by direct verification that it is convex on $(-\infty, \varepsilon_0)$, using Lemma 2.3. Moreover, by (7) applied to $\bar{\Psi}_\kappa$, and (9), we have $\Phi(\varepsilon) = \mathcal{Z}^{\bar{\Psi}_\kappa + \varepsilon} \geq V^{\bar{\Psi}_\kappa} = V^\Psi > -\infty$. Therefore, Φ is a convex, and finite function on $(-\infty, \varepsilon_0)$, hence it is continuous (see e.g. Corollary 10.1.1 in [40]), in particular at $\varepsilon = 0$. As a consequence $\mathcal{Z}^{\bar{\Psi}_\kappa} = \inf_{\varepsilon > 0} \mathcal{Z}^{\bar{\Psi}_\kappa + \varepsilon}$, and by Theorem 2.2 applied to $\bar{\Psi}_\kappa$, we obtain $\mathcal{Z}^{\bar{\Psi}_\kappa} = V^{\bar{\Psi}_\kappa}$. Then recalling that $\mathcal{Z}^\Psi = \mathcal{Z}^{\bar{\Psi}_\kappa}$, $V^\Psi = V^{\bar{\Psi}_\kappa}$, the result follows. Concerning the controls, take $\varepsilon > 0$, and consider an ε -optimal control $\alpha^\varepsilon \in \mathcal{A}$ such that

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

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

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

- 3 Hence $J(X_0, \alpha^\varepsilon) \leq \mathcal{Z}^\Psi + \varepsilon = V^\Psi + \varepsilon$ and $\Psi(s, \mathbb{P}_{X_s^{\alpha^\varepsilon}}) \leq \varepsilon, \quad \forall s \in [0, T]$. □

4 2.3. Proofs.

Proof of Proposition 1. 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 $z, z' \in \mathbb{R}$

$$\begin{aligned} & |\mathcal{Y}^\Psi(z) - \mathcal{Y}^\Psi(z')| \\ &= \left| \inf_{\alpha \in \mathcal{A}} \left[\{\hat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z, \alpha}\}_+ + \sup_{s \in [t, T]} \{\Psi(s, \mathbb{P}_{X_s^\alpha})\}_+ \right] \right. \\ & \quad \left. - \inf_{\alpha \in \mathcal{A}} \left[\{\hat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z', \alpha}\}_+ + \sup_{s \in [t, T]} \{\Psi(s, \mathbb{P}_{X_s^\alpha})\}_+ \right] \right| \\ & \leq \sup_{\alpha \in \mathcal{A}} |\{\hat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z, \alpha}\}_+ - \{\hat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z', \alpha}\}_+| \\ & \quad + \sup_{s \in [t, T]} \{\Psi(s, \mathbb{P}_{X_s^\alpha})\}_+ - \sup_{s \in [t, T]} \{\Psi(s, \mathbb{P}_{X_s^\alpha})\}_+| \\ & \leq \sup_{\alpha \in \mathcal{A}} |Z_T^{z, \alpha} - Z_T^{z', \alpha}| = |z - z'|, \end{aligned}$$

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

2) Denote by

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

1 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

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

2

□

Proof of Theorem 2.2. 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

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

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

4

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

8

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

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

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

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

Hence $J(\alpha^{\varepsilon'}) \leq \mathcal{Z}^{\Psi+\varepsilon} + \varepsilon'$ and $\Psi(s, \mathbb{P}_{X_s^{\alpha^{\varepsilon'}}}) \leq 0$, $\forall s \in [0, T]$.

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

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

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

Proof of Lemma 2.3. Under Assumption 1 on the linear dynamics of the controlled state process, we have for all $\alpha \in \mathcal{A}$, $z \in \mathbb{R}$, $\varepsilon \in \mathbb{R}_+$,

$$\begin{aligned} L^{\Psi+\varepsilon}(z, \alpha) &= \left\{ \mathbb{E}[\tilde{g}(X_T^\alpha, \mathbb{E}[X_T^\alpha]) + \int_0^T \tilde{f}(s, X_s^\alpha, \alpha_s, \mathbb{E}[X_s^\alpha], \mathbb{E}[\alpha_s]) \, ds] - z \right\}_+ \\ &\quad + \sup_{s \in [0, T]} \{\tilde{\Psi}(s, \mathbb{E}[X_s^\alpha]) + \varepsilon\}_+. \end{aligned}$$

Let α^1, α^2 be two arbitrary controls in \mathcal{A} , $z^1, z^2 \in \mathbb{R}$, $\varepsilon^1, \varepsilon^2 \in \mathbb{R}_+$, and $\lambda \in (0, 1)$. Define $\alpha = \lambda\alpha^1 + (1-\lambda)\alpha^2$, and notice by the linear mean-field dynamics in Assumption 1 that $X^\alpha = \lambda X^{\alpha^1} + (1-\lambda)X^{\alpha^2}$. Then, by the convexity assumption on \tilde{f} , \tilde{g} , and $\tilde{\Psi}$ in Assumption 1, and the convexity of $x \mapsto \{x\}_+$, we have

$$L^{\Psi+\lambda\varepsilon^1+(1-\lambda)\varepsilon^2}(\lambda z^1 + (1-\lambda)z^2, \alpha) \leq \lambda L^{\Psi+\varepsilon^1}(z^1, \alpha^1) + (1-\lambda)L^{\Psi+\varepsilon^2}(z^2, \alpha^2).$$

By taking the infimum over α^1, α^2 in the r.h.s. of the above inequality, we deduce the required convexity result:

$$\begin{aligned} \mathcal{Y}^{\Psi+\lambda\varepsilon^1+(1-\lambda)\varepsilon^2}(\lambda z^1 + (1-\lambda)z^2) &\leq L^{\Psi+\lambda\varepsilon^1+(1-\lambda)\varepsilon^2}(\lambda z^1 + (1-\lambda)z^2, \alpha) \\ &\leq \lambda \mathcal{Y}^{\Psi+\varepsilon^1}(z^1) + (1-\lambda) \mathcal{Y}^{\Psi+\varepsilon^2}(z^2). \end{aligned}$$

1

□

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

$$\begin{aligned} X_s^{t,\xi,\alpha} &:= \xi + \int_t^s b(u, X_u^{t,\xi,\alpha}, \alpha_u, \mathbb{P}_{(X_u^{t,\xi,\alpha}, \alpha_u)}) du + \int_t^s \sigma(u, X_u^{t,\xi,\alpha}, \alpha_u, \mathbb{P}_{(X_u^{t,\xi,\alpha}, \alpha_u)}) dW_u, \end{aligned}$$

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

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

coming from existence and pathwise uniqueness in (1). We thus consider the cost function

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

and the value function

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

Then we introduce the auxiliary state variable

$$\begin{aligned} Z_r^{t,\xi,z,\alpha} &:= z - \mathbb{E} \left[\int_t^r f(s, X_s^{t,\xi,\alpha}, \alpha_s, \mathbb{P}_{(X_s^{t,\xi,\alpha}, \alpha_s)}) ds \right] \\ &= z - \int_t^r \hat{f}(s, \mathbb{P}_{(X_s^{t,\xi,\alpha}, \alpha_s)}) ds, \quad t \leq r \leq T, \end{aligned} \quad (10)$$

and the auxiliary value function is given by

$$\begin{aligned} \mathcal{Y}^\Psi(t, \xi, z) &= \inf_{\alpha \in \mathcal{A}} \left[\hat{g}(\mathbb{P}_{X_T^{t,\xi,\alpha}}) - Z_{t,\xi,z}^\alpha(T) \right]_+ + \sup_{s \in [t, u]} \{ \Psi(s, \mathbb{P}_{X_s^{t,\xi,\alpha}}) \}_+ \\ &=: \inf_{\alpha \in \mathcal{A}} L^\Psi(t, \xi, z, \alpha). \end{aligned} \quad (11)$$

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

$$\tilde{\mathcal{Y}}^\Psi(t, \xi, z, m) = \inf_{\alpha \in \mathcal{A}} \left[\hat{g}(\mathbb{P}_{X_T^{t,\xi,\alpha}}) - Z_{t,\xi,z}^\alpha(T) \right]_+ + Y_T^{t,\xi,\alpha,m} =: \inf_{\alpha \in \mathcal{A}} \bar{L}^\Psi(t, \xi, z, m, \alpha).$$

The two problems are related by

$$\mathcal{Y}^\Psi(t, \xi, z) = \tilde{\mathcal{Y}}^\Psi(t, \xi, z, \{ \Psi(t, \mathbb{P}_{X_t^{t,\xi,\alpha}}) \}_+).$$

- 2 With this formulation, the problem (11) 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

1 mentioned in [6], this procedure is used for instance for hedging lookback options
 2 in finance, see e.g. [27]. Now the infimum of the zero level-set is given by

$$\mathcal{Z}^\Psi(t, \xi) := \inf\{z \in \mathbb{R} \mid \tilde{\mathcal{Y}}^\Psi(t, \xi, z, 0) = 0\}. \quad (12)$$

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

4 The Lipschitz and convexity properties of the value function are proven exactly
 5 as in Proposition 1 but we detail here the continuity in space and in the running
 6 maximum variable m .

Assumption 2. Ψ, f, g, b, σ are Lipschitz continuous uniformly with respect 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 $t \in [0, T]$, $x, x' \in \mathbb{R}^d, \mu \in \mathcal{P}_2(\mathbb{R}^d), \nu, \nu' \in \mathcal{P}_2(\mathbb{R}^d \times A), a \in A$

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

Proposition 3. Under Assumption 2, $\tilde{\mathcal{Y}}^\Psi$ is Lipschitz continuous: there exists $C > 0$ such that for any $t \in [0, T]$, $\xi, \xi' \in L^2(\mathcal{F}_t, \mathbb{R}^d), m, m' \in \mathbb{R}$

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

Proof of Proposition 3. By the inequalities $|\inf_u A(u) - \inf_u B(u)| \leq \sup_u |A(u) - B(u)|$, $|\sup_u A(u) - \sup_u B(u)| \leq \sup_u |A(u) - B(u)|$, and $|a \vee b - c \vee d| \leq |a - c| \vee |b - d| \leq |a - c| + |b - d|$ 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}
& |\tilde{\mathcal{Y}}^\Psi(t, \xi, z, m) - \tilde{\mathcal{Y}}^\Psi(t, \xi', z', m')| \\
& \leq \sup_{\alpha \in \mathcal{A}} \{|\hat{g}(\mathbb{P}_{X_T^{t, \xi, \alpha}}) - Z_T^{t, \xi, z, \alpha}\}_+ - \{\hat{g}(\mathbb{P}_{X_T^{t, \xi', \alpha}}) - Z_T^{t, \xi', z', \alpha}\}_+ \\
& \quad + \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' \\
& \leq \sup_{\alpha \in \mathcal{A}} (|\hat{g}(\mathbb{P}_{X_T^{t, \xi, \alpha}}) - \hat{g}(\mathbb{P}_{X_T^{t, \xi', \alpha}})| + |Z_T^{t, \xi, z, \alpha} - Z_T^{t, \xi', z', \alpha}| \\
& \quad + |\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'|) \\
& \leq \sup_{\alpha \in \mathcal{A}} \left(|\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'| \right. \\
& \quad \left. + \left| \mathbb{E} \left[\int_t^T f(s, X_s^{t, \xi, \alpha}, \alpha_s, \mathbb{P}_{(X_s^{t, \xi, \alpha}, \alpha_s)}) ds - \int_t^T f(s, X_s^{t, \xi', \alpha}, \alpha_s, \mathbb{P}_{(X_s^{t, \xi', \alpha}, \alpha_s)}) ds \right] \right| \right) \\
& \quad + \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'| \\
& \leq [\hat{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'| \\
& \quad + [\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}}) \\
& \quad + 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

$$\sup_{s \in [t, T]} \mathcal{W}_2(\mathbb{P}_{X_s^{t, \xi, \alpha}}, \mathbb{P}_{X_s^{t, \xi', \alpha}}) = \sqrt{\sup_{s \in [t, T]} \mathcal{W}_2(\mathbb{P}_{X_s^{t, \xi, \alpha}}, \mathbb{P}_{X_s^{t, \xi', \alpha}})^2} \leq C \sqrt{\mathbb{E}|\xi - \xi'|^2} \quad (13)$$

$$\mathbb{E}[\sup_{s \in [t, T]} |X_s^{t, \xi, \alpha} - X_s^{t, \xi', \alpha}|] \leq C \mathbb{E}|\xi - \xi'| \leq C \sqrt{\mathbb{E}|\xi - \xi'|^2}, \quad (14)$$

- 1 obtained by standard arguments (see e.g. the proof of Proposition 3.3 in [21]). Then
2 the result follows. \square

Proposition 4 (Law invariance properties). *Under Assumption 2, we have law invariance of $\tilde{\mathcal{Y}}^\Psi$ and \mathcal{Z}^Ψ , namely if ξ, η are \mathcal{F}_t -adapted square integrable with the same law, for any $(t, z, m) \in [0, T] \times \mathbb{R} \times \mathbb{R}$*

$$\tilde{\mathcal{Y}}^\Psi(t, \xi, z, m) = \tilde{\mathcal{Y}}^\Psi(t, \eta, z, m)$$

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

- 3 Therefore we can define the lifted functions y^Ψ, z^Ψ on $[0, T] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}$ (re-
4 spectively $[0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ by $y^\Psi(t, \mathbb{P}_\xi, z, m) := \tilde{\mathcal{Y}}^\Psi(t, \xi, z, m)$ and $z^\Psi(t, \mathbb{P}_\xi, z, m) :=$
5 $\mathcal{Z}^\Psi(t, \xi, z, m)$.

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

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

- 1 holds true with a_s defined in Lemma B.2. from [20] (verifying in particular the
 2 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
 3 a \mathcal{F}_t -adapted uniform random variable on $[0, 1]$ independent of η (respectively ξ).
 4 Then use the definition 6 to obtain the same law invariance property for \mathcal{Z}^Ψ too. \square

Theorem 2.2 and Theorem 2.4 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 $-\infty < V^\Psi(t, \xi) < \infty$ then

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

Similarly, arguments like in Theorem 2.4 prove that

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

- 5 if $-\infty < V^\Psi(t, \xi) < \infty$.
 6 If the value function is law invariant (see Proposition 4) and Theorem 2.4 holds
 7 true, we expect y to be formally (by combining arguments from [6, 21]) characterized
 8 by a Master Bellman equation in Wassertein space with oblique derivative boundary
 9 conditions.

3. An alternative auxiliary problem. We study the constrained McKean-Vlasov control problem

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

where we now assume that the running constraint Ψ is continuous (hence, no discrete time constraints, see Remark 5), and with a terminal constraint function φ . We now consider an alternative auxiliary control problem as in [7]:

$$w(z) := \inf_{\alpha \in \mathcal{A}} \left[\{\hat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z, \alpha}\}_+ + \int_0^T \{\Psi(s, \mathbb{P}_{X_s^\alpha})\}_+ ds + \{\varphi(\mathbb{P}_{X_T^\alpha})\}_+ \right]. \quad (15)$$

Compared to the control problem (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 Markovian with respect to the variables X_t , \mathbb{P}_{X_t} and Z_t , and we shall show that it also provides a similar representation of the value function by its zero level set:

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

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

- 12 The mean-field control problem (15) is Markovian with respect to the state vari-
 13 ables $(X_t^\alpha, \mathbb{P}_{X_t^\alpha}, Z_t^{z, \alpha})$, and it is known from [20] that the infimum over open-loop
 14 controls α in \mathcal{A} can be taken equivalently over randomized feedback policies, i.e.
 15 controls α in the form: $\alpha_t = \mathbf{a}(t, X_t^\alpha, \mathbb{P}_{X_t^\alpha}, Z_t^{z, \alpha}, U)$, for some deterministic function
 16 \mathbf{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
 17 random variable on $[0, 1]$.

- 18 Let us now discuss conditions under which the infimum in (15) can be taken
 19 equivalently over (deterministic) feedback policies, i.e. for controls α in the form: α_t
 20 $= \mathbf{a}(t, X_t^\alpha, \mathbb{P}_{X_t^\alpha}, Z_t^{z, \alpha})$, for some deterministic function \mathbf{a} from $[0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}$
 21 into A . This will be helpful for numerical purpose in Section 5. We assume on top of
 22 Assumption 2 that the running cost f , the drift b and the volatility coefficient σ do
 23 not depend on the law of the control process. We also assume that the running cost

- 1 $f = f(t, x, \mu)$ does not depend on the control argument. The terminal constraint
 2 function φ should also verify the same assumptions as the terminal cost function g ,
 3 namely Lipschitz continuity and local boundedness (see Assumption 2).

In this case, the corresponding dynamic auxiliary problem of (15) is written as

$$\begin{aligned} w(t, \mu, z) &= \inf_{\alpha \in \mathcal{A}} \left[\{\widehat{g}(\mathbb{P}_{X_T^{t, \xi, \alpha}}) - Z_T^{t, \xi, z, \alpha}\}_+ + \int_0^T \{\Psi(s, \mathbb{P}_{X_s^{t, \xi, \alpha}})\}_+ ds \right. \\ &\quad \left. + \{\varphi(\mathbb{P}_{X_T^{t, \xi, \alpha}})\}_+ \right] \\ X_r^{t, \xi, \alpha} &= \xi + \int_t^r b(s, X_s^{t, \xi, \alpha}, \alpha_s, \mathbb{P}_{X_s^{t, \xi, \alpha}}) ds + \int_t^r \sigma(s, X_s^{t, \xi, \alpha}, \alpha_t, \mathbb{P}_{X_s^{t, \xi, \alpha}}) dW_s, \\ Z_r^{t, \xi, z, \alpha} &= z - \int_t^r \bar{f}(s, \mathbb{P}_{X_s^{t, \xi, \alpha}}) ds, \quad r \geq t, \quad \xi \sim \mu \end{aligned} \quad (16)$$

where \bar{f} is the function defined on $[0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ by $\bar{f}(t, \mu) = \int_{\mathbb{R}^d} f(t, x, \mu) \mu(dx)$. 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 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 (16) is:

$$\begin{aligned} &\partial_t w(t, \mu, z) + \mathbb{E}[\inf_{a \in A} \{b(t, \xi, a, \mu) \partial_\mu w(t, \mu, z)(\xi) \\ &\quad - \bar{f}(t, \mu) \partial_z w(t, \mu, z) + \frac{1}{2} \text{Tr}(\sigma \sigma^\top(t, \xi, a, \mu) \partial_x \partial_\mu w(t, \mu, z)(\xi))\}] \\ &\quad + \{\Psi(t, \mu)\}_+ = 0 \quad \text{for } (t, \mu, z) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R} \\ w(T, \mu, z) &= \{\widehat{g}(\mu) - z\}_+ + \{\varphi(\mu)\}_+ \quad \text{for } (\mu, z) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}, \end{aligned} \quad (17)$$

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

$$\inf_{a \in A} \{b(t, x, a, \mu) \partial_\mu w(t, \mu, z)(x) + \frac{1}{2} \text{Tr}(\sigma \sigma^\top(t, x, a, \mu) \partial_x \partial_\mu w(t, \mu, z)(x))\}$$

- 4 is attained for some measurable function $\hat{a}(t, x, \mu, z)$ on $[0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}$, we
 5 get an optimal control for (15) given in feedback form by $\alpha_t^* = \hat{a}(t, X_t^{\alpha^*}, \mathbb{P}_{X_t^{\alpha^*}}, Z_t^{z, \alpha^*})$,
 6 $0 \leq t \leq T$, which shows that one can restrict in (15) to deterministic feedback poli-
 7 cies.

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:

$$\begin{aligned} J(\alpha) &:= \mathbb{E} \left[\int_0^T f(t, X_t^\alpha, \alpha_t, \mathbb{P}_{(X_t^\alpha, \alpha_t)}^{W^0}) dt + g(X_T^\alpha, \mathbb{P}_{X_T^\alpha}^{W^0}) \right] \\ dX_t^\alpha &= b(t, X_t^\alpha, \alpha_t, \mathbb{P}_{(X_t^\alpha, \alpha_t)}^{W^0}) dt + \sigma(t, X_t^\alpha, \alpha_t, \mathbb{P}_{(X_t^\alpha, \alpha_t)}^{W^0}) dW_t \\ &\quad + \sigma^0(t, X_t^\alpha, \alpha_t, \mathbb{P}_{(X_t^\alpha, \alpha_t)}^{W^0}) dW_t^0. \end{aligned} \quad (18)$$

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

The controlled McKean-Vlasov process X is constrained to verify $\Psi(t, \mathbb{P}_{X_t^\alpha}^{W^0}) \leq 0$
 and $\varphi(\mathbb{P}_{X_T^\alpha}^{W^0}) \leq 0$. 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 condi-
 tional law with respect to the common noise. We refer to [37, 25] for the dynamic
 programming approach to these problems. The problem of interest is

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

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. $\mathbb{E}[\int_0^T |\beta_t|^2 dt] < \infty$, define

$$Z_t^{z, \alpha, \beta} := z - \int_0^t \widehat{f}(s, \mathbb{P}_{X_s^\alpha}^{W^0, \alpha_s}) ds + \int_0^t \beta_s dW_s^0, \quad 0 \leq t \leq T. \quad (19)$$

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

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

- 4 Lemma 4.1 is proven in Section 4.2.

Define the **auxiliary unconstrained** control problem

$$\mathcal{U}(z) := \inf_{(\alpha, \beta) \in \mathcal{A} \times L^2(\mathbb{F}^0, \mathbb{R}^p)} \mathbb{E} \left[\{\widehat{g}(\mathbb{P}_{X_T^\alpha}^{W^0}) - Z_T^{z, \alpha, \beta}\}_+ + \int_0^T \{\Psi(s, \mathbb{P}_{X_s^\alpha}^{W^0})\}_+ ds \right. \\ \left. + \{\varphi(\mathbb{P}_{X_T^\alpha}^{W^0})\}_+ \right] \quad (20)$$

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

Proposition 5. *\mathcal{U} is 1-Lipschitz. For any $z, z' \in \mathbb{R}$*

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

- 6 Proposition 5 is proven exactly as Proposition 1.

- 7 **Assumption 3.** *Problem (20) admits an optimal control for any $z \in \mathbb{R}$ and the*
 8 *constraint function $(t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \mapsto \Psi(t, \mu)$ is continuous.*

- 9 **Remark 5.** Notice that the integral penalization in (20) does not allow to consider
 10 discrete times constraints (except at terminal time) because the contribution to
 11 the integral would be null and the constraint function Ψ would be discontinuous in
 12 time. We could consider discrete time constraints in the objective of the auxiliary
 13 problem by adding a sum of functions of $\mathbb{P}_{X_{t_i}^\alpha}^{W^0}$ for some $(t_i)_i \in [0, T]$ but it would
 14 lose its standard Bolza form.

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

- 16 **Theorem 4.2.** *1. If $\exists (\alpha, \beta) \in \mathcal{A} \times L^2(\mathbb{F}^0, \mathbb{R}^p)$, $\widehat{g}(\mathbb{P}_{X_T^\alpha}^{W^0}) \leq Z_T^{z, \alpha, \beta}$, $\Psi(s, \mathbb{P}_{X_s^\alpha}^{W^0}) \leq$
 17 $0, \forall s \in [0, T]$, and $\varphi(\mathbb{P}_{X_T^\alpha}^{W^0}) \leq 0, \mathbb{P} \text{ a.s.}$ then $\mathcal{U}(z) = 0$. Hence $\mathcal{Z} \leq V^0$.*

1 2. The value function verifies $V^0 \leq \mathcal{Z}$ thus $V^0 = \mathcal{Z}$. Moreover optimal controls
2 for the problem $\mathcal{U}(\mathcal{Z}) = 0$ are optimal for the original problem.

3 Theorem 4.2 is proven in Section 4.2.

4 4.2. Proofs in the common noise framework.

Proof of Lemma 4.1. We first observe that

$$V^0 = \inf\{z \in \mathbb{R} \mid \exists \alpha \in \mathcal{A} \text{ s.t. } \mathbb{E}\left[\int_0^T \widehat{f}(s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}^{W^0}) ds + \widehat{g}(\mathbb{P}_{X_T^\alpha}^{W^0})\right] \leq z, \\ \Psi(s, \mathbb{P}_{X_s^\alpha}^{W^0}) \leq 0, \forall s \in [0, T], \mathbb{P}^0 \text{ a.s.}\}.$$

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

$$\exists (\alpha, \beta) \in \mathcal{A} \times L^2(\mathbb{F}^0, \mathbb{R}^p) \text{ s.t. } \widehat{g}(\mathbb{P}_{X_T^\alpha}^{W^0}) \leq Z_T^{z, \alpha, \beta}, \Psi(s, \mathbb{P}_{X_s^\alpha}^{W^0}) \leq 0, \forall s \in [0, T], \\ \varphi(\mathbb{P}_{X_T^\alpha}^{W^0}) \leq 0, \mathbb{P}^0 \text{ a.s.}, \quad (21)$$

and

$$\exists \alpha \in \mathcal{A} \text{ s.t. } \mathbb{E}\left[\int_0^T \widehat{f}(s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}^{W^0}) ds + \widehat{g}(\mathbb{P}_{X_T^\alpha}^{W^0})\right] \leq z, \Psi(s, \mathbb{P}_{X_s^\alpha}^{W^0}) \leq 0, \forall s \in [0, T], \\ \varphi(\mathbb{P}_{X_T^\alpha}^{W^0}) \leq 0, \mathbb{P}^0 \text{ a.s.}, \quad (22)$$

are equivalent. It is immediate to see that (21) \implies (22) by taking the expectation and noticing that the Itô integral is a true martingale. Conversely, assuming (22), the martingale representation theorem provides a process $\widehat{\beta}$ such that

$$z \geq \mathbb{E}\left[\int_0^T \widehat{f}(s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}^{W^0}) ds + \widehat{g}(\mathbb{P}_{X_T^\alpha}^{W^0})\right] \\ = \int_0^T \widehat{f}(s, \mathbb{P}_{(X_s^\alpha, \alpha_s)}^{W^0}) ds + \widehat{g}(\mathbb{P}_{X_T^\alpha}^{W^0}) - \int_0^T \widehat{\beta}_s dW_s^0.$$

Thus by (19)

$$Z_T^{z, \alpha, \widehat{\beta}} \geq \widehat{g}(\mathbb{P}_{X_T^\alpha}^{W^0}), \mathbb{P}^0 \text{ a.s.},$$

5 and we see that (22) \implies (21). Then the result follows. \square

Proof of Theorem 4.2. 1) $\exists (\alpha, \beta) \in \mathcal{A} \times L^2(\mathbb{F}^0, \mathbb{R}^p)$, $\widehat{g}(\mathbb{P}_{X_T^\alpha}^{W^0}) \leq Z_T^{z, \alpha, \beta}$, $\Psi(s, \mathbb{P}_{X_s^\alpha}^{W^0}) \leq 0$, $\forall s \in [0, T]$ and $\varphi(\mathbb{P}_{X_T^\alpha}^{W^0}) \leq 0$, \mathbb{P}^0 a.s.. Therefore

$$\{\widehat{g}(\mathbb{P}_{X_T^\alpha}^{W^0}) - Z_T^{z, \alpha, \beta}\}_+ + \int_0^T \{\Psi(s, \mathbb{P}_{X_s^\alpha}^{W^0})\}_+ ds + \{\varphi(\mathbb{P}_{X_T^\alpha}^{W^0})\}_+ = 0, \mathbb{P}^0 \text{ a.s.}$$

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

8

2) By 1) and the continuity given by Proposition 5, we obtain $\mathcal{U}(\mathcal{Z}) = 0$. Then by Assumption 3 $\exists (\alpha, \beta) \in \mathcal{A} \times L^2(\mathbb{F}^0, \mathbb{R}^p)$ such that

$$\mathbb{E}\left[\{\widehat{g}(\mathbb{P}_{X_T^\alpha}^{W^0}) - Z_T^{z, \alpha, \beta}\}_+ + \sum_{l=1}^k \int_0^T \{\Psi(s, \mathbb{P}_{X_s^\alpha}^{W^0})\}_+ ds + \{\varphi(\mathbb{P}_{X_T^\alpha}^{W^0})\}_+\right] = 0.$$

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

$$(\mathbb{P}_{X_T^\alpha}^{W^0}, Z_T^{z, \alpha, \beta}) \in \text{Epi}(\widehat{g}), \Psi(s, \mathbb{P}_{X_s^\alpha}^{W^0}) \leq 0 \forall s \in [0, T], \text{ and } \varphi(\mathbb{P}_{X_T^\alpha}^{W^0}) \leq 0 \quad \mathbb{P} \text{ a.s.}$$

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

4 **5. Applications and numerical tests.** We design several machine learning meth-
 5 ods to solve this problem. We discretize the problem in time, parametrize the con-
 6 trol by a neural network and directly minimize the cost. When the constraints
 7 are almost sure, we can sometimes enforce them by choosing an appropriate neural
 8 network architecture, for instance in storage problems. A more adaptive alternative
 9 is to solve the unconstrained auxiliary problem. We propose an extension of the
 10 first algorithm from [17] to achieve this task. Thus we obtain a machine learning
 11 method able to solve state constrained mean field control problems.

12 **5.1. Algorithms.** We solve the auxiliary problem in the simpler case without com-
 13 mon noise with a first algorithm. We fix a relevant line segment K of \mathbb{R} on which we
 14 are going to explore the potential values of the problem. For instance we know that
 15 the value is greater than the value of the unconstrained problem V therefore it is
 16 useless to compute the auxiliary value function for $z \leq V$. We discretize the prob-
 17 lem in time on the grid $t_k := \frac{kT}{N}$. We call $\Delta t := \frac{kT}{N}$ and the Brownian increment
 18 $\Delta W_i := W_{t_{i+1}} - W_{t_i}$. For $j = 1, \dots, N$, ΔW_i^j (respectively X_0^j) correspond to sam-
 19 ples from N independent Brownian motions W^j (respectively from N independent
 20 random variables with law μ_0). For training we discretize K by using N points. We
 21 choose ε as a small parameter, typically smaller than 10^{-8} . In our tests we took
 22 $\varepsilon = 10^{-8}$ but we notice that with the level of discretization we chose for K , with 25
 23 points, the obtained values for w decrease to around 10^{-5} and 10^{-6} before reaching
 24 exactly zero so any value of $\varepsilon < 10^{-8}$ yields the same result on our examples. We
 25 refer to [5] for results on the numerical approximation of level sets with a given
 26 threshold in the context of constrained deterministic optimal control. We propose
 27 the following extension of the Method 1 from [17]. It is tested in Subsection 5.2. It
 28 can indeed also be used to solve unconstrained problem.

Remark 6. 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

$$\mathcal{Y}_\Lambda^\Psi := z \in \mathbb{R}$$

$$\mapsto \inf_{\alpha \in \mathcal{A}} \left[\{\hat{g}(\mathbb{P}_{X_T^\alpha}) - Z_T^{z,\alpha}\}_+ + \Lambda \int_0^T \{\Psi(s, \mathbb{P}_{X_s^\alpha})\}_+ ds + \Lambda \{\varphi(\mathbb{P}_{X_T^\alpha})\}_+ \right]. \quad (23)$$

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 17
 with the approach of [33] would provide another numerical method for mean-field
 control with state constraints. The extension to the common noise case, where we

aim to solve the auxiliary problem

$$\begin{aligned} \mathcal{Y}_\Lambda^\Psi &:= z \in \mathbb{R} \\ \mapsto &\inf_{(\alpha, \beta) \in \mathcal{A} \times L^2(\mathbb{F}^0, \mathbb{R}^p)} \left[\widehat{\mathcal{G}}(\mathbb{P}_{X_T^0}^{W^0}) - Z_T^{z, \alpha} \right]_+ + \Lambda \int_0^T \{ \Psi(s, \mathbb{P}_{X_s^0}^{W^0}) \}_+ ds + \Lambda \{ \varphi(\mathbb{P}_{X_T^0}^{W^0}) \}_+. \end{aligned} \quad (24)$$

1 is given in Algorithm 2 where the neural network for the control at each time step
 2 t_i takes in addition as input the current value of the common noise $W_{t_i}^0$. Notice
 3 that in general, the control may depend on the past values of the common noise,
 4 which could be taken into account in the neural network by taking as inputs the
 5 past increments of the common noise $\Delta W_0^0, \dots, \Delta W_{i-1}^0$, where $\Delta W_i^0 = W_{t_{i+1}}^0 - W_{t_i}^0$.
 6 The neural network for the auxiliary control β at each time t_i takes as inputs the
 7 current state $Z_i^{z, \alpha}$ and the current value of the common noise. An illustration is
 8 given in Section 5.3.

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

$$dX_t = \alpha_t r dt + \alpha_t \sigma dW_t, \quad 0 \leq t \leq T, \quad X_0 = x_0 \in \mathbb{R}.$$

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

$$\inf_{\alpha} J(\alpha) = \lambda \text{Var}(X_T^\alpha) - \mathbb{E}[X_T^\alpha] \quad (25)$$

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

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

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

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

The auxiliary deterministic unconstrained control problem is therefore

$$\begin{aligned} \mathcal{Y}_\Lambda(z) &:= \inf_{\alpha \in \mathcal{A}} \left[\lambda \text{Var}(X_T^\alpha) - \mathbb{E}[X_T^\alpha] - z \right]_+ \\ &\quad + \Lambda \int_0^T \{ (\delta - E[X_s^\alpha \mid X_s^\alpha \leq \theta]) \mathbb{P}(X_s^\alpha \leq \theta) \}_+ ds \end{aligned} \quad (26)$$

10 with the dynamics $dX_s^\alpha = \alpha_s r ds + \alpha_s \sigma dW_s$, which corresponds to the constraint
 11 function $\Psi(t, \mu) \mapsto (\delta - E_\mu[\xi \mid \xi \leq \theta]) \mu((-\infty, \theta])$. We have the representation
 12 $J(\alpha^*) = \mathcal{Z} = \inf\{z \in \mathbb{R} \mid \mathcal{Y}_\Lambda(z) = 0\}$. Indeed we see that the null control
 13 is admissible with the modified constraint $E[X_t^\alpha \mid X_t^\alpha \leq \theta] \mathbb{P}(X_t^\alpha \leq \theta) = 0 \geq$
 14 $(\delta + \varepsilon) \mathbb{P}(X_t^\alpha \leq \theta) = 0, \forall t \in [0, T]$ for any $0 < \varepsilon < \theta - \delta$ because $x_0 \geq \theta$ hence
 15 $\mathbb{P}(X_t^\alpha \leq \theta) = 0$ so we can apply Theorem 2.4. For practical application, other
 16 constraints could be considered like almost sure constraints on the portfolio weights
 17 as in [42]. Instead of the dualization method used by [34], constraints on the law of
 18 the tracking error with respect to a reference portfolio could be enforced.

Algorithm 1: Algorithm to solve the mean-field control problem (23)

Input parameters: $\Lambda, M, N, N_T, \varepsilon$.

For a discretization $z_1 < \dots < z_M$ of K , minimize over neural networks

$(\alpha_i)_{i \in 0, \dots, N_T-1}: \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^q$ the loss function

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

with w_Λ defined by

$$\begin{aligned} w_\Lambda(z) := & \mathbb{E} \left[\left\{ \frac{1}{N} \sum_{l=1}^N g \left(X_{N_T}^l, \frac{1}{N} \sum_{j=1}^N \delta_{X_{N_T}^j} \right) - Z_{N_T}^{z, \alpha} \right\}_+ \right. \\ & \left. + \Lambda \sum_{m=1}^k \sum_{i=1}^{N_T} \left\{ \Psi^m \left(t_i, \frac{1}{N} \sum_{j=1}^N \delta_{X_i^j} \right) \right\}_+ + \Lambda \left\{ \varphi \left(\frac{1}{N} \sum_{j=1}^N \delta_{X_{N_T}^j} \right) \right\}_+ \right]. \end{aligned}$$

/* Auxiliary problem */

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

$$\begin{aligned} X_{i+1}^j &= X_i^j + b(t_i, X_i^j, \alpha_i(X_i^j, Z_i^{z, \alpha}), \bar{\mu}_i) \Delta t \\ &+ \sigma(t_i, X_i^j, \alpha_i(X_i^j, Z_i^{z, \alpha}), \bar{\mu}_i) \Delta W_i^j, \quad X_0^j \sim \mu_0 \\ Z_{i+1}^{z, \alpha} &= Z_i^{z, \alpha} - \frac{1}{N} \sum_{l=1}^N f(t_i, X_i^l, \alpha_i(X_i^l, Z_i^{z, \alpha}), \bar{\mu}_i) \Delta t, \quad Z_0^{z, \alpha} = z \\ \bar{\mu}_i &= \frac{1}{N} \sum_{j=1}^N \delta_{(X_i^j, \alpha_i(X_i^j, Z_i^{z, \alpha}))} \end{aligned}$$

/* Particle approximations */

Define α^* as the solution to this minimization problem.

Then, compute $V_0 = \inf\{z_i, i \in \llbracket 1, M \rrbracket \mid w_\Lambda(z_i) \leq \varepsilon\}$ with $\alpha = \alpha^*$ in the dynamics.

/* Recovering the cost of the original problem */

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

/* Recovering the control of the original problem */

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

$$\begin{aligned} \inf_{\alpha} J(\alpha) &= \lambda \text{Var}(X_T^\alpha) - \mathbb{E}[X_T^\alpha] \\ dX_t &= \alpha_t r dt + \alpha_t \sigma dW_t, \\ (0.8 - \mathbb{E}[X_t^\alpha \mid X_t^\alpha \leq 0.9]) \mathbb{P}(X_t^\alpha \leq 0.9) &\leq 0, \quad \forall t \in [0, T]. \end{aligned} \quad (27)$$

- 1 We compare the controls from Algorithm 1 with the exact optimal ones in the
- 2 unconstrained case for which we have an analytical value. We also solve without
- 3 constraints for comparison and plot the final time histograms. We solve the uncon-
- 4 strained case with algorithm 1 and the one from [17] for comparison. We take 50

Algorithm 2: Algorithm to solve mean-field control problem (24)

Input parameters: $\Lambda, M, N, N_T, \varepsilon$.

For a discretization $z_1 < \dots < z_M$ of K , minimize over neural networks

 $(\alpha_i)_{i \in 0, \dots, N_T-1}: \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^p \rightarrow \mathbb{R}^q$ and $(\beta_i)_{i \in 0, \dots, N_T-1}: \mathbb{R} \times \mathbb{R}^p \rightarrow \mathbb{R}^p$ the loss function

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

with w_Λ defined by

$$\begin{aligned} w_\Lambda(z) := & \mathbb{E} \left[\left\{ \frac{1}{N} \sum_{l=1}^N g \left(X_{N_T}^l, \frac{1}{N} \sum_{j=1}^N \delta_{X_{N_T}^j} \right) - Z_{N_T}^{z, \alpha, \beta} \right\}_+ \right. \\ & \left. + \Lambda \sum_{i=1}^{N_T} \left\{ \Psi \left(t_i, \frac{1}{N} \sum_{j=1}^N \delta_{X_i^j} \right) \right\}_+ \Delta t + \Lambda \left\{ \varphi \left(\frac{1}{N} \sum_{j=1}^N \delta_{X_{N_T}^j} \right) \right\}_+ \right]. \end{aligned}$$

/* Auxiliary problem */

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

$$\begin{aligned} X_{i+1}^j &= X_i^j + b(t_i, X_i^j, \alpha_i(X_i^j, Z_i^{z, \alpha, \beta}, W_{t_i}^0), \bar{\mu}_i) \Delta t \\ &+ \sigma(t_i, X_i^j, \alpha_i(X_i^j, Z_i^{z, \alpha, \beta}, W_{t_i}^0), \bar{\mu}_i) \Delta W_i^j \\ &+ \sigma_0(t_i, X_i^j, \alpha_i(X_i^j, Z_i^{z, \alpha, \beta}, W_{t_i}^0), \bar{\mu}_i) \Delta W_i^0, \quad X_0^j \sim \mu_0 \\ Z_{i+1}^{z, \alpha, \beta} &= Z_i^{z, \alpha, \beta} - \frac{1}{N} \sum_{l=1}^N f(t_i, X_i^l, \alpha_i(X_i^l, Z_i^{z, \alpha, \beta}, W_{t_i}^0), \bar{\mu}_i) \Delta t \\ &+ \beta_i(Z_i^{z, \alpha, \beta}, W_{t_i}^0) \Delta W_i^0, \quad Z_0^{z, \alpha, \beta} = z \\ \bar{\mu}_i &= \frac{1}{N} \sum_{j=1}^N \delta_{(X_i^j, \alpha_i(X_i^j, Z_i^{z, \alpha, \beta}, W_{t_i}^0))} \end{aligned}$$

/* Particle approximations */

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

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

 $\beta = \beta^*$ in the dynamics.

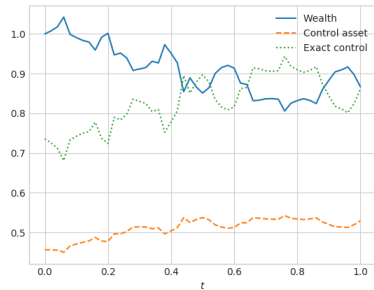
/* Recovering the cost of the original problem */

Return the value V_0 and the optimal controls $\hat{\alpha}_i: x \mapsto \alpha_i^*(x, Z_i^{V_0, \alpha^*, \beta^*}, W_{t_i}^0)$

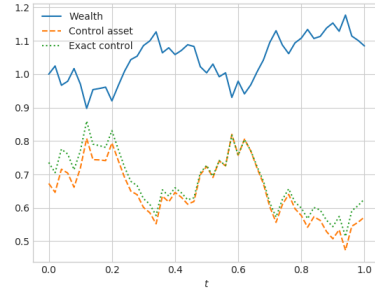
for $i = 0, \dots, N_T - 1$.

/* Recovering the control of the original problem */

- 1 time steps for the time discretization and a batch size of 20000. We use an feedfor-
- 2 ward architecture with two hidden layers of 15 neurons. We perform 15000 gradient
- 3 descent iterations with Adam optimizer (see [31]) thanks to the Tensorflow library.
- 4 The true value $v = J(\alpha^*)$ is -1.05041 without constraints. We also have the upper
- 5 bound -1 . for the value in the constrained case, corresponding to the identically
- 6 null control and wealth process $X_t = 1 \forall t \in [0, T]$. With constraint we choose
- 7 $K = [-1.047, -1.041]$, without constraint we take $K = [-1.07, -1.03]$, discretized
- 8 by regular grids with 25 points.

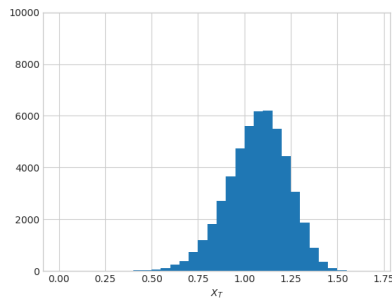


Problem (27)

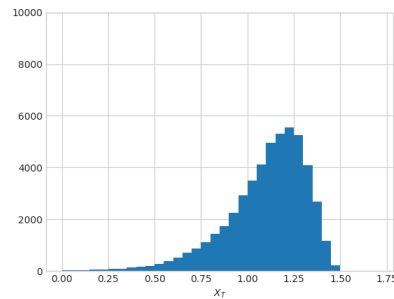


No constraint, problem (25)

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 27

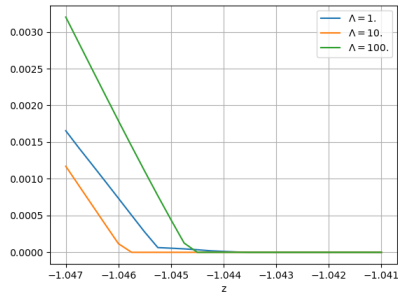


No constraint, problem 25

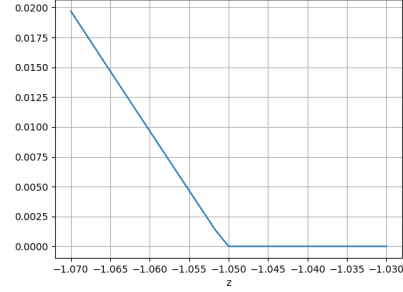
FIGURE 2. Histogram of X_T^α for 50000 samples. Here $\Lambda = 100$.

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

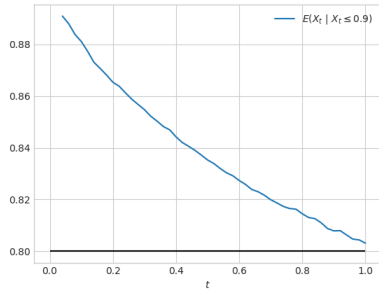
Our method can also handle directly the primal of the mean-variance problem, that is to maximize over portfolio control α the expected terminal wealth under a



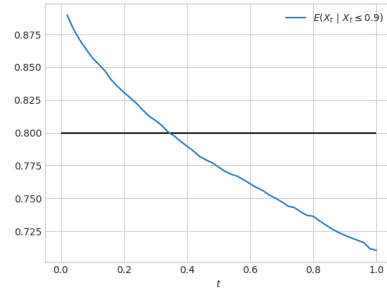
Problem (27)



No constraint, problem (25)

 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 (27)



No constraint, problem (25)

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

terminal variance constraint:

$$\begin{aligned} \inf_{\alpha} \bar{J}(\alpha) &= -\mathbb{E}[X_T^\alpha] \\ dX_t &= \alpha_t r dt + \alpha_t \sigma dW_t, \\ \text{Var}(X_T^\alpha) &\leq \vartheta. \end{aligned} \tag{28}$$

which give the same optimal control as Problem (25) under the correspondence $\lambda = \sqrt{\frac{\exp(\sigma^{-2}r^2T)-1}{4\vartheta}}$. This problem allows us to consider a constrained problem with an analytical solution. In this case $\text{Var}(X_T^{\alpha^*}) = \vartheta$ thus $J(\alpha^*) = \lambda \text{Var}(X_T^{\alpha^*}) + \bar{J}(\alpha^*) = \lambda\vartheta + \bar{J}(\alpha^*)$. For comparison with Problem (25) we thus report $\lambda\vartheta + \bar{J}(\alpha^*)$ for Problem (28) in Table 1 and choose $\vartheta = \frac{\exp(\sigma^{-2}r^2T)-1}{4\lambda^2} = 0.0504$. In this case

the auxiliary deterministic unconstrained control problem is now

$$\mathcal{U}_\Lambda(z) = \inf_{\alpha \in \mathcal{A}} \left[\{-\mathbb{E}[X_T^\alpha] - z\}_+ + \Lambda \{\text{Var}(X_T^\alpha) - \vartheta\}_+ \right]$$

$$dX_t = \alpha_t r dt + \alpha_t \sigma dW_t, \quad (29)$$

- 1 which corresponds to the constraint function $\Psi(t, \mu) \mapsto (\text{Var}(\mu) - \vartheta)_+ \mathbb{1}_{t=T}$ and the
- 2 modified constraint function $\bar{\Psi}_\eta(t, \mu) \mapsto (\text{Var}(\mu) - \vartheta)_+ \mathbb{1}_{t=T} - \eta \mathbb{1}_{t < T}$ (see Remark 3).
- 3 Theorem 2.4 still applies as far as the null control is admissible with the modified
- 4 constraint $(\text{Var}(\mu) - \vartheta)_+ \mathbb{1}_{t=T} + \varepsilon - \eta \mathbb{1}_{t < T} \leq 0$ for any $0 < \varepsilon < \eta$ and any $t \in [0, T]$.

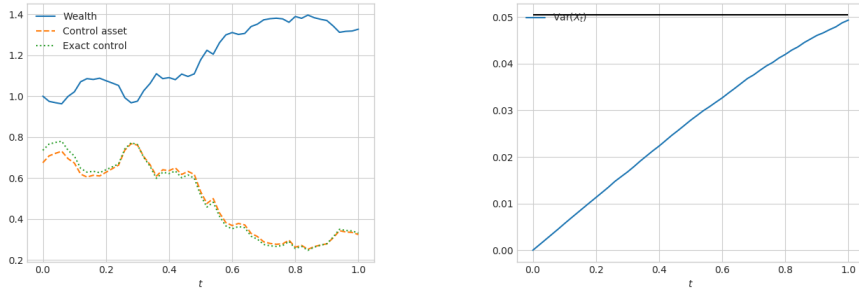


FIGURE 5. Sample trajectory of the controlled process X_t^α and the control for problem (28) (left). Variance $\text{Var}(X_t)$ estimated with 50000 samples for problem (28) (right) with $\Lambda = 10$

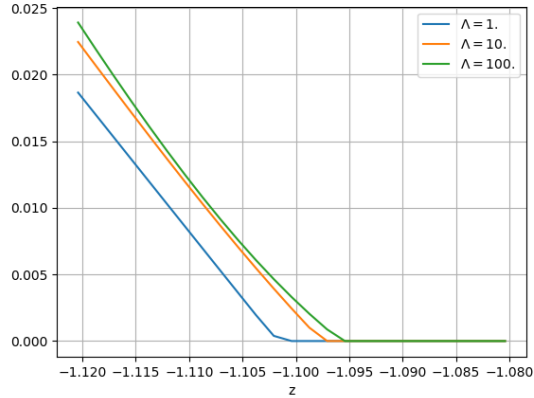


FIGURE 6. Auxiliary value function $\mathcal{U}_\Lambda(z)$ for several values of Λ

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

Problem	Average	Std	Tr. val.	Error	$\mathbb{E}[X_T^{\alpha^*}]$	Tr. $\mathbb{E}[X_T^{\alpha^*}]$	$\text{Var}(X_T^{\alpha^*})$	Tr. $\text{Var}(X_T^{\alpha^*})$
(27) $\Lambda = 1.$	-1.044	0.0010	Not avail.	Not avail.	1.07	Not avail.	0.026	Not avail.
(27) $\Lambda = 10.$	-1.044	0.0005	Not avail.	Not avail.	1.07	Not avail.	0.026	Not avail.
(27) $\Lambda = 100.$	-1.045	0.0005	Not avail.	Not avail.	1.07	Not avail.	0.027	Not avail.
(28) $\Lambda = 10.$	-1.048	0.0017	-1.050	0.22	1.10	1.10	0.049	0.050
(25)	-1.050	0.0009	-1.050	0.07	1.10	1.10	0.050	0.050
(25) [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. 'Not avail.' means that we don't have a reference value and 'Tr.' means true. For problem (28), we take $\Lambda = 10$ and for problem (27) we illustrate the values obtained for $\Lambda \in \{1., 10., 100\}$.

1 In Table 1 we observe that our method gives a small variance for the results over
2 several runs. In the case where an analytical solution is known, the value of the
3 control problem is computed accurately with less than 0.5% of relative error. The
4 expectation and variance of the terminal value of the optimally controlled process
5 are also very close to their theoretical values. In the case of a conditional expectation
6 constraint, even though we don't have an exact solution we notice that the value is
7 close to the unconstrained value hence since our solution is admissible, we expect
8 to be near optimality. On the unconstrained problem (25) our scheme and the one
9 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 range¹. We consider the following constraints

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

with $X_{\max} \geq 0$, $\underline{\alpha} \leq 0 \leq \bar{\alpha}$. Define the spot price of electricity S_t without wind power, \tilde{S}_t the price with wind production. Selling a quantity $P_t^i - \alpha_t^i$ on the market, producer i obtains a revenue $\tilde{S}_t(P_t^i - \alpha_t^i)$ (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

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

modeling the impact of intermittent renewable production on the market. Θ is positive, non-decreasing and bounded. We call $\Theta_\infty = \lim_{N \rightarrow \infty} \Theta(N) < \infty$. We

¹<https://css.umich.edu/factsheets/us-grid-energy-storage-factsheet>

consider $N + 2$ independent Brownian motions $W_t^0, B_t^0, W_t^1, \dots, W_t^N$ and the following dynamics for the producers $i = 1, \dots, N$ state processes

$$\begin{cases} dX_t^i = \alpha_t^i dt \\ dP_t^i = \kappa(\phi P_{\max} - P_t^i) dt + \sigma_p(P_t \wedge \{P_{\max} - P_t^i\})_+ (\rho dW_t^0 + \sqrt{1 - \rho^2} dW_t^i) \\ dF(t, T) = F(t, T) \sigma_f e^{-a(T-t)} dB_t^0 \\ S_t = F(t, t), \end{cases}$$

for some positive constants $\kappa, \phi, P_{\max}, \sigma_p, \sigma_f$, and $\rho \in [-1, 1]$. 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 \tilde{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

$$J^i(\alpha_1, \dots, \alpha_N) = \mathbb{E} \left[\int_0^T \left\{ S_t (P_t^i - \alpha_t^i) - \frac{\Theta(N)}{N} (P_t^i - \alpha_t^i) \sum_{j=1}^N (P_t^j - \alpha_t^j) \right\} dt \right].$$

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

$$\begin{aligned} & - \inf_{\alpha \in \mathcal{A}} \mathbb{E} \left[\int_0^T \left\{ -S_t (P_t - \alpha_t) + \Theta_\infty (P_t - \alpha_t) \mathbb{E}[P_t - \alpha_t | \mathbb{F}^0] \right\} dt \right] \\ & dX_t = \alpha_t dt \\ & dP_t = \kappa(\phi P_{\max} - P_t) dt + \sigma_p(P_t \wedge \{P_{\max} - P_t\})_+ (\rho dW_t^0 + \sqrt{1 - \rho^2} dW_t^1) \\ & dF(t, T) = F(t, T) \sigma_f e^{-a(T-t)} dB_t^0 \\ & S_t = F(t, t) \\ & 0 \leq X_t \leq X_{\max} \mathbb{P} \text{ a.s.} \end{aligned}$$

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 $[\underline{\alpha}, \bar{\alpha}]$. A similar problem is solved by [1] without any storage constraints by Pontryagin principle. With constraints but without mean-field interaction, a close problem is solved by [39]. 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

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

where μ_1 is the first marginal law of the measure μ . The auxiliary unconstrained control problem is therefore

$$\begin{aligned}
& w(z) \\
& := - \inf_{\alpha, \beta^0 \in \mathcal{A} \times L^2(\mathbb{F}^0, \mathbb{R}^2)} \mathbb{E} \left[\left\{ \int_0^T \mathbb{E}[-S_t(P_t - \alpha_t) + \Theta_\infty(P_t - \alpha_t) \mathbb{E}[P_t - \alpha_t | \mathbb{F}^0] | \mathbb{F}^0] dt \right. \right. \\
& \quad \left. \left. - z - \int_0^T \beta_s^{0,1} dW_s^0 - \int_0^T \beta_s^{0,2} dB_s^0 \right\}_+ + \frac{1}{\epsilon} \int_0^T \mathbb{E}[(-X_s)_+^2 + (X_s - X_{\max})_+^2] ds \right] \\
& dX_t = \alpha_t dt \\
& dP_t = \kappa(\phi P_{\max} - P_t) dt + \sigma_p(P_t \wedge \{P_{\max} - P_t\})_+ (\rho dW_t^0 + \sqrt{1 - \rho^2} dW_t^1) \\
& dF(t, T) = F(t, T) \sigma_f e^{-a(T-t)} dB_t^0 \\
& S_t = F(t, t), \tag{30}
\end{aligned}$$

1 where ϵ is a small term used to force the a.s. constraints and $\beta^{0,1}, \beta^{0,2}$ are the
2 coordinates of β^0 .

3

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

$$\begin{aligned}
& - \inf_{\alpha \in \mathcal{A}} \mathbb{E} \left[\int_0^T \{(-S_t + \Theta_\infty(P_t - \alpha_t))(P_t - \alpha_t)\} dt \right] \\
& dX_t = \alpha_t dt \\
& dP_t = \kappa(\phi P_{\max} - P_t) dt + \sigma_p(P_t \wedge \{P_{\max} - P_t\})_+ dW_t^0 \\
& dF(t, T) = F(t, T) \sigma_f e^{-a(T-t)} dB_t^0 \\
& S_t = F(t, t) \\
& 0 \leq X_t \leq X_{\max} \mathbb{P} \text{ a.s.} \tag{31}
\end{aligned}$$

and equation (30) gives

$$\begin{aligned}
& w(z) := - \inf_{\alpha, \beta \in \mathcal{A} \times L^2} \mathbb{E} \left[((Y^{\alpha, \beta} - z)_+ + \frac{1}{\epsilon} \int_0^T \mathbb{E}[(-X_s)_+^2 + (X_s - X_{\max})_+^2] ds \right] \\
& dX_t = \alpha_t dt \\
& dP_t = \kappa(\phi P_{\max} - P_t) dt + \sigma_p(P_t \wedge \{P_{\max} - P_t\})_+ dW_t^0 \\
& dF(t, T) = F(t, T) \sigma_f e^{-a(T-t)} dB_t^0 \\
& S_t = F(t, t)
\end{aligned}$$

where

$$Y^{\alpha, \beta} = \int_0^T (-S_t + \Theta_\infty(P_t - \alpha_t))(P_t - \alpha_t) dt - \int_0^T \beta_s^{0,1} dW_s^0 - \int_0^T \beta_s^{0,2} dB_s^0. \tag{32}$$

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 (32) will be estimated discretizing the integral $\int_0^T \beta_s^{0,1} dW_s^0$ and $\int_0^T \beta_s^{0,2} dB_s^0$ 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 :

$$\begin{aligned} Y^{\alpha, \beta} &= \int_0^T (-S_t + \Theta_\infty(P_t - \alpha_t))(P_t - \alpha_t) dt - \int_0^T (-S_t + \Theta_\infty(P_t - \hat{\alpha}_t))(P_t - \hat{\alpha}_t) dt \\ &+ \mathbb{E} \left[\int_0^T (-S_t + \Theta_\infty(P - \hat{\alpha}_t))(P_t - \hat{\alpha}_t) dt \right] - \int_0^T \beta_s^{0,1} dW_s^0 - \int_0^T \beta_s^{0,2} dB_s^0 \end{aligned}$$

1 where $\hat{\alpha}_t$ is the rough estimation of the optimal **deterministic** command maximiz-
2 ing the gain.

3 We take $T = 40$, $N_T = 40$ time steps, $X_{\max} = 1$, $X_0 = 0.5$, $P_0 = 0.12$, $F(0, t) =$
4 $30 + 5 \cos(\frac{2\pi t}{N}) + \cos(\frac{2\pi t}{7})$, $\sigma_f = 0.3$, $a = 0.16$, $\kappa = 0.2$, $\sigma_p = 0.2$, $\phi = 0.3$, $P_{\max} = 0.2$,
5 $-0.2 \leq \alpha \leq 0.2$, $\Theta(N) = 10$.

6 The network depends on P_t, S_t, X_t and z where z takes some deterministic values
7 on a grid with the same spacing. The global curve is therefore approximated by a
8 single run.

9 The grid is taken from 107 to 127 with a spacing of 0.5. The neural networks have
10 two hidden layers with 14 neurons on each layer. We take a ϵ parameter equal to
11 10^{-4} . The number of gradient iterations is set to 50000 with a learning rate equal
to 2×10^{-3} . We give the \hat{w} function on figure 7. Using Dynamic Programming

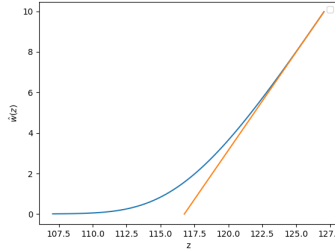


FIGURE 7. \hat{w} function value for the storage problem

12 with the StOpt library [29], we get an optimal value equal to 117.28 while a direct
13 optimization of (31) using some neural networks as in [43], [17] we get a value of
14 117.11. Encouraged by Remark 6, Figure 3, Figure 6 and the related comments,
15 we empirically estimate the value function by the point where the linear part of
16 the auxiliary function reaches zero when $\Lambda = \frac{1}{\epsilon}$ is sufficiently large. The estimated
17 value is 116.75, close to the reference solutions. On figure 8, we compare trajecto-
18 ries obtained by Dynamic Programming and by the Level Set approach : they are
19 accurately calculated.
20

21 **Acknowledgments.** We thank Marianne Akian, Olivier Bokanowski, and Nadia
22 Oudjane for useful comments.

23

REFERENCES

- 24 [1] C. Alasseur, I. Ben Taher and A. Matoussi, An extended mean field game for storage in smart
25 grids, *Journal of Optimization Theory and Applications*, **184** (2020), 644–670.
26 [2] A. Altarovici, O. Bokanowski and H. Zidani, A general hamilton-jacobi framework for non-
27 linear state-constrained control problems, *ESAIM: COCV*, **19** (2013), 337–357.

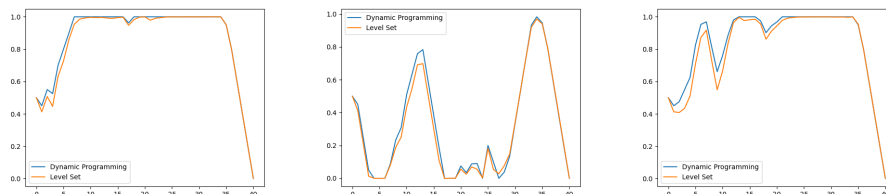


FIGURE 8. Storage trajectories with the Level Set and Dynamic Programming method.

- 1 [3] S. S. Arjmand and G. Mazanti, Nonsmooth mean field games with state constraints,
2 *arXiv:2110.15713*.
- 3 [4] A. Balata, M. Ludkovski, A. Maheshwari and J. Palczewski, Statistical learning for
4 probability-constrained stochastic optimal control, *European Journal of Operational Re-*
5 *search*, **290** (2021), 640–656.
- 6 [5] O. Bokanowski, N. Gammoudi and H. Zidani, Optimistic Planning Algorithms For State-
7 Constrained Optimal Control Problems, *Computers and Mathematics with applications*, **109**
8 (2022), 158–179
- 9 [6] O. Bokanowski, A. Picarelli and H. Zidani, Dynamic programming and error estimates for
10 stochastic control problems with maximum cost, *Appl Math Optim*, **74** (2014), 125–163.
- 11 [7] O. Bokanowski, A. Picarelli and H. Zidani, State-constrained stochastic optimal control prob-
12 lems via reachability approach, *SIAM Journal on Control and Optimization*, **54** (2016), 2568–
13 2593.
- 14 [8] B. Bonnet, A pontryagin maximum principle in wasserstein spaces for constrained optimal
15 control problems, *ESAIM: COCV*, **25** (2019), 52.
- 16 [9] B. Bonnet and H. Frankowska, Necessary optimality conditions for optimal control problems
17 in wasserstein spaces, *Appl Math Optim*, **84** (2021), 1281–1330
- 18 [10] B. Bouchard, B. Djehiche and I. Kharroubi, Quenched Mass Transport of Particles Toward a
19 Target, *Journal of Optimization Theory and Applications*, **186** (2020), 345–374.
- 20 [11] B. Bouchard, R. Elie and C. Imbert, Optimal control under stochastic target constraints,
21 *SIAM Journal on Control and Optimization*, **48** (2010), 3501–3531.
- 22 [12] P. Cannarsa and R. Capuani, Existence and uniqueness for mean field games with state
23 constraints, in *PDE Models for Multi-Agent Phenomena* (eds. C. Pierre, P. A. and S. F.),
24 vol. 28 of Springer INdAM Series, Springer, Cham, 2018.
- 25 [13] P. Cannarsa, R. Capuani and P. Cardaliaguet, Mean field games with state constraints: from
26 mild to pointwise solutions of the pde system, *Calculus of Variations and Partial Differential*
27 *Equations*, **60** (2021), article number 108.
- 28 [14] R. Carmona and F. Delarue, Forward–backward stochastic differential equations and controlled
29 McKean–Vlasov dynamics, *The Annals of Probability*, **43** (2015), 2647 – 2700.
- 30 [15] R. Carmona and F. Delarue, *Probabilistic Theory of Mean Field Games: vol. I, Mean Field*
31 *FBSDEs, Control, and Games*, Springer, 2018.
- 32 [16] R. Carmona and F. Delarue, *Probabilistic Theory of Mean Field Games: vol. II, Mean Field*
33 *game with common noise and Master equations*, Springer, 2018.
- 34 [17] R. Carmona and M. Laurière, Convergence analysis of machine learning algorithms for the nu-
35 merical solution of mean-field control and games: II the finite horizon case, *arXiv:1908.01613*,
36 *to appear in The Annals of Applied Probability*.
- 37 [18] L. Chen and J. Wang, Maximum principle for delayed stochastic mean-field control problem
38 with state constraint, *Advances in Difference Equations*, (2019), article number 348.
- 39 [19] Y.-L. Chow, X. Yu and C. Zhou, On Dynamic Programming Principle for Stochastic Control
40 under Expectation Constraints, *Journal of Optimization Theory and Applications*, **185**
41 (2020), 803–818.
- 42 [20] A. Cosso, F. Gozzi, I. Kharroubi, H. Pham and M. Rosestolato, Optimal control of path-
43 dependent mckean-vlasov sdes in infinite dimension, *arXiv:2012.14772*, to appear in *Annals*
44 *of Applied Probability*.

- 1 [21] A. Cosso and H. Pham, Zero-sum stochastic differential games of generalized mckean–vlasov
2 type, *Journal de Mathématiques Pures et Appliquées*, **129** (2019), 180 – 212.
- 3 [22] N. Curin, M. Kettler, X. Kleinsinger-Yu, V. Komaric, T. Krabichler, J. Teichmann and
4 H. Wutte, A deep learning model for gas storage optimization, *arXiv:2102.01980*.
- 5 [23] S. Daudin, Optimal control of diffusion processes with terminal constraint in law,
6 *arXiv:2012.10707*.
- 7 [24] S. Daudin, Optimal control of the fokker-planck equation under state constraints in the wasser-
8 stein space, *arXiv:2109.14978*.
- 9 [25] M. F. Djete, D. Possamaï and X. Tan, Mckean-Vlasov optimal control: the dynamic program-
10 ming principle, *arXiv:1907.08860*, to appear in *Annals of Probability*.
- 11 [26] G. Fu and U. Horst, Mean-field leader-follower games with terminal state constraint, *SIAM*
12 *Journal on Control and Optimization*, **58** (2020), 2078–2113.
- 13 [27] A. Galichon, P. Henry-Labordère and N. Touzi, A stochastic control approach to no-arbitrage
14 bounds given marginals, with an application to lookback options, *The Annals of Applied*
15 *Probability*, **24** (2014), 312 – 336.
- 16 [28] A. Geletu, M. Klöppel, H. Zhang and P. Li, Advances and applications of chance-constrained
17 approaches to systems optimisation under uncertainty, *International Journal of Systems Sci-*
18 *ence*, **44** (2013), 1209–1232.
- 19 [29] H. Gevret, N. Langrené, J. Lelong, R. Lobato, T. Ouillon, X. Warin and A. Maheshwari,
20 Stochastic optimization library in c++, 2018, URL [https://hal.archives-ouvertes.fr/
21 hal-01361291](https://hal.archives-ouvertes.fr/hal-01361291).
- 22 [30] J. Graber and S. Mayorga, A note on mean field games of controls with state constraints:
23 existence of mild solutions, *arXiv:2109.11655*.
- 24 [31] D.P. Kingma and J. Ba, Adam: A Method for Stochastic Optimization *3rd International*
25 *Conference for Learning Representations* (2014)
- 26 [32] M. Laurière and O. Pironneau, Dynamic programming for mean-field type control, *Comptes*
27 *Rendus Mathématique*, **352** (2014), 707–713.
- 28 [33] M. Germain, M. Laurière, H. Pham and X. Warin, DeepSets and derivative networks for
29 solving symmetric PDEs, *Journal of Scientific Computing*, **91** (2022), article number 63
- 30 [34] W. Lefebvre, G. Loeper and H. Pham, Mean-variance portfolio selection with tracking error
31 penalization, *Mathematics*, **8** (2020).
- 32 [35] S. M. Pesenti and S. Jaimungal, Portfolio optimisation within a wasserstein ball, *Available at*
33 *SSRN: <https://ssrn.com/abstract=3744994>*.
- 34 [36] L. Pfeiffer, X. Tan and Y.-L. Zhou, Duality and approximation of stochastic optimal control
35 problems under expectation constraints, *SIAM Journal on Control and Optimization*, **59**
36 (2021), 3231–3260.
- 37 [37] H. Pham and X. Wei, Dynamic programming for optimal control of stochastic mckean-vlasov
38 dynamics, *SIAM Journal on Control and Optimization*, **55** (2017), 1069–1101.
- 39 [38] H. Pham and X. Wei, Bellman equation and viscosity solutions for mean-field stochastic
40 control problem, *ESAIM: COCV*, **24** (2018), 437–461.
- 41 [39] A. Picarelli and T. Vargiolu, Optimal management of pumped hydroelectric production with
42 state constrained optimal control, *Journal of Economic Dynamics and Control*, **126** (2021),
43 103940.
- 44 [40] T. Rockafellar, *Convex Analysis* Princeton University Press, 1970.
- 45 [41] H. M. Soner and N. Touzi, Stochastic target problems, dynamic programming, and viscosity
46 solutions, *SIAM Journal on Control and Optimization*, **41** (2002), 404–424.
- 47 [42] X. Warin, Deep learning for efficient frontier calculation in finance, *arXiv:2101.02044*.
- 48 [43] X. Warin, Reservoir optimization and machine learning methods, *arXiv:2106.08097*.

49 Received xxxx 20xx; revised xxxx 20xx; early access xxxx 20xx.

50 *E-mail address:* mgermain@lpsm.paris

51 *E-mail address:* pham@lpsm.paris

52 *E-mail address:* xavier.warin@edf.fr