



**HAL**  
open science

# Entropic optimal planning for path-dependent mean field games

Zhenjie Ren, Xiaolu Tan, Nizar Touzi, Junjian Yang

► **To cite this version:**

Zhenjie Ren, Xiaolu Tan, Nizar Touzi, Junjian Yang. Entropic optimal planning for path-dependent mean field games. 2022. hal-03903130

**HAL Id: hal-03903130**

**<https://inria.hal.science/hal-03903130v1>**

Preprint submitted on 16 Dec 2022

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

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

# Entropic optimal planning for path-dependent mean field games

Zhenjie Ren <sup>\*</sup>    Xiaolu Tan <sup>†</sup>    Nizar Touzi <sup>‡</sup>    Junjian Yang <sup>§</sup>

December 16, 2022

## Abstract

In the context of mean field games, with possible control of the diffusion coefficient, we consider a path-dependent version of the planning problem introduced by P.L. Lions: given a pair of marginal distributions  $(\mu_0, \mu_1)$ , find a specification of the game problem starting from the initial distribution  $\mu_0$ , and inducing the target distribution  $\mu_1$  at the mean field game equilibrium. Our main result reduces the path-dependent planning problem into an embedding problem, that is, constructing a McKean-Vlasov dynamics with given marginals  $(\mu_0, \mu_1)$ . Some sufficient conditions on  $(\mu_0, \mu_1)$  are provided to guarantee the existence of solutions. We also characterize, up to integrability, the minimum entropy solution of the planning problem. In particular, as uniqueness does not hold anymore in our path-dependent setting, one can naturally introduce an optimal planning problem which would be reduced to an optimal transport problem along controlled McKean-Vlasov dynamics.

**MSC 2010 Subject Classification:** 49N70, 91A13, 91B40, 93E20

**Key words:** Mean field games, planning problem, McKean-Vlasov dynamic, optimal transport.

## 1 Introduction

During his courses at Collège de France [16], P.-L. Lions introduced the following planning problem for a class mean field games (MFG hereafter): *given two marginal distributions  $\mu_0$  and  $\mu_1$  on  $\mathbb{R}^d$ , find a solution  $(u, m)$  of the following MFG system:*

$$-\partial_t u - \frac{\sigma^2}{2} \Delta u - H(x, \nabla u) + F(x, m) = 0, \quad \text{in } (0, 1) \times \mathbb{R}^d, \quad (1.1)$$

$$\partial_t m - \frac{\sigma^2}{2} \Delta m + \nabla \cdot (m \nabla_z H(x, \nabla u)) = 0, \quad \text{in } (0, 1) \times \mathbb{R}^d, \quad (1.2)$$

$$m(0, \cdot) = \mu_0, \quad m(1, \cdot) = \mu_1, \quad \text{in } \mathbb{R}^d. \quad (1.3)$$

Namely, let  $c(x, b)$  denote the Legendre transform of the Hamiltonian  $H(x, z)$  in  $z$ , then

- $u$  in (1.1) corresponds to the value function of the stochastic optimal control problem

$$\sup_{\beta} \mathbb{E} \left[ u(1, X_1^\beta) - \int_0^1 \left( c(X_t^\beta, \beta_t) \right) + F(X_t^\beta, m_t) dt \right], \quad \text{subject to } X_t^\beta = X_0 + \int_0^t \beta_s ds + W_t,$$

---

<sup>\*</sup>CEREMADE, Université Paris-Dauphine, PSL Research University. ren@ceremade.dauphine.fr.

<sup>†</sup>Department of Mathematics, The Chinese University of Hong Kong. xiaolu.tan@cuhk.edu.hk.

<sup>‡</sup>CMAP, Ecole Polytechnique, nizar.touzi@polytechnique.edu

<sup>§</sup>FAM, Fakultät für Mathematik und Geoinformation, Vienna University of Technology, A-1040 Vienna, Austria. junjian.yang@tuwien.ac.at

where  $W$  is a Brownian motion, and one optimizes over all progressively measurable processes  $\beta$ ;

- the Fokker-Plank equation (1.2) characterizes the marginal distribution of the state process  $X^*$  under optimal control  $\beta_t^* = \nabla_z H(X_t^*, \nabla u(t, X_t^*))$ ;
- and (1.3) collects the initial distribution of the Fokker-Plank equation, as standard, and in addition a final condition which conditions the choice of the final reward function  $u(1, \cdot)$

In other wrds, unlike the standard MFG formulation, the HJB equation (1.1) is not complemented with a terminal condition for  $u|_{t=1}$ , and instead the Fokker-Planck equation (1.3) is equipped with a terminal condition on  $m|_{t=1}$  in addition to the initial condition  $m|_{t=0}$ . In other words, the planning problem consists in finding an appropriate reward function which stands as the terminal condition for the HJB equation (1.1):

$$g := u|_{t=1}.$$

At the level of the control problem, this can be interpreted as an incentive for the population so that the classical MFG problem has a solution satisfying the marginal constraint  $m|_{t=0} = \mu_0$  and  $m|_{t=1} = \mu_1$ . For this reason,  $g$  is usually referred to as the incentive function.

In the quadratic Hamiltonian setting, Lions [16] proved an existence and uniqueness result for a large class of initial and target measures. Various extensions have been achieved since then essentially allowing for Hamiltonians with quadratic growth in the gradient, and using weak solutions for the MFG equation, see Achdou, Camilli, and Capuzzo-Dolcetta [1], Porretta [22], Graber, Mészáros, Silva, and Tonon [11], Orrieri, Porretta, and Savaré [21], Benamou, Carlier, Di Marino, and Nenna [4], among others.

The main objective of this paper is to extend the formulation of the planning problem to the path-dependent setting. More precisely, the HJB equation is replaced by a possibly path-dependent stochastic control problem, and the Fokker-Planck equation is replaced by the path-dependent stochastic differential equation characterizing the dynamics of the underlying state under the optimal action induced by the control problem. As another extension that we consider in the present paper, we allow for the control of diffusion coefficient which means that, unlike (1.1), the HJB equation in the corresponding Markovian setting is allowed to be fully nonlinear.

By allowing for path dependency, we are considering a much larger class of incentives  $\xi$  which may now be chosen as the set of all functionals of the path of the underlying state. Therefore, existence of solution should be easier, but we lose the uniqueness feature of the initial planning problem under the monotone condition in Porretta [22, Theorem 1.3]. On the other hand, the multiplicity of solutions in our path-dependent extension raises naturally the planner's optimization problem over all possible  $\xi$  according to some performance or loss criterion. This point of view is in fact very popular in the literature on contract theory which sets the rules of the so-called delegation problem between a principal and an agent subject to moral hazard. The nature of the incentive salary of the principal to the agent in compensation for the management of some given output is modeled by means of a leader-follower stochastic game: the leader choses the best incentive compensation given the follower's optimal response. The seminal paper by Holström & Milgrom [12] introduces the continuous time modeling of this problem as a Stackelberg stochastic differential game, and obtains the best incentive compensation as a linear function of the output value at the terminal time. The primal inspiration of our results are from Sannikov [26] and Cvitanic, Possamai, and Touzi [5], where the optimal incentive contract falls naturally in the more general class of path dependent functions of the output process. We

also refer to Elie, Mastrolia & Possamaï [8] for the extension to the multiple agents in Nash equilibrium context.

Our main results are first stated in the context where the diffusion is not controlled, a similar situation to the semilinear HJB equation (1.1) in the Markovian case. Under appropriate integrability conditions on the starting and target measures, we provide a complete characterization of the set of all solutions to the path-dependent planning problem in terms of a controlled auxiliary process. Our result then reduces the planning problem into an embedding problem, that is, to find a good controlled McKean-Vlasov dynamic satisfying the marginal constraints.

Technically, our approach is adapted from the contract theory literature, such as Cvitanic, Possamaï, and Touzi [5], Elie, Mastrolia, and Possamaï [8]. Nevertheless, it consists in a non-trivial adaptation as the drift coefficient of the controlled process is allowed to be unbounded in our setting. This corresponds to the Markovian case with quadratically growing Hamiltonians in terms of the gradient component in the literature of mean-field planning problem.

When only the drift of the state process is controlled, we exhibit an explicit planning solution which, up to some integrability requirement, coincides with the (unique) minimum entropy solution of the planning problem. When both the drift and the diffusion of the state process are controlled, the distributions of different controlled processes may not be equivalent, which brings some technical difficulties. By using similar quasi-sure analysis techniques as in the 2nd order BSDE theory by Soner, Touzi and Zhang [27], we are still able to provide a similar description of the set of all solutions of the planning problem in terms of a controlled auxiliary process. This would reduce the optimal planning problem to an optimal transport problem along controlled McKean-Vlasov dynamics.

A remarkable feature of the extension to the controlled diffusion setting is that it encompasses other classes of optimal transport problems, as for instance the martingale optimal transport of Beiglböck, Henry-Labordère, and Penkner [3] and Galichon, Henry-Labordère, and Touzi [10], and its connection to the Skorokhod Embedding problem (see Obłój [20] for a review), the martingale Benamou-Brenier problem in Huesmann and Trevisan [13], Backhoff-Veraguas, Beiglböck, Huesmann, and Källblad [2], and the semimartingale optimal transport problem in Mikami and Thieullen [17], Tan and Touzi [29], etc.

The rest of the paper is organized as follows. Section 2 provides our minimum entropy solution of the path-dependent planning problem in the purely quadratic setting. This is exactly the path-dependent analogue of (1.1). The extension to a larger class of drift control problems is reported in Section 3. Finally, Section 4 contains our results for the general case when both the drift and the diffusion coefficients are controlled.

**Notations.** Denote by  $\Omega = C([0, 1], \mathbb{R}^d)$  the canonical space of all  $\mathbb{R}^d$ -valued paths on  $[0, 1]$ , equipped with canonical filtration  $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, 1]}$  and canonical process  $X$ .

Let  $\mathcal{P}(\mathbb{R}^d)$  be the space of all (Borel) probability measures on  $\mathbb{R}^d$ , and denote by  $\mathbb{M}$  the collection of all flows of probability measure  $(m_t)_{t \in [0, 1]}$  with  $m_t \in \mathcal{P}(\mathbb{R}^d)$  for all  $t \in [0, 1]$ .

Throughout this paper, we fix some initial distribution  $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ , and we denote by  $\mathbb{P}_0$  the Wiener measure on  $\Omega$  with initial distribution  $\mu_0$ , i.e.,  $\mathbb{P}_0 \circ X_0^{-1} = \mu_0$  and the process  $(X_t - X_0)_{t \in [0, 1]}$  is a Brownian motion independent of  $X_0$  under  $\mathbb{P}_0$ .

Finally, for a probability measure  $\mathbb{P}$  on  $(\Omega, \mathcal{G})$ , we denote by  $\mathbb{L}^p(\mathbb{P}, \mathcal{G})$  the collection of all  $\mathcal{G}$ -measurable random variable with finite  $p$ -th moment, by  $\mathbb{H}^2(\mathbb{P})$  the collection of all progressively measurable processes  $Z$  such that  $\mathbb{E}^{\mathbb{P}}[\int_0^1 |Z_t|^2 dt] < \infty$ , and by  $\mathbb{H}_{\text{loc}}^2(\mathbb{P})$  the collection of all progressively measurable processes  $Z$  such that  $\int_0^1 |Z_t|^2 dt < \infty$ ,  $\mathbb{P}$ -a.s.

## 2 MFG planning problem: the linear quadratic setting

In this section, we introduce a path-dependent version of the Lions' MFG planning problem in the context of the simplest linear quadratic setting, and then provide a constructive solution to the planning problem.

### 2.1 The path-dependent linear-quadratic MFG problem

Recall that  $\mathbb{P}_0$  is the Wiener measure on the canonical space  $\Omega$  with initial distribution  $\mu_0$ . Let  $\mathcal{P}(\mu_0)$  denote the collection of all (Borel) probability measures  $\mathbb{P}$  on  $\Omega$  equivalent to  $\mathbb{P}_0$  with starting measure  $\mathbb{P} \circ X_0^{-1} = \mu_0$ . For an arbitrary  $\mathbb{P} \in \mathcal{P}(\mu_0)$ , we may find a unique process  $\beta^\mathbb{P} \in \mathbb{H}_{\text{loc}}^2(\mathbb{P}_0)$  such that the density of  $\mathbb{P}$  with respect to  $\mathbb{P}_0$  has a representation as the Doléans-Dade exponential

$$\frac{d\mathbb{P}}{d\mathbb{P}_0} = \mathcal{E}(\beta^\mathbb{P} \cdot X)_1 := \exp \left( \int_0^1 \beta_s^\mathbb{P} \cdot dX_s - \frac{1}{2} \int_0^1 |\beta_s^\mathbb{P}|^2 ds \right). \quad (2.1)$$

Indeed, it follows from [24, Proposition VIII.1.6] that the density process  $D_t^\mathbb{P} := \frac{d\mathbb{P}}{d\mathbb{P}_0} \Big|_{\mathcal{F}_t}$ , which is a strictly positive continuous martingale, can be represented as

$$D_t^\mathbb{P} = \mathcal{E}(L^\mathbb{P})_t := \exp \left( L_t^\mathbb{P} - \frac{1}{2} \langle L^\mathbb{P}, L^\mathbb{P} \rangle_t \right),$$

with a unique continuous local martingale  $L^\mathbb{P}$  satisfying  $L_0^\mathbb{P} = 0$ ,  $\mathbb{P}_0$ -a.s. By the predictable representation of the Brownian motion, see, e.g., [24, Theorem V.3.4], the local martingale  $L$  can be represented as a stochastic integral

$$L_t^\mathbb{P} = \int_0^t \beta_s^\mathbb{P} \cdot dX_s.$$

Therefore, the desired assertion follows.

It follows from the Girsanov theorem that the canonical process  $X$  satisfies the dynamics

$$X_t = X_0 + \int_0^t \beta_s^\mathbb{P} ds + W_t^\mathbb{P}, \quad \mathbb{P}\text{-a.s.}, \quad (2.2)$$

for some  $\mathbb{P}$ -Brownian motion  $W^\mathbb{P}$ . We further define the subspace

$$\mathcal{P}_2(\mu_0) := \left\{ \mathbb{P} \in \mathcal{P}(\mu_0) : \ln \left( \frac{d\mathbb{P}}{d\mathbb{P}_0} \right) \in \mathbb{L}^1(\mathbb{P}_0) \text{ and } \frac{d\mathbb{P}}{d\mathbb{P}_0} \in \mathbb{L}^2(\mathbb{P}_0) \right\}.$$

Let  $f : [0, 1] \times \Omega \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}$  be such that  $(t, \omega, \mu) \mapsto f_t(\omega, \mu)$  is  $\mathbb{F}$ -progressively measurable for every fixed  $\mu \in \mathcal{P}(\mathbb{R}^d)$  and

$$\mathbb{E}^\mathbb{P} \left[ \int_0^1 |f_t(m_t)| dt \right] < \infty, \quad \text{for all } \mathbb{P} \in \mathcal{P}_2(\mu_0) \text{ and } m \in \mathbb{M}.$$

Let  $\Xi$  define the set of all admissible (path-dependent and measurable) reward function  $\xi : \Omega \rightarrow \mathbb{R}$  such that  $\mathbb{E}^\mathbb{P}[\xi^+] < \infty$  for all  $\mathbb{P} \in \mathcal{P}_2(\mu_0)$ .

For  $\xi \in \Xi$  and  $m \in \mathbb{M}$  with  $m_0 = \mu_0$ , we consider the optimal control problem:

$$V_0(\xi, m) := \sup_{\mathbb{P} \in \mathcal{P}_2(\mu_0)} J(\xi, m, \mathbb{P}), \quad (2.3)$$

where

$$J(\xi, m, \mathbb{P}) := \mathbb{E}^\mathbb{P} \left[ \xi - \int_0^1 \left( \frac{1}{2} |\beta_s^\mathbb{P}|^2 + f_s(m_s) \right) ds \right].$$

**Remark 2.1.** An equivalent formulation of the last control problem is to introduce the set of admissible controls

$$\mathcal{U} := \left\{ \beta \in \mathbb{H}^2(\mathbb{P}_0) : \mathbb{E}^{\mathbb{P}_0} [\mathcal{E}(\beta \cdot X)] = 1 \text{ and } \mathcal{E}(\beta \cdot X) \in \mathbb{L}^2(\mathbb{P}_0) \right\}.$$

Then, each  $\beta \in \mathcal{U}$  induces a unique equivalent probability measure  $\mathbb{P}^\beta$  defined by the density  $\frac{d\mathbb{P}^\beta}{d\mathbb{P}_0} = \mathcal{E}(\beta \cdot X)$ , and we have therefore

$$V_0(\xi, m) := \sup_{\beta \in \mathcal{U}} J(\xi, m, \mathbb{P}^\beta).$$

See Lemma 2.4 below for the exact correspondence between  $\mathcal{P}_2(\mu_0)$  and  $\mathcal{U}$ .

**Definition 2.2** (Mean field game). *A probability measure  $\widehat{\mathbb{P}} \in \mathcal{P}_2(\mu_0)$  is a solution of the MFG with reward function  $\xi \in \Xi$  if*

$$V_0(\xi, m) = J(\xi, m, \widehat{\mathbb{P}}) \in \mathbb{R}, \text{ and } m_t := \widehat{\mathbb{P}} \circ X_t^{-1}, \text{ for all } t \in [0, 1].$$

We denote by  $\text{MFG}(\xi, \mu_0)$  the collection of all such solutions of the MFG problem.

Our main focus in this paper is on the following mean field game planning problem.

**Definition 2.3** (MFG planning). *An admissible reward function  $\xi \in \Xi$  is a solution to the MFG planning problem with starting and target distributions  $\mu_0, \mu_1 \in \mathcal{P}(\mathbb{R}^d)$  if*

$$\widehat{\mathbb{P}} \circ X_1^{-1} = \mu_1, \text{ for some } \widehat{\mathbb{P}} \in \text{MFG}(\xi, \mu_0).$$

We denote by  $\text{MFP}(\mu_0, \mu_1)$  the collection of all such solutions of the MFG planning problem.

## 2.2 Characterization of the solutions of mean field planning problem

In this section, we provide a characterization of all MFG planning solutions by using a decomposition induced by the dynamic programming principle. This characterization follows the idea of the representation of the agent problem in the so called Principal-Agent problem as introduced by Sannikov [26], and further extended in Cvitanic, Possamaï, and Touzi [5], and Elie, Mastrolia, and Possamaï [8]. Denote

$$\mathcal{P}_2(\mu_0, \mu_1) := \left\{ \mathbb{P} \in \mathcal{P}_2(\mu_0) : \mathbb{P} \circ X_1^{-1} = \mu_1 \right\}, \quad \mu_0, \mu_1 \in \mathcal{P}(\mathbb{R}^d).$$

**Lemma 2.4.** *For each  $\mathbb{P} \in \mathcal{P}_2(\mu_0)$ , we have  $\beta^\mathbb{P} \in \mathbb{H}^2(\mathbb{P}_0)$ .*

**Proof.** Set  $\zeta := \frac{d\mathbb{P}}{d\mathbb{P}_0}$ , and

$$Y_t := \ln \mathbb{E}^{\mathbb{P}_0}[\zeta | \mathcal{F}_t] = \int_0^t \beta_s^\mathbb{P} \cdot dX_s - \frac{1}{2} \int_0^t |\beta_s^\mathbb{P}|^2 ds, \quad t \in [0, 1], \quad \mathbb{P}_0\text{-a.s.}$$

By the Jensen inequality,  $Y_t \geq \mathbb{E}^{\mathbb{P}_0}[\ln \zeta | \mathcal{F}_t] \geq -\mathbb{E}^{\mathbb{P}_0}[(\ln \zeta)^- | \mathcal{F}_t] \geq -\mathbb{E}^{\mathbb{P}_0}[|\ln \zeta| | \mathcal{F}_t]$ . Then, introducing the stopping times  $\tau_n := 1 \wedge \inf \{ t > 0 : \mathbb{E}^{\mathbb{P}_0}[|\ln \zeta| | \mathcal{F}_t] > n, \int_0^t |\beta_s^\mathbb{P}|^2 ds > n \}$ ,  $n \in \mathbb{N}$ , it follows the tower property that

$$\mathbb{E}^{\mathbb{P}_0} \left[ \int_0^{\tau_n} |\beta_s^\mathbb{P}|^2 ds \right] \leq 2\mathbb{E}^{\mathbb{P}_0} \left[ \mathbb{E}^{\mathbb{P}_0}[|\ln \zeta| | \mathcal{F}_{\tau_n}] \right] = 2\mathbb{E}^{\mathbb{P}_0}[|\ln \zeta|] < \infty.$$

Since  $\tau_n \nearrow 1$  as  $n \rightarrow \infty$ , the assertion follows by the monotone convergence theorem. ■

**Theorem 2.5.** For all pair of starting and target measures  $(\mu_0, \mu_1) \in \mathcal{P}(\mathbb{R}^d) \times \mathcal{P}(\mathbb{R}^d)$ , we have:

$$\text{MFP}(\mu_0, \mu_1) = \mathbb{L}^1(\mathcal{F}_0, \mathbb{P}_0) + \left\{ \int_0^1 \beta_t^{\mathbb{P}} \cdot dX_t - \int_0^1 \left( \frac{1}{2} |\beta_t^{\mathbb{P}}|^2 - f_t(\mathbb{P} \circ X_t^{-1}) \right) dt : \mathbb{P} \in \mathcal{P}_2(\mu_0, \mu_1) \right\}.$$

**Proof.** “ $\supseteq$ ”: We first prove that  $\text{MFP}(\mu_0, \mu_1)$  contains the right hand side set. For arbitrary  $Y_0 \in \mathbb{L}^1(\mathcal{F}_0, \mathbb{P}_0)$  and  $\widehat{\mathbb{P}} \in \mathcal{P}_2(\mu_0, \mu_1)$ , denote  $m = (m_t)_{t \in [0,1]}$  with  $m_t := \widehat{\mathbb{P}} \circ X_t^{-1}$ , and

$$\xi := Y_0 + \int_0^1 \beta_t^{\widehat{\mathbb{P}}} \cdot dX_t - \int_0^1 \left( \frac{1}{2} |\beta_t^{\widehat{\mathbb{P}}}|^2 - f_t(m_t) \right) dt.$$

Let us verify that  $V_0(\xi, m) = J(\xi, m, \widehat{\mathbb{P}})$ . This would show that  $\widehat{\mathbb{P}} \in \text{MFG}(\xi, \mu_0)$  and therefore  $\xi \in \text{MFP}(\mu_0, \mu_1)$ .

We directly compute for all  $\mathbb{P} \in \mathcal{P}_2(\mu_0)$  that

$$\begin{aligned} J(\xi, m, \mathbb{P}) &= \mathbb{E}^{\mathbb{P}} \left[ \xi - \int_0^1 \left( \frac{1}{2} |\beta_s^{\mathbb{P}}|^2 + f_s(m_s) \right) ds \right] \\ &= \mathbb{E}^{\mathbb{P}_0} [Y_0] + \mathbb{E}^{\mathbb{P}} \left[ \int_0^1 \beta_s^{\widehat{\mathbb{P}}} \cdot (dW_s^{\mathbb{P}} + \beta_s^{\mathbb{P}} ds) - \int_0^1 \left( \frac{1}{2} |\beta_s^{\widehat{\mathbb{P}}}|^2 + \frac{1}{2} |\beta_s^{\mathbb{P}}|^2 \right) ds \right]. \end{aligned}$$

We next observe that the stochastic integral above is a true martingale under  $\mathbb{P}$ , i.e.,

$$\mathbb{E}^{\mathbb{P}} \left[ \int_0^1 \beta_s^{\widehat{\mathbb{P}}} \cdot dW_s^{\mathbb{P}} \right] = 0,$$

which is due to the following application of the Burkholder-Davis-Gundy inequality together with the Cauchy-Schwarz inequality, and  $\beta^{\widehat{\mathbb{P}}} \in \mathbb{H}(\mathbb{P}_0)$  by Lemma 2.4:

$$\begin{aligned} \mathbb{E}^{\mathbb{P}} \left[ \sup_{0 \leq t \leq 1} \left| \int_0^t \beta_s^{\widehat{\mathbb{P}}} \cdot dW_s^{\mathbb{P}} \right| \right] &\leq C_1 \mathbb{E}^{\mathbb{P}} \left[ \left( \int_0^1 |\beta_s^{\widehat{\mathbb{P}}}|^2 ds \right)^{\frac{1}{2}} \right] \\ &\leq C_1 \mathbb{E}^{\mathbb{P}_0} \left[ \left( \frac{d\mathbb{P}}{d\mathbb{P}_0} \right)^2 \right]^{\frac{1}{2}} \mathbb{E}^{\mathbb{P}_0} \left[ \int_0^1 |\beta_s^{\widehat{\mathbb{P}}}|^2 ds \right]^{\frac{1}{2}} < \infty. \end{aligned}$$

Then,

$$J(\xi, m, \mathbb{P}) = \mathbb{E}^{\mathbb{P}_0} [Y_0] - \frac{1}{2} \mathbb{E}^{\mathbb{P}} \left[ \int_0^1 |\beta_s^{\mathbb{P}} - \beta_s^{\widehat{\mathbb{P}}}|^2 ds \right],$$

so that  $J(\xi, m, \mathbb{P}) \leq \mathbb{E}^{\mathbb{P}_0} [Y_0]$  for all  $\mathbb{P} \in \mathcal{P}_2(\mu_0)$ , and  $J(\xi, m, \widehat{\mathbb{P}}) = \mathbb{E}^{\mathbb{P}_0} [Y_0]$ , as required.

“ $\subseteq$ ”: Let  $\xi \in \text{MFP}(\mu_0, \mu_1)$ , with a corresponding MFG solution  $\widehat{\mathbb{P}} \in \text{MFG}(\xi, \mu_0)$ , so that  $\widehat{\mathbb{P}} \in \mathcal{P}_2(\mu_0)$  is solution of the optimal control problem  $V_0(\xi, m)$  with  $m_t := \widehat{\mathbb{P}} \circ X_t^{-1}$  for all  $t \in [0, 1]$ . Then, it is clear that  $\widehat{\mathbb{P}} \in \mathcal{P}_2(\mu_0, \mu_1)$ . We aim to show that one can represent  $\xi$  as

$$\xi = V_0 + \int_0^1 \beta_t^{\widehat{\mathbb{P}}} \cdot dX_t - \int_0^1 \left( \frac{1}{2} |\beta_t^{\widehat{\mathbb{P}}}|^2 - f_t(m_t) \right) dt, \quad (2.4)$$

for some random variable  $V_0 \in \mathbb{L}^1(\mathcal{F}_0, \mathbb{P}_0)$ . To see this, we introduce the process

$$V_t := \text{ess sup}_{\mathbb{P} \in \mathcal{P}_2(\mu_0)} \mathbb{E}^{\mathbb{P}} \left[ \xi - \int_t^1 c_s^{\mathbb{P}} ds \mid \mathcal{F}_t \right], \quad \text{with } c_t^{\mathbb{P}} := \frac{1}{2} |\beta_t^{\mathbb{P}}|^2 + f_t(m_t), \quad t \in [0, 1].$$

Then, it is clear that  $\mathbb{E}^{\mathbb{P}_0}[V_0] = V_0(\xi, m) < \infty$ , so that  $V_0 \in \mathbb{L}^1(\mathcal{F}_0, \mathbb{P}_0)$ . Moreover, it follows by the dynamic programming principle (see e.g. Djete, Possamai and Tan [6, Definition 2.1, Remark 2.3 and Theorem 3.1]) that

$$V_t = \operatorname{ess\,sup}_{\mathbb{P} \in \mathcal{P}_2(\mu_0)} \mathbb{E}^{\mathbb{P}} \left[ V_u - \int_t^u c_s^{\mathbb{P}} ds \mid \mathcal{F}_t \right], \quad u \in [t, 1].$$

Moreover, it follows by

$$V_t - \int_0^t c_s^{\mathbb{P}} ds = \operatorname{ess\,sup}_{\mathbb{P} \in \mathcal{P}_2(\mu_0)} \mathbb{E}^{\mathbb{P}} \left[ V_u - \int_0^u c_s^{\mathbb{P}} ds \mid \mathcal{F}_t \right], \quad u \in [t, 1],$$

the following martingale optimal principle:

- For any  $\mathbb{P} \in \mathcal{P}_2(\mu_0)$ , the process  $\{V_t - \int_0^t c_s^{\mathbb{P}} ds\}_{t \in [0,1]}$  is  $\mathbb{P}$ -supermartingale. By the Doob-Meyer decomposition together with the predictable representation property of the Brownian motion, we have

$$V_t - \int_0^t c_s^{\mathbb{P}_0} ds = V_0 + \int_0^t Z_s \cdot dX_s - A_t^{\mathbb{P}_0}, \quad \mathbb{P}_0\text{-a.s.},$$

for some  $Z \in \mathbb{H}_{\text{loc}}^2(\mathbb{P}_0)$  and non-decreasing process  $A^{\mathbb{P}_0}$  starting from zero. By the change of measure from  $\mathbb{P}_0$  to  $\mathbb{P}$ , we have

$$V_t = \int_0^t c_s^{\mathbb{P}} ds + V_0 + \int_0^t Z_s \cdot dW_s^{\mathbb{P}} - A_t^{\mathbb{P}}, \quad \mathbb{P}\text{-a.s.},$$

with

$$A_t^{\mathbb{P}} = A_t^{\mathbb{P}_0} + \int_0^t (c_s^{\mathbb{P}} - c_s^{\mathbb{P}_0} - Z_s \cdot \beta_s^{\mathbb{P}}) ds.$$

Moreover, by uniqueness of the Doob-Meyer decomposition under each  $\mathbb{P}$ , the processes  $A^{\mathbb{P}}$  are also non-decreasing.

- The process  $\{V_t - \int_0^t c_s^{\widehat{\mathbb{P}}} ds\}_{t \in [0,1]}$  is a  $\widehat{\mathbb{P}}$ -martingale, i.e.,

$$0 = A_t^{\widehat{\mathbb{P}}} = A_t^{\mathbb{P}_0} + \int_0^t (c_s^{\widehat{\mathbb{P}}} - c_s^{\mathbb{P}_0} - Z_s \cdot \beta_s^{\widehat{\mathbb{P}}}) ds.$$

This shows that  $A^{\mathbb{P}_0}$  is absolutely continuous with respect to the Lebesgue measure, and provides the expression for the non-decreasing process  $A^{\mathbb{P}}$  which inherits the absolute continuity property with respect to the Lebesgue measure with density:

$$\frac{dA_t^{\mathbb{P}}}{dt} = c_t^{\mathbb{P}} + Z_t \cdot \beta_t^{\widehat{\mathbb{P}}} - c_t^{\widehat{\mathbb{P}}} - Z_t \cdot \beta_t^{\mathbb{P}} \geq 0, \quad \text{for all } \mathbb{P} \in \mathcal{P}_2(\mu_0).$$

In particular,  $\beta^{\widehat{\mathbb{P}}}$  is the maximizer of the optimization problem  $\max_{\beta^{\mathbb{P}}} (Z \cdot \beta^{\mathbb{P}} - c^{\mathbb{P}}) = \max_{\beta^{\mathbb{P}}} (Z \cdot \beta^{\mathbb{P}} - \frac{1}{2} |\beta^{\mathbb{P}}|^2 - f)$ . By the first-order condition, we obtain that  $\beta^{\widehat{\mathbb{P}}} = Z$ .

The previous analysis shows that  $\xi$  has the representation

$$\xi = V_1 = V_0 + \int_0^1 \left( \frac{1}{2} |\beta_s^{\widehat{\mathbb{P}}}|^2 + f_s(m_s) \right) ds + \int_0^1 \beta_s^{\widehat{\mathbb{P}}} \cdot dW_s^{\widehat{\mathbb{P}}},$$

which induces (2.4) by substituting  $dW_s^{\widehat{\mathbb{P}}} = dX_s - \beta_s^{\widehat{\mathbb{P}}} ds$ . ■



**Remark 2.6.** *Theorem 2.5 provides a systematic description of the class of all (possibly path-dependent) solutions of the MFG planning problem. Nevertheless, it can not be used to give a precise description of the Markovian solution, i.e.,  $\xi(X_\cdot) = u(T, X_T)$  for some function  $u$ . Indeed, to find some  $\beta^\mathbb{P}$  and initial condition  $u(0, X_0)$  such that*

$$u(0, X_0) + \int_0^1 \beta_t^\mathbb{P} \cdot dX_t - \int_0^1 \frac{1}{2} |\beta_t^\mathbb{P}|^2 dt = u(T, X_T), \quad \mathbb{P}\text{-a.s. (or equivalently } \mathbb{P}_0\text{-a.s.)},$$

one can formally use Itô's formula to identify that

$$\beta_t^\mathbb{P} = \nabla u(t, X_t), \quad \text{and} \quad \frac{1}{2} |\beta_t^\mathbb{P}|^2 = \partial_t u(t, X_t) + \frac{1}{2} \Delta u(t, X_t), \quad d\mathbb{P} \otimes dt\text{-a.e.}$$

This reduces exactly to Lions' original PDE formulation of the MFG planning problem in (1.1). On the other hand, an advantage of the description in Theorem 2.5 of all solutions allows the planner to choose an optimal one, see more discussions in Section 2.4.

### 2.3 A constructive solution to the mean field planning problem

Theorem 2.5 provides a characterization of all solutions of the MFG planning problem by means of the probability measures in  $\mathcal{P}_2(\mu_0, \mu_1)$ . We now use this characterization in order to derive an explicit construction of a particular solution.

Let  $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$  denote the set of probability measures on the product space  $\mathbb{R}^d \times \mathbb{R}^d$  and let  $\Pi(\mu_0, \mu_1)$  denote the set of all probability measures  $\pi \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$  with marginal distributions  $\mu_0$  and  $\mu_1$ , i.e.,  $\pi(dx, \mathbb{R}^d) = \mu_0(dx)$  and  $\pi(\mathbb{R}^d, dy) = \mu_1(dy)$ . We say that  $\pi$  is a coupling measure between the starting and target probability measures  $\mu_0$  and  $\mu_1$ . We also introduce a reference measure

$$\rho := \mathbb{P}_0 \circ (X_0, X_1)^{-1} \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d).$$

Let  $\pi \in \Pi(\mu_0, \mu_1)$  be some coupling measure equivalent to the reference measure  $\rho$ , and consider the corresponding density function  $\frac{d\pi}{d\rho}$  on  $\mathbb{R}^d \times \mathbb{R}^d$ . We define the following positive random variable on the canonical space  $\Omega$

$$\zeta := \frac{d\pi}{d\rho}(X_0, X_1),$$

and observe that

$$\mathbb{E}^{\mathbb{P}_0}[\zeta] = \mathbb{E}^{\mathbb{P}_0} \left[ \frac{d\pi}{d\rho}(X_0, X_1) \right] = \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{d\pi}{d\rho}(x_0, x_1) d\rho(x_0, x_1) = 1.$$

By the martingale representation theorem (see, e.g., [14, Theorem III.4.33]), there exists a  $\mathbb{F}$ -progressively measurable process  $\hat{\beta}$  such that

$$M_t := \mathbb{E}^{\mathbb{P}_0}[\zeta | \mathcal{F}_t] = M_0 \mathcal{E}(\hat{\beta} \cdot X)_t = M_0 \exp \left( \int_0^t \hat{\beta}_s \cdot dX_s - \frac{1}{2} \int_0^t |\hat{\beta}_s|^2 ds \right).$$

In particular, as  $\pi(dx, \mathbb{R}^d) = \rho(dx, \mathbb{R}^d) = \mu_0(dx)$ , we have  $M_0 = 1$ ,  $\mathbb{P}_0$ -a.s.

Before stating the main result of this section, we recall the notion of entropy of a probability  $\mathbb{Q}_1$  with respect to a reference probability  $\mathbb{Q}_0$ :

$$\mathbf{H}(\mathbb{Q}_1 | \mathbb{Q}_0) := \begin{cases} \mathbb{E}^{\mathbb{Q}_1} \left[ \ln \left( \frac{d\mathbb{Q}_1}{d\mathbb{Q}_0} \right) \right] = \int_{\Omega} \ln \left( \frac{d\mathbb{Q}_1}{d\mathbb{Q}_0} \right) d\mathbb{Q}_1, & \text{whenever } \mathbb{Q}_1 \ll \mathbb{Q}_0, \\ \infty, & \text{otherwise,} \end{cases}$$

where  $\mathbb{Q}_1 \ll \mathbb{Q}_0$  means that  $\mathbb{Q}_1$  is absolutely continuous with respect to  $\mathbb{Q}_0$ .

**Proposition 2.7.** Let  $\pi \in \Pi(\mu_0, \mu_1)$  be equivalent to  $\rho$ , such that the random variable  $\zeta := \frac{d\pi}{d\rho}(X_0, X_1)$  satisfies

$$\mathbb{E}^{\mathbb{P}_0} [|\ln \zeta| + \zeta^2] < \infty.$$

Then, the probability measure  $\widehat{\mathbb{P}}$  defined by  $\frac{d\widehat{\mathbb{P}}}{d\mathbb{P}_0} = \zeta$  is an element in  $\mathcal{P}_2(\mu_0, \mu_1)$ . Moreover,  $\widehat{\mathbb{P}}$  is the unique minimizer of  $\mathbf{H}(\cdot | \mathbb{P}_0)$  on  $\mathcal{P}_\pi := \{\mathbb{P} \in \mathcal{P}_2(\mu_0, \mu_1) : \mathbb{P} \circ (X_0, X_1)^{-1} = \pi\}$ .

**Proof.** (i). First, by its definition and the transformation formula, it is clear that  $\widehat{\mathbb{P}} \circ X_0^{-1} = \mu_0$  and  $\widehat{\mathbb{P}} \circ X_1^{-1} = \mu_1$ . Indeed, let  $B \in \mathcal{B}(\mathbb{R}^d)$ , i.e., Borel-measurable subset in  $\mathbb{R}^d$ , be arbitrary. Then,

$$\begin{aligned} \widehat{\mathbb{P}} \circ X_0^{-1}(B) &= \mathbb{E}^{\widehat{\mathbb{P}}} [\mathbf{1}_{X_0^{-1}(B)}] = \mathbb{E}^{\mathbb{P}_0} \left[ \frac{d\pi}{d\rho}(X_0, X_1) \mathbf{1}_{X_0^{-1}(B)} \right] \\ &= \mathbb{E}^{\mathbb{P}_0} \left[ \frac{d\pi}{d\rho}(X_0, X_1) \mathbf{1}_{X_0^{-1}(B)} \mathbf{1}_{X_1^{-1}(\mathbb{R}^d)} \right] \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{d\pi}{d\rho}(x_0, x_1) \mathbf{1}_{B \times \mathbb{R}^d}(x_0, x_1) \rho(x_0, x_1) \\ &= \pi(B \times \mathbb{R}^d) = \mu_0(B). \end{aligned}$$

Similarly, one can prove that  $\widehat{\mathbb{P}} \circ X_1^{-1} = \mu_1$ . Moreover, by the integrability assumption on  $\zeta$ , we have

$$\mathbb{E}^{\mathbb{P}_0} \left[ \left| \ln \left( \frac{d\widehat{\mathbb{P}}}{d\mathbb{P}_0} \right) \right| + \left( \frac{d\widehat{\mathbb{P}}}{d\mathbb{P}_0} \right)^2 \right] = \mathbb{E}^{\mathbb{P}_0} [|\ln \zeta| + \zeta^2] < \infty,$$

and therefore  $\widehat{\mathbb{P}} \in \mathcal{P}_2(\mu_0, \mu_1)$ .

(ii). Let us denote by  $K^{\mathbb{P}}(\cdot; x_0, x_1)$  the kernel function of  $\mathbb{P}$  conditional on  $(X_0, X_1) = (x_0, x_1)$ , for any  $\mathbb{P} \in \mathcal{P}$ . We observe from the definition of  $\widehat{\mathbb{P}}$  and Bayes formula that for any  $B \in \mathcal{B}(\Omega)$

$$\begin{aligned} K^{\widehat{\mathbb{P}}}(B; x_0, x_1) &= \mathbb{E}^{\widehat{\mathbb{P}}} [\mathbf{1}_{\{X \in B\}} | X_0 = x_0, X_1 = x_1] \\ &= \frac{\mathbb{E}^{\mathbb{P}_0} \left[ \frac{d\pi}{d\rho}(X_0, X_1) \mathbf{1}_{\{X \in B\}} | X_0 = x_0, X_1 = x_1 \right]}{\mathbb{E}^{\mathbb{P}_0} \left[ \frac{d\pi}{d\rho}(X_0, X_1) | X_0 = x_0, X_1 = x_1 \right]} \\ &= \frac{\frac{d\pi}{d\rho}(x_0, x_1) \mathbb{E}^{\mathbb{P}_0} [\mathbf{1}_{\{X \in B\}} | X_0 = x_0, X_1 = x_1]}{\frac{d\pi}{d\rho}(x_0, x_1)} \\ &= \mathbb{E}^{\mathbb{P}_0} [\mathbf{1}_{\{X \in B\}} | X_0 = x_0, X_1 = x_1] = K^{\mathbb{P}_0}(B; x_0, x_1), \end{aligned}$$

as  $\frac{d\pi}{d\rho}$  is strictly positive due to the equivalence. Therefore,

$$K^{\widehat{\mathbb{P}}}(\cdot; x_0, x_1) = K^{\mathbb{P}_0}(\cdot; x_0, x_1), \quad \text{for } \pi\text{-a.e. } (x_0, x_1) \in \mathbb{R}^d \times \mathbb{R}^d.$$

Denote by  $C([0, 1])|_{x_0, x_1}$  the set of all continuous functions  $\omega$  on  $[0, 1]$  with  $\omega(0) = x_0$  and  $\omega(1) = x_1$ . Denote  $\omega^{(-1)} := \omega|_{(0,1)}$  for  $\omega \in C([0, 1])$ . Further, for any  $\mathbb{P} \in \mathcal{P}_\pi$ , one has

$$\mathbb{P}(d\omega) = K^{\mathbb{P}}(d\omega^{(-1)}; x_0, x_1) \pi(dx_0, dx_1),$$

and

$$\mathbb{P}_0(d\omega) = K^{\mathbb{P}_0}(d\omega^{(-1)}; x_0, x_1) \rho(dx_0, dx_1).$$

By disintegration theorem and non-negativity of the entropy, we have

$$\begin{aligned}
& \mathbf{H}(\mathbb{P}|\mathbb{P}_0) \\
&= \int_{\Omega} \ln \left( \frac{d\mathbb{P}}{d\mathbb{P}_0} \right) d\mathbb{P} \\
&= \int_{\mathbb{R}^d \times \mathbb{R}^d} \int_{C([0,1])|_{x_0, x_1}} \ln \left( \frac{K^{\mathbb{P}}(d\omega^{(-1)}; x_0, x_1) \pi(dx_0, dx_1)}{K^{\mathbb{P}_0}(d\omega^{(-1)}; x_0, x_1) \rho(dx_0, dx_1)} \right) K^{\mathbb{P}}(d\omega^{(-1)}; x_0, x_1) \pi(dx_0, dx_1) \\
&= \int_{\mathbb{R}^d \times \mathbb{R}^d} \int_{C([0,1])|_{x_0, x_1}} \ln \left( \frac{K^{\mathbb{P}}(d\omega^{(-1)}; x_0, x_1)}{K^{\mathbb{P}_0}(d\omega^{(-1)}; x_0, x_1)} \right) K^{\mathbb{P}}(d\omega^{(-1)}; x_0, x_1) \pi(dx_0, dx_1) \\
&\quad + \int_{\mathbb{R}^d \times \mathbb{R}^d} \int_{C([0,1])|_{x_0, x_1}} \ln \left( \frac{\pi(dx_0, dx_1)}{\rho(dx_0, dx_1)} \right) K^{\mathbb{P}}(d\omega^{(-1)}; x_0, x_1) \pi(dx_0, dx_1) \\
&= \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathbf{H} \left( K^{\mathbb{P}}(\cdot; x_0, x_1) \middle| K^{\mathbb{P}_0}(\cdot; x_0, x_1) \right) \pi(dx_0, dx_1) + \mathbf{H}(\pi|\rho).
\end{aligned}$$

It follows that  $\mathbf{H}(\widehat{\mathbb{P}}|\mathbb{P}_0) = \mathbf{H}(\pi|\rho) \leq \mathbf{H}(\mathbb{P}|\mathbb{P}_0)$  for all  $\mathbb{P} \in \mathcal{P}_{\pi}$ .

(iii). Finally, the uniqueness follows directly from the strict convexity of  $\mathbb{P} \mapsto \mathbf{H}(\mathbb{P}|\mathbb{P}_0)$ .  $\blacksquare$

## 2.4 Entropic MFG planning, and further optimal MFG planning solutions

Proposition 2.7 reduces the problem of minimum entropy MFG planning to the standard static Schrödinger bridge problem, i.e., minimize the entropy  $\mathbf{H}(\pi|\rho)$  among the set  $\Pi(\mu_0, \mu_1)$  of all joint measures with marginals  $\mu_0$  and  $\mu_1$  (see e.g. the lecture note of Nutz [19]). However, due to the integrability requirements in Proposition 2.7, we need to restrict this set of coupling measures to the following subset

$$\Pi_2(\mu_0, \mu_1) := \left\{ \pi \in \Pi(\mu_0, \mu_1) : \mathbb{E}^{\mathbb{P}_0} \left[ \left| \ln \left( \frac{d\pi}{d\rho}(X_0, X_1) \right) \right| \right] + \mathbb{E}^{\mathbb{P}_0} \left[ \left( \frac{d\pi}{d\rho}(X_0, X_1) \right)^2 \right] < \infty \right\}.$$

This set  $\Pi_2(\mu_0, \mu_1)$  is convex, but fails to be closed so that the Schrödinger bridge problem may not have a solution in  $\Pi_2(\mu_0, \mu_1)$ . If the solution of the Schrödinger bridge problem happens to satisfy the required integrability conditions, then our construction in Proposition 2.7 provides the minimum entropy solution of the MFG planning problem. This is stated in the following Corollary which is an immediate consequence of our previous results.

**Corollary 2.8.** *Let  $\rho = \mathbb{P}_0 \circ (X_0, X_1)^{-1}$  be the reference measure on  $\mathbb{R}^d \times \mathbb{R}^d$ , and assume that the Schrödinger bridge problem  $\min_{\pi \in \Pi(\mu_0, \mu_1)} \mathbf{H}(\pi|\rho)$  has a unique solution  $\pi^* \in \Pi_2(\mu_0, \mu_1)$ . Then, the probability measure  $\widehat{\mathbb{P}}$  defined by*

$$\frac{d\widehat{\mathbb{P}}}{d\mathbb{P}_0} = \frac{d\pi^*}{d\rho}(X_0, X_1)$$

*is the unique minimizer of  $\mathbf{H}(\cdot|\mathbb{P}_0)$  on  $\mathcal{P}_2(\mu_0, \mu_1)$ .*

As is standard in the literature, in particular for the Schrödinger bridge problem, the relative entropy  $H(\mathbb{P}|\mathbb{P}_0)$  can be considered as a distance between the measure  $\mathbb{P}$  and the reference measure  $\mathbb{P}_0$ . The above result implies that the corresponding solution induces a population distribution for the output process with smallest departure from the reference distribution  $\mathbb{P}_0$ , in terms of the entropy.

We conclude this section by a formal discussion on the selection among optimal planning solutions. Given our characterization of all solutions to the MFG planning problem in Theorem 2.5, Corollary 2.8 selects a solution of the MFG planning which has minimum entropy with respect to the Wiener measure. One may consider other optimization criteria which can be seen as the planner problem whose task is to implement the optimal solution of the MFG planning problem in view of some collective objective. Except for the constraint on the target distribution of the population, this point of view is close to the spirit of contract theory in the economics literature, where the planner, called principal, faces a population of agents in Nash equilibrium, see Sannikov [26] and Cvitanović, Possamaï and Touzi [5] for the one-agent setting, and Elie, Mastrolia, and Possamaï [8] for the corresponding MFG problem.

### 3 MFG planning problem under uncontrolled diffusion

In this section, we show that the mean field planning solution of the linear quadratic MFG, as derived in the previous section, can be adapted to a general class of nonlinear MFG problems whose corresponding Hamiltonian has quadratic growth in the gradient.

#### 3.1 Formulation of the mean field planning problem

Throughout this section,  $U$  is a given closed subset of  $\mathbb{R}^d$ , and we denote by  $\mathcal{P}_2^U(\mu_0)$  the subset of all measures  $\mathbb{P} \in \mathcal{P}_2(\mu_0)$  such that  $\beta^\mathbb{P} \in U$ , Leb $\otimes\mathbb{P}$ -a.s.

Let  $c : [0, 1] \times \Omega \times U \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}$  be an  $\mathbb{F}$ -progressively measurable map with

$$\mathbb{E}^\mathbb{P} \left[ \int_0^1 |c_s(\beta_s^\mathbb{P}, m_s)| ds \right] < \infty, \quad \text{for all } m \in \mathbb{M}, \mathbb{P} \in \mathcal{P}_2^U(\mu_0).$$

Similarly, we introduce the subset  $\Xi^U$  of all measurable reward functions  $\xi \in \Xi$  such that  $\mathbb{E}^\mathbb{P}[\xi^+] < \infty$  for all  $\mathbb{P} \in \mathcal{P}_2^U(\mu_0)$ .

For all  $m \in \mathbb{M}$  and  $\xi \in \Xi^U$ , we consider the control problem

$$V_0(\xi, m) := \sup_{\mathbb{P} \in \mathcal{P}_2^U(\mu_0)} J(\xi, m, \mathbb{P}), \quad \text{where } J(\xi, m, \mathbb{P}) := \mathbb{E}^\mathbb{P} \left[ \xi - \int_0^1 c_s(\beta_s^\mathbb{P}, m_s) ds \right]. \quad (3.1)$$

The notions of mean field game and mean field planning are defined as in Definitions 2.2 and 2.3, up to the substitution of  $\mathcal{P}_2$  and  $\Xi$  by  $\mathcal{P}_2^U$  and  $\Xi^U$ .

#### 3.2 Characterization of the solutions of the mean field planning problem

We introduce the Hamiltonian  $H$  defined on  $[0, 1] \times \Omega \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$  by

$$H_s(z, m) := H_s(\omega, z, m) := \sup_{b \in U} \{b \cdot z - c_s(\omega, b, m)\}. \quad (3.2)$$

This defines a convex map in  $z$ . The following Assumption 3.1 guarantees that it is finite, so that the supremum in (3.2) is attained at any point of the partial sub-gradient  $\partial_z H_s(z, m)$  of the convex function  $H$  in  $z$ , i.e.,

$$\partial_z H_s(z, m) := \left\{ y \in \mathbb{R}^d : H_s(z', m) - H_s(z, m) \geq y \cdot (z' - z), \forall z' \in \mathbb{R}^d \right\}.$$

Our main result holds on the following condition which restricts the Hamiltonian to have quadratic growth on terms of the gradient.

**Assumption 3.1.** *The Hamiltonian satisfies the quadratic growth condition:*

$$\operatorname{ess\,inf}_{(s,m) \in [0,1] \times \mathcal{P}(\mathbb{R}^d)} \min_{(s,m) \in [0,1] \times \mathcal{P}(\mathbb{R}^d)} |\partial_z H_s(z, m)| \geq C_1 |z| - C_2, \quad \text{for all } z \in \mathbb{R}^d,$$

for some constants  $C_1, C_2 > 0$ .

Let  $\widehat{b} : [0, 1] \times \Omega \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}^d$  be a measurable function such that  $\widehat{b}_s(\omega, z, m) \in \partial_z H_s(\omega, z, m)$  for all  $(s, \omega, z, m) \in [0, 1] \times \Omega \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$ , and  $Z \in \mathbb{H}^2(\mathbb{P}_0)$  be a control process, we next consider the controlled McKean-Vlasov SDE

$$X_t = X_0 + \int_0^t \widehat{b}_s(Z_s, \mathbb{P} \circ X_s^{-1}) ds + W_t^\mathbb{P}, \quad t \in [0, 1], \quad \mathbb{P}\text{-a.s.}, \quad (3.3)$$

where a solution is a probability  $\mathbb{P} \in \mathcal{P}(\Omega)$  on the canonical space  $\Omega$ , such that for some  $\mathbb{P}$ -Brownian motion  $W^\mathbb{P}$  the equality (3.3) holds. Further, let us denote

$$\begin{aligned} \text{MKV}(\mu_0, \mu_1) &:= \{(Z, \mathbb{P}) \in \mathbb{H}^2(\mathbb{P}_0) \times \mathcal{P}_2^U(\mu_0, \mu_1) : \mathbb{P} \text{ solution of (3.3)}\}, \\ \Xi(\mu_0, \mu_1) &:= \mathbb{L}^1(\mathcal{F}_0, \mathbb{P}_0) + \{Y_1^Z : (Z, \mathbb{P}) \in \text{MKV}(\mu_0, \mu_1)\}, \end{aligned}$$

with

$$Y_t^Z := \int_0^t Z_s \cdot dX_s - \int_0^t H_s(Z_s, \mathbb{P} \circ X_s^{-1}) ds, \quad t \in [0, 1].$$

**Theorem 3.2.** *For all pairs of starting and target measures  $(\mu_0, \mu_1) \in \mathcal{P}(\mathbb{R}^d) \times \mathcal{P}(\mathbb{R}^d)$ , we have  $\Xi(\mu_0, \mu_1) \subseteq \text{MFP}(\mu_0, \mu_1)$ .*

*Moreover, the equality  $\Xi(\mu_0, \mu_1) = \text{MFP}(\mu_0, \mu_1)$  holds under Assumption 3.1.*

**Proof.** (i). Let  $\xi := Y_0 + Y_1^Z$  be an arbitrary element in  $\Xi(\mu_0, \mu_1)$  with corresponding  $(Z, \widehat{\mathbb{P}}) \in \text{MKV}(\mu_0, \mu_1)$ , and denote  $m_s := \widehat{\mathbb{P}} \circ X_s^{-1}$ . To show that  $\xi \in \text{MFP}(\mu_0, \mu_1)$ , it is enough to show that  $\xi \in \Xi^U$ , and moreover,  $\widehat{\mathbb{P}}$  is a solution of the optimization problem  $V_0(\xi, m)$  in (3.1), so that  $\widehat{\mathbb{P}} \in \text{MFG}(\xi)$ .

For an arbitrary  $\mathbb{P} \in \mathcal{P}_2^U(\mu_0)$ , we first check that, by the Burkholder-Davis-Gundy inequality and Cauchy-Schwarz inequality,

$$\mathbb{E}^\mathbb{P} \left[ \sup_{0 \leq t \leq 1} \left| \int_0^t Z_s \cdot dW_s^\mathbb{P} \right| \right] \leq C_1 \mathbb{E}^{\mathbb{P}_0} \left[ \left( \frac{d\mathbb{P}}{d\mathbb{P}_0} \right)^2 \right]^{1/2} \mathbb{E}^{\mathbb{P}_0} \left[ \int_0^1 |Z_s|^2 ds \right]^{1/2} < \infty.$$

Then, the stochastic integral  $(Z \cdot W^\mathbb{P})$  is a true martingale and therefore  $\mathbb{E}^\mathbb{P}[(Z \cdot W^\mathbb{P})_1] = 0$ . We compute that

$$\begin{aligned} J(\xi, m, \mathbb{P}) &= \mathbb{E}^\mathbb{P} \left[ \xi - \int_0^1 c_s(\beta_s^\mathbb{P}, m_s) ds \right] \\ &= \mathbb{E}^{\mathbb{P}_0} [Y_0] + \mathbb{E}^\mathbb{P} \left[ \int_0^1 Z_s \cdot (dW_s^\mathbb{P} + \beta_s^\mathbb{P} ds) - \int_0^1 H_s(Z_s, m_s) ds - \int_0^1 c_s(\beta_s^\mathbb{P}, m_s) ds \right] \\ &= \mathbb{E}^{\mathbb{P}_0} [Y_0] + \mathbb{E}^\mathbb{P} \left[ \int_0^1 (Z_s \cdot \beta_s^\mathbb{P} - c_s(\beta_s^\mathbb{P}, m_s) - H_s(Z_s, m_s)) ds \right]. \end{aligned}$$

By the definition of the Hamiltonian  $H$ , it follows that  $J(\xi, m, \mathbb{P}) \leq \mathbb{E}^{\mathbb{P}_0} [Y_0]$  for all  $\mathbb{P} \in \mathcal{P}_2^U(\mu_0)$ . As  $\widehat{\mathbb{P}}$  is the solution to (3.3) and

$$\beta_s^{\widehat{\mathbb{P}}} = \widehat{b}(Z_s, m_s) \in \partial_z H_s(Z_s, m_s),$$

it deduces by [25, Theorem 23.5] that  $\beta_s^{\widehat{\mathbb{P}}}$  is the optimizer of the Hamiltonian, hence

$$Z_s \cdot \beta_s^{\widehat{\mathbb{P}}} - c_s(\beta_s^{\widehat{\mathbb{P}}}, m_s) = H_s(Z_s, m_s), \quad \widehat{\mathbb{P}}\text{-a.s.}$$

Therefore,  $J(\xi, m, \widehat{\mathbb{P}}) = \mathbb{E}^{\mathbb{P}_0}[Y_0]$ .

(ii). Under the additional conditions in Assumption 3.1, we consider  $\xi \in \text{MFP}(\mu_0, \mu_1)$ , together with  $\widehat{\mathbb{P}} \in \text{MFG}(\xi, \mu_0)$  such that  $\widehat{\mathbb{P}} \circ X_1^{-1} = \mu_1$  so that  $J(\xi, m, \widehat{\mathbb{P}}) = V(\xi, m)$ , for  $m_s := \widehat{\mathbb{P}} \circ X_s^{-1}$ ,  $s \in [0, 1]$ . Let us define

$$V_t := \text{ess sup}_{\mathbb{P} \in \mathcal{P}^U(\mu_0)} \mathbb{E}^{\mathbb{P}} \left[ \xi - \int_t^1 c_s(\beta_s^{\mathbb{P}}, m_s) ds \middle| \mathcal{F}_t \right], \quad t \in [0, 1].$$

Then,  $\mathbb{E}^{\mathbb{P}_0}[V_0] = V_0(\xi, m) \in \mathbb{R}$ , so that  $V_0 \in \mathbb{L}^1(\mathcal{F}_0, \mathbb{P}_0)$ . Moreover, by the dynamic programming principle, we argue as in Step (ii) of the proof of Theorem 2.5 to show the existence of some  $Z \in \mathbb{H}_{\text{loc}}^2(\mathbb{P}_0)$  such that

$$V_t = V_0 + \int_0^t Z_s \cdot dX_s - \int_0^t (Z_s \cdot \beta_s^{\widehat{\mathbb{P}}} - c_s^{\widehat{\mathbb{P}}}) ds,$$

and

$$Z_t \cdot \beta_t^{\widehat{\mathbb{P}}} - c_t^{\widehat{\mathbb{P}}} = \max_{\mathbb{P} \in \mathcal{P}_2^U(\mu_0)} \left\{ Z_t \cdot \beta_t^{\mathbb{P}} - c_t^{\mathbb{P}} \right\} = H_t(Z_t, m_t).$$

Moreover, since  $\beta^{\mathbb{P}} \in \mathbb{H}^2(\mathbb{P}_0)$  by the definition of  $\mathcal{P}_2^U(\mu_0)$ , it follows by Assumption 3.1 that  $Z \in \mathbb{H}^2(\mathbb{P}_0)$ . This concludes the proof that  $(Z, \widehat{\mathbb{P}}) \in \text{MKV}(\mu_0, \mu_1)$ , and hence  $\xi \in \Xi(\mu_0, \mu_1)$ .  $\blacksquare$

### 3.3 Existence of solution to the mean field planning problem

Under Assumption 3.1, the last theorem reduces the construction of a solution of the MFG planning problem to the construction of a solution of the McKean-Vlasov SDE (3.3) with given starting and target marginals. To do this, we adapt the same arguments as in Section 2.3 under the following additional condition.

**Assumption 3.3.** *The Hamiltonian satisfies the full range condition  $\partial_z H_t(\omega, \mathbb{R}^d, m) = \mathbb{R}^d$  for all  $(t, \omega, m) \in [0, 1] \times \Omega \times \mathcal{P}(\mathbb{R}^d)$ .*

The condition in Assumption 3.3 ensures that for any  $b \in \mathbb{R}^d$ , there exists  $z \in \mathbb{R}^d$  such that  $b \in \partial_z H_t(\omega, z, m)$  (or equivalently  $b$  is an optimizer in the definition of  $H$  in (3.2)). As application, let us consider a probability measure  $\pi \in \Pi(\mu_0, \mu_1)$  (i.e. a measure on  $\mathbb{R}^d \times \mathbb{R}^d$  with marginals  $\mu_0$  and  $\mu_1$ ), assume that  $\pi$  is equivalent to the reference measure  $\rho := \mathbb{P}_0 \circ (X_0, X_1)^{-1}$  so that one can define  $\zeta := \frac{d\pi}{d\rho}(X_0, X_1)$ . Recall the Doléans-Dade exponential  $\mathcal{E}(\cdot)$  defined in (2.1), we can then choose a progressively measurable process  $Z$  satisfying

$$\widehat{\beta}_s \in \partial_z H_s(Z_s, m_s), \quad \text{Leb} \otimes \mathbb{P}_0\text{-a.s.}, \quad \text{with } \widehat{\beta} \text{ defined by } \zeta = \mathcal{E}(\widehat{\beta} \cdot X)_1. \quad (3.4)$$

Then following the same argument as in Proposition 2.7, one can prove that this provides a solution of the MFG planning problem.

**Proposition 3.4.** *Let Assumptions 3.1 and 3.3 hold true, and suppose in addition that there exists  $\pi \in \Pi(\mu_0, \mu_1)$  equivalent to the reference measure  $\rho = \mathbb{P}_0 \circ (X_0, X_1)^{-1}$  such that the density  $\zeta := \frac{d\pi}{d\rho}(X_0, X_1)$  satisfies  $\mathbb{E}^{\mathbb{P}_0}[|\ln \zeta| + \zeta^2] < \infty$ . Define the measure  $\widehat{\mathbb{P}}$  equivalent to  $\mathbb{P}_0$  by  $\frac{d\widehat{\mathbb{P}}}{d\mathbb{P}_0} = \zeta$  and let  $Z$  be as defined in (3.4).*

*Then, the pair  $(Z, \widehat{\mathbb{P}}) \in \text{MKV}(\mu_0, \mu_1)$ , and consequently  $Y_1^Z \in \text{MFP}(\mu_0, \mu_1)$ . Moreover,  $\widehat{\mathbb{P}}$  is the unique minimizer of  $\mathbf{H}(\cdot | \mathbb{P}_0)$  on  $\mathcal{P}_\pi := \{\mathbb{P} \in \mathcal{P}_2(\mu_0, \mu_1) : \mathbb{P} \circ (X_0, X_1)^{-1} = \pi\}$ .*

**Proof.** As in Proposition 2.7, we have  $\widehat{\mathbb{P}} \circ X_0^{-1} = \mu_0$ ,  $\widehat{\mathbb{P}} \circ X_1^{-1} = \mu_1$ . By Lemma 2.4, we obtain  $\beta^{\widehat{\mathbb{P}}} \in \mathbb{H}^2(\mathbb{P}_0)$ . Due to Assumptions 3.1 and 3.3, the process  $Z$  defined in (3.4) with  $m_s = \widehat{\mathbb{P}} \circ X_s^{-1}$  satisfies  $Z \in \mathbb{H}^2(\mathbb{P}_0)$ . Finally, our construction immediately yields  $\widehat{\mathbb{P}} \in \mathcal{P}_2^U(\mu_0)$  and

$$X_t = X_0 + \int_0^t \widehat{b}_s(Z_s, \widehat{\mathbb{P}} \circ X_s^{-1}) ds + W_t^{\widehat{\mathbb{P}}}, \quad \widehat{\mathbb{P}}\text{-a.s.},$$

for some measurable selection  $\widehat{b} \in \partial_z H$ ,  $\text{Leb} \otimes \widehat{\mathbb{P}}$ -a.s. Therefore,  $(Z, \widehat{\mathbb{P}}) \in \text{MKV}(\mu_0, \mu_1)$  and  $Y_1^Z \in \text{MFP}(\mu_0, \mu_1)$ .

Following the proof of Proposition 2.7, we obtain the entropy minimality.  $\blacksquare$

**Remark 3.5.** The discussion of Subsection 2.4 fully applies to the present setting. Therefore, under the appropriate integrability condition on the solution of the static Schrödinger bridge problem, Proposition 3.4 provides the minimum entropy solution of the MFG planning problem.

## 4 MFG planning problem under controlled diffusion

### 4.1 Formulation of the mean field planning problem

Let  $\mathbb{S}^d$  denote the space of all symmetric matrices, and  $\mathbb{S}_+^d$  the subspace of all positive semidefinite symmetric matrices. Let  $\mathcal{P}$  denote the collection of all probability measures  $\mathbb{P}$  on the canonical space  $\Omega$ , under which the canonical process  $X$  is a diffusion process with the following decomposition

$$X_t = X_0 + \int_0^t \widehat{b}_s^{\mathbb{P}} ds + \int_0^t \widehat{\sigma}_s dW_s^{\mathbb{P}}, \quad t \in [0, 1], \quad \mathbb{P}\text{-a.s.},$$

for some  $\mathbb{P}$ -Brownian motion  $W^{\mathbb{P}}$ . We recall from Karandikar [15, Theorem 3 and below] that the quadratic variation process  $\langle X \rangle$  can be defined independently of  $\mathbb{P} \in \mathcal{P}$ , so that  $\widehat{\sigma}_t$  can be defined as the unique square root matrix of  $\widehat{\sigma}_t^2$  in  $\mathbb{S}_+^d$ , with

$$\widehat{\sigma}_t^2 := \lim_{\varepsilon \searrow 0} \frac{\langle X \rangle_t - \langle X \rangle_{(t-\varepsilon) \vee 0}}{\varepsilon}, \quad t \in [0, 1].$$

Let  $U$  be a closed convex subset of  $\mathbb{R}^d \times \mathbb{S}_+^d$ , with the given two marginal distributions  $\mu_0$  and  $\mu_1$ , we will introduce the set  $\mathcal{P}^U(\mu_0)$  in two different settings:

- Setting 1: let

$$\mathcal{P}^U(\mu_0) := \left\{ \mathbb{P} \in \mathcal{P} : \mathbb{P} \circ X_0^{-1} = \mu_0 \text{ and } \left( \widehat{b}_s^{\mathbb{P}}, \frac{1}{2} \widehat{\sigma}_s^2 \right) \in U, \text{ Leb} \otimes \mathbb{P}\text{-a.e.} \right\}. \quad (4.1)$$

- Setting 2: let  $U$  satisfy

$$(b, a) \in U \implies (0, a) \in U \text{ and } b = a^{1/2} \beta \text{ for some } \beta \in \mathbb{R}^d, \quad (4.2)$$

and define

$$\mathcal{Q}(\mu_0) := \{ \mathbb{Q} \in \mathcal{P} : \mathbb{Q} \circ X_0^{-1} = \mu_0, X \text{ is a } \mathbb{Q}\text{-martingale} \},$$

and

$$\mathcal{P}^U(\mu_0) := \left\{ \mathbb{P} \in \mathcal{P} : \frac{d\mathbb{P}}{d\mathbb{Q}} = \mathcal{E} \left( \int_0^\cdot \beta_s^{\mathbb{P}} \cdot dW_s^{\mathbb{Q}} \right), \text{ for some } \beta^{\mathbb{P}}, \mathbb{Q} \in \mathcal{Q}(\mu_0), \text{ s.t.} \right. \\ \left. \int_0^1 |\beta_s^{\mathbb{P}}|^2 ds < \infty, \mathbb{Q}\text{-a.s.}, \text{ and } \left( \widehat{\sigma}_s \beta_s^{\mathbb{P}}, \frac{1}{2} \widehat{\sigma}_s^2 \right) \in U, \text{ Leb} \otimes \mathbb{P}\text{-a.e.} \right\}. \quad (4.3)$$

With  $\mathcal{P}^U(\mu_0)$ , we introduce

$$\mathcal{P}^U(\mu_0, \mu_1) := \{\mathbb{P} \in \mathcal{P}^U(\mu_0) : \mathbb{P} \circ X_1^{-1} = \mu_1\}.$$

We next consider a cost function  $c : [0, 1] \times \Omega \times \mathbb{R}^d \times \mathbb{S}_+^d \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\infty\}$  such that  $(t, \omega) \mapsto c_t(\omega, \cdot)$  is  $\mathbb{F}$ -progressively measurable, in particular  $c_t(\omega, \cdot) = c_t(\omega_{t \wedge \cdot}, \cdot)$ , for all  $(t, \omega) \in [0, 1] \times \Omega$ , and we assume that

$$\mathbb{E}^{\mathbb{P}} \left[ \int_0^1 c_s^- (\widehat{b}_s^{\mathbb{P}}, \widehat{\sigma}_s^2, m_s) ds \right] < \infty, \text{ for all } \mathbb{P} \in \mathcal{P}^U(\mu_0), m \in \mathbb{M}, \text{ with } c^-(\cdot) := \max(-c(\cdot), 0).$$

We now introduce the following control problem in weak formulation

$$V_0(\xi, m) := \sup_{\mathbb{P} \in \mathcal{P}^U(\mu_0)} J(\xi, m, \mathbb{P}), \text{ with } J(\xi, m, \mathbb{P}) := \mathbb{E}^{\mathbb{P}} \left[ \xi - \int_0^1 c_s(\widehat{b}_s^{\mathbb{P}}, \widehat{\sigma}_s^2, m_s) ds \right], \quad (4.4)$$

where the reward function  $\xi : \Omega \rightarrow \mathbb{R} \cup \{-\infty\}$  is restricted to the set

$$\Xi^U := \left\{ \xi : \Omega \rightarrow \mathbb{R} : \mathbb{E}^{\mathbb{P}}[\xi^+] < \infty, \text{ for all } \mathbb{P} \in \mathcal{P}^U(\mu_0) \right\}.$$

**Definition 4.1.** (i) For  $\xi \in \Xi^U$  and  $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ , we denote by

$$\text{MFG}(\xi, \mu_0) := \left\{ \widehat{\mathbb{P}} \in \mathcal{P}^U(\mu_0) : J(\xi, m, \widehat{\mathbb{P}}) = V_0(\xi, m) \in \mathbb{R} \text{ with } m_s = \widehat{\mathbb{P}} \circ X_s^{-1}, s \in [0, 1] \right\}$$

the set of all solutions to the MFG problem with reward function  $\xi$ .

(ii) Given a pair  $(\mu_0, \mu_1)$  of starting and target marginals, we denote by

$$\text{MFP}(\mu_0, \mu_1) := \left\{ \xi \in \Xi^U : \widehat{\mathbb{P}} \circ X_1^{-1} = \mu_1 \text{ for some } \widehat{\mathbb{P}} \in \text{MFG}(\xi, \mu_0) \right\}$$

the collection of all reward functions  $\xi \in \Xi^U$  which induce some MFG solution  $\widehat{\mathbb{P}}$  with marginals  $\widehat{\mathbb{P}} \circ X_0^{-1} = \mu_0, \widehat{\mathbb{P}} \circ X_1^{-1} = \mu_1$ .

## 4.2 Characterization of the solutions of the mean field planning problem

Let  $\mathcal{H}^2(\mu_0) := \bigcap_{\mathbb{P} \in \mathcal{P}^U(\mu_0)} \mathbb{H}^2(\mathbb{P})$ , where  $\mathbb{H}^2(\mathbb{P})$  denotes the collection of all  $\mathbb{F}$ -progressively measurable processes  $Z : [0, 1] \times \Omega \rightarrow \mathbb{R}^d$  such that  $\mathbb{E}^{\mathbb{P}} \left[ \int_0^1 |\widehat{\sigma}_s Z_s|^2 ds \right] < \infty$ . The Hamiltonian of the last stochastic control problem is defined by:

$$H_s(\omega, z, \gamma, m) := \sup_{(b, a) \in U} \left\{ b \cdot z + \frac{1}{2} a : \gamma - c_s(\omega, b, a, m) \right\}. \quad (4.5)$$

Let us denote the domain of  $H$  by

$$D_H(s, \omega, m) := \{(z, \gamma) \in \mathbb{R}^d \times \mathbb{S}^d : H_s(\omega, z, \gamma, m) < \infty\},$$

and by  $\partial_{(z, \gamma)} H_s(z, \gamma, m) \subseteq U$  the sub-gradient of the convex function  $(z, \gamma) \mapsto H_s(z, \gamma, m)$  in  $U$ .

Given  $\mathbb{F}$ -progressively measurable processes  $(Z, \Gamma)$  on  $\Omega$  taking value in  $\mathbb{R}^d \times \mathbb{S}^d$ , we introduce the McKean-Vlasov SDE

$$X_t = X_0 + \int_0^t \bar{b}_s(Z_s, \Gamma_s, \widehat{\mathbb{P}} \circ X_s^{-1}) ds + \int_0^t \bar{\sigma}_s(Z_s, \Gamma_s, \widehat{\mathbb{P}} \circ X_s^{-1}) dW_s^{\widehat{\mathbb{P}}}, \quad \widehat{\mathbb{P}}\text{-a.s.} \quad (4.6)$$

for some measurable selection  $(\bar{b}_s, \frac{1}{2} \bar{\sigma}_s^2)(z, \gamma, m) \in \partial_{(z, \gamma)} H_s(z, \gamma, m) \subseteq U$ .



Let  $\text{MKV}_0(\mu_0, \mu_1)$  be the collection of all triples  $(Z, \Gamma, \widehat{\mathbb{P}})$  such that  $(Z, \Gamma) \in D_H(\cdot, \widehat{m}_\cdot)$  with  $\widehat{m}_s := \widehat{\mathbb{P}} \circ X_s^{-1}$ , and  $\widehat{\mathbb{P}} \in \mathcal{P}^U(\mu_0, \mu_1)$  is a (weak) solution of the last McKean-Vlasov SDE. We next define

$$\text{MKV}(\mu_0, \mu_1) := \left\{ (Z, \Gamma, \widehat{\mathbb{P}}) \in \text{MKV}_0(\mu_0, \mu_1) : Z \in \mathcal{H}^2(\mu_0) \right\}.$$

Finally, we introduce for all  $(Z, \Gamma, \widehat{\mathbb{P}}) \in \text{MKV}(\mu_0, \mu_1)$  the  $\mathcal{F}_1$ -measurable random variable

$$Y_1^{Z, \Gamma, \widehat{\mathbb{P}}} := \int_0^1 Z_s \cdot dX_s + \int_0^1 \left( \frac{1}{2} \Gamma_s : \widehat{\sigma}_s^2 - H_s(Z_s, \Gamma_s, \widehat{\mathbb{P}} \circ X_s^{-1}) \right) ds.$$

**Remark 4.2.** We implicitly work here under the ZFC set-theoretic axioms and the continuum hypothesis. Then, for  $Z \in \mathcal{H}^2(\mu_0)$ , the stochastic integral  $\int_0^t Z_s \cdot dX_s$  is well-defined under each  $\mathbb{P} \in \mathcal{P}^U(\mu_0)$  and it can be aggregated as a universal process independent of  $\mathbb{P} \in \mathcal{P}^U(\mu_0)$ , see Nutz [18, Theorem 2.2, Lemma 2.5]. Further, for each  $\mathbb{P} \in \mathcal{P}^U(\mu_0)$ , as  $\int_0^t Z_s \cdot \widehat{\sigma}_s \beta_s^\mathbb{P} ds$  is well defined, and

$$Z_s \cdot \widehat{\sigma}_s \beta_s^\mathbb{P} + \frac{1}{2} \Gamma_s : \widehat{\sigma}_s^2 - H_s(Z_s, \Gamma_s, \mathbb{P} \circ X_s^{-1}) \leq c_s(\widehat{\sigma}_s \beta_s^\mathbb{P}, \widehat{\sigma}_s^2, \mathbb{P} \circ X_s^{-1}),$$

it follows that

$$\int_0^1 \left( \frac{1}{2} \Gamma_s : \widehat{\sigma}_s^2 - H_s(Z_s, \Gamma_s, \widehat{\mathbb{P}} \circ X_s^{-1}) \right) ds$$

is pathwisely well-defined under each  $\mathbb{P} \in \mathcal{P}^U(\mu_0)$ . Consequently,  $Y_1^{Z, \Gamma, \widehat{\mathbb{P}}}$  can be aggregated as a universal random variable on  $\Omega$  taking value in  $\mathbb{R} \cup \{-\infty\}$ .

For the main result of this section, we denote  $\pi(U) := \{a : (b, a) \in U \text{ for some } b \in \mathbb{R}^d\}$ , and we define  $\mathcal{L}_0^1(\mu_0) := \bigcap_{\mathbb{P} \in \mathcal{P}^U(\mu_0)} \mathbb{L}^1(\mathcal{F}_0^+, \mathbb{P})$ , and

$$\Xi(\mu_0, \mu_1) := \mathcal{L}_0^1(\mu_0) + \left\{ Y_1^{Z, \Gamma, \widehat{\mathbb{P}}} : (Z, \Gamma, \widehat{\mathbb{P}}) \in \text{MKV}(\mu_0, \mu_1) \right\}.$$

**Theorem 4.3.** *The following holds true:*

(i) *In both settings (4.1) and (4.3) for the definition of  $\mathcal{P}^U(\mu_0)$ , one has*

$$\Xi(\mu_0, \mu_1) \subseteq \text{MFP}(\mu_0, \mu_1).$$

(ii) *In the setting (4.3) for the definition of  $\mathcal{P}^U(\mu_0)$ , let  $\xi \in \text{MFP}(\mu_0, \mu_1)$ . Assume in addition that  $\sup_\gamma \left\{ \frac{1}{2} a : \gamma - H(\cdot, \gamma) \right\}$  has a maximizer in the domain of  $H$  for all  $a \in \pi(U)$ , and that*

$$\sup_{\mathbb{P} \in \mathcal{P}^U(\mu_0)} \mathbb{E}^\mathbb{P} \left[ \left| \xi \right|^{2+\kappa} + \int_0^1 |c_s(\widehat{\sigma}_s \beta_s^\mathbb{P}, \widehat{\sigma}_s^2, m_s)|^{2+\kappa} ds \right] < \infty, \quad \text{for some } \kappa > 0. \quad (4.7)$$

*Then,  $\xi \in \Xi(\mu_0, \mu_1)$ . Moreover, let  $\widehat{m}_s := \widehat{\mathbb{P}} \circ X_s^{-1}$  for some  $\widehat{\mathbb{P}} \in \text{MFG}(\xi, \mu_0)$ , then we may choose the corresponding  $Y_0 + Y_1^{Z, \Gamma, \widehat{\mathbb{P}}} \in \Xi(\mu_0, \mu_1)$  such that*

$$\arg \max_{\mathbb{P} \in \mathcal{P}^U(\mu_0)} J(\xi, \widehat{m}, \mathbb{P}) = \arg \max_{\mathbb{P} \in \mathcal{P}^U(\mu_0)} J\left(Y_0 + Y_1^{Z, \Gamma, \widehat{\mathbb{P}}}, \widehat{m}, \mathbb{P}\right). \quad (4.8)$$

**Proof.** (i) In order to prove the inclusion, we only need to verify that, for all  $(Z, \Gamma, \widehat{\mathbb{P}}) \in \text{MKV}(\mu_0, \mu_1)$  such that  $Y_1^{Z, \Gamma, \widehat{\mathbb{P}}} \in \Xi(\mu_0, \mu_1)$ , one has  $\widehat{\mathbb{P}} \in \text{MFG}(Y_1^{Z, \Gamma, \widehat{\mathbb{P}}}, \mu_0)$ , i.e.,

$$J\left(Y_1^{Z, \Gamma, \widehat{\mathbb{P}}}, \widehat{m}, \widehat{\mathbb{P}}\right) = V_0\left(Y_1^{Z, \Gamma, \widehat{\mathbb{P}}}, \widehat{m}\right), \quad \text{where } \widehat{m}_s := \widehat{\mathbb{P}} \circ X_s^{-1}. \quad (4.9)$$

Indeed, as  $J(Y_0 + Y_1^{Z, \Gamma, \widehat{\mathbb{P}}}, \widehat{m}, \mathbb{P}) = \mathbb{E}^{\mathbb{P}_0}[Y_0] + J(Y_1^{Z, \Gamma, \widehat{\mathbb{P}}}, \widehat{m}, \mathbb{P})$ , for all  $\mathbb{P} \in \mathcal{P}^U(\mu_0)$ , this implies that  $Y_0 + Y_1^{Z, \Gamma, \widehat{\mathbb{P}}} \in \text{MFP}(\mu_0, \mu_1)$ . To prove (4.9), we first compute for  $\mathbb{P} \in \mathcal{P}^U(\mu_0)$  that

$$\begin{aligned} J\left(Y_1^{Z, \Gamma, \widehat{\mathbb{P}}}, \widehat{m}, \mathbb{P}\right) &= \mathbb{E}^{\mathbb{P}} \left[ \int_0^1 Z_s \cdot dX_s + \int_0^1 \left( \frac{1}{2} \Gamma_s : \widehat{\sigma}_s^2 - H_s(Z_s, \Gamma_s, \widehat{m}_s) - c_s(\widehat{b}_s^{\mathbb{P}}, \widehat{\sigma}_s^2, \widehat{m}_s) \right) ds \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[ \int_0^1 \left( Z_s \cdot \widehat{b}_s^{\mathbb{P}} + \frac{1}{2} \Gamma_s : \widehat{\sigma}_s^2 - c_s(\widehat{b}_s^{\mathbb{P}}, \widehat{\sigma}_s^2, \widehat{m}_s) - H_s(Z_s, \Gamma_s, \widehat{m}_s) \right) ds \right], \end{aligned}$$

as  $Z \in \mathcal{H}^2(\mu_0)$ . Then, it follows from the definition of the Hamiltonian that  $J(Y_1^{Z, \Gamma, \widehat{\mathbb{P}}}, \widehat{m}, \mathbb{P}) \leq 0$  for all  $\mathbb{P} \in \mathcal{P}^U(\mu_0)$ . On the other hand, as  $(Z, \Gamma, \widehat{\mathbb{P}}) \in \text{MKV}(\mu_0, \mu_1)$ , it follows from (4.6) that  $(\widehat{b}_s^{\widehat{\mathbb{P}}}, \frac{1}{2} \widehat{\sigma}_s^2) \in \partial_{z, \gamma} H_s(Z_s, \Gamma_s, \widehat{m}_s)$ , and therefore the supremum in the Hamiltonian is attained:

$$Z_s \cdot \widehat{b}_s^{\widehat{\mathbb{P}}} + \frac{1}{2} \Gamma_s : \widehat{\sigma}_s^2 - c_s(\widehat{b}_s^{\widehat{\mathbb{P}}}, \widehat{\sigma}_s^2, \widehat{m}_s) = H_s(Z_s, \Gamma_s, \widehat{m}_s), \quad \text{Leb} \otimes \widehat{\mathbb{P}}\text{-a.e.}$$

Hence  $J(Y_1^{Z, \Gamma, \widehat{\mathbb{P}}}, \widehat{m}, \widehat{\mathbb{P}}) = 0 = V_0(Y_1^{Z, \Gamma, \widehat{\mathbb{P}}}, \widehat{m})$ , which concludes the proof of (4.9).

(ii) Next, let us consider  $\xi \in \text{MFP}(\mu_0, \mu_1)$  in the setting (4.3) for the definition of  $\mathcal{P}^U(\mu_0)$ , and under the additional conditions in Item (ii) of the statement. We observe that by condition (4.2), one has  $\mathcal{Q}(\mu_0) \subseteq \mathcal{P}^U(\mu_0)$ .

Let  $\widehat{\mathbb{P}} \in \text{MFG}(\xi, \mu_0)$  satisfying  $\widehat{\mathbb{P}} \circ X_1^{-1} = \mu_1$ , then  $\widehat{\mathbb{P}}$  is an optimal control of the stochastic control problem  $V_0(\xi, \widehat{m})$  defined in (4.4) with  $\widehat{m}_s := \widehat{\mathbb{P}} \circ X_s^{-1}$ ,  $s \in [0, 1]$ .

(ii-1) Let  $\mathcal{P}^U(t, \omega)$  denote the dynamic version of  $\mathcal{P}^U(\mu_0)$  by considering the dynamics on  $[t, 1]$  starting at time  $t \in [0, T]$  from the path  $\omega \in \Omega$ , and consider the dynamic version of the control problem  $V_0$ :

$$V_t(\omega) := \sup_{\mathbb{P} \in \mathcal{P}^U(t, \omega)} \mathbb{E}^{\mathbb{P}} \left[ \xi - \int_t^1 c_s(\widehat{\sigma}_s \beta_s^{\mathbb{P}}, \widehat{\sigma}_s^2, \widehat{m}_s) ds \right], \quad \text{for all } (t, \omega) \in [0, 1] \times \Omega.$$

Since  $\xi$  and  $c$  satisfy the addition integrability condition in (4.7), it follows that (see e.g. Soner, Touzi, and Zhang [27; 28] and Possamaï, Tan, and Zhou [23]) the process  $\{V_t, t \in [0, 1]\}$  satisfies the dynamic programming principle:

$$V_t = \sup_{\mathbb{P} \in \mathcal{P}^U(t, \omega)} \mathbb{E}^{\mathbb{P}} \left[ V_{t+h} - \int_t^{t+h} c_s(\widehat{\sigma}_s \beta_s^{\mathbb{P}}, \widehat{\sigma}_s^2, \widehat{m}) ds \right], \quad \mathbb{P}\text{-a.s.}, \quad (4.10)$$

for all  $\mathbb{P} \in \mathcal{P}^U(\mu_0)$  and  $0 \leq h \leq T - t$ , so that  $V + \int_0^\cdot c_s(\widehat{\sigma}_s \beta_s^{\mathbb{P}}, \widehat{\sigma}_s^2, \widehat{m}) ds$  is  $\mathbb{P}$ -supermartingale for all  $\mathbb{P} \in \mathcal{P}^U(\mu_0)$ , and we may introduce the corresponding right-continuous limit  $V_t^+(\omega) := \lim_{s \searrow t} V_s(\omega)$ , which inherits the dynamic programming principle (4.10). Moreover, one has

$$\sup_{\mathbb{P} \in \mathcal{P}^U(\mu_0)} \mathbb{E}^{\mathbb{P}} \left[ \sup_{0 \leq s \leq 1} |V_s^+|^{2+\kappa'} \right] < \infty, \quad \text{for some } 0 < \kappa' < \kappa. \quad (4.11)$$

(ii-2) As  $\mathcal{Q}(\mu_0) \subseteq \mathcal{P}^U(\mu_0)$ , for any  $\mathbb{P} \in \mathcal{P}^U(\mu_0)$ , the process  $V_t^+ - \int_0^t c_s(\hat{\sigma}_s \beta_s^{\mathbb{P}}, \hat{\sigma}_s^2, \hat{m}_s) ds$  is a càdlàg supermartingale under any martingale measure equivalent to  $\mathbb{P}$ . By the optional decomposition theorem (see El Karoui and Quenez [7, Theorem 2.4.2], or Föllmer and Kramkov [9]), there is a non-decreasing process  $K^{\mathbb{P}}$  starting from  $K_0^{\mathbb{P}} = 0$ , and a process  $Z^{\mathbb{P}}$  such that for  $t \in [0, T]$

$$V_t^+ = \xi + \int_t^1 \left( \hat{\sigma}_s \beta_s^{\mathbb{P}} \cdot Z_s^{\mathbb{P}} - c_s(\hat{\sigma}_s \beta_s^{\mathbb{P}}, \hat{\sigma}_s^2, \hat{m}_s) \right) ds - \int_t^1 Z_s^{\mathbb{P}} \cdot dX_s - K_t^{\mathbb{P}} + K_T^{\mathbb{P}}, \quad \mathbb{P}\text{-a.s.}$$

Let  $Z_t := \frac{d\langle V^+, X \rangle_t}{dt}$ , then  $Z^{\mathbb{P}} = Z$ ,  $\mathbb{P}$ -a.s., for all  $\mathbb{P} \in \mathcal{P}^U(\mu_0)$ , and then one can find a process  $K$  such that  $K^{\mathbb{P}} = K$ ,  $\mathbb{P}$ -a.s., for all  $\mathbb{P} \in \mathcal{P}^U(\mu_0)$ . Moreover, by standard estimates, it follows from (4.11) that  $Z \in \mathcal{H}^2(\mu_0)$ .

Further, since the optimal control  $\hat{\mathbb{P}}$  is a maximizer in the dynamic programming principle, we have  $K = 0$ ,  $\hat{\mathbb{P}}$ -a.s. Then, by the same arguments as in Theorem 2.5, we that

$$V_t^+ = \xi + \int_t^1 F_s(Z_s, \hat{\sigma}_s^2, \hat{m}_s) ds - \int_t^1 Z_s \cdot dX_s, \quad t \in [0, T], \quad \hat{\mathbb{P}}\text{-a.s.},$$

where  $F_s(\omega, z, a, m) := \sup_{b \in U_a} \{b \cdot z - c_s(\omega, b, a, m)\}$  with  $U_a := \{b \in \mathbb{R}^d : (b, a) \in U\}$ , and

$$F_s(Z_s, \hat{\sigma}_s^2, \hat{m}_s) = \hat{\sigma}_s \beta_s^{\hat{\mathbb{P}}} \cdot Z_s - c_s(\hat{\sigma}_s \beta_s^{\hat{\mathbb{P}}}, \hat{\sigma}_s^2, \hat{m}_s), \quad \text{Leb} \otimes \hat{\mathbb{P}}\text{-a.e.}$$

(ii-3) Notice that

$$H_s(\omega, z, \gamma, m) = \sup_{a \in \pi(U)} \left\{ \frac{1}{2} a : \gamma + F_s(\omega, z, a, m) \right\},$$

with  $\pi(U) := \{a : (b, a) \in U \text{ for some } b \in \mathbb{R}^d\}$ .

Then, under the additional conditions in Item (ii) of the present Theorem, we can find a (measurable) process  $\Gamma$  in the sub-gradient of  $\frac{1}{2}a \mapsto c(\cdot, a)$ , such that  $(Z, \Gamma) \in D_H(\cdot, \hat{m}_\cdot)$  and

$$H_s(\cdot, Z_s, \Gamma_s, \hat{m}_s) = \frac{1}{2} \hat{\sigma}_s^2 : \Gamma_s + F_s(\cdot, Z_s, \hat{\sigma}_s^2, \hat{m}_s), \quad \omega\text{-wisely},$$

and a (measurable) process  $\Gamma'$  such that one has strict inequality “>” if  $\Gamma$  is replaced by  $\Gamma'$  in the above formula. Therefore, for the optimal control  $\hat{\mathbb{P}}$ , one has

$$H_s(\cdot, Z_s, \Gamma_s, \hat{m}_s) = \hat{\sigma}_s \beta_s^{\hat{\mathbb{P}}} \cdot Z_s + \frac{1}{2} \hat{\sigma}_s^2 : \Gamma_s - c_s(\hat{\sigma}_s \beta_s^{\hat{\mathbb{P}}}, \hat{\sigma}_s^2, \hat{m}_s), \quad \text{Leb} \otimes \hat{\mathbb{P}}\text{-a.e.}$$

By standard convex duality, one has

$$\left( \hat{\sigma}_s \beta_s^{\hat{\mathbb{P}}}, \frac{1}{2} \hat{\sigma}_s^2 \right) \in \partial_{(z, \gamma)} H_s(\cdot, Z_t, \Gamma_t) \subseteq U, \quad \text{Leb} \otimes \hat{\mathbb{P}}\text{-a.e.}$$

We then define  $\bar{\Gamma}_s := \Gamma_s \mathbf{1}_{\{K_s=0\}} + \Gamma'_s \mathbf{1}_{\{K_s>0\}}$ . Consequently,  $\Gamma = \bar{\Gamma}$ ,  $\text{Leb} \otimes \hat{\mathbb{P}}$ -a.e., and  $(Z, \bar{\Gamma}, \hat{\mathbb{P}}) \in \text{MKV}(\mu_0, \mu_1)$ , and the random variable  $\xi$  has the required decomposition  $\xi = V_0^+ + Y_1^{Z, \bar{\Gamma}, \hat{\mathbb{P}}}$ ,  $\hat{\mathbb{P}}$ -a.s.

(ii-4) Finally, the equality in (4.8) follows immediately from the last construction of  $\bar{\Gamma}$  as a probability measure  $\mathbb{P}$  is optimal if and only if the corresponding process  $K^{\mathbb{P}} = 0$ ,  $\mathbb{P}$ -a.s.  $\blacksquare$

### 4.3 Existence of solutions to the planning problem

The first inclusion in Theorem 4.3 provides a systematic way to construct solutions to the mean field planning problem with given marginals  $\mu_0$  and  $\mu_1$ , that is, it is enough to construct a McKean-Vlasov dynamics  $X$  in (4.6) with given marginal distributions. Under further conditions, our next result ensures that it is enough to construct a semi-martingale measure in  $\mathcal{P}^U(\mu_0, \mu_1)$ .

**Proposition 4.4.** (i) *In both settings (4.1) and (4.3) for the definition of  $\mathcal{P}^U(\mu_0)$ , assume that, for all  $(b, \frac{1}{2}a) \in U$  and  $(s, \omega, m) \in [0, 1] \times \Omega \times \mathcal{P}(\mathbb{R}^d)$ , there exists a maximizer for  $\sup_{(z, \gamma) \in D_H(\cdot)} \{b \cdot z + \frac{1}{2}a : \gamma - H(\cdot, z, \gamma)\}$ . Then, for all  $\widehat{\mathbb{P}} \in \mathcal{P}^U(\mu_0, \mu_1)$ , there exists a measurable selection  $(\bar{b}, \frac{1}{2}\bar{\sigma}^2)(\cdot)$  in  $\partial_{(z, \gamma)} H_s(\cdot) \subseteq U$ , together with  $\mathbb{F}$ -progressively measurable processes  $(Z, \Gamma) \in D_m(\cdot)$  such that*

$$X_t = X_0 + \int_0^t \bar{b}_s(Z_s, \Gamma_s, \widehat{\mathbb{P}} \circ X_s^{-1}) ds + \int_0^t \bar{\sigma}_s(Z_s, \Gamma_s, \widehat{\mathbb{P}} \circ X_s^{-1}) dW_s^{\widehat{\mathbb{P}}}, \quad \widehat{\mathbb{P}}\text{-a.s.},$$

where  $W^{\widehat{\mathbb{P}}}$  is  $\widehat{\mathbb{P}}$ -Brownian motion, i.e.  $(Z, \Gamma, \widehat{\mathbb{P}}) \in \text{MKV}_0(\mu_0, \mu_1)$ .

(ii) *Consequently, if  $Z \in \mathcal{H}^2(\mu_0)$ , then we obtain that  $(Z, \Gamma, \widehat{\mathbb{P}}) \in \text{MKV}(\mu_0, \mu_1)$  and hence  $Y_1^{Z, \Gamma, \widehat{\mathbb{P}}} \in \text{MFP}(\mu_0, \mu_1)$ .*

**Proof.** We notice that the Hamiltonian  $H(\cdot, z, \gamma)$  is convex in  $(z, \gamma)$ , then by standard result in convex analysis, see, e.g., [25, Theorem 23.5],  $(\widehat{z}, \widehat{\gamma})$  is maximizer of

$$\sup_{(z, \gamma) \in D_H(\cdot)} \left\{ b \cdot z + \frac{1}{2}a : \gamma - H(\cdot, z, \gamma) \right\} =: c^{**}(\cdot, b, a)$$

is equivalent to

$$\left( b, \frac{1}{2}a \right) \in \partial_{(z, \gamma)} H(\cdot, \widehat{z}, \widehat{\gamma}) \iff \left( \widehat{z}, \frac{1}{2}\widehat{\gamma} \right) \in \partial_{(b, a)} c^{**}(\cdot, b, a).$$

Let  $\widehat{\mathbb{P}} \in \mathcal{P}^U(\mu_0, \mu_1)$  so that the dynamic of the canonical process  $X$  is given by

$$dX_t = \widehat{b}_t^{\widehat{\mathbb{P}}} dt + \widehat{\sigma}_t^{\widehat{\mathbb{P}}} dW_t^{\widehat{\mathbb{P}}}, \quad \widehat{\mathbb{P}}\text{-a.s.} \quad \text{and} \quad (\widehat{b}^{\widehat{\mathbb{P}}}, \widehat{\sigma}^{\widehat{\mathbb{P}}}) \in U, \quad \text{Leb} \otimes \widehat{\mathbb{P}}\text{-a.e.}$$

Let  $\widehat{m}_s := \widehat{\mathbb{P}} \circ X_s^{-1}$ ,  $s \in [0, 1]$ . One can use the measurable selection theorem to choose a version of sub-gradient  $(\bar{b}, \frac{1}{2}\bar{\sigma}^2)(\cdot)$  in  $\partial_{(z, \gamma)} H(\cdot)$ , together with  $\mathbb{F}$ -progressively measurable processes  $(Z, \Gamma)$  such that

$$(\beta_t^{\widehat{\mathbb{P}}}, \widehat{\sigma}_t^{\widehat{\mathbb{P}}}) = (\bar{b}, \bar{\sigma}^2)(Z_t, \Gamma_t, \widehat{m}_t), \quad \text{Leb} \otimes \widehat{\mathbb{P}}\text{-a.e.},$$

and hence  $(Z, \Gamma, \widehat{\mathbb{P}}) \in \text{MKV}_0(\mu_0, \mu_1)$ . ■

**Construction of martingales with given marginals** Let  $U = \{0\} \times \mathbb{S}_+^d$ , and

$$c_s(\omega, b, a, m) := \frac{1}{4} C_m |a|^2, \quad \text{with} \quad C_m := \int_{\mathbb{R}^d} (1 + |x|^2) m(dx).$$

Then, the Hamiltonian has domain  $D_H(\cdot) = \mathbb{R}^d \times \mathbb{S}^d$ , and

$$H_s(\omega, z, \gamma, m) = \frac{1}{4} C_m^{-1} |\gamma|^2, \quad \partial_{(z, \gamma)} H(\cdot, z, \gamma, m) = \left\{ \left( 0, \frac{1}{2} C_m^{-1} \gamma \right) \right\}.$$

Clearly, all the conditions in Proposition 4.4 hold true. Therefore, to find a solution to the mean field planning problem  $\text{MFP}(\mu_0, \mu_1)$ , a first approach consists in finding  $\widehat{\mathbb{P}} \in \mathcal{P}^U(\mu_0, \mu_1)$  so that

$$X_t = X_0 + \int_0^t \widehat{\sigma}_s dW_s^{\widehat{\mathbb{P}}}, \quad \widehat{\mathbb{P}}\text{-a.s.},$$

and hence with  $Z \equiv 0$ ,  $\Gamma_t := C_{\widehat{m}_t} \widehat{\sigma}_t^2$ ,  $\widehat{m}_t := \widehat{\mathbb{P}} \circ X_t^{-1}$ , one has  $(\widehat{\mathbb{P}}, Z, \Gamma) \in \text{MKV}(\mu_0, \mu_1)$ .

In this setting and when  $d = 1$ , the problem of finding an element  $\mathcal{P}^U(\mu_0, \mu_1)$  is equivalent to the so-called Skorokhod embedding problem, which consists in finding a stopping time  $\tau$  in some filtered probability space equipped with a Brownian motion  $W$  such that

$$W_0 \sim \mu_0, \quad W_\tau \sim \mu_1.$$

Indeed, given  $\widehat{\mathbb{P}} \in \mathcal{P}^U(\mu_0, \mu_1)$ , under which  $X$  is a diffusion martingale with marginals  $\mu_0$  and  $\mu_1$ . By Dambis-Dubins-Schwarz theorem, one can represent  $X$  as a time-changed Brownian motion, i.e.,

$$X_t = W_{\langle X \rangle_t}, \quad t \in [0, 1],$$

where  $W$  is some Brownian motion and  $\langle X \rangle_t$  are stopping times w.r.t. the time-changed filtration. Thus  $(W, \langle X \rangle_1)$  provides a solution to the Skorokhod embedding problem with marginals  $(\mu_0, \mu_1)$ . Conversely, given a solution  $(W, \tau)$ , to the Skorokhod embedding problem with marginals  $\mu_0$  and  $\mu_1$ , let us define

$$X_t := W_{\tau \wedge \frac{t}{1-t}}, \quad t \in [0, 1].$$

Then, it is easy to check that  $X$  is a martingale diffusion process such that  $X_0 = W_0 \sim \mu_0$  and  $X_1 = W_\tau \sim \mu_1$ , and hence  $\mathbb{P} \circ X^{-1} \in \mathcal{P}^U(\mu_0)$ .

We also notice that the Skorokhod embedding problem has a solution if and only if  $\mu_0$  and  $\mu_1$  have finite first order moment and  $\int_{\mathbb{R}} \phi(x) \mu_0(dx) \leq \int_{\mathbb{R}} \phi(x) \mu_1(dx)$  for all convex function  $\phi$ . Moreover, there are various constructions of solutions to the Skorokhod embedding problem, and many of them enjoy some optimal property, see e.g. Oblój [20] for a survey. Consequently, the induced solution of the mean field planning problem enjoys the same optimal property among all possible solutions, and hence solves the corresponding optimal planning problem as discussed in Section 2.4.

**Construction of semi-martingales with given marginals** When  $U = \mathbb{R}^d \times \mathbb{S}^d$ , the problem of construction a semi-martingale measure in  $\mathcal{P}^U(\mu_0, \mu_1)$  is very easy. Let us report the construction by using the so called Bass solution of the Skorokhod embedding problem, see e.g. Oblój [20]. Assume that both  $\mu_0$  and  $\mu_1$  have finite first order moment, and let  $B$  be a standard Brownian motion in a probability space  $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ , together with a random variable  $\widehat{X}_0$  independent of  $B$  such that  $\mathbb{P}^* \circ \widehat{X}_0^{-1} = \mu_0$ . We can then find some measurable function  $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that  $\mathbb{P}^* \circ (T(B_1))^{-1} \sim \mu_1$ . Next, by martingale representation, there exists a constant  $c \in \mathbb{R}$  and a predictable process  $\tilde{\sigma}$  (w.r.t. the Brownian filtration generated by  $B$ ) such that

$$T(B_1) = c + \int_0^1 \tilde{\sigma}_t dB_t.$$

We then define a process  $\widehat{X}$  by

$$\widehat{X}_t := \widehat{X}_0 + \int_0^t (c - \widehat{X}_0) ds + \int_0^t \tilde{\sigma}_s dB_s.$$

It is immediate to check that

$$\widehat{\mathbb{P}} := \mathbb{P}^* \circ \widehat{X}^{-1} \in \mathcal{P}^U(\mu_0, \mu_1).$$

**Optimal transport along controlled McKean-Vlasov dynamic** As in the discussion in Section 2.4, one can consider an optimal mean field planning problem, by choosing an optimal solution  $\xi$  in the class  $\Xi(\mu_0, \mu_1)$  w.r.t. some criteria. The problem can be reduced to an optimal transport problem along controlled McKean-Vlasov dynamic: for some reward function  $\Psi$ , one solves

$$\sup_{(Z, \Gamma, \widehat{\mathbb{P}}) \in \text{MKV}(\mu_0, \mu_1)} \Psi(Z, \Gamma, \widehat{\mathbb{P}}),$$

where we recall from (4.6) that  $\text{MKV}(\mu_0, \mu_1)$  is the set of all  $(Z, \Gamma, \widehat{\mathbb{P}})$  such that, with a version of sub-gradient  $(\bar{b}_s, \frac{1}{2}\bar{\sigma}_s^2)(z, \gamma, m) \in \partial_{(z, \gamma)} H_s(z, \gamma, m) \subseteq U$ ,  $\widehat{\mathbb{P}}$  is weak solution to the McKean-Vlasov equation:

$$X_t = X_0 + \int_0^t \bar{b}_s(Z_s, \Gamma_s, \widehat{\mathbb{P}} \circ X_s^{-1}) ds + \int_0^t \bar{\sigma}_s(Z_s, \Gamma_s, \widehat{\mathbb{P}} \circ X_s^{-1}) dW_s^{\widehat{\mathbb{P}}}, \quad \widehat{\mathbb{P}}\text{-a.s.},$$

under the marginal constraints:

$$\widehat{\mathbb{P}} \circ X_0^{-1} = \mu_0 \quad \text{and} \quad \widehat{\mathbb{P}} \circ X_1^{-1} = \mu_1.$$

Such a problem extends the classical optimal transport problem studied in the literature, such as the martingale optimal transport in Beiglböck, Henry-Labordère, and Penkner [3] and Galichon, Henry-Labordère, and Touzi [10], the martingale Benamou-Brenier problem in Huesmann and Trevisan [13], Backhoff-Veraguas, Beiglböck, Huesmann, and Källblad [2], or the semimartingale optimal transport problem in Mikami and Thieullen [17], Tan and Touzi [29], etc.

## References

- [1] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta. Mean field games: numerical methods for the planning problem. *SIAM J. Control Optim.*, 50(1):77–109, 2012. ISSN 0363-0129. doi: 10.1137/100790069. URL <https://doi.org/10.1137/100790069>.
- [2] J. Backhoff-Veraguas, M. Beiglböck, M. Huesmann, and S. Källblad. Martingale Benamou-Brenier: a probabilistic perspective. *Ann. Probab.*, 48(5):2258–2289, 2020. ISSN 0091-1798. doi: 10.1214/20-AOP1422. URL <https://doi.org/10.1214/20-AOP1422>.
- [3] M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices—a mass transport approach. *Finance Stoch.*, 17(3):477–501, 2013. ISSN 0949-2984. doi: 10.1007/s00780-013-0205-8. URL <https://doi.org/10.1007/s00780-013-0205-8>.
- [4] J.-D. Benamou, G. Carlier, S. Di Marino, and L. Nenna. An entropy minimization approach to second-order variational mean-field games. *Math. Models Methods Appl. Sci.*, 29(8):1553–1583, 2019. ISSN 0218-2025. doi: 10.1142/S0218202519500283. URL <https://doi.org/10.1142/S0218202519500283>.
- [5] J. Cvitanić, D. Possamaï, and N. Touzi. Dynamic programming approach to principal-agent problems. *Finance Stoch.*, 22(1):1–37, 2018. ISSN 0949-2984. doi: 10.1007/s00780-017-0344-4. URL <https://doi.org/10.1007/s00780-017-0344-4>.
- [6] M. F. Djete, D. Possamaï, and X. Tan. McKean-Vlasov optimal control: the dynamic programming principle. *Preprint*, 2020.
- [7] N. El Karoui and M.-C. Quenez. Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM J. Control Optim.*, 33(1):29–66, 1995. ISSN 0363-0129. doi: 10.1137/S0363012992232579. URL <https://doi.org/10.1137/S0363012992232579>.

- [8] R. Elie, T. Mastrolia, and D. Possamai. A tale of a principal and many, many agents. *Math. Oper. Res.*, 44(2):440–467, 2019. ISSN 0364-765X. doi: 10.1287/moor.2018.0931. URL <https://doi.org/10.1287/moor.2018.0931>.
- [9] H. Föllmer and D. Kramkov. Optional decompositions under constraints. *Probab. Theory Related Fields*, 109(1):1–25, 1997. ISSN 0178-8051. doi: 10.1007/s004400050122. URL <https://doi.org/10.1007/s004400050122>.
- [10] A. Galichon, P. Henry-Labordère, and N. Touzi. A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann. Appl. Probab.*, 24(1):312–336, 2014. ISSN 1050-5164. doi: 10.1214/13-AAP925. URL <https://doi.org/10.1214/13-AAP925>.
- [11] P. J. Graber, A. R. Mészáros, F. J. Silva, and D. Tonon. The planning problem in mean field games as regularized mass transport. *Calc. Var. Partial Differential Equations*, 58(3):Paper No. 115, 28, 2019. ISSN 0944-2669. doi: 10.1007/s00526-019-1561-9. URL <https://doi.org/10.1007/s00526-019-1561-9>.
- [12] B. Holmström and P. Milgrom. Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica*, 55(2):303–328, 1987. ISSN 00129682, 14680262. URL <http://www.jstor.org/stable/1913238>.
- [13] M. Huesmann and D. Trevisan. A Benamou-Brenier formulation of martingale optimal transport. *Bernoulli*, 25(4A):2729–2757, 2019. ISSN 1350-7265. doi: 10.3150/18-BEJ1069. URL <https://doi.org/10.3150/18-BEJ1069>.
- [14] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003. ISBN 3-540-43932-3. doi: 10.1007/978-3-662-05265-5. URL <https://doi.org/10.1007/978-3-662-05265-5>.
- [15] R. L. Karandikar. On pathwise stochastic integration. *Stochastic Process. Appl.*, 57(1):11–18, 1995. ISSN 0304-4149. doi: 10.1016/0304-4149(95)00002-O. URL [https://doi.org/10.1016/0304-4149\(95\)00002-O](https://doi.org/10.1016/0304-4149(95)00002-O).
- [16] P.-L. Lions. Cours au Collège de France (Lectures on Nov 27th, Dec 4th-11th, 2009). [www.college-de-france.fr](http://www.college-de-france.fr), 2009.
- [17] T. Mikami and M. Thieullen. Optimal transportation problem by stochastic optimal control. *SIAM J. Control Optim.*, 47(3):1127–1139, 2008. ISSN 0363-0129. doi: 10.1137/050631264. URL <https://doi.org/10.1137/050631264>.
- [18] M. Nutz. Pathwise construction of stochastic integrals. *Electron. Commun. Probab.*, 17:no. 24, 7, 2012. doi: 10.1214/ECP.v17-2099. URL <https://doi.org/10.1214/ECP.v17-2099>.
- [19] M. Nutz. Introduction to Entropic Optimal Transport. *Preprint*, 2022.
- [20] J. Oblój. The Skorokhod embedding problem and its offspring. *Probab. Surv.*, 1:321–390, 2004. doi: 10.1214/154957804100000060. URL <https://doi.org/10.1214/154957804100000060>.
- [21] C. Orrieri, A. Porretta, and G. Savaré. A variational approach to the mean field planning problem. *J. Funct. Anal.*, 277(6):1868–1957, 2019. ISSN 0022-1236. doi: 10.1016/j.jfa.2019.04.011. URL <https://doi.org/10.1016/j.jfa.2019.04.011>.

- [22] A. Porretta. On the planning problem for the mean field games system. *Dyn. Games Appl.*, 4(2):231–256, 2014. ISSN 2153-0785. doi: 10.1007/s13235-013-0080-0. URL <https://doi.org/10.1007/s13235-013-0080-0>.
- [23] D. Possamaï, X. Tan, and C. Zhou. Stochastic control for a class of nonlinear kernels and applications. *Ann. Probab.*, 46(1):551–603, 2018. ISSN 0091-1798. doi: 10.1214/17-AOP1191. URL <https://doi.org/10.1214/17-AOP1191>.
- [24] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999. ISBN 3-540-64325-7. doi: 10.1007/978-3-662-06400-9. URL <https://doi.org/10.1007/978-3-662-06400-9>.
- [25] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.
- [26] Y. Sannikov. A continuous-time version of the principal-agent problem. *Rev. Econom. Stud.*, 75(3):957–984, 2008. ISSN 0034-6527. doi: 10.1111/j.1467-937X.2008.00486.x. URL <https://doi.org/10.1111/j.1467-937X.2008.00486.x>.
- [27] H. M. Soner, N. Touzi, and J. Zhang. Wellposedness of second order backward SDEs. *Probab. Theory Related Fields*, 153(1-2):149–190, 2012. ISSN 0178-8051. doi: 10.1007/s00440-011-0342-y. URL <https://doi.org/10.1007/s00440-011-0342-y>.
- [28] H. M. Soner, N. Touzi, and J. Zhang. Dual formulation of second order target problems. *Ann. Appl. Probab.*, 23(1):308–347, 2013. ISSN 1050-5164. doi: 10.1214/12-AAP844. URL <https://doi.org/10.1214/12-AAP844>.
- [29] X. Tan and N. Touzi. Optimal transportation under controlled stochastic dynamics. *Ann. Probab.*, 41(5):3201–3240, 2013. ISSN 0091-1798. doi: 10.1214/12-AOP797. URL <https://doi.org/10.1214/12-AOP797>.