Steering the distribution of agents in mean-field games

Yongxin Chen, Tryphon Georgiou and Michele Pavon

Abstract— **In this work we pose and solve the problem to guide a collection of weakly interacting dynamical systems e.g., agents, to a specified target distribution. The problem is formulated using the mean-field game theory where each agent seeks to minimize its own cost. The underlying dynamics is assumed to be linear and the cost is assumed to be quadratic. In our framework a terminal cost is added as an incentive term to accomplish the task; we establish that the map between terminal costs and terminal probability distributions is onto. By adding a proper terminal cost/incentive, the agents will reach any desired terminal distribution providing they are adopting the Nash equilibrium strategy. A similar problem is considered in the cooperative game setting where the agents work together to minimize a total cost. Our approach relies on and extends the theory of optimal mass transport and its generalizations.**

I. INTRODUCTION

Mean-field game (MFG) theory provides a nice approach to study games involving a large number of agents. The basic paradigm requires agents to follow identical dynamics and seek to minimize an individual cost function that is also the same for all. When the number of agents is large, the behavior of the systems can be captured by the empirical distribution of their states. The desire to minimize individual cost drives this empirical distribution. The purpose of this work is to study the control problem to steer the collective response of agents over a finite window of time between two specified end-point marginal distributions by suitable choice of cost (i.e., incentives). We also compare it with a similar problem in a cooperative games setting where a centralized control is established to jointly optimize a common performance index. Our viewpoint is influenced by recent development of optimal mass transport (OMT) theory that deals with the control and modeling problems of a collective (agents, particles, resources) [1].

The study of MFG's was introduced into the engineering literature by Huang, Malhamé and Caines [2] and, independently, by Lasry and Lions [3]. Earlier, in the economics literature, similar models were considered by Jovanovic and Rosental [4]. The importance of the subject stems from the wide range of applications that include modeling and control of multi-agent dynamical systems, stock market dynamics, crowd dynamics and more; see [2], [3], [5], [6], and also see [7], [8], [9], [10], [11] in the special case of linear dynamics and quadratic cost. On the other hand, OMT originates in the work of Monge [12] and solved, in the relaxed form, in Kantorovich [13]. In recent years, a fast developing phase was spurred by a wide range of applications of OMT to probability theory, economics, biology and mathematical physics [14], [15], [16], [17], [18], [19], [20], [21]. The connection between OMT [22] and stochastic control has been explored in our work, e.g. [23], [24], where the focus has been on regulating uncertainty of stochastically driven dynamical systems by suitable control action. These stochastic control problems, in turn, relate to a classical maximum entropy problem on path space known as the Schrödinger bridge problem, see e.g., [25], [26], [27], [28], [29], [30], [31].

The goal of the present work is to study density steering problems in an MFG framework. In particular, we are interested in how to design an added terminal cost so as to provide incentives for agents, under a Nash equilibrium strategy, to move collectively the agents to a specified target distribution. To this end, we establish that the map between terminal costs and terminal probability distributions is an onto map. Thereby, we develop an MFG-based synthesis framework for OMT-type stochastic control problems with or without stochastic excitation.

The rest of the paper is structured as follows. First, we discuss the motivation and problem formulation in Section II. The solution is provided in Section III. In Section IV we study similar problems with less or no disturbance. Section V is dedicated to the special case with Gaussian marginal distributions. In Section VI, we developed the cooperative game counterpart of the density steering problem. This follows by a numerical example in Section VII and a brief concluding remark in Section VIII.

II. PROBLEM FORMULATION

We study the collective dynamics of a group of agents that interact weakly with each other. The terminology "weakly" refers to the agents being statistically indistinguishable (anonymous) and affecting each other's response only through their empirical distribution [32]. Thus, we consider

Y. Chen is with the Department of Electrical and Computer Engineering, Iowa State University, IA 50011, Ames, Iowa, USA; yongchen@iastate.edu
T.T. Georgiou is with the Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA 92697; tryphon@uci.edu
M. Pavon is with the Dipartimento di Matematica "Tullio Levi Civita", Università di Padova, via Trieste 63, 35121 Padova, Italy; pavon@math.unipd.it
Supported in part by the NSF under grant ECCS-1509387, the AFOSR under grants FA9550-15-1-0045 and FA9550-17-1-0435, and by the University of Padova Research Project CPDA 140897.

such a system of N agents with dynamics[1] specified by

$$\begin{aligned} dx_i(t) &= Ax_i(t)dt + \frac{1}{N-1}\sum_{j\neq i}\bar{A}x_j(t)dt \\ &\quad +Bu_i(t)dt + Bdw_i(t), \\ &\quad x_i(0) = x_0^i, \quad i = 1, \ldots, N. \end{aligned} \tag{1}$$

Here, x_i, u_i, w_i represent the state, control input, white noise disturbance, respectively, for the ith agent, and the model parameters are the same for all. We further assume that their initial conditions $x_0^1, x_0^2, \ldots, x_0^N$ are all independent with the same probability density ρ_0. The ith agent interacts with the rest through the averaged position. The matrices $A, \bar{A} \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$ are continuous functions of time; for notational simplicity we often use e.g., A instead of $A(t)$. The pair (A, B) is assumed to be controllable in the sense that the reachability Gramian

$$M(t, s) = \int_s^t \Phi(t, \tau)B(\tau)B(\tau)'\Phi(t, \tau)'d\tau$$

is invertible for all $s < t$. Here, $\Phi(\cdot, \cdot)$ denotes the state transition matrix that is defined via

$$\frac{\partial \Phi(t, s)}{\partial t} = A\Phi(t, s), \quad \Phi(s, s) = I.$$

In MFG [2], each agent searches for an optimal control strategy to minimize its own cost[2]

$$J_i(u_i) = \mathbb{E}\left\{\int_0^1 f(t, x_i(t), u_i(t), \mu^N(t))dt + g(x_i(1), \mu^N(1))\right\} \tag{2}$$

where

$$\mu^N(t) = \frac{1}{N}\sum_{i=1}^N \delta_{x_i(t)} \tag{3}$$

is the empirical distribution of the states of the N agents at time t. This is a non-cooperative game and the cost of the ith agent is affected by the strategies of others only through the empirical distribution μ^N. An optimal control corresponds to a Nash equilibrium for the game. We follow the arguments in [6], and restrict ourselves to equilibria that correspond to symmetric Markovian control strategies (state feedback)

$$u_i(t) = \phi(t, x_i(t)), i = 1, \ldots, N. \tag{4}$$

When N is large, the empirical distribution μ^N is indifferent to small perturbations of control strategy of a single agent. This observation leads to the following approach [6] to obtain an approximate Nash equilibrium: fix a family $(\mu(t))_{0\leq t\leq 1}$ of probability measures and solve the standard stochastic control problem

$$\min_\phi \mathbb{E}\left\{\int_0^1 f(t, x(t), \phi(t, x(t)), \mu(t))dt + g(x(1), \mu(1))\right\} \tag{5}$$

subject to the dynamics

$$\begin{aligned} dx(t) &= Ax(t)dt + \bar{A}\bar{x}_\mu(t)dt + B\phi(t, x(t))dt + Bdw(t), \\ x(0) &= x_0 \end{aligned} \tag{6}$$

where

$$\bar{x}_\mu(t) := \langle x, \mu(t)\rangle$$

denotes the mean[3] of the distribution $\mu(t)$, and x_0 is a random vector with probability density ρ_0. The optimal control is denoted by $\phi^\star$. Treating $(\mu(t))_{0\leq t\leq 1}$ as a parameter, the remaining issue is to choose this distribution flow so that the actual distribution of the solution $x(t)$ of (6) with optimal control strategy

$$u^\star(t) = \phi^\star(t, x(t)) \tag{7}$$

coincides with $\mu(t)$. The solution to the MFG problem involves establishing the existence and uniqueness of the solution to two coupled partial differential equations (PDEs) [6]. It has been shown that a Nash equilibrium point for this mean-field game exists under mild assumptions on the cost function [2], [3], [5], [6], [7], [8], [9], [10]. That is, there exists a family $(\mu(t))_{0\leq t\leq 1}$ such that the distribution flow of the solution $x(t)$ of (6) under optimal control strategy $\phi^\star$ coincides with this same μ. In addition, this optimal control $\phi^\star$ is proven to be an ε-Nash equilibrium to the N-player-game for N large [6], [33].

Departing from previous literature, this paper deals with the density steering problem of the N-player-game system. More specifically, we are interested in introducting a suitable cost incentive so that the system is driven to a specific distribution ρ_1 at time $t = 1$ under (7). In fact, it turns out that under mild conditions, a quadratic running cost in both the control and state (i.e., group linear tracking as in the work of Huang, Malamé and Caines [2]), can be enhanced by a suitable terminal cost g as follows

$$\begin{aligned} J_i(u_i) \;=\; \mathbb{E}\left\{\int_0^1 \left(\frac{1}{2}\|u_i(t)\|^2 + \frac{1}{2}\|x_i(t) - \bar{x}(t)\|_Q^2\right)dt \right. \\ \left. +g(x_i(1), \mu^N(1))\right\} \end{aligned} \tag{8}$$

so as to acomplish the task of steering the initial distribution to any desired terminal one. In other words, we show that the mapping between a choice of g and the terminal distribution ρ_1 is onto. Formally, the problem we are interested in can be stated as follows.

Problem 1: Given N agents governed by (1) with initial probability density ρ_0, find a terminal cost g such that, in the Nash equilibrium with cost functional (8), the agents will reach a given terminal density ρ_1 at time $t = 1$, in the limit as N goes to ∞.

[1]This type of weakly coupled system of linear stochastic models has been studied in [7], [8], [9]. In our setting we further assume that the noise dw_i and control action u affect the dynamics in a similar manner, through the same matrix B. The more general case, where this is not so, is more demanding and will be pursued in future publication, cf. [29], [30].

[2]For simplicity of notation and without loss in generality we take the end point to be $t = 1$.

[3]Throughout, we use the expressions $\bar{x}_\mu(t)$ or $\langle x, \mu(t)\rangle$ interchangeably.

III. General approach and solution

Without loss of generality and for simplicity of exposition we consider a running cost depending only on the control actuation (i.e., taking the matrix Q in (8) to be zero). We begin with the optimal steering problem [1], [29], [30], [34] without terminal cost. In particular, for a fixed density flow $(\mu(t))_{0\le t\le 1}$, we consider the control problem to minimize

$$J(u)=\mathbb{E}\left\{\int_0^1 \frac{1}{2}\|u(t)\|^2 dt\right\}$$

subject to the dynamics

$$\begin{aligned} dx(t) &= Ax(t)dt+\bar{A}\bar{x}_\mu(t)dt+Bu(t)dt+Bdw(t), \quad (9)\\ & x(0)=x_0\sim\rho_0 \end{aligned}$$

and the constraint that $x(1)$ has probability density ρ_1. This problem can be posed as

$$\inf_{\rho,u} \quad \int_0^1\int_{\mathbb{R}^n}\frac{1}{2}\rho(t,x)\|u(t,x)\|^2 dxdt, \tag{10a}$$

$$\frac{\partial\rho}{\partial t}+\nabla\cdot((Ax+\bar{A}\bar{x}_\mu+Bu)\rho)-\frac{1}{2}\operatorname{tr}(BB'\nabla^2\rho)=0, \tag{10b}$$

$$\rho(0,\cdot)=\rho_0, \quad \rho(1,\cdot)=\rho_1. \tag{10c}$$

Following a similar argument as in [24], we establish the following sufficient condition for optimality.

Proposition 1: If there exist functions $\rho^\star,\lambda$ satisfying

$$\frac{\partial\lambda}{\partial t}+\nabla\lambda\cdot Ax+\nabla\lambda\cdot\bar{A}\bar{x}_\mu+\frac{1}{2}\operatorname{tr}(BB'\nabla^2\lambda)+\frac{1}{2}\nabla\lambda\cdot BB'\nabla\lambda=0, \tag{11a}$$

$$\frac{\partial\rho^\star}{\partial t}+\nabla\cdot((Ax+\bar{A}\bar{x}_\mu+BB'\nabla\lambda)\rho^\star)-\frac{1}{2}\operatorname{tr}(BB'\nabla^2\rho^\star)=0, \tag{11b}$$

and boundary conditions

$$\rho^\star(0,\cdot)=\rho_0, \quad \rho^\star(1,\cdot)=\rho_1, \tag{11c}$$

then $(\rho^\star, u^\star=B'\nabla\lambda)$ is a solution to (10).

Replacing μ in (11) by $\rho^\star$ we obtain the system of (nonlinear) PDE's

$$\frac{\partial\lambda}{\partial t}+\nabla\lambda\cdot Ax+\nabla\lambda\cdot\bar{A}\bar{x}_{\rho^\star}+\frac{1}{2}\operatorname{tr}(BB'\nabla^2\lambda)+\frac{1}{2}\nabla\lambda\cdot BB'\nabla\lambda=0, \tag{12a}$$

$$\frac{\partial\rho^\star}{\partial t}+\nabla\cdot((Ax+\bar{A}\bar{x}_{\rho^\star}+BB'\nabla\lambda)\rho^\star)-\frac{1}{2}\operatorname{tr}(BB'\nabla^2\rho^\star)=0, \tag{12b}$$

$$\rho^\star(0,\cdot)=\rho_0, \quad \rho^\star(1,\cdot)=\rho_1. \tag{12c}$$

Remark 2: Note that the coupled PDEs (12a-12b) are the same as the PDEs that arise in classic MFG problems corresponding to (1) and (8). However, the usual boundary conditions

$$\rho^\star(0,\cdot)=\rho_0, \quad \lambda(1,x)=-g(x,\rho^\star(1,\cdot)),$$

are now different and given by (12c). Evidently, the Lagrange multiplier $-\lambda$ is the value (cost-to-go) function of the associated optimal control problem.

It can be shown that (12) has a solution under the assumption that the "coupled" Gramian

$$\bar{M}_{10}=\int_0^1\bar{\Phi}(1,\tau)BB'\Phi(1,\tau)'d\tau \tag{13}$$

is invertible, where $\bar{\Phi}$ is the state transition matrices for the pair $(A+\bar{A},B)$. The proof relies on the standard Schrödinger bridge theory [23], [24] and can be found in a forth coming paper [35].

Finally, back to Problem 1, we assert that with terminal cost

$$g(x,\mu)=-\lambda(1,x), \tag{14}$$

we can guide the agents to have terminal distribution ρ_1. To this extent, we follow the strategy in [6] as mentioned in Section II. First fix $\mu=\rho^\star$ with $\rho^\star$ being the solution to (12), and then solve the optimal control problem (5). Since $g(x,\rho^\star(1,\cdot))=g(x,\rho_1)=-\lambda(1,x)$, we have

$$\begin{aligned} &\mathbb{E}\left\{\int_0^1\frac{1}{2}\|u(t)\|^2dt+g(x,\rho^\star(1,\cdot))\right\}\\ =\;& \mathbb{E}\left\{\int_0^1[\frac{1}{2}\|u(t)\|^2dt-d\lambda(t,x(t))]-\lambda(0,x(0))\right\}\\ =\;& \mathbb{E}\left\{\int_0^1[\frac{1}{2}\|u(t)\|^2dt-\frac{\partial\lambda}{\partial t}dt-\nabla\lambda\cdot dx(t)\right.\\ &\left.-\frac{1}{2}\operatorname{tr}(BB'\nabla^2\lambda)dt]-\lambda(0,x(0))\right\}\\ =\;& \mathbb{E}\left\{\int_0^1\frac{1}{2}\|u(t)-B'\nabla\lambda(t,x(t))\|^2dt\right\}-\mathbb{E}\{\lambda(0,x(0)\}. \end{aligned}$$

Hence, the unique optimal control strategy is $u^\star(t)=B'\nabla\lambda(t,x(t))$. It follows from (12) that the probability distribution of the controlled state $x(t)$ is $\rho^\star$. Therefore, with terminal cost g as in (14) we are able to steer the system to terminal distribution ρ_1. Thus, we have established the following result.

Theorem 3: Consider N agents governed by (1) with initial density ρ_0. Suppose the terminal cost in (8) is as in (14). Then, in the Nash equilibrium, the agents will reach density ρ_1 at time $t=1$, in the mean-field game limit.

IV. Zero-noise limit

In this section, we study the same problem (Problem 1), with however reduced disturbance. More specifically, we consider a system of N agents with dynamics

$$\begin{aligned} dx_i(t) &= Ax_i(t)dt+\frac{1}{N-1}\sum_{j\neq i}\bar{A}x_j(t)dt+Bu_i(t)dt \quad (15)\\ &\quad +\sqrt{\epsilon}Bdw_i(t),\\ &\quad x_i(0)=x_0^i, \quad i=1,\ldots,N, \end{aligned}$$

where $\epsilon > 0$ represents the variance of the noise. We are especially interested in the limit behavior of the solution to Problem 1 with dynamics (15) when ϵ goes to 0. Following the same arguments as in Section III, we arrive at the coupled PDEs

$$\frac{\partial\lambda}{\partial t}+\nabla\lambda\cdot Ax+\nabla\lambda\cdot\bar{A}\bar{x}_{\rho^\star}+\frac{\epsilon}{2}\operatorname{tr}(BB'\nabla^2\lambda)+\frac{1}{2}\nabla\lambda\cdot BB'\nabla\lambda=0, \tag{16a}$$

$$\frac{\partial\rho^\star}{\partial t}+\nabla\cdot((Ax+\bar{A}\bar{x}_{\rho^\star}+BB'\nabla\lambda)\rho^\star)-\frac{\epsilon}{2}\operatorname{tr}(BB'\nabla^2\rho^\star)=0, \tag{16b}$$

$$\rho^\star(0,\cdot)=\rho_0, \quad \rho^\star(1,\cdot)=\rho_1. \tag{16c}$$

The optimal control strategy is given by $u(t) = B'\nabla\lambda(t,x(t))$ and terminal cost g is as in (14) with adjusted diffusivity.

Taking the limit of (16) as $\epsilon\to 0$ gives

$$\frac{\partial\lambda}{\partial t}+\nabla\lambda\cdot Ax+\nabla\lambda\cdot\bar{A}\bar{x}_{\rho^\star}+\frac{1}{2}\nabla\lambda\cdot BB'\nabla\lambda=0, \tag{17a}$$

$$\frac{\partial\rho^\star}{\partial t}+\nabla\cdot((Ax+\bar{A}\bar{x}_{\rho^\star}+BB'\nabla\lambda)\rho^\star)=0, \tag{17b}$$

$$\rho^\star(0,\cdot)=\rho_0, \quad \rho^\star(1,\cdot)=\rho_1. \tag{17c}$$

Again it can be shown that the above PDEs system has a (viscosity) solution [36], [35]. The solution to (17) in fact solves the following problem.

Problem 4: Given N agents governed by (15) with $\epsilon = 0$, and initial probability density ρ_0, find a function g such that, in the Nash equilibrium with cost function (8), the agents would reach a specified density ρ_1 at time $t = 1$, in the limit as N goes to ∞.

With the solution to (17), we can choose a terminal cost as in (14). The corresponding equilibrium control strategy is again $u(t,x)=B'\nabla\lambda(t,x)$.

Theorem 5: Consider N agents governed by (15) with $\epsilon = 0$ and initial density ρ_0. Suppose the terminal cost g in (8) is as in (14), then, in the Nash equilibrium, the agents will reach density ρ_1 at time $t = 1$, in the mean-field game limit.

V. GAUSSIAN CASE

In case when ρ_0 and ρ_1 are normal (Gaussian) distributions, the solutions have a nice linear structure. Let the two marginal distributions be

$$\rho_0\sim\mathcal{N}[m_0,\Sigma_0], \quad \rho_1\sim\mathcal{N}[m_1,\Sigma_1],$$

i.e., Gaussian distributions with, respectively, means m_0, m_1 and covariances Σ_0, Σ_1. The solution of (12) has the form

$$\lambda(t,x)=\hat{\lambda}(t,x-y(t))+m(t)\cdot x+\gamma(t), \tag{18}$$

where

$$m(t)=\Phi(1,t)'\bar{M}_{10}^{-1}(m_1-\bar{\Phi}_{10}m_0),$$

$y(t)$ is the solution to

$$\dot{y}(t)=(A+\bar{A})y(t)+BB'm(t), \quad y(0)=m_0,$$

and

$$\gamma(t)=-\int_0^t(\bar{A}y(s)\cdot m(s)+\frac{1}{2}m(s)\cdot BB'm(s))ds.$$

Recall the definition of $\bar{M}_{10}$ and $\bar{\Phi}_{10}$ in (13).

When $\epsilon = 1$, $\hat{\lambda}$ equals

$$\hat{\lambda}(t,x)=-\frac{1}{2}x'\Pi(t)x+\frac{1}{2}\int_0^t\operatorname{tr}(BB'\Pi(s))ds,$$

where $\Pi(t)$ is the solution to the Riccati equation

$$\dot{\Pi}(t)=-A'\Pi(t)-\Pi(t)A+\Pi(t)BB'\Pi(t) \tag{19}$$

with boundary condition

$$\Pi(0)=\Sigma_0^{-1/2}\left[\frac{I}{2}+\Sigma_0^{1/2}\Phi_{10}'M_{10}^{-1}\Phi_{10}\Sigma_0^{1/2}\right.$$
$$\left.-(\frac{I}{4}+\Sigma_0^{1/2}\Phi_{10}'M_{10}^{-1}\Sigma_1M_{10}^{-1}\Phi_{10}\Sigma_0^{1/2})^{1/2}\right]\Sigma_0^{-1/2}.$$

where $\Phi_{10}=\Phi(1,0), M_{10}=M(1,0)$. And so, in view of (14), one choice of terminal cost is

$$g(x,\mu)=\frac{1}{2}(x-m_1)'\Pi(1)(x-m_1)-m(1)\cdot x. \tag{20}$$

In the above we have discarded some constant terms $-\lambda(1,x)$ as it doesn't affect the final result.

Theorem 6: Consider N agents governed by (1) with initial density $\rho_0\sim\mathcal{N}[m_0,\Sigma_0]$. Suppose the terminal cost in (8) is (20). Then, in the Nash equilibrium, the agents will reach density $\rho_1\sim\mathcal{N}[m_1,\Sigma_1]$ at time $t = 1$, in the mean-field game limit.

Following the discussion in Section IV, the solution to the problem with noise intensity ϵ is almost identical to the above except that, the initial condition of the Riccati equation (19) becomes

$$\Pi_\epsilon(0)=\Sigma_0^{-1/2}\left[\frac{\epsilon I}{2}+\Sigma_0^{1/2}\Phi_{10}'M_{10}^{-1}\Phi_{10}\Sigma_0^{1/2}\right.$$
$$\left.-(\frac{\epsilon^2 I}{4}+\Sigma_0^{1/2}\Phi_{10}'M_{10}^{-1}\Sigma_1M_{10}^{-1}\Phi_{10}\Sigma_0^{1/2})^{1/2}\right]\Sigma_0^{-1/2}.$$

Taking the limit as $\epsilon\to 0$ we obtain the solution to the deterministic problem, which corresponds to the initial condition

$$\Pi_0(0)=\Sigma_0^{-1/2}\left[\Sigma_0^{1/2}\Phi_{10}'M_{10}^{-1}\Phi_{10}\Sigma_0^{1/2}\right.$$
$$\left.-(\Sigma_0^{1/2}\Phi_{10}'M_{10}^{-1}\Sigma_1M_{10}^{-1}\Phi_{10}\Sigma_0^{1/2})^{1/2}\right]\Sigma_0^{-1/2}.$$

VI. Cooperative game

We next shift to a slightly different problem. Given the same interacting agents' system (1), we would like to investigate the density steering problem in the cooperative game setting. How to select an optimal controller to drive the agents from given initial distribution ρ_0 to terminal distribution ρ_1? Again, we restrict ourself to equilibriums given by symmetric Markovian strategies in closed-loop feedback form

$$u_i(t) = \phi(t, x_i(t)), i = 1, \ldots, N. \tag{21}$$

The cost function we attempt to minimize is the average control energy

$$J(u) = \mathbb{E}\left\{\frac{1}{N}\sum_{i=1}^{N}\int_0^1 \frac{1}{2}\|u_i(t)\|^2 dt\right\}. \tag{22}$$

We are interested in the mean-field limit, namely, the asymptotical behavior of the solution when $N \to \infty$.

Problem 7: Given N agents governed by (1) with initial density ρ_0, find a control strategy (21) with minimum control energy (22) so that the agents will reach density ρ_1 at time $t = 1$, as N goes to ∞.

The major difference between this problem and the mean-field game is that all the agents always use the same control strategy. A small perturbation on the control will affect the probability density flow as the perturbation is applied to the controllers of all the agents, see [6], [37] for more discussions on their differences. The average control energy (22) is equivalent to relative entropy of the controller system with respect to the uncontrolled system [38], [39], [32]. Therefore, the above problem can also be viewed as an Schrödinger bridge problem for interacting particle systems.

Problem 7 can be formulated as an optimal control problem over the McKean-Vlasov model

$$\begin{aligned} dx(t) &= Ax(t)dt + \bar{A}\bar{x}(t)dt + Bu(t)dt + Bdw(t), \\ & x(0) = x_0 \sim \rho_0. \end{aligned} \tag{23}$$

It has the following fluid dynamic formulation. Let $\rho(t, \cdot)$ be the probability density of the controlled process $x(t)$, then the optimal control problem can be stated as

$$\inf_{\rho,u} \quad \int_0^1 \int_{\mathbb{R}^n} \frac{1}{2}\rho(t,x)\|u(t,x)\|^2 dxdt, \tag{24a}$$

$$\frac{\partial \rho}{\partial t} + \nabla\cdot((Ax + \bar{A}\bar{x}_\rho + Bu)\rho) - \frac{1}{2}\operatorname{tr}(BB'\nabla^2\rho) = 0, \tag{24b}$$

$$\rho(0,\cdot) = \rho_0, \quad \rho(1,\cdot) = \rho_1. \tag{24c}$$

Proposition 2: If there exists $(\lambda, \rho^\star)$ satisfying

$$\begin{aligned}\frac{\partial \lambda}{\partial t} &+ \nabla\lambda\cdot Ax + \nabla\lambda\cdot\bar{A}\bar{x}_{\rho^\star} + \frac{1}{2}\operatorname{tr}(BB'\nabla^2\lambda) \\ &+ \frac{1}{2}\nabla\lambda\cdot BB'\nabla\lambda + \bar{A}x\cdot\langle\nabla\lambda, \rho^\star\rangle = 0, \end{aligned} \tag{25a}$$

$$\frac{\partial \rho^\star}{\partial t} + \nabla\cdot((Ax + \bar{A}\bar{x}_{\rho^\star} + BB'\nabla\lambda)\rho^\star) - \frac{1}{2}\operatorname{tr}(BB'\nabla^2\rho^\star) = 0, \tag{25b}$$

with boundary conditions

$$\rho^\star(0,\cdot) = \rho_0, \quad \rho^\star(1,\cdot) = \rho_1, \tag{25c}$$

then $(\rho^\star, u^\star = B'\nabla\lambda)$ is a solution to (24).

Equations (25) are highly coupled. In general, one may not expect a solution to exist. But interestingly, it can be shown that (25) always has a solution, under the assumption that the pair $(A + \bar{A}, B)$ is controllable. The proof that this pair $(\rho^\star, u^\star)$ provides a solution to the optimal control problem (24) can be found in a forth coming paper [35].

Remark 8: The sufficient condition, the pair $(A+\bar{A}, B)$ is controllable, that (25) has a solution is different to that of (12), which is the "coupled" Gramian (13) being invertible.

VII. Examples

Consider N agents with dynamics

$$dx_i(t) = x_i(t)dt - \frac{2}{N-1}\sum_{j\neq i} x_j(t)dt + dw_i(t), \; 1 \le i \le N.$$

The two marginal distributions ρ_0 and ρ_1 are two normal distributions

$$\rho_0 \sim \mathcal{N}[1, 4], \quad \rho_1 \sim \mathcal{N}[-4, 1].$$

A. Noncooperative game

One choice of terminal cost that will steer the agents from ρ_0 to ρ_1 is

$$g(x,\mu) = 0.9805(x - \bar{x}_\mu)^2 + 4.3679x.$$

Figure 1 showcases the evolution of the probability density in the Nash equilibrium. To show that the distribution of the agents would evolve according to Figure 1, we simulated the dynamics for a system with $N = 20000$ agents under the optimal strategy. Figure 2 and Figure 3 depict the empirical distributions of the particles at time $t = 0$ and $t = 1$. They match with the theoretical distributions ρ_0 and ρ_1 very well. We also show the empirical mean of these particles in Figure 4, which perfectly matches the theoretical result.

B. Cooperative game

Figure 5 depicts the time evolution of the probability densities with these two marginal distributions in the cooperative game setting. Similarly, we ran some simulations for a particle system with $N = 20000$ and obtained Figure 6 and Figure 7 as the empirical distributions of the agents at time $t = 0$ and $t = 1$. We also show the empirical mean of these particles in Figure 8. Clearly the mean is different to the Nash equilibrium in the noncooperative game setting.

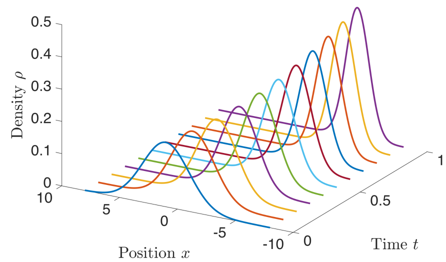

Fig. 1: Time evolution of probability densities

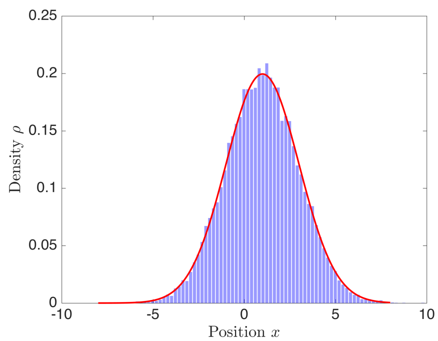

Fig. 2: Empirical distribution of $x(0)$

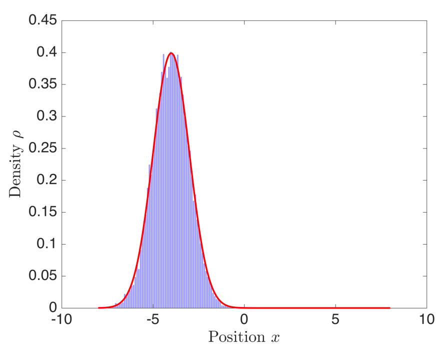

Fig. 3: Empirical distribution of $x(1)$

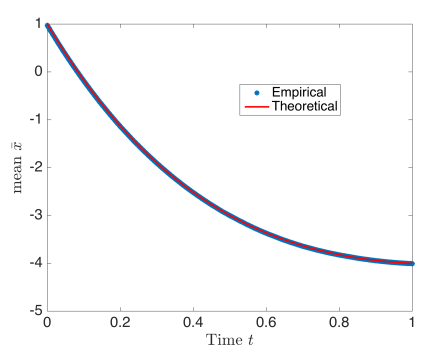

Fig. 4: Time evolution of mean $\bar{x}(t)$

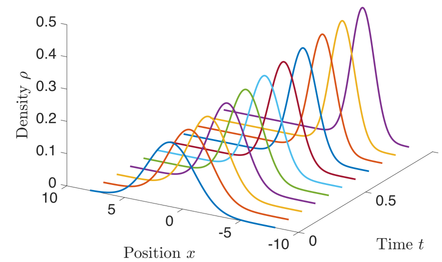

Fig. 5: Time evolution of probability densities

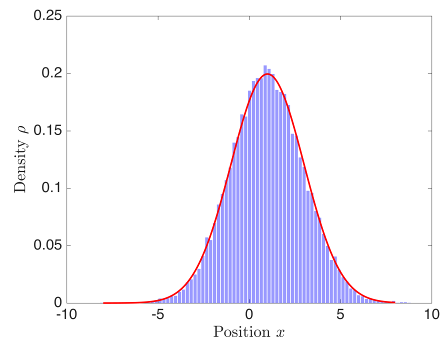

Fig. 6: Empirical distribution of $x(0)$

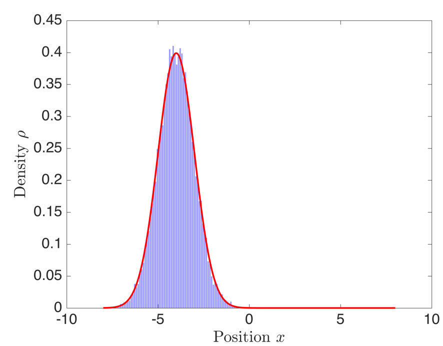

Fig. 7: Empirical distribution of $x(1)$

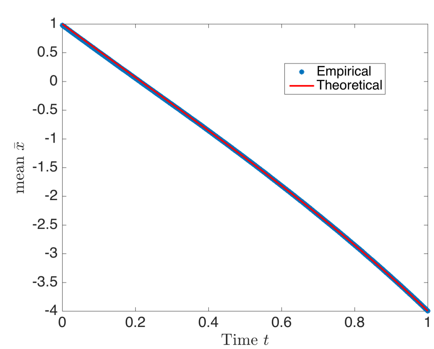

Fig. 8: Time evolution of mean $\bar{x}(t)$

VIII. Conclusion

We introduce a paradigm to steer a large number of agents from one distribution to another. The problem lies in the intersection of MFG, OMT and optimal control. We study such problems for linearly weakly interacting agents under quadratic running cost. Results for the cooperative game counterpart are also presented. We expect this paradigm to bring in a new dimension to the study of MFG and OMT and to be useful in social science such as economics.

References

[1] Y. Chen, "Modeling and control of collective dynamics: From Schrödinger bridges to optimal mass transport," Ph.D. dissertation, University of Minnesota, 2016.

[2] M. Huang, R. P. Malhamé, and P. E. Caines, "Large population stochastic dynamic games: closed-loop mckean-vlasov systems and the nash certainty equivalence principle," *Communications in Information & Systems*, vol. 6, no. 3, pp. 221–252, 2006.

[3] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.

[4] B. Jovanovic and R. W. Rosenthal, "Anonymous sequential games," *Journal of Mathematical Economics*, vol. 17, no. 1, pp. 77–87, 1988.

[5] M. Nourian and P. E. Caines, "ε-Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents," *SIAM Journal on Control and Optimization*, vol. 51, no. 4, pp. 3302–3331, 2013.

[6] R. Carmona, F. Delarue, and A. Lachapelle, "Control of McKean–Vlasov dynamics versus mean field games," *Mathematics and Financial Economics*, vol. 7, no. 2, pp. 131–166, 2013.

[7] M. Huang, P. E. Caines, and R. P. Malhamé, "Individual and mass behaviour in large population stochastic wireless power control problems: centralized and nash equilibrium solutions," in *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, vol. 1. IEEE, 2003, pp. 98–103.

[8] ——, "Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε-Nash equilibria," *Automatic Control, IEEE Transactions on*, vol. 52, no. 9, pp. 1560–1571, 2007.

[9] A. Bensoussan, K. Sung, S. C. P. Yam, and S.-P. Yung, "Linear-quadratic mean field games," *Journal of Optimization Theory and Applications*, vol. 169, no. 2, pp. 496–529, 2016.

[10] M. Bardi, "Explicit solutions of some linear-quadratic mean field games," *Networks and Heterogeneous Media*, vol. 7, no. 2, pp. 243–261, 2012.

[11] J. Moon and T. Başar, "Linear quadratic risk-sensitive and robust mean field games," *IEEE Transactions on Automatic Control*, vol. 62, no. 3, pp. 1062–1077, 2017.

[12] G. Monge, *Mémoire sur la théorie des déblais et des remblais*. De l'Imprimerie Royale, 1781.

[13] L. V. Kantorovich, "On the transfer of masses," in *Dokl. Akad. Nauk. SSSR*, vol. 37, no. 7-8, 1942, pp. 227–229.

[14] W. Gangbo and R. J. McCann, "The geometry of optimal transportation," *Acta Mathematica*, vol. 177, no. 2, pp. 113–161, 1996.

[15] L. C. Evans, "Partial differential equations and Monge-Kantorovich mass transfer," *Current developments in mathematics*, vol. 1997, no. 1, pp. 65–126, 1997.

[16] L. C. Evans and W. Gangbo, *Differential equations methods for the Monge-Kantorovich mass transfer problem*. American Mathematical Soc., 1999, vol. 653.

[17] C. Villani, *Topics in Optimal Transportation*. American Mathematical Soc., 2003, no. 58.

[18] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows: in metric spaces and in the space of probability measures*. Springer, 2006.

[19] C. Villani, *Optimal Transport: Old and New*. Springer, 2008, vol. 338.

[20] L. Ambrosio and N. Gigli, "A user's guide to optimal transport," in *Modelling and optimisation of flows on networks*. Springer, 2013, pp. 1–155.

[21] F. Santambrogio, "Optimal transport for applied mathematicians," *Birkäuser, NY*, 2015.

[22] J.-D. Benamou and Y. Brenier, "A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem," *Numerische Mathematik*, vol. 84, no. 3, pp. 375–393, 2000.

[23] Y. Chen, T. T. Georgiou, and M. Pavon, "On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint," *Journal of Optimization Theory and Applications*, vol. 169, no. 2, pp. 671–691, 2016.

[24] ——, "Optimal transport over a linear dynamical system," *IEEE Transactions on Automatic Control*, vol. 62, no. 5, pp. 2137–2152, 2017.

[25] P. Dai Pra, "A stochastic control approach to reciprocal diffusion processes," *Applied mathematics and Optimization*, vol. 23, no. 1, pp. 313–329, 1991.

[26] C. Léonard, "From the Schrödinger problem to the Monge–Kantorovich problem," *Journal of Functional Analysis*, vol. 262, no. 4, pp. 1879–1920, 2012.

[27] ——, "A survey of the Schrödinger problem and some of its connections with optimal transport," *Dicrete Contin. Dyn. Syst. A*, vol. 34, no. 4, pp. 1533–1574, 2014.

[28] I. Gentil, C. Léonard, and L. Ripani, "About the analogy between optimal transport and minimal entropy," *arXiv:1510.08230*, 2015.

[29] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal steering of a linear stochastic system to a final probability distribution, Part I," *IEEE Trans. on Automatic Control*, vol. 61, no. 5, pp. 1158–1169, 2016.

[30] ——, "Optimal steering of a linear stochastic system to a final probability distribution, Part II," *IEEE Trans. on Automatic Control*, vol. 61, no. 5, pp. 1170–1180, 2016.

[31] ——, "Fast cooling for a system of stochastic oscillators," *Journal of Mathematical Physics*, vol. 56, no. 11, p. 113302, 2015.

[32] M. Fischer, "On the form of the large deviation rate function for the empirical measures of weakly interacting systems," *Bernoulli*, vol. 20, no. 4, pp. 1765–1801, 2014.

[33] P. Cardaliaguet, "Notes on mean field games," Technical report, Tech. Rep., 2010.

[34] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal steering of a linear stochastic system to a final probability distribution, Part III," *IEEE Trans. on Automatic Control*, in print, 2017.

[35] ——, "Steering the distribution of agents in mean-field and cooperative games," *arXiv preprint arXiv:1712.03578*, 2017.

[36] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*. Springer Science & Business Media, 2006, vol. 25.

[37] R. Carmona and F. Delarue, "Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics," *The Annals of Probability*, vol. 43, no. 5, pp. 2647–2700, 2015.

[38] D. A. Dawson and J. Gärtner, "Large deviations from the McKean-Vlasov limit for weakly interacting diffusions," *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 20, no. 4, pp. 247–308, 1987.

[39] J. Feng and T. G. Kurtz, *Large deviations for stochastic processes*. American Mathematical Society Providence, 2006, vol. 131.