



Steering the Distribution of Agents in Mean-Field Games System

Yongxin Chen¹ · Tryphon T. Georgiou² · Michele Pavon³

Received: 10 December 2017 / Accepted: 28 July 2018 / Published online: 7 August 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

The purpose of this work is to pose and solve the problem to guide a collection of weakly interacting dynamical systems (agents, particles, etc.) to a specified terminal distribution. This is formulated as a mean-field game problem, and is discussed in both non-cooperative games and cooperative games settings. In the non-cooperative games setting, a terminal cost is used to accomplish the task; we establish that the map between terminal costs and terminal probability distributions is onto. In the cooperative games setting, the goal is to find a common optimal control that would drive the distribution of the agents to a targeted one. We focus on the cases when the underlying dynamics is linear and the running cost is quadratic. Our approach relies on and extends the theory of optimal mass transport and its generalizations.

Keywords Mean-field games · Linear stochastic systems · Weakly interacting particle system · McKean–Vlasov dynamics · Optimal control

Mathematics Subject Classification 91A13 · 60J60 · 49J20

Communicated by Qianchuan Zhao.

✉ Tryphon T. Georgiou
tryphon@uci.edu

Yongxin Chen
yongchen@iastate.edu

Michele Pavon
pavon@math.unipd.it

¹ Iowa State University, IA 50011 Ames, IA, USA

² University of California, Irvine, CA 92697, USA

³ Università di Padova, via Trieste 63, 35121 Padova, Italy

1 Introduction

Mean-field game (MFG) theory is the study of games involving a large number of agents. In the non-cooperative games setting, the basic model requires agents to abide by identical dynamics and seek to minimize an individual cost function that is also the same for all. As the number of agents increases to infinity, the empirical distribution of their states becomes indifferent to the strategy of any single agent and yet it couples their individual responses. Thus, the aggregate response of the agents (mean field) drives individual responses while the action of individual agents is insignificant. On the flip side, in the cooperative games setting, the agents seek to jointly optimize a common performance index. Either way, the desire to minimize cost, individually or collectively, drives the empirical distribution of agents in suitable ways. The purpose of this work to study for both, non-cooperative and cooperative MFGs, 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) in non-cooperative games and centralized control with cooperative agents, and also the problem to ensure a desired stationary distribution under similar conditions. This viewpoint is influenced by optimal mass transport (OMT) theory that deals with the flow of time-marginal densities for a collective (agents, particles, resources) and corresponding control and modeling problems.

The study of MFG was introduced into the engineering literature by Huang et al. [1] and, independently, by Lasry and Lions [2]. Earlier, in the economics literature, similar models were considered by Jovanovic and Rosental [3]. 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, power systems and more; see [1,2,4,5], and also see [6–12] in the special case of linear dynamics and quadratic cost. The cooperative games counterparts of MFG have been studied in [5,13,14], and are usually known as the optimal control of McKean–Vlasov dynamics.

The other important component, OMT, originates in the work of Monge [15] and aims directly at relating/transporting distributions under minimum cost. Kantorovich [16] introduced linear programming and duality theory for solving OMT resource allocation problems and, in recent years, a fast developing phase was spurred by a wide range of applications of OMT to probability theory, economics, weather modeling, biology and mathematical physics [17–24]. The connection between dynamic OMT [25] and stochastic control has been explored in our work, e.g., [26,27], where the focus has been on regulating stochastic uncertainty of diffusion processes and 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., [28–34].

One goal of the present work is to study density-steering problems in an MFG framework or, equivalently, explore the role of interaction potential and decentralized strategies when steering an initial distribution to a terminal one. In particular, we are interested on how to design an added terminal cost so as to provide incentives for agents, under a Nash equilibrium strategy, to move collectively as specified. To this end, we establish that the map between terminal costs and terminal probability

distributions is onto, under the assumptions that the underlying dynamics is linear and the running cost is quadratic. Thereby, we develop an MFG-based synthesis framework for OMT-type stochastic control problems with or without stochastic excitation.

The paper evolves along the following lines. First, we discuss the motivation and problem formulation in Sect. 2. The solution is provided in Sect. 3. In Sect. 4 we study similar problems with less or no disturbance. Section 5 is dedicated to the special case with Gaussian marginal distributions. Similar problems in the stationary setting are investigated in Sect. 6. In Sect. 7, we developed the cooperative game counterpart of the density-steering problem. This follows by a simple academic example in Sect. 8 and a brief concluding remark in Sect. 9.

2 Problem Formulation

We herein investigate the collective dynamical response of a group of agents (also thought of as particles, players, and so on) 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 [35]. Thus, we consider such a system of N agents with dynamics¹ specified by

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

Here, x_i , u_i , w_i represent the state, control input, white noise disturbance, respectively, for the i th agent, and the model parameters are the same for all. We further assume that their initial conditions $x_0^1, x_0^2, \dots, x_0^N$ are all independent with the same probability density ρ_0 . The i th 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 $A(t)$. It is assumed that the pair is controllable in the sense that the reachability Gramian

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

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

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

Clearly, in case when A is time-invariant, $\Phi(t, s) = e^{A(t-s)}$.

¹ This type of weakly coupled system of linear stochastic models has been studied in [6,7,9]. 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. [32,33].

In non-cooperative MFG [1], each agent searches for an optimal control strategy to minimize its own cost²

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

where

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

is the empirical distribution of the states of the N agents at time t . Thus, this is a non-cooperative game and the cost of the i th agent is affected by the strategies of others only through the empirical distribution. An optimal control corresponds to a Nash equilibrium for the game. We follow the arguments in [5] and restrict ourselves to equilibria that correspond to symmetric Markovian control strategies (state feedback)

$$u_i(t) = \phi(t, x_i(t)), i = 1, \dots, N. \quad (4)$$

When N is large, the empirical distribution μ^N is indifferent to small perturbations of control strategy of a single agent. This points to the following approach [5] 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

$$\phi^* = \operatorname{argmin}_{\phi} \mathbb{E} \left\{ \int_0^1 f(t, x(t), \phi(t, x(t)), \mu(t)) dt + g(x(1), \mu(1)) \right\} \quad (5)$$

subject to the dynamics

$$dx(t) = Ax(t)dt + \bar{A}\bar{x}_{\mu}(t)dt + B\phi(t, x(t))dt + Bd w(t), \quad x(0) = x_0 \quad (6)$$

where

$$\bar{x}_{\mu} := \langle x, \mu(t) \rangle = \int_{\mathbb{R}^n} x \mu(dx)$$

denotes the mean³ of the distribution $\mu(t)$, and x_0 is a random vector with probability density ρ_0 .

Considering the choice $(\mu(t))_{0 \leq t \leq 1}$ 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

$$u^*(t) = \phi^*(t, x(t)) \quad (7)$$

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

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 $\bar{x}_{\mu}$ or $\langle x, \mu(t) \rangle$ interchangeably.

(PDEs) [5]. It has been shown that a Nash equilibrium point for this mean-field game exists under rather mild assumptions on the cost function [1,2,4–9]. 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 ϕ^* coincides with this same μ . In addition, this optimal control ϕ^* is proven to be an ε -Nash equilibrium to the N -player-game for N large [5,36].

Departing from previous literature, this paper deals with the density-steering problem of the N -player-game system. More specifically, we are interested in introducing a suitable cost incentive so that the system is driven to a specific distribution μ_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 et al. [1]), can be enhanced by a suitable terminal cost g as follows

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

so as to accomplish the task of steering the initial distribution to the desired terminal one. In other words, we show that the mapping between a choice of g and the terminal distribution μ_1 is onto, under the assumption that μ_0, μ_1 have finite second-order moments. Formally, the problem we are interested in can be stated as follows.

Problem 2.1 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 ∞ .

After this work was completed, it was brought to our attention by a referee that our framework coincides with the mean-field game *planning problem* (MFGP), which was introduced by P-L Lions in his lecture notes at College de France and later studied in [37,38]. However, the precise problem we study differs from those in [37,38] in two ways. First, the diffusion term in (1) can be degenerate, in the sense that B is not required to be square and nonsingular. The presence of a degenerate diffusion violates the assumptions in [37,38] where the proofs rely on the regularity of the Laplacian operator Δ . Second, Eq. (1) includes a weak interaction term which is not covered in [37,38]. Last but not least, our solution is “explicit” and does not require an iterative scheme. This makes it suitable for direct implementation.

3 General Approach and Solution

For the simplicity of exposition we focus on the case where the running cost depends only on the control actuation (i.e., taking the matrix Q in (8) to be zero). The extension to general $Q \geq 0$ is straightforward, as can be seen in the discussion at the end of this section.

We begin by considering the optimal steering problem [32,33,39,40] without terminal cost, i.e., for a fixed density flow $(\mu(t))_{0 \leq t \leq 1}$, 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

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

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

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

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

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

Following a similar argument as in [27], we establish the following sufficient condition for optimality.

Proposition 3.1 *If there exists a function λ such that ρ^*, λ satisfy*

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

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

and boundary conditions

$$\rho^*(0, \cdot) = \rho_0, \quad \rho^*(1, \cdot) = \rho_1, \quad (11c)$$

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

Replacing μ in (11) by ρ^* we obtain the system of (nonlinear) PDE's

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

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

$$\rho^*(0, \cdot) = \rho_0, \quad \rho^*(1, \cdot) = \rho_1. \quad (12c)$$

We will show below that, under mild assumptions, (12) has a solution. In fact, we will construct such a solution using standard Schrödinger bridge theory [26,27].

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

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

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.

To solve (12), we first consider the Schrödinger bridge problem with prior dynamics

$$dx(t) = Ax(t)dt + Bdw(t). \quad (13)$$

We remark that the Schrödinger bridge can be viewed as a regularized OMT problem [27,29,30]. Let

$$\hat{\rho}_0(x) = \rho_0(x + \bar{x}_{\rho_0})$$

and

$$\hat{\rho}_1(x) = \rho_1(x + \bar{x}_{\rho_1}),$$

then

$$\langle x, \hat{\rho}_0 \rangle = 0, \quad \langle x, \hat{\rho}_1 \rangle = 0.$$

The Schrödinger bridge with prior dynamics (13) and marginal distributions $\hat{\rho}_0$ and $\hat{\rho}_1$ is [26,27]

$$dx(t) = Ax(t)dt + BB'\nabla\hat{\lambda}(t, x(t))dt + Bdw(t), \quad (14)$$

where $\hat{\lambda}$ satisfies

$$\frac{\partial\hat{\lambda}}{\partial t} + \nabla\hat{\lambda} \cdot Ax + \frac{1}{2} \text{tr} \left(BB'\nabla^2\hat{\lambda} \right) + \frac{1}{2} \nabla\hat{\lambda} \cdot BB'\nabla\hat{\lambda} = 0. \quad (15)$$

The boundary condition $\hat{\lambda}(1, \cdot)$ for $\hat{\lambda}$ is chosen in a way so that the resulting density flow $\hat{\rho}(t, x)$ of (14), which is

$$\frac{\partial\hat{\rho}}{\partial t} + \nabla \cdot \left((Ax + BB'\nabla\hat{\lambda}) \hat{\rho} \right) - \frac{1}{2} \text{tr} \left(BB'\nabla^2\hat{\rho} \right) = 0,$$

matches the marginal distributions $\hat{\rho}_0$ and $\hat{\rho}_1$.

The Schrödinger bridge corresponds to an optimal control problem [26,27] that minimizes

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

subject to

$$dx(t) = Ax(t)dt + Bu(t)dt + Bdw$$

and the constraints on the marginal distributions $x(0) \sim \hat{\rho}_0$ and $x(1) \sim \hat{\rho}_1$. The optimal control is given by $u(t, x) = B' \nabla \hat{\lambda}(t, x)$. The pair $(\hat{\lambda}, \hat{\rho})$ satisfies that

$$\langle \nabla \hat{\lambda}(t, \cdot), \hat{\rho}(t, \cdot) \rangle = 0 \quad (16)$$

and therefore

$$\langle x, \hat{\rho}(t, \cdot) \rangle = 0 \quad (17)$$

for all $0 \leq t \leq 1$. The intuition for (16) is that if the expectation of the control, i.e., $\langle \nabla \hat{\lambda}(t, \cdot), \hat{\rho}(t, \cdot) \rangle$, is not constantly 0, then one can always shift the control by its mean to achieve a smaller cost.

Now let

$$m(t) = \Phi(1, t)' \bar{M}_{10}^{-1} (\bar{x}_{\rho_1} - \bar{\Phi}_{10} \bar{x}_{\rho_0}), \quad (18)$$

$y(t)$ the solution to

$$\dot{y}(t) = (A + \bar{A})y(t) + BB'm(t), \quad y(0) = \bar{x}_{\rho_0},$$

and

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

Here $\bar{\Phi}_{10} := \bar{\Phi}(1, 0)$ with $\bar{\Phi}$ being the state transition matrices for the pair $(A + \bar{A}, B)$ and the “coupled” Gramian

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

is assumed to be invertible. Note that $y(1) = \bar{x}_{\rho_1}$.

With these ingredients, we construct a solution to (12) as follows. Define (λ, ρ^*) by

$$\lambda(t, x) = \hat{\lambda}(t, x - y(t)) + m(t) \cdot x + \gamma(t), \quad (19a)$$

and

$$\rho^*(t, x) = \hat{\rho}(t, x - y(t)). \quad (19b)$$

In so doing, (λ, ρ^*) is a solution of (12). To see this, on the one hand, substituting (19) into (12b), in view of (17), we obtain

$$\begin{aligned} & \frac{\partial \rho^*}{\partial t} + \nabla \cdot ((Ax + \bar{A}\bar{x}_{\rho^*} + BB' \nabla \lambda) \rho^*) - \frac{1}{2} \text{tr} (BB' \nabla^2 \rho^*) \\ &= \frac{\partial \hat{\rho}}{\partial t} - \nabla \hat{\rho} \cdot ((A + \bar{A})y + BB'm) \\ & \quad + \nabla \cdot ((Ax + BB' \nabla \lambda) \hat{\rho}) + \bar{A} \langle \xi, \hat{\rho}(t, \xi - y) \rangle \cdot \nabla \hat{\rho} - \frac{1}{2} \text{tr} (BB' \nabla^2 \hat{\rho}) \\ &= -\nabla \hat{\rho} \cdot (\bar{A}y + BB'm) + \nabla \cdot (BB'm \hat{\rho}) + \bar{A} \langle \xi, \hat{\rho}(t, \xi - y) \rangle \cdot \nabla \hat{\rho} = 0, \end{aligned}$$

where we referred to (17) in the last step. The fact that ρ^* matches the boundary conditions (12c) follows directly from the definition (19b). On the other hand, plugging (19) into (12a) yields

$$\begin{aligned} & \frac{\partial \lambda}{\partial t} + \nabla \lambda \cdot Ax + \nabla \lambda \cdot \bar{A} \bar{x}_{\rho^*} + \frac{1}{2} \operatorname{tr} (BB' \nabla^2 \lambda) + \frac{1}{2} \nabla \lambda \cdot BB' \nabla \lambda \\ &= \frac{\partial \hat{\lambda}}{\partial t} - \nabla \hat{\lambda} \cdot ((A + \bar{A})y + BB'm) + \dot{m} \cdot x + \dot{\gamma} \\ & \quad + (\nabla \hat{\lambda} + m) \cdot (Ax + \bar{A} \bar{x}_{\rho^*}) + \frac{1}{2} \operatorname{tr} (BB' \nabla^2 \hat{\lambda}) \\ & \quad + \frac{1}{2} (\nabla \hat{\lambda} + m) \cdot BB' (\nabla \hat{\lambda} + m) \\ &= -\nabla \hat{\lambda} \cdot \bar{A} y + \dot{m} \cdot x + \dot{\gamma} + (\nabla \hat{\lambda} + m) \cdot \bar{A} \bar{x}_{\rho^*} + m \cdot Ax + \frac{1}{2} m \cdot BB'm = 0. \end{aligned}$$

Therefore (λ, ρ^*) in (19) is indeed a solution to (12).

The above construction of a solution to (12) is easy to implement numerically. The only nontrivial requirement is to compute a Schrödinger bridge, which can be accomplished using an efficient algorithm [41]. The rest follows explicitly from (19). This construction is very different to that in [37, 38] which relies on an iterative scheme between ρ and λ . The key factor leading to this simplification is that the mean-field effect can be decoupled when the dynamics is linear and the running cost quadratic.

Finally, back to Problem 2.1, we assert that with terminal cost

$$g(x, \mu) = -\hat{\lambda}(1, x - \bar{x}_\mu) - m(1) \cdot x - \gamma(1), \quad (20)$$

we can lead the agents to have terminal distribution ρ_1 . To this extent, we follow the strategy in [5] as mentioned earlier in Sect. 2. First fix $\mu = \rho^*$ with ρ^* as in (19b), and then solve the optimal control problem (5). Since $g(x, \rho^*(1, \cdot)) = g(x, \rho_1) = -\lambda(1, x)$, we have

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

Hence, the unique optimal control strategy is $u^*(t) = B' \nabla \lambda(t, x(t))$. It follows from (12) that the probability distribution of the controlled state $x(t)$ is ρ^* . Therefore, with

terminal cost g as in (20) we are able to steer the system to terminal distribution ρ_1 providing that the agents follow the equilibrium strategy $u = B'\nabla\lambda$. Thus, we have established the following result.

Theorem 3.1 *Consider N agents governed by (1) with initial density ρ_0 . Suppose the terminal cost in (8) is as in (20). Then, in the Nash equilibrium, the agents will reach density ρ_1 at time $t = 1$, in the mean-field limit.*

Even though Theorem 3.1 is a result in the mean-field limit, it is straightforward to extend it to N agents systems to establish ε optimality. After all, once the terminal cost g is fixed, the problem becomes a standard MFG problem. Thus, the rest follows directly from standard MFG theory [5,36].

In fact, the dependence of g on μ is not necessary. One can simply take $g(x, \mu) = g(x) = -\lambda(1, x)$. With this terminal cost, we can still conclude that (λ, ρ^*) corresponds to a Nash equilibrium as well. This is due to the fact that we fix the density flow first when we derive a Nash equilibrium. We might need the dependence of g on μ to conclude the uniqueness of the equilibrium. It is unclear to us if this is the case.

For general $Q \geq 0$ in (8), the solution is almost the same. The only difference is that $\hat{\lambda}, \hat{\rho}$ now solve the equations

$$\begin{aligned} \frac{\partial \hat{\lambda}}{\partial t} + \nabla \hat{\lambda} \cdot Ax + \frac{1}{2} \text{tr}(BB'\nabla^2 \hat{\lambda}) + \frac{1}{2} \nabla \hat{\lambda} \cdot BB'\nabla \hat{\lambda} - \frac{1}{2} x'Qx &= 0, \\ \frac{\partial \hat{\rho}}{\partial t} + \nabla \cdot ((Ax + BB'\nabla \hat{\lambda})\hat{\rho}) - \frac{1}{2} \text{tr}(BB'\nabla^2 \hat{\rho}) &= 0, \\ \hat{\rho}(0, \cdot) &= \hat{\rho}_0, \quad \hat{\rho}(1, \cdot) = \hat{\rho}_1. \end{aligned}$$

This corresponds to a Schrödinger bridge for diffusion processes with killing [40]. Again, the same efficient algorithm is applicable [41].

4 Zero-Noise Limit

In this section, we study the same problem (Problem 2.1), with however vanishing disturbance. More specifically, we consider a system of N agents with dynamics

$$\begin{aligned} dx_i(t) &= Ax_i(t)dt + \frac{1}{N-1} \sum_{j \neq i} \bar{A}x_j(t)dt + Bu_i(t)dt + \sqrt{\epsilon} Bdw_i(t), \\ x_i(0) &= x_0^i, \quad i = 1, \dots, N, \end{aligned} \quad (21)$$

where $\epsilon > 0$ represents the variance of the noise. We are especially interested in the limit behavior of the solution to Problem 2.1 with dynamics (21) when ϵ goes to 0.

Following the same arguments as in Sect. 3, we arrive at the coupled PDEs

$$\frac{\partial \lambda}{\partial t} + \nabla \lambda \cdot Ax + \nabla \lambda \cdot \bar{A}\bar{x}_{\rho^*} + \frac{\epsilon}{2} \text{tr}(BB'\nabla^2 \lambda) + \frac{1}{2} \nabla \lambda \cdot BB'\nabla \lambda = 0, \quad (22a)$$

$$\frac{\partial \rho^*}{\partial t} + \nabla \cdot ((Ax + \bar{A}\bar{x}_{\rho^*} + BB'\nabla\lambda)\rho^*) - \frac{\epsilon}{2} \text{tr}(BB'\nabla^2\rho^*) = 0, \quad (22b)$$

$$\rho^*(0, \cdot) = \rho_0, \quad \rho^*(1, \cdot) = \rho_1. \quad (22c)$$

The optimal control strategy is given by $u(t) = B'\nabla\lambda(t, x(t))$ and terminal cost g is as in (20) with adjusted diffusivity.

Taking the limit of (22) as $\epsilon \rightarrow 0$ gives

$$\frac{\partial \lambda}{\partial t} + \nabla\lambda \cdot Ax + \nabla\lambda \cdot \bar{A}\bar{x}_{\rho^*} + \frac{1}{2}\nabla\lambda \cdot BB'\nabla\lambda = 0, \quad (23a)$$

$$\frac{\partial \rho^*}{\partial t} + \nabla \cdot ((Ax + \bar{A}\bar{x}_{\rho^*} + BB'\nabla\lambda)\rho^*) = 0, \quad (23b)$$

$$\rho^*(0, \cdot) = \rho_0, \quad \rho^*(1, \cdot) = \rho_1. \quad (23c)$$

With similar analysis as in Sect. 3 we conclude that the above PDEs system has a (viscosity) solution [42]. However, in this case, we need to resort to OMT instead of Schrödinger bridge. Indeed, the former is the zero-noise limit of the latter [27,29,30]. The solution (λ, ρ^*) to (23) has the form (19) with $\hat{\lambda}$ being

$$\hat{\lambda}(t, x) = \inf_y \left\{ \hat{\lambda}(0, y) + \frac{1}{2} (x - \Phi(t, 0)y)' M(t, 0)^{-1} (x - \Phi(t, 0)y) \right\},$$

where

$$\hat{\lambda}(0, x) = \psi \left(M_{10}^{-1/2} \Phi_{10} x \right) - \frac{1}{2} x' \Phi_{10}' M_{10}^{-1} \Phi_{10} x.$$

The function ψ corresponds to the optimal transport map with prior dynamics $\dot{x} = Ax + Bu$, and marginal distributions $\hat{\rho}_0$ and $\hat{\rho}_1$ after coordinate transformation, see [27]. It can be computed efficiently. It turns out that the solution to (23) solves the following problem.

Problem 4.1 Given N agents governed by (21) 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 (23), we can choose a terminal cost as in (20). The corresponding equilibrium control strategy is $u(t, x) = B'\nabla\lambda(t, x)$. The argument is the same as that in Sect. 3 and is therefore omitted.

Theorem 4.1 Consider N agents governed by (21) with $\epsilon = 0$ and initial density ρ_0 . Suppose the terminal cost g in (8) is as in (20), then, in the Nash equilibrium, the agents will reach density ρ_1 at time $t = 1$, in the mean-field limit.

5 Gaussian Case

In the special 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 .
When $\epsilon = 1$, $\hat{\lambda}$ in (15) equals

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

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

$$\dot{\Pi}(t) = -A'\Pi(t) - \Pi(t)A + \Pi(t)BB'\Pi(t) \quad (24)$$

with boundary condition

$$\begin{aligned} \Pi(0) = \Sigma_0^{-1/2} \left[\frac{I}{2} + \Sigma_0^{1/2} \Phi'_{10} M_{10}^{-1} \Phi_{10} \Sigma_0^{1/2} \right. \\ \left. - \left(\frac{I}{4} + \Sigma_0^{1/2} \Phi'_{10} M_{10}^{-1} \Sigma_1 M_{10}^{-1} \Phi_{10} \Sigma_0^{1/2} \right)^{1/2} \right] \Sigma_0^{-1/2}. \end{aligned}$$

where $\Phi_{10} = \Phi(1, 0)$, $M_{10} = M(1, 0)$. And so, in view of (20), one choice of terminal cost is

$$g(x, \mu) = \frac{1}{2}(x - \bar{x}_\mu)'\Pi(1)(x - \bar{x}_\mu) - m(1) \cdot x, \quad (25)$$

where $m(t)$ is as in (18). In the above we have discarded some constant terms as it does not affect the final result. The resulting optimal strategy for each agent is $u(t, x) = B'\nabla\lambda(t, x)$, which is a linear state feedback.

Theorem 5.1 *Consider N agents governed by (1) with initial distribution $\rho_0 \sim \mathcal{N}[m_0, \Sigma_0]$. Suppose the terminal cost in (8) is (25). 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 limit.*

Following the discussion in Sect. 4, the solution to the problem with noise intensity ϵ is almost identical to the above except that, the initial condition of the Riccati Eq. (24) becomes

$$\begin{aligned} \Pi_\epsilon(0) = \Sigma_0^{-1/2} \left[\frac{\epsilon I}{2} + \Sigma_0^{1/2} \Phi'_{10} M_{10}^{-1} \Phi_{10} \Sigma_0^{1/2} \right. \\ \left. - \left(\frac{\epsilon^2 I}{4} + \Sigma_0^{1/2} \Phi'_{10} M_{10}^{-1} \Sigma_1 M_{10}^{-1} \Phi_{10} \Sigma_0^{1/2} \right)^{1/2} \right] \Sigma_0^{-1/2}. \end{aligned}$$

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

$$\begin{aligned} \Pi_0(0) = \Sigma_0^{-1/2} \left[\Sigma_0^{1/2} \Phi'_{10} M_{10}^{-1} \Phi_{10} \Sigma_0^{1/2} \right. \\ \left. - \left(\Sigma_0^{1/2} \Phi'_{10} M_{10}^{-1} \Sigma_1 M_{10}^{-1} \Phi_{10} \Sigma_0^{1/2} \right)^{1/2} \right] \Sigma_0^{-1/2}. \end{aligned}$$

Since the terminal cost g in (25) is quadratic, once g is fixed, the resulting problem falls within the class of linear-quadratic mean-field games [6–11]. The optimal policy, which is a linear state feedback, now follows from standard theory. Our contribution is to obtain the appropriate value for the terminal cost g in (25).

6 Stationary Case and Invariant Measure

We now turn to the stationary counterpart of Problem 2.1. We would like to design a cost function that will lead the agents to achieve a given invariant measure ρ , if the agents follows the equilibrium strategy. In particular, given N agents with identical dynamics (1) that attempt to minimize their control effort

$$J_i(u_i) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left\{ \int_0^T \frac{1}{2} \|u_i(t)\|^2 dt \right\},$$

we look for an extra cost $g(x, \mu)$ term added to the above such that, in the equilibrium state, the agents have some specified distribution. The new cost function is

$$J_i(u_i) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left\{ \int_0^T \left[\frac{1}{2} \|u_i(t)\|^2 + g(x_i(t), \mu^N(t)) \right] dt \right\} \quad (26)$$

where μ^N is the empirical distribution (3). Again we are interested in the mean-field limit of the problem, that is, the case when N goes to ∞ .

In this stationary problem an extra running cost is being considered. This is different from the finite horizon case discussed earlier where a terminal cost is needed instead. Adding a running cost to dictate the stationary distribution, in an optimal control setting, has been studied in [33].

Let us first recall some relevant results in the stationary mean-field game problems. Suppose the N agents with dynamics (1) attempt to minimize the cost function

$$J_i(u_i) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left\{ \int_0^T f(x_i(t), u_i(t), \mu^N(t)) dt \right\}.$$

We restrict ourselves to equilibria with symmetric stationary Markovian strategies

$$u_i(t) = \phi(x_i(t)).$$

In the mean-field limit, one can adapt the following steps [5]. First, fix a probability measure μ and then solve the standard stochastic control problem (parametrized by μ)

$$\phi^* = \operatorname{argmin}_{\phi} \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left\{ \int_0^T f(x(t), \phi(x(t)), \mu(t)) dt \right\} \quad (27)$$

subject to the dynamics

$$dx(t) = Ax(t)dt + \bar{A}\bar{x}_{\mu}dt + B\phi(x(t))dt + Bd w(t). \quad (28)$$

Once this standard optimal control problem is solved, the remaining issue is finding the correct distribution μ such that the stationary distribution of (28) with optimal control strategy

$$u^*(t) = \phi^*(x(t))$$

coincides with μ .

The solution to this MFG problem involves the coupled PDEs [2,5]

$$\frac{1}{2} \operatorname{tr} \left(B B' \nabla^2 \lambda \right) + \eta - H^{\rho^*} (x, -\nabla \lambda) = 0, \quad (29a)$$

$$\nabla \cdot \left((A x + \bar{A} \bar{x}_{\rho^*} + B \phi^*) \rho^* \right) - \frac{1}{2} \operatorname{tr} \left(B B' \nabla^2 \rho^* \right) = 0, \quad (29b)$$

$$\rho^* \geq 0, \quad \int \rho^* = 1, \quad (29c)$$

where η is an unknown constant, $u^* = \phi^*(x)$ is the minimizer of

$$H^{\rho}(x, p) = \min_{u \in \mathbb{R}^m} \left\{ p'(A x + \bar{A} \bar{x}_{\rho} + B u) + f(x, u, \rho) \right\}.$$

When the cost function is of the form (26), $\phi^* = B' \nabla \lambda$ and the PDEs boil down to

$$\frac{1}{2} \operatorname{tr} \left(B B' \nabla^2 \lambda \right) + \eta + \nabla \lambda \cdot A x + \nabla \lambda \cdot \bar{A} \bar{x}_{\rho^*} + \frac{1}{2} \nabla \lambda \cdot B B' \nabla \lambda - g(x, \rho^*) = 0, \quad (30a)$$

$$\nabla \cdot \left((A x + \bar{A} \bar{x}_{\rho^*} + B B' \nabla \lambda) \rho^* \right) - \frac{1}{2} \operatorname{tr} \left(B B' \nabla^2 \rho^* \right) = 0, \quad (30b)$$

$$\rho^* \geq 0, \quad \int \rho^* = 1. \quad (30c)$$

The existence of a solution (ρ^*, λ) can be shown under some proper assumptions on g , see [2,5].

Back to our problem, the cost function g in (26) becomes a design parameter, which is different from the classic MFG setting. Our goal is to choose a function g such that the corresponding stationary distribution in Nash equilibrium is ρ^* . The solution relies on the same PDEs (30), but with different variables. Given a distribution ρ^* , we need to find λ and the proper cost g that solve (30). It turns out that (30) has solution only for a small class of distributions ρ^* , which we call the feasible distributions. We next focus on Gaussian distributions. In this case, the feasible distributions can be characterized by some algebraic equations. The cases of general distributions will be investigated in future study.

Let ρ^* be a Gaussian distribution with mean m and covariance Σ . Plugging the ansatz

$$\lambda(x) = -\frac{1}{2} x' \Pi x + n' x$$

with $\Pi = \Pi'$, into (30) yields (after discarding constant terms)

$$-\frac{1}{2} \operatorname{tr}(BB'\Pi) + \eta + \frac{1}{2}x'(-\Pi A - A'\Pi + \Pi BB'\Pi)x + n'(A - BB'\Pi)x - m'\bar{A}'\Pi x - g(x, \rho^*) = 0 \quad (31a)$$

$$(A - BB'\Pi)\Sigma + \Sigma(A - BB'\Pi)' + BB' = 0 \quad (31b)$$

$$(A + \bar{A} - BB'\Pi)m + BB'n = 0. \quad (31c)$$

In order for a solution to exist, in view of (31b), it is necessary that Σ satisfies

$$A\Sigma + \Sigma A' \in \operatorname{range}(f_B), \quad (32a)$$

where $f_B(X) = BX' + XB'$ is a map from $\mathbb{R}^{n \times m}$ to the space of symmetric matrices (see [33] for other equivalent algebraic conditions). Likewise, by (31c), the mean m has to satisfy

$$(A + \bar{A})m \in \operatorname{range}(B). \quad (32b)$$

On the other hand, given (m, Σ) satisfying (32), assuming B has full column rank, then (31b) has a unique symmetric solution [33]. Therefore, these two conditions (32) are also sufficient. Now from (31a) it is easy to conclude that a possible cost function is

$$g(x, \rho) = \frac{1}{2}x'Qx + n \cdot (A - BB'\Pi)x - \bar{x}_\rho \cdot \bar{A}'\Pi x, \quad (33a)$$

where

$$Q = -\Pi A - A'\Pi + \Pi BB'\Pi, \quad (33b)$$

with Π being the unique solution to (31b) and n being a solution to (31c). Therefore, we have established the following result.

Theorem 6.1 *Consider N agents governed by (1). Assuming the Gaussian distribution $\mathcal{N}(m, \Sigma)$ satisfies (32). If g function in the cost (26) is as in (33), then, in the Nash equilibrium, the agents have stationary Gaussian distribution $\mathcal{N}(m, \Sigma)$ in the mean-field limit.*

7 Cooperative Game

In this section we 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 ourselves to equilibriums given by symmetric Markovian strategies in closed-loop feedback form

$$u_i(t) = \phi(t, x_i(t)), i = 1, \dots, N. \quad (34)$$

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

$$J(u) = \mathbb{E} \left\{ \frac{1}{N} \sum_{i=1}^N \int_0^1 \frac{1}{2} \|u_i(t)\|^2 dt \right\}. \quad (35)$$

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

Problem 7.1 Given N agents governed by (1) with initial density ρ_0 , find a control strategy (34) with minimum control energy (35) so that the agents will reach density ρ_1 at time $t = 1$, as N goes to ∞ .

The above problem is closely related to the optimal control problem of McKean–Vlasov dynamics [5,13,14]. The goal in the latter is to minimize a cost function with possibly a terminal cost while in our problem it becomes to achieve a specified target distribution with minimum running cost. Our problem reduces to an optimal control problem of McKean–Vlasov dynamics with some proper terminal cost. Therefore, one contribution of this section is a method to construct such a terminal cost.

The major difference between Problems 7.1 and 2.1 is that the agents are cooperative. 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 [5,14] for more discussions on their differences. The average control energy (35) turns out to be equivalent to relative entropy of the controller system with respect to the uncontrolled system [35,43,44]. Therefore, the above problem can also be viewed as an Schrödinger bridge problem for weakly interacting particle systems.

Problem 7.1 is formulated as an optimal control problem over the McKean–Vlasov model

$$dx(t) = Ax(t)dt + \bar{A}\bar{x}(t)dt + Bu(t)dt + BdW(t), \quad x(0) = x_0 \sim \rho_0. \quad (36)$$

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

$$\inf_{\rho, u} \int_0^1 \int_{\mathbb{R}^n} \frac{1}{2} \rho(t, x) \|u(t, x)\|^2 dx dt, \quad (37a)$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot ((Ax + \bar{A}\bar{x}_\rho + Bu)\rho) - \frac{1}{2} \text{tr}(BB'\nabla^2 \rho) = 0, \quad (37b)$$

$$\rho(0, \cdot) = \rho_0, \quad \rho(1, \cdot) = \rho_1. \quad (37c)$$

Proposition 7.1 *If there exists (λ, ρ^*) satisfying*

$$\begin{aligned} \frac{\partial \lambda}{\partial t} + \nabla \lambda \cdot Ax + \nabla \lambda \cdot \bar{A}\bar{x}_{\rho^*} + \frac{1}{2} \text{tr}(BB'\nabla^2 \lambda) \\ + \frac{1}{2} \nabla \lambda \cdot BB'\nabla \lambda + \bar{A}x \cdot \langle \nabla \lambda, \rho^* \rangle = 0, \end{aligned} \quad (38a)$$

$$\frac{\partial \rho^*}{\partial t} + \nabla \cdot ((Ax + \bar{A}\bar{x}_{\rho^*} + BB'\nabla\lambda)\rho^*) - \frac{1}{2} \operatorname{tr}(BB'\nabla^2\rho^*) = 0, \quad (38b)$$

with boundary conditions

$$\rho^*(0, \cdot) = \rho_0, \quad \rho^*(1, \cdot) = \rho_1, \quad (38c)$$

then $(\rho^*, u^* = B'\nabla\lambda)$ is a solution to (37).

The above optimality conditions (38) are highly coupled. In general, one may not expect a solution to exist. But interestingly, as we will see below, (38) always has a solution. Again, we can construct a solution based on standard Schrödinger bridge theory.

Let $(\hat{\rho}, \hat{\lambda})$ be as in (13)–(16),

$$m(t) = \bar{\Phi}(1, t)' \hat{M}_{10}^{-1} (\bar{x}_{\rho_1} - \bar{\Phi}_{10} \bar{x}_{\rho_0}) \quad (39a)$$

and $y(t)$ the solution to

$$\dot{y}(t) = (A + \bar{A})y(t) + BB'm(t), \quad y(0) = \bar{x}_{\rho_0}, \quad (39b)$$

where $\hat{M}_{10} = \hat{M}(1, 0)$ with

$$\hat{M}(t, s) = \int_s^t \bar{\Phi}(t, \tau) BB' \bar{\Phi}(t, \tau)' d\tau.$$

Define

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

$$\lambda(t, x) = \hat{\lambda}(t, x - y(t)) + m(t) \cdot x + \gamma(t), \quad (40a)$$

and

$$\rho^*(t, x) = \hat{\rho}(t, x - y(t)), \quad (40b)$$

then (λ, ρ^*) solves (38).

Apparently, (40b) satisfies the boundary conditions (38c). To verify (38b), substitute (40) into (38b), which gives

$$\begin{aligned} & \frac{\partial \rho^*}{\partial t} + \nabla \cdot ((Ax + \bar{A}\bar{x}_{\rho^*} + BB'\nabla\lambda)\rho^*) - \frac{1}{2} \operatorname{tr}(BB'\nabla^2\rho^*) \\ &= \frac{\partial \hat{\rho}}{\partial t} - \nabla \hat{\rho} \cdot ((A + \bar{A})y + BB'm) + \nabla \cdot ((Ax + BB'\nabla\lambda)\hat{\rho}) \\ & \quad + \bar{A}\langle \xi, \hat{\rho}(\xi - y) \rangle \cdot \nabla \hat{\rho} - \frac{1}{2} \operatorname{tr}(BB'\nabla^2\hat{\rho}) \\ &= -\nabla \hat{\rho} \cdot (\bar{A}y + BB'm) + \nabla \cdot (BB'm\hat{\rho}) + \bar{A}\langle \xi, \hat{\rho}(\xi - y) \rangle \cdot \nabla \hat{\rho} = 0. \end{aligned}$$

Similarly, Combing (40) and (38a) yields

$$\begin{aligned}
 & \frac{\partial \lambda}{\partial t} + \nabla \lambda \cdot Ax + \nabla \lambda \cdot \bar{A} \bar{x}_{\rho^*} + \frac{1}{2} \operatorname{tr} \left(BB' \nabla^2 \lambda \right) \\
 & \quad + \frac{1}{2} \nabla \lambda \cdot BB' \nabla \lambda + \bar{A} x \cdot \langle \nabla \lambda, \rho^* \rangle \\
 & = \frac{\partial \hat{\lambda}}{\partial t} - \nabla \hat{\lambda} \cdot ((A + \bar{A})y + BB'm) \\
 & \quad + \dot{m} \cdot x + \dot{y} + \left(\nabla \hat{\lambda} + m \right) \cdot (Ax + \bar{A} \bar{x}_{\rho^*}) + \frac{1}{2} \operatorname{tr} \left(BB' \nabla^2 \hat{\lambda} \right) \\
 & \quad + \frac{1}{2} \left(\nabla \hat{\lambda} + m \right) \cdot BB' \left(\nabla \hat{\lambda} + m \right) + \bar{A} x \cdot \langle \nabla \hat{\lambda} + m, \rho^* \rangle \\
 & = -\nabla \hat{\lambda} \cdot \bar{A} y + \dot{m} \cdot x + \dot{y} + \left(\nabla \hat{\lambda} + m \right) \cdot \bar{A} \bar{x}_{\rho^*} \\
 & \quad + m \cdot Ax + \frac{1}{2} m \cdot BB'm + \bar{A} x \cdot m + \bar{A} y \cdot \langle \nabla \hat{\lambda}, \rho^* \rangle \\
 & = \bar{A} y \cdot \langle \nabla \hat{\lambda} (\xi - y), \hat{\rho} (\xi - y) \rangle = 0.
 \end{aligned}$$

Therefore, the pair $(\rho^*, u^* = B' \nabla \lambda)$ is indeed a solution to (38).

Next we prove that this pair (ρ^*, u^*) provides a solution to the optimal control problem (37). Let $\bar{u}(t) = \mathbb{E}\{u(t)\}$, then, by (36), we have

$$d\bar{x}(t) = (A + \bar{A})\bar{x}(t)dt + B\bar{u}(t)dt \quad (41)$$

and

$$d(\tilde{x}(t)) = A\tilde{x}(t)dt + B\tilde{u}(t)dt + Bdw, \quad (42)$$

where $\tilde{x} = x - \bar{x}$ and $\tilde{u} = u - \bar{u}$. The control energy can then be decomposed into two parts as

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

These two parts of control energy, corresponding to $\bar{u}$ and $u - \bar{u}$, respectively, can be minimized independently since the dynamics (41) and (42) are decoupled. We next show that

(i) $\bar{u}^*$ minimizes

$$\int_0^1 \frac{1}{2} \|\bar{u}(t)\|^2 dt \quad (43)$$

(ii) $\tilde{u}^*$ minimizes

$$\mathbb{E} \left\{ \int_0^1 \frac{1}{2} \|\tilde{u}(t)\|^2 dt \right\}. \quad (44)$$

For (i), recalling

$$u^*(t) = B' \nabla \lambda = B' \nabla \hat{\lambda}(t, x(t) - y(t)) + B'm(t),$$

we have

$$\begin{aligned}\bar{u}^*(t) &= B' \mathbb{E} \{ \nabla \hat{\lambda}(t, x(t) - y(t)) \} + B' m(t) \\ &= B' \langle \nabla \hat{\lambda}(t, x - y(t)), \rho^*(t, x) \rangle + B' m(t) = B' m(t).\end{aligned}$$

Using standard optimal control, it is easy to see that $\bar{u}^*(t)$ minimizes (43) subject to (41) and boundary conditions

$$\bar{x}(0) = \bar{x}_{\rho_0}, \quad \bar{x}(1) = \bar{x}_{\rho_1}.$$

For (ii), we note

$$\tilde{u}^*(t) = B' \nabla \hat{\lambda}(t, x(t) - y(t)) = B' \nabla \hat{\lambda}(t, x(t) - \bar{x}(t)) = B' \nabla \hat{\lambda}(t, \tilde{x}(t)).$$

The rest is an immediate consequence of the Schrödinger bridge theory [26,27]. Hence we have established the following result.

Theorem 7.1 *The pair $(\rho^*, u^* = B' \nabla \lambda)$ with ρ^*, λ as in (40) solves (37).*

As discussed earlier, by adding a proper terminal cost to (35), our problem reduces to an optimal control problem of McKean–Vlasov dynamics. One such choice of terminal cost is $-\lambda(1, \cdot)$.

7.1 Gaussian Case

When both of the marginals ρ_0 and ρ_1 are normal distributions, the optimal control u^* is a linear-state feedback control. Let the two marginal distributions ρ_0 and ρ_1 be

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

By Theorem 7.1, we need only to compute λ as in (40a), which is

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

The function $\hat{\lambda}$ corresponds to the Schrödinger bridge (14), which satisfies [32]

$$\nabla \hat{\lambda}(t, x) = -\Pi(t)x,$$

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

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

with boundary condition

$$\begin{aligned} \Pi(0) = \Sigma_0^{-1/2} & \left[\frac{I}{2} + \Sigma_0^{1/2} \Phi'_{10} M_{10}^{-1} \Phi_{10} \Sigma_0^{1/2} \right. \\ & \left. - \left(\frac{I}{4} + \Sigma_0^{1/2} \Phi'_{10} M_{10}^{-1} \Sigma_1 M_{10}^{-1} \Phi_{10} \Sigma_0^{1/2} \right)^{1/2} \right] \Sigma_0^{-1/2}. \end{aligned}$$

It follows that the optimal control is

$$\begin{aligned} u^*(t) &= B' \nabla \lambda(t, x(t)) = -B' \Pi(t)(x(t) - y(t)) + B' m(t) \\ &= -B' \Pi(t)x(t) + B' n(t), \end{aligned}$$

where

$$\begin{aligned} n(t) &= \Pi(t)y(t) + m(t) \\ &= \Pi(t)\bar{\Phi}(t, 1)\hat{M}(1, t)\hat{M}_{10}^{-1}\bar{\Phi}_{10}m_0 \\ &\quad + \Pi(t)\hat{M}(t, 0)\bar{\Phi}(1, t)'\hat{M}_{10}^{-1}m_1 + \bar{\Phi}(1, t)'\hat{M}_{10}^{-1}(m_1 - \bar{\Phi}_{10}m_0). \end{aligned}$$

7.2 Zero-Noise Limit

We study the optimal steering problem for McKean–Vlasov model (36) with reduced disturbance $\sqrt{\epsilon}dw(t)$, namely,

$$dx(t) = Ax(t)dt + \bar{A}\bar{x}(t)dt + Bu(t)dt + \sqrt{\epsilon}Bdw(t), \quad x(0) = x_0. \quad (45)$$

In particular, we are interested in the zero-noise limit of this problem. That is, optimal steering problem for dynamics

$$dx(t) = Ax(t)dt + \bar{A}\bar{x}(t)dt + Bu(t)dt, \quad x(0) = x_0. \quad (46)$$

We show that the probability flow of the solution to the latter is the limit of that of the former as ϵ goes to 0. This is achieved in a constructive manner.

Let us start with the steering problem for the dynamics

$$dx(t) = Ax(t)dt + B\tilde{u}(t)dt$$

with marginals $\hat{\rho}_0$ and $\hat{\rho}_1$. This problem has been studied in [27]. The solution is

$$\tilde{u}(t, x) = B' \Phi(1, t)' M_{10}^{-1} (T \circ T_t^{-1}(x) - \Phi_{10}(T_t^{-1}(x))) \quad (47)$$

where T is the generalized optimal mass transport map [27] with marginals $\hat{\rho}_0, \hat{\rho}_1$, and

$$T_t(x) = \Phi(t, 1)M(1, t)M_{10}^{-1}\Phi_{10}x + M(t, 0)\Phi(1, t)'M_{10}^{-1}T(x).$$

The corresponding distribution flow is

$$\hat{\rho}(t, \cdot) = (T_t)_\# \hat{\rho}_0.$$

Note that $\hat{\rho}$ and $\tilde{u}$ satisfy the continuity equation

$$\frac{\partial \hat{\rho}}{\partial t} + \nabla \cdot ((Ax + B\tilde{u})\hat{\rho}) = 0.$$

We claim that

$$u^*(t, x) = \tilde{u}(t, x - y(t)) + B'm(t) \quad (48)$$

with y, m in (39), is the optimal control strategy for the steering problem for the dynamics (46) and the corresponding distribution flow is

$$\rho^*(t, x) = \hat{\rho}(t, x - y(t)). \quad (49)$$

We shall skip the proof as it is similar to the case with disturbance [(see (40)–(44)].

The solution to the density-steering problem with dynamics (45), weakly converges to ρ^* as ϵ goes to 0. This follows directly from the fact a Schrödinger bridge converges to the corresponding optimal transport solution [27] as ϵ goes to 0.

8 Examples

Consider N agents with dynamics

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

The two marginal distributions ρ_0 and ρ_1 are two normal distributions

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

8.1 Non-cooperative 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 Fig. 1, we simulated the dynamics for a system with $N = 20,000$ agents under the optimal strategy. Figures 2 and 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 Fig. 4, which perfectly matches the theoretical result.

Fig. 1 Time evolution of probability densities

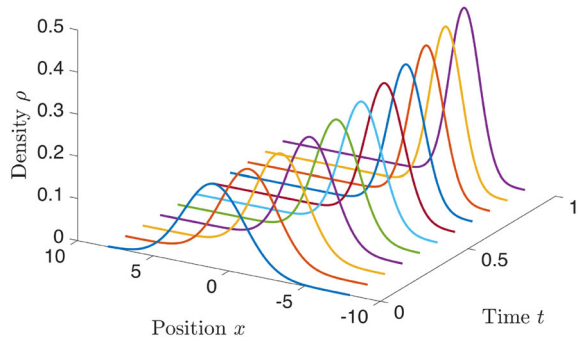


Fig. 2 Empirical distribution of $x(0)$

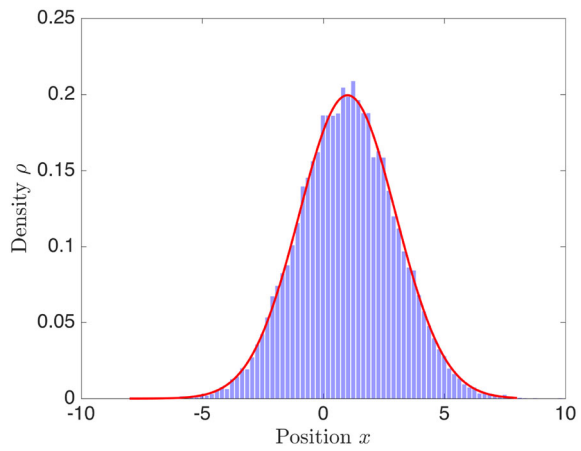


Fig. 3 Empirical distribution of $x(1)$

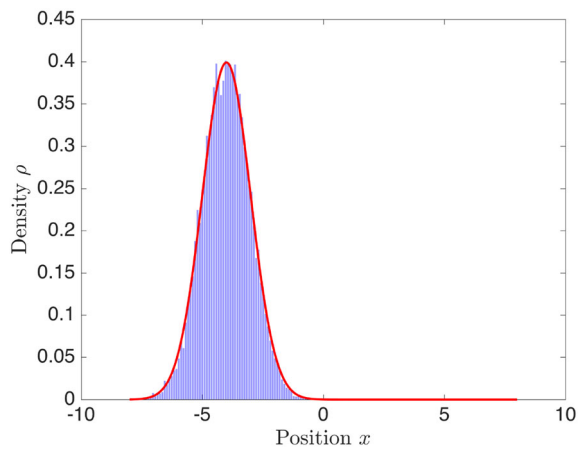


Fig. 4 Time evolution of mean $\bar{x}(t)$

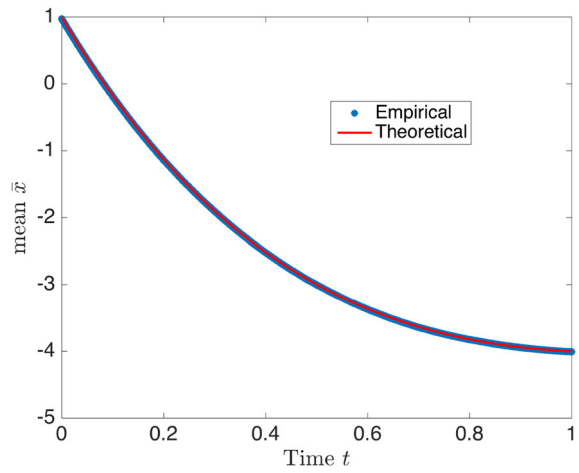


Fig. 5 Time evolution of probability densities

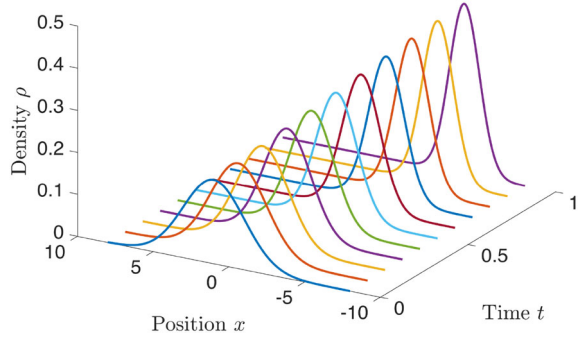


Fig. 6 Empirical distribution of $x(0)$

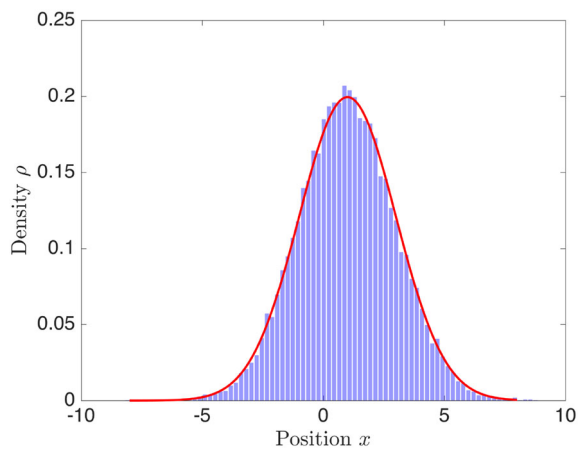


Fig. 7 Empirical distribution of $x(1)$

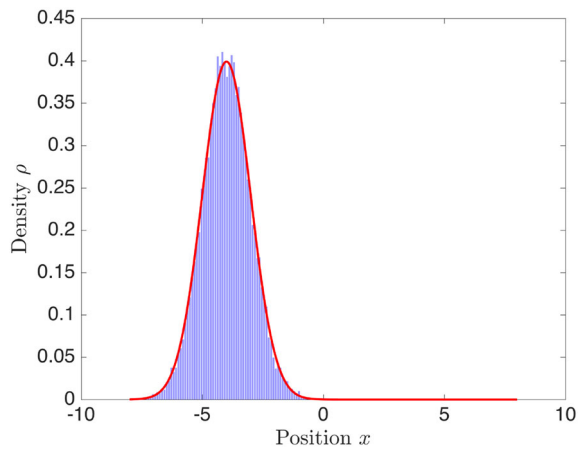
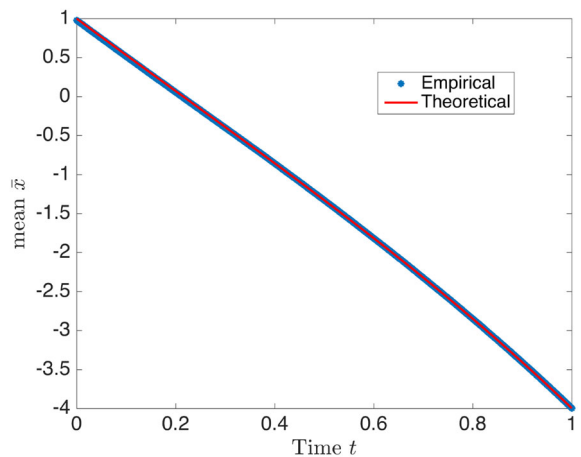


Fig. 8 Time evolution of mean $\bar{x}(t)$



8.2 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 $N = 20,000$ and obtained Figs. 6 and 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 Fig. 8. Clearly the mean is different to the Nash equilibrium in the non-cooperative game setting.

9 Conclusions

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 with quadratic running

cost and solve the problem using tools from all these three areas. Results for several extensions such as stationary and cooperative game problems are also presented. In the future, we will focus on extending the results in this paper to more general running cost functions. We expect this paradigm to bring in a new dimension to the study and applications of MFG and OMT.

Acknowledgements Research supported in part by the National Science Foundation under Grant ECCS-1509387, the Air Force Office for Scientific Research under Grants FA9550-15-1-0045 and FA9550-17-1-0435, the Army Office of Research under W911NF-17-1-0429, and by the University of Padova Research Project CPDA 140897.

References

1. Huang, M., Malhamé, R.P., Caines, P.E.: Large population stochastic dynamic games: closed-loop McKean–Vlasov systems and the nash certainty equivalence principle. *Commun. Inf. Syst.* **6**(3), 221–252 (2006)
2. Lasry, J.M., Lions, P.L.: Mean field games. *Jpn. J. Math.* **2**(1), 229–260 (2007)
3. Jovanovic, B., Rosenthal, R.W.: Anonymous sequential games. *J. Math. Econ.* **17**(1), 77–87 (1988)
4. Nourian, M., Caines, P.E.: ε -Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents. *SIAM J. Control Optim.* **51**(4), 3302–3331 (2013)
5. Carmona, R., Delarue, F., Lachapelle, A.: Control of McKean–Vlasov dynamics versus mean field games. *Math. Financ. Econ.* **7**(2), 131–166 (2013)
6. Huang, M., Caines, P.E., Malhamé, R.P.: Individual and mass behaviour in large population stochastic wireless power control problems: centralized and nash equilibrium solutions. In: *Proceedings of the 42nd IEEE Conference on Decision and Control*, 2003, vol. 1, pp. 98–103. IEEE (2003)
7. Huang, M., Caines, P.E., Malhamé, R.P.: Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε -Nash equilibria. *IEEE Trans. Autom. Control* **52**(9), 1560–1571 (2007)
8. Bardi, M.: Explicit solutions of some linear-quadratic mean field games. *Netw. Heterog. Media* **7**(2), 243–261 (2012)
9. Bensoussan, A., Sung, K., Yam, S.C.P., Yung, S.P.: Linear-quadratic mean field games. *J. Optim. Theory Appl.* **169**(2), 496–529 (2016)
10. Bardi, M., Priuli, F.S.: Linear-quadratic N-person and mean-field games with ergodic cost. *SIAM J. Control Optim.* **52**(5), 3022–3052 (2014)
11. Graber, P.J.: Linear quadratic mean field type control and mean field games with common noise, with application to production of an exhaustible resource. *Appl. Math. Optim.* **74**(3), 459–486 (2016)
12. Moon, J., Başar, T.: Linear quadratic risk-sensitive and robust mean field games. *IEEE Trans. Autom. Control* **62**(3), 1062–1077 (2017)
13. Andersson, D., Djehiche, B.: A maximum principle for sdes of mean-field type. *Appl. Math. Optim.* **63**(3), 341–356 (2011)
14. Carmona, R., Delarue, F.: Forward-backward stochastic differential equations and controlled McKean–Vlasov dynamics. *Ann. Probab.* **43**(5), 2647–2700 (2015)
15. Monge, G.: *Mémoire sur la théorie des déblais et des remblais*. De l’Imprimerie Royale, Paris (1781)
16. Kantorovich, L.V.: On the transfer of masses. *Dokl. Akad. Nauk. SSSR* **37**(7–8), 227–229 (1942)
17. Gangbo, W., McCann, R.J.: The geometry of optimal transportation. *Acta Math.* **177**(2), 113–161 (1996)
18. Evans, L.C.: Partial differential equations and Monge–Kantorovich mass transfer. *Current Dev. Math.* **1997**(1), 65–126 (1997)
19. Evans, L.C., Gangbo, W.: *Differential equations methods for the Monge–Kantorovich mass transfer problem*, vol. 653. American Mathematical Soc, Providence (1999)
20. Villani, C.: *Topics in Optimal Transportation*. No. 58 in GSM. American Mathematical Soc., Providence (2003)
21. Ambrosio, L., Gigli, N., Savaré, G.: *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*. Springer, Berlin (2006)

22. Villani, C.: Optimal Transport: Old and New, vol. 338. Springer, Berlin (2008)
23. Ambrosio, L., Gigli, N.: A user's guide to optimal transport. In: *Modelling and Optimisation of Flows on Networks*, pp. 1–155. Springer, Berlin (2013)
24. Santambrogio, F.: Optimal Transport for Applied Mathematicians. Birkhäuser, New York (2015)
25. Benamou, J.D., Brenier, Y.: A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numer. Math.* **84**(3), 375–393 (2000)
26. Chen, Y., Georgiou, T.T., Pavon, M.: On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint. *J. Optim. Theory Appl.* **169**(2), 671–691 (2016)
27. Chen, Y., Georgiou, T.T., Pavon, M.: Optimal transport over a linear dynamical system. *IEEE Trans. Autom. Control* **62**(5), 2137–2152 (2017)
28. Dai Pra, P.: A stochastic control approach to reciprocal diffusion processes. *Appl. Math. Optim.* **23**(1), 313–329 (1991)
29. Léonard, C.: From the Schrödinger problem to the Monge–Kantorovich problem. *J. Funct. Anal.* **262**(4), 1879–1920 (2012)
30. Léonard, C.: A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst. A* **34**(4), 1533–1574 (2014)
31. Gentil, I., Léonard, C., Ripani, L.: About the analogy between optimal transport and minimal entropy. *Annales de la Faculté des Sciences de Toulouse. Mathématiques. Série 6* **3**, 569–600 (2017)
32. Chen, Y., Georgiou, T.T., Pavon, M.: Optimal steering of a linear stochastic system to a final probability distribution, Part I. *IEEE Trans. Autom. Control* **61**(5), 1158–1169 (2016)
33. Chen, Y., Georgiou, T.T., Pavon, M.: Optimal steering of a linear stochastic system to a final probability distribution, Part II. *IEEE Trans. Autom. Control* **61**(5), 1170–1180 (2016)
34. Chen, Y., Georgiou, T.T., Pavon, M.: Fast cooling for a system of stochastic oscillators. *J. Math. Phys.* **56**(11), 113,302 (2015)
35. Fischer, M.: On the form of the large deviation rate function for the empirical measures of weakly interacting systems. *Bernoulli* **20**(4), 1765–1801 (2014)
36. Cardaliaguet, P.: Notes on mean field games. Tech. rep., Technical report (2010)
37. Achdou, Y., Camilli, F., Capuzzo-Dolcetta, I.: Mean field games: numerical methods for the planning problem. *SIAM J. Control Optim.* **50**(1), 77–109 (2012)
38. Porretta, A.: On the planning problem for the mean field games system. *Dyn. Games Appl.* **4**(2), 231–256 (2014)
39. Chen, Y.: Modeling and control of collective dynamics: from Schrödinger bridges to optimal mass transport. Ph.D. thesis, University of Minnesota (2016)
40. Chen, Y., Georgiou, T.T., Pavon, M.: Optimal steering of a linear stochastic system to a final probability distribution, Part III. *IEEE Trans. Autom. Control* (in print 2017) (2017)
41. Chen, Y., Georgiou, T.T., Pavon, M.: Entropic and displacement interpolation: a computational approach using the hilbert metric. *SIAM J. Appl. Math.* **76**(6), 2375–2396 (2016)
42. Fleming, W.H., Soner, H.M.: *Controlled Markov Processes and Viscosity Solutions*, vol. 25. Springer Science & Business Media, Berlin (2006)
43. Dawson, D.A., Gärtner, J.: Large deviations from the McKean–Vlasov limit for weakly interacting diffusions. *Stoch. An Int. J. Probab. Stoch. Process.* **20**(4), 247–308 (1987)
44. Feng, J., Kurtz, T.G.: *Large Deviations for Stochastic Processes*, vol. 131. American Mathematical Society Providence, Providence (2006)