

Density control of interacting agent systems

Yongxin Chen

Abstract—We consider the problem of controlling the group behavior of a large number of dynamic systems that are constantly interacting with each other. These systems are assumed to have identical dynamics (e.g., flocks of birds, UAV swarms) and their group behavior can be modeled by a distribution. Thus, this problem can be viewed as an optimal control problem over the space of distributions. We propose a novel algorithm to compute a feedback control strategy so that, when adopted by the agents, the distribution of them would be transformed from an initial one to a target one over a finite time window. Our method is built on the optimal transport theory but differs significantly from existing work in this area in that our method models the interactions among agents explicitly. From an algorithmic point of view, our algorithm is based on the generalized proximal gradient descent algorithm and has a convergence guarantee with a sublinear rate. We further extend our framework to account for the scenarios where the agents are from multiple species. In the linear quadratic setting, the solution is characterized by a system of coupled Riccati equations which can be solved in closed form. Finally, several numerical examples are presented to illustrate our framework.

I. INTRODUCTION

Consider the swarm control [1], [2] task to establish and regulate a formation of N drones (“agents”). There are two substantially different angles from which one may consider this problem. A first (straightforward) approach is to concatenate the states of all the drones into one state vector and then formulate a control problem over this joint state space. Suppose the state dimension of each individual drone is d , then the dimension of the joint state space becomes Nd , which scales linearly as the group size N increases. An alternative approach is to treat the distribution of the drones as *the* state of a system, and formulate a corresponding optimal control problem over the space of distributions. One major difference between the two approaches is that, in the former, each individual has a label and the controller aims to jointly optimize the performance of each individual, while in the latter, the *individuals are indistinguishable, and only the group behavior matters*. Thus, when the optimality criteria only involves the group behavior of individuals, the problem reduces to a density control problem. In this formulation, the state is a probability distribution and is independent of the group size N .

The density control problem is an optimal control problem over distributions where the objective/cost function is fully determined by the evolution of the distribution. The dynamics of the distribution of the group follows the Liouville equation if the individuals have deterministic dynamics, or the Fokker-Planck equation if the individual dynamics are stochastic, or the McKean-Vlasov equation if the individuals interact with each other [3], [4]. In addition to controlling the group behavior of a large number of individuals, density control can also be used to deal with controlling the state uncertainty of a single dynamical system. When a dynamical system either has uncertainty in its initial state or is disturbed by random process noise, the state remains uncertain and can be captured by a probability distribution at each time point. Regulating the state uncertainties of such systems is thus equivalent to controlling their state distribution. The density control problem provides a more direct

approach to achieving this goal than standard stochastic control theory where an indirect cost needs to be properly handcrafted. The density control framework has found applications in a range of areas [5], [6], [7], [8], [9].

Two popular tools for density control are the optimal transport (OT) theory [10] and the Schrödinger bridge theory [11], [12]. The latter can be viewed as a regularized version of the former. The minimum effort control between two specified distributions over a finite time interval can be addressed using the OT theory if the dynamics is deterministic and the Schrödinger bridge theory if the dynamics are stochastic. This paradigm is suitable for controlling the uncertainty of a single dynamical system. An important instance of it is covariance control [13], [14], [15], [16], [17], [18], [6], [7] where the distribution at each time point is parametrized by a Gaussian distribution determined by its mean and covariance. The connections between covariance control and the Schrödinger bridge have been extensively studied in [14]. One major limitation of these existing methods of density control for controlling group behaviors is that the individuals in the group are assumed to be independent from each other. Thus, important properties for swarm control such as collision avoidance are not explicitly modeled in these methods. Indeed, when these methods are applied for swarm control [19], [20], [21], [22], [23], [9], the collision avoidance requirement is often ignored.

The goal of this work is to address density control problems for controlling group behaviors when the individuals in the group are constantly interacting with each other. We consider the scenario where the interactions between every two agents are through an interaction potential that is the same for each pair of agents. When the number of individuals is large, the distribution of the individuals solves the McKean-Vlasov equation. We propose to build on recent work on mean-field Schrödinger bridges [24] which is a generalization of the Schrödinger bridge theory when the prior dynamics are modeled by a McKean-Vlasov equation, and reformulate this problem into an optimization over the space of path measures. With proper discretization over space and time the problem becomes a nonlinear version of the multi-marginal optimal transport (MOT) [25], [26], [27]. To numerically solve this problem, we adopt the proximal gradient algorithm [28], [29] to sequentially linearize the nonlinear MOT and then take advantage of existing algorithms [26], [27] for MOT for each iteration. We further extend our method to the setting when multiple species are involved. Finally, in the linear quadratic setting where the dynamics are linear and the cost function is quadratic, we characterize the optimal solution via a system of coupled Riccati equations and obtain the closed form solution.

The density control problem we consider is related to but differs from three lines of research: mean-field games [4], [30], mean-field type control [31], and mean-field optimal control [32]. In mean-field games, each individual aims to minimize her own cost while in density control the individuals have a common objective to optimize. Mean-field type control is similar to mean-field optimal control; they differ from each other in the type of dynamical systems they cover. The former focuses on stochastic individual dynamics while the latter focuses on deterministic dynamics. They can both be formulated

This work was supported by the NSF under grant 2206576, 1942523 and 2008513. Y. Chen is with the School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, USA. {yongchen}@gatech.edu

as optimal control problems over the space of distributions, just like the density control problem we study. The existing work on mean-field type control and mean-field optimal control focuses on theoretical questions such as the well-posedness of the problem and the characterization of the optimal solution. Numerical studies were only conducted under very strong assumptions such as linear dynamics, quadratic cost, or first-order moment-based interaction among the individuals. Also, in mean-field optimal control [32], the control strategy is parametrized in a specific way so that the problem becomes finding the optimal parameters. In contrast, the control strategy we search for can be any function. Another major difference between the density control problem and the three other problems mentioned above is the control objective. In the density control problem, an explicit target distribution constraint is imposed while in the other three problems, a terminal cost is used to regularize the collective dynamics. The target distribution constraint adds another layer of complexity; this constraint is normally associated with an unknown terminal cost that needs to be recovered in order to compute the optimal control [14], [33].

The rest of the paper is organized as follows. In Section II we briefly introduce the several tools that will be used in this work. The main results on density control for interacting agent system are presented in Section III. An extension to density control problems involving multiple species is provided in Section IV. We investigate the problem in the linear quadratic setting in Section V and obtain a closed-form solution. Several numerical examples are presented in Section VI to illustrate the proposed framework. This is followed by a short concluding remark in VII.

II. BACKGROUND

In this section we introduce several mathematical tools on which our density control framework is based, including optimal transport and its generalizations, and the proximal gradient algorithm.

A. Optimal transport

Given two nonnegative measures μ, ν on \mathbb{R}^d having equal total mass (often assumed to be probability distributions), the Monge's formulation of optimal transport seeks a transport map $T : \mathbb{R}^d \rightarrow \mathbb{R}^d : x \mapsto T(x)$ from μ to ν in the sense $T_\# \mu = \nu$, that incurs minimum cost of transportation $\int c(x, T(x)) \mu(dx)$. Here, $c(x, y)$ stands for the transportation cost per unit mass from point x to y . The dependence of the total transportation cost on T is highly nonlinear, complicating early analyses to the problem [10]. This problem was later relaxed by Kantorovich, where, instead of a transport map, a joint distribution π on the product space $\mathbb{R}^d \times \mathbb{R}^d$ is sought. Let $\Pi(\mu, \nu)$ be the set of joint distributions of μ and ν , then the Kantorovich formulation of OT reads

$$\inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y) \pi(dx dy). \quad (1)$$

Both the Monge's and the Kantorovich's formulations are "static" focusing on "what goes where." It turns out that the OT problem can also be cast as a dynamical problem with a temporal dimension. In particular, when $c(x, y) = \frac{1}{2} \|x - y\|^2$, OT can be formulated as a stochastic control problem

$$\inf_{v \in \mathcal{V}} \mathbb{E} \left\{ \int_0^1 \frac{1}{2} \|v(t, x^v(t))\|^2 dt \right\}, \quad (2a)$$

$$\dot{x}^v(t) = v(t, x^v(t)), \quad (2b)$$

$$x^v(0) \sim \mu, \quad x^v(1) \sim \nu. \quad (2c)$$

Here \mathcal{V} represents the family of admissible state feedback control laws. Note that this control problem (2) differs from standard ones in that the terminal constraint $x^v(1) \sim \nu$, meaning $x^v(1)$ follows distribution ν , is unconventional. In (2), the goal is to find an optimal control policy to drive the system (2b) from an uncertain initial state $x^v(0) \sim \mu$ to an uncertain target state $x^v(1) \sim \nu$. The solution to (2) specifies how to move mass over time from configuration μ to ν , providing more resolution to the optimal transport plan.

Assuming $x^v(t)$ has a absolutely continuous distribution with density ρ_t , ρ_t satisfies weakly¹ the continuity equation

$$\partial_t \rho_t + \nabla \cdot (v \rho_t) = 0, \quad (3)$$

and the total transport cost becomes

$$\mathbb{E} \left\{ \int_0^1 \frac{1}{2} \|v(t, x^v(t))\|^2 dt \right\} = \int_{\mathbb{R}^d} \int_0^1 \frac{1}{2} \|v(t, x)\|^2 \rho_t(x) dt dx.$$

Thus, (2) is equivalent to [34]

$$\inf_{\rho, v} \int_{\mathbb{R}^d} \int_0^1 \frac{1}{2} \|v(t, x)\|^2 \rho_t(x) dt dx, \quad (4a)$$

$$\partial_t \rho_t + \nabla \cdot (v \rho_t) = 0, \quad (4b)$$

$$\rho_0 = \mu, \quad \rho_1 = \nu. \quad (4c)$$

The minimum is taken over all pairs (ρ, v) satisfying (4b)-(4c) and some technical conditions, see [10, Theorem 8.1], [35, Chapter 8].

Suppose we have a large number of individuals that share the same dynamics (2b) but are independent from each other. Assume their initial states follow the same distribution μ , then (4) can be viewed as an density control problem for this group whose objective is to find a common control strategy so that it would reach target distribution/configuration ν at time $t = 1$. The solution to (4) is characterized by the coupled partial differential equations (PDEs)

$$\partial_t \lambda + \frac{1}{2} \nabla \lambda^T \nabla \lambda = 0$$

$$\partial_t \rho_t + \nabla \cdot (\rho_t \nabla \lambda) = 0$$

$$\rho_0 = \mu, \quad \rho_1 = \nu,$$

and the optimal control is $v(t, x) = \nabla \lambda(t, x)$.

B. Schrödinger bridges

In 1931/32, Schrödinger [36], [37] posed the following problem: A large number N of independent Brownian particles (whose randomness is scaled by $\sqrt{\epsilon}$) in \mathbb{R}^d is observed to have an empirical distribution approximately equal to μ at time $t = 0$, and at some later time $t = 1$ an empirical distribution approximately equal to ν . Suppose that ν differs from what it should be according to the law of large numbers, namely

$$\int q_\epsilon(0, x, 1, y) \mu(dx),$$

where $q_\epsilon(s, x, t, y) = (2\pi)^{-d/2} [\epsilon(t-s)]^{-d/2} \exp\left(-\frac{\|x-y\|^2}{2\epsilon(t-s)}\right)$ denotes the scaled Brownian transition probability density. It is apparent that the particles have been transported in an unlikely way. But of the many unlikely ways in which this could have happened, which one is the most likely?

This problem can be understood in the modern language of large deviation theory as a problem [38] of determining a probability law

¹In the sense that $\int_{[0,1] \times \mathbb{R}^d} [(\partial_t f + v \cdot \nabla f) \rho_t] dt dx = 0$ for smooth functions f with compact support.

\mathcal{P} on the path space $\Omega = C([0, 1], \mathbb{R}^d)$ that minimizes the relative entropy (a.k.a., Kullback-Leibler divergence)²

$$\text{KL}(\mathcal{P} \parallel \mathcal{Q}) := \int_{\Omega} \log \left(\frac{d\mathcal{P}}{d\mathcal{Q}} \right) d\mathcal{P}. \quad (5)$$

Here \mathcal{Q} is the probability law induced by the Brownian motion and \mathcal{P} is chosen among probability laws that are absolutely continuous with respect to \mathcal{Q} and have the prescribed marginals. The solution to this optimization is referred to as the *Schrödinger bridge*. Existence and uniqueness of the minimizer have been proven in various degrees of generality by Fortet [39], Beurling [40], Jamison [41], Föllmer [38].

It has been shown that the above Schrödinger's problem can be reformulated as the stochastic control problem [42]

$$\inf_{v \in \mathcal{V}} \mathbb{E} \left\{ \int_0^1 \frac{1}{2} \|v(t, X_t)\|^2 dt \right\}, \quad (6a)$$

$$dX_t = v(t, X_t)dt + \sqrt{\epsilon} dB_t, \quad (6b)$$

$$X_0 \sim \mu, \quad X_1 \sim \nu. \quad (6c)$$

Here \mathcal{V} is the class of finite energy Markov controls. This reformulation relies on the fact that the relative entropy between distributions induced by the controlled and uncontrolled processes is

$$\text{KL}(\mathcal{P} \parallel \mathcal{Q}) = \mathbb{E} \left\{ \int_0^1 \frac{1}{2\epsilon} \|v(t, X_t)\|^2 dt \right\}.$$

The proof is based on Girsanov theorem, see [43], [44]. The problem (6) has the following density control reformulation [45], [46]

$$\inf_{\rho, v} \int_{\mathbb{R}^d} \int_0^1 \frac{1}{2} \|v(t, x)\|^2 \rho_t dt dx, \quad (7a)$$

$$\partial_t \rho_t + \nabla \cdot (v \rho_t) - \frac{\epsilon}{2} \Delta \rho_t = 0, \quad (7b)$$

$$\rho_0 = \mu, \quad \rho_1 = \nu, \quad (7c)$$

where the infimum is over smooth fields v and ρ that solve weakly of the corresponding Fokker-Planck equation (7b).

Formulation (7) resembles the OT problem (4) except for the presence of the Laplacian in (7b). It has been shown [47], [48], [49], [11] that the OT problem is, in a suitable sense, indeed the limit of the Schrödinger problem when the diffusion coefficient ϵ of the reference Brownian motion goes to zero. On the other hand, the Schrödinger bridge can be viewed as a regularized version of OT. Similar to (4), the solution to (7) is characterized by the coupled PDEs

$$\partial_t \lambda + \frac{1}{2} \nabla \lambda^T \nabla \lambda + \frac{\epsilon}{2} \Delta \lambda = 0 \quad (8a)$$

$$\partial_t \rho_t + \nabla \cdot (\rho_t \nabla \lambda) - \frac{\epsilon}{2} \Delta \rho_t = 0 \quad (8b)$$

$$\rho_0 = \mu, \quad \rho_1 = \nu, \quad (8c)$$

and the corresponding optimal control policy is $v(t, x) = \nabla \lambda(t, x)$.

C. Multi-marginal optimal transport

Next we introduce discrete OT where the distributions have discrete supports. In this setting, the marginals $\mu_1 \in \mathbb{R}_+^{d_1}, \mu_2 \in \mathbb{R}_+^{d_2}$ are nonnegative vectors with equal sum. The transport cost can be rewritten in the matrix form $\mathbf{C} = [C(x_1, x_2)] \in \mathbb{R}^{d_1 \times d_2}$ where $C(x_1, x_2)$ represents the cost of moving a unit mass from x_1 to x_2 . Similarly, a transport plan is encoded in a joint probability matrix $\mathbf{B} = [B(x_1, x_2)] \in \mathbb{R}_+^{d_1 \times d_2}$ of μ_1, μ_2 . The total transport cost is $\sum_{x_1, x_2} C(x_1, x_2) B(x_1, x_2) = \text{Tr}(\mathbf{C}^T \mathbf{B})$ and the problem becomes

$$\min_{\mathbf{B} \in \mathbb{R}_+^{d_1 \times d_2}} \text{Tr}(\mathbf{C}^T \mathbf{B}) \quad \text{subject to} \quad \mathbf{B} \mathbf{1} = \mu_1, \quad \mathbf{B}^T \mathbf{1} = \mu_2, \quad (9)$$

² $\frac{d\mathcal{P}}{d\mathcal{Q}}$ denotes the Radon-Nikodym derivative between \mathcal{P} and \mathcal{Q} .

where $\mathbf{1}$ denotes a vector of ones of proper dimension.

Multi-marginal optimal transport (MOT) extends OT to the setting involving multiple distributions. In particular, in MOT, one seeks a transport plan among a set of marginals μ_1, \dots, μ_J with $J \geq 2$. In the discrete setting, the transport cost is encoded in a tensor $\mathbf{C} = [C(x_1, x_2, \dots, x_J)] \in \mathbb{R}^{d_1 \times d_2 \times \dots \times d_J}$ where $C(x_1, x_2, \dots, x_J)$ denotes the unit cost associated with (x_1, x_2, \dots, x_J) , and the transport plan is described by a tensor $\mathbf{B} \in \mathbb{R}_+^{d_1 \times d_2 \times \dots \times d_J}$. For a transport plan \mathbf{B} , the total cost is $\langle \mathbf{C}, \mathbf{B} \rangle := \sum_{x_1, x_2, \dots, x_J} C(x_1, \dots, x_J) B(x_1, \dots, x_J)$. Thus, similar to (9), MOT has a linear programming formulation

$$\min_{\mathbf{B} \in \mathbb{R}_+^{d_1 \times \dots \times d_J}} \langle \mathbf{C}, \mathbf{B} \rangle \quad \text{subject to} \quad P_j(\mathbf{B}) = \mu_j, \quad \text{for } j \in \Gamma, \quad (10)$$

where $\Gamma \subset \{1, \dots, J\}$ is an index set specifying given marginal distributions, and the projection on the j -th marginal is defined by

$$P_j(\mathbf{B}) = \sum_{x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_J} B(x_1, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_J). \quad (11)$$

A popular method to solve the OT problem is entropy regularization, which adds an entropy term

$$\mathcal{H}(\mathbf{B}) = - \sum_{x_1, \dots, x_J} B(x_1, \dots, x_J) \log B(x_1, \dots, x_J) \quad (12)$$

to (10), resulting in the strictly convex optimization problem

$$\min_{\mathbf{B} \in \mathbb{R}_+^{d_1 \times \dots \times d_J}} \langle \mathbf{C}, \mathbf{B} \rangle - \epsilon \mathcal{H}(\mathbf{B}) \quad \text{subject to} \quad P_j(\mathbf{B}) = \mu_j, \quad \text{for } j \in \Gamma \quad (13)$$

with $\epsilon > 0$ being a regularization parameter. Invoking Lagrangian duality, one can show that the optimal solution to (13) is

$$\mathbf{B} = \mathbf{K} \odot \mathbf{U}, \quad (14)$$

where \odot denotes element-wise multiplication,

$$\mathbf{K} = \exp(-\mathbf{C}/\epsilon), \quad (15)$$

and $\mathbf{U} = \mathbf{u}_1 \otimes \mathbf{u}_2 \otimes \dots \otimes \mathbf{u}_J$ with the vectors $\mathbf{u}_j \in \mathbb{R}^{d_j}$ being associated with the Lagrange multipliers. The Sinkhorn algorithm [50], [51], [52] iteratively updates the vectors \mathbf{u}_j according to

$$\mathbf{u}_j \leftarrow \mathbf{u}_j \odot \mu_j ./ P_j(\mathbf{K} \odot \mathbf{U}), \quad (16)$$

for all $j \in \Gamma$. Here $./$ denotes element-wise division. The Sinkhorn algorithm has a global linear convergence guarantee [53], [54], [55], [26]. Nevertheless, its complexity still scales exponentially as J grows. The bottleneck of it lies in the calculation of the projections $P_j(\mathbf{B})$, $j \in \Gamma$ in (11).

Recently it was discovered that the computation of MOT can be greatly accelerated if the cost tensor \mathbf{C} has a graphical structure [27], that is, the cost tensor \mathbf{C} can be decomposed as

$$C(\mathbf{x}) = C(x_1, x_2, \dots, x_J) = \sum_{(i,j) \in E} C_{ij}(x_i, x_j), \quad (17)$$

where E denotes the set of edges of an undirected graph $G = (V, E)$.

The marginal constraints $P_j(\mathbf{B}) = \mu_j$ for the graphical OT problem can be imposed on any variable node $j \in V$. Consider the entropy regularized MOT problem (13). When the cost \mathbf{C} has form (17), \mathbf{K} in (15) equals

$$\mathbf{K} = [K(\mathbf{x})] = \left[\prod_{(i,j) \in E} K_{ij}(x_i, x_j) \right]$$

with $K_{ij}(x_i, x_j) = \exp(-C_{ij}(x_i, x_j)/\epsilon)$. It follows that the optimal solution (14) to the entropy regularized MOT problem (13) has a graphical representation as

$$\mathbf{B} = \mathbf{K} \odot \mathbf{U} = [K(\mathbf{x})U(\mathbf{x})] = \left[\left(\prod_{(i,j) \in E} K_{ij}(x_i, x_j) \right) \left(\prod_{j \in V} u_j(x_j) \right) \right].$$

This is nothing but a probabilistic graphical model [56], and calculating the projection $P_j(\mathbf{K} \odot \mathbf{U})$ is exactly a Bayesian inference [56] problem of inferring the j -th variable node over this graphical model.

When the graphical structure of \mathbf{C} is a tree, we arrive at the Sinkhorn belief propagation [27] algorithm (Algorithm 1), to solve the MOT problem (13): applying the Sinkhorn algorithm and utilizing the Belief Propagation algorithm [57] to carry out the computation of $P_j(\mathbf{K} \odot \mathbf{U})$ with the current multiplier \mathbf{U} . Here we have assumed, without loss of generality, Γ is a subset of the leaf nodes [27]. The algorithm relies on messages (known as beliefs) $m_{i \rightarrow j}$ passed from a node i to its neighbor $j \in N(i)$ with $N(i)$ denoting the set of all the neighbors of i . Let j_1, j_2, \dots be a sequence taking values in Γ in cyclic order and suppose the Sinkhorn algorithm is carried out in this order, then after the k -th iteration, \mathbf{u}_{j_k} is updated, and the only projection required in the next iteration is $P_{j_{k+1}}(\mathbf{K} \odot \mathbf{U})$. It suffices to update the messages on the path from j_k to j_{k+1} to evaluate $P_{j_{k+1}}(\mathbf{K} \odot \mathbf{U})$. Compared with standard Sinkhorn algorithm, the acceleration of SBP is tremendous for MOT problems with a large number of marginals; the Belief Propagation algorithm scales well for large problem while the complexity of the brute force projection using the definition (11) grows exponentially as the number of marginals increases. Upon convergence of Algorithm 1, the solution to graphical OT problem can be obtained through $\mathbf{B} = \mathbf{K} \odot \mathbf{U}$ with $\mathbf{U} = \mathbf{u}_1 \otimes \mathbf{u}_2 \otimes \dots \otimes \mathbf{u}_J$, where $\mathbf{u}_j = \boldsymbol{\mu}_j / \mathbf{m}_{i \rightarrow j}$, $i \in N(j)$ for $j \in \Gamma$, and $\mathbf{u}_j = \mathbf{1}$ otherwise. For general graphical structures, we convert it into a tree first using the junction tree algorithm [56] and then apply Algorithm 1 on the resulting junction tree. The complexity of the algorithm depends on the node size of the junction tree, which scales exponentially as the tree-width of the graph. For graphical structure with small tree width, this algorithm is still efficient.

D. Proximal gradient algorithm

The proximal gradient algorithm [29] is an popular algorithm for the composite optimization

$$\min_{y \in \mathcal{Y}} F(y) + G(y), \quad (19)$$

where \mathcal{Y} denotes the feasibility set. The function F is assumed to be smooth. The function G is usually a regularizer that is possibly nonsmooth. The algorithm reads

$$y^{k+1} = \operatorname{argmin}_{y \in \mathcal{Y}} G(y) + \frac{1}{2\eta} \|y - (y^k - \eta \nabla F(y^k))\|^2 \quad (20a)$$

$$= \operatorname{argmin}_{y \in \mathcal{Y}} G(y) + \frac{1}{2\eta} \|y - y^k\|^2 + \langle \nabla F(y^k), y - y^k \rangle \quad (20b)$$

where $\eta > 0$ is the stepsize. One advantage of the proximal gradient algorithm is that it only evaluates the gradient of F and doesn't require G to be differentiable. In many applications, G is a regularizer of simple form, e.g., 1-norm, and the minimization (20) can be implemented efficiently.

The proximal gradient algorithm has been generalized to the non-Euclidean setting. It is built upon the mirror descent method [28], [29]. Let $D(\cdot, \cdot)$ be a Bregman divergence, then the generalized non-Euclidean proximal gradient algorithm reads

$$y^{k+1} = \operatorname{argmin}_{y \in \mathcal{Y}} G(y) + \frac{1}{\eta} D(y, y^k) + \langle \nabla F(y^k), y - y^k \rangle. \quad (21)$$

A popular choice of $D(\cdot, \cdot)$ is the Kullback-Leibler divergence $\text{KL}(\cdot \| \cdot)$, which is suitable for optimization over probability vectors/distributions.

The (generalized) proximal gradient algorithm has nice convergence properties. When both F and G are convex, the algorithm is guaranteed to converge to the global minimum with rate $\mathcal{O}(1/k)$ [28], [29]. When F is nonconvex, one can only expect for convergence to local solutions. It turns out that objective function $F(y) + G(y)$ is monotonically decreasing along the updates, and the updates converge to some stationary points with sublinear rate $\mathcal{O}(1/k)$ with respect to some suitable criteria [58].

III. DENSITY CONTROL OF INTERACTING AGENT SYSTEMS

Consider a collection of N dynamical systems

$$dX_t^i = -\frac{1}{N} \sum_{j=1}^N \nabla W(X_t^i - X_t^j) dt + u_t^i dt + \sqrt{\epsilon} dB_t^i, \quad i = 1, \dots, N, \quad (22)$$

where $X_t^i \in \mathbb{R}^d$, $u_t^i \in \mathbb{R}^d$ denote the state and control of agent i respectively. The stochastic disturbance is modeled by a standard Wiener process B_t and $\epsilon \geq 0$ captures the level of stochasticity in the dynamics. The N agents interact with each other through an interaction potential W , which is assumed to be continuously differentiable and symmetric, i.e., $W(x) = W(-x)$, $\forall x$. Clearly, $\nabla W(0) = 0$. The Hessian of W is assumed to be bounded, from both above and below [24]. We are interested in controlling the collective dynamics of the individuals (22). Our goal is to find a common feedback strategy for the N agents to steer them from an initial group configuration to a target configuration over a finite time interval $[0, 1]^3$ with minimum effort. Let $\xi_t(x)$ be the feedback strategy of the agents, meaning $u_t^i = \xi_t(X_t^i)$. The cost function to minimize is the average quadratic control effort

$$\mathbb{E} \left\{ \int_0^1 \frac{1}{2N} \sum_i \|\xi_t(X_t^i)\|^2 dt \right\}.$$

In the mean field limit as $N \rightarrow \infty$, the group behavior can be captured by a probability distribution

$$\rho_t \approx \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}$$

with δ_x denoting the Dirac distribution, and this density evolves according to the McKean-Vlasov equation [3]

$$\partial_t \rho_t + \nabla \cdot (\rho_t (-\nabla W * \rho_t + \xi_t)) - \frac{\epsilon}{2} \Delta \rho_t = 0. \quad (23)$$

The average control effort is approximately

$$\int_0^1 \int_{\mathbb{R}^d} \frac{1}{2} \|\xi_t(x)\|^2 \rho_t(x) dx dt.$$

The initial and target configurations can both be modeled by probability distributions. Thus, in the mean field limit, our density/distribution problem can be formulated as

$$\inf_{\rho, \xi} \int_0^1 \int_{\mathbb{R}^d} \frac{1}{2} \|\xi_t(x)\|^2 \rho_t(x) dx dt \quad (24a)$$

$$\partial_t \rho_t + \nabla \cdot (\rho_t (-\nabla W * \rho_t + \xi_t)) - \frac{\epsilon}{2} \Delta \rho_t = 0 \quad (24b)$$

$$\rho_0 = \mu, \quad \rho_1 = \nu. \quad (24c)$$

³We use the unit time interval $[0, 1]$ to simplify the notation. A general time interval can be transformed into $[0, 1]$ by rescaling.

Algorithm 1 Sinkhorn Belief Propagation (SBP) Algorithm

Input: $K_{ij}(x_i, x_j) = \exp(-C_{ij}(x_i, x_j)/\epsilon)$, $(i, j) \in E$, marginals μ_j for $j \in \Gamma$. Initialize the messages $m_{i \rightarrow j}(x_j)$ to be 1

Let j_1, j_2, \dots be a sequence taking values in Γ in cyclic order

Output: $\mathbf{u}_j = \mu_j / \mathbf{m}_{i \rightarrow j}$, $i \in N(j)$ for $j \in \Gamma$, and $\mathbf{u}_j = \mathbf{1}$ otherwise. $P_j(\mathbf{K} \odot \mathbf{U}) \propto \prod_{k \in N(j)} \mathbf{m}_{k \rightarrow j}$, $j \notin \Gamma$

while not converged **do**

Update $m_{j_k \rightarrow i}(x_i)$, $i \in N(j_k)$ using

$$m_{j \rightarrow i}(x_i) \propto \sum_{x_j} K_{ij}(x_i, x_j) \frac{\mu_j(x_j)}{m_{i \rightarrow j}(x_j)}, \quad \forall x_i \quad (18a)$$

Update the rest of messages on the path from node j_k to node j_{k+1} according to

$$m_{i \rightarrow j}(x_j) \propto \sum_{x_i} K_{ij}(x_i, x_j) \prod_{k \in N(i) \setminus j} m_{k \rightarrow i}(x_i), \quad \forall x_j \quad (18b)$$

end while

One can view this as an optimal control problem for a dynamical system over the space of probability distributions with ρ_t being the state. We note that while the control policy ξ is closed-loop in individual level, it is open-loop for (24) where the distribution is the state. We also note that one can in principle consider a more general control strategy of the form

$$u_t^i = \xi_t(X_t^i, \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}),$$

implying the control for each individual should not only depend on its state, but also the population distribution. In the mean field limit, $1/N \sum_{i=1}^N \delta_{X_t^i} \approx \rho_t$ is determined by (23) and is a function of time in the end. Thus, this general form of control does not make any difference in the mean field limit. The dynamics is (24b) with state ρ_t and control ξ_t . The constraints (24c) specify the initial and terminal states. We seek an optimal strategy with minimum control effort to steer the agents from an initial distribution μ to a target distribution ν . A major difference between our density control problem (24) and the mean-field type control problem [31] or the mean-field optimal control problem [32] is the nonstandard terminal constraint $\rho_1 = \nu$.

Using the Lagrangian duality method, one can derive a characterization of the solutions to (24). In particular, the optimal solution to (24) can be characterized by the coupled PDEs

$$\begin{aligned} \partial_t \lambda + \frac{1}{2} \nabla \lambda^T \nabla \lambda - \nabla \lambda^T \nabla W * \rho_t \\ - \int_{\mathbb{R}^d} \rho_t(y) \nabla \lambda(y)^T \nabla W(y - x) dy + \frac{\epsilon}{2} \Delta \lambda = 0 \end{aligned} \quad (25a)$$

$$\begin{aligned} \partial_t \rho_t + \nabla \cdot (\rho_t (-\nabla W * \rho_t + \nabla \lambda)) - \frac{\epsilon}{2} \Delta \rho_t = 0 \quad (25b) \\ \rho_0 = \mu, \quad \rho_1 = \nu, \quad (25c) \end{aligned}$$

where λ is the Lagrange multiplier associated with the continuity constraint (24b). Under the assumptions that W is symmetric, smooth, and has bounded Hessian, and μ, ν have finite free energy, it is shown in [24] that (25) admits a solution. The optimal control policy is a state feedback

$$\xi_t(x) = \nabla \lambda(t, x).$$

There are several potential approaches to compute an optimal solution to the density control problem (24). For instance, the optimality condition (25) can be viewed as the Pontryagin's principle for (24) when (24) is treated as an optimal control problem with state ρ_t [59]. The multiplier λ then becomes the costate in the Pontryagin's principle [60], [61]. To get a solution to (25), one can use indirect method such as shooting method [60], [61] that is widely adopted for optimal control problems. However, due to the coupling between the

state ρ_t and the costate λ , the nonstandard terminal constraint $\rho_1 = \nu$, and more importantly the fact that they are of infinite dimension, the shooting method maybe unstable and is not guaranteed to converge. Next we present a completely different approach to solve (24) based on a reformulation.

A. Reformulation and Discretization

For a given feedback policy ξ_t , in the mean field limit, the distribution ρ_t of the individuals follows the McKean-Vlasov equation (23) and is deterministic. Moreover, the interaction between agents is of the form $-\frac{1}{N} \sum_{j=1}^N \nabla W(X_t^i - X_t^j) \approx -\nabla W * \rho_t$, which only depends on the group behavior. Thus, when N is sufficiently large, the interactions between an agent with other agents becomes the interaction between the agent and the deterministic group distribution ρ_t . By the theory of propagation of chaos [62], the N agents become effectively independent to each other and each of them follows the same stochastic dynamics

$$dX_t = -[\nabla W * \rho_t](X_t)dt + \xi_t(X_t)dt + \sqrt{\epsilon}dB_t. \quad (26)$$

Denote by \mathcal{P} the distribution induced by (26) over the path space $\Omega = C([0, 1], \mathbb{R}^d)$, and by $\mathcal{Q}(\mathcal{P})$ be distribution induced by the process

$$dX_t = -[\nabla W * \rho_t](X_t)dt + \sqrt{\epsilon}dB_t, \quad (27)$$

then by the Girsanov theorem [61], [44], following a similar argument as in the Schrödinger bridge problem (5)-(7), we obtain

$$\text{KL}(\mathcal{P} \| \mathcal{Q}(\mathcal{P})) = \int_0^1 \int_{\mathbb{R}^d} \frac{1}{2\epsilon} \|\xi_t(x)\|^2 \rho_t(x) dx dt.$$

Note that we used $\mathcal{Q}(\mathcal{P})$ to emphasize the fact that \mathcal{Q} depends on the marginal flow of \mathcal{P} , denoted by $\rho_t = (X_t)_\# \mathcal{P} = \mathcal{P}_t$.

Consequently, the density control problem (24) can be reformulated as

$$\min_{\mathcal{P}} \text{KL}(\mathcal{P} \| \mathcal{Q}(\mathcal{P})) \quad (28a)$$

$$(X_0)_\# \mathcal{P} = \mu, \quad (X_1)_\# \mathcal{P} = \nu. \quad (28b)$$

This formulation (28) coincides with the mean field Schrödinger bridge problem [24]. The equivalence between (24) and (28) is rigorously justified in [24], extending the large deviation theory to interacting particle systems. The major difference between (28) and the standard Schrödinger bridge problem (5) lies in the fact that the prior distribution \mathcal{Q} in the former depends on the solution \mathcal{P} , rendering a nonconvex optimization, in general, over the space of path distributions.

The optimization variable \mathcal{P} of (28) is of infinite dimension. To develop an implementable algorithm for (28), we first discretize the problem in time $t_i = i/T$, $i = 0, 1, \dots, T$ as well as in space over a grid. With this discretization, the path distribution \mathcal{P} becomes a $(T+1)$ -dimensional tensor \mathbf{M} with $M(x_0, x_1, \dots, x_T)$ representing the probability of the process \mathcal{P} goes through a grid neighborhood of $(X_0 = x_0, X_{1/T} = x_1, \dots, X_1 = x_T)$. In terms of \mathbf{M} , the objective function $\text{KL}(\mathcal{P} \parallel \mathcal{Q}(\mathcal{P}))$ becomes

$$\langle \mathbf{C}(\mathbf{M}), \mathbf{M} \rangle + \epsilon \langle \mathbf{M}, \log \mathbf{M} \rangle$$

where

$$\langle \mathbf{M}, \log \mathbf{M} \rangle = \sum_{x_0, x_1, \dots, x_T} M(x_0, x_1, \dots, x_T) \log M(x_0, x_1, \dots, x_T),$$

is the negative entropy $-\mathcal{H}(\mathbf{M})$, and $C(\mathbf{M})(x_0, x_1, \dots, x_T)$ is the minimum control effort to drive the deterministic version ($\epsilon = 0$) of (26) to go through the state x_0, x_1, \dots, x_T . More explicitly, when the discretization grid is sufficiently fine,

$$C(\mathbf{M})(x_0, x_1, \dots, x_T) = \sum_{i=0}^{T-1} \frac{T}{2} \|x_{i+1} - x_i + \frac{1}{T} [\nabla W * P_i(\mathbf{M})](x_i)\|^2,$$

where $P_i(\mathbf{M})$ denotes the marginal of \mathbf{M} over x_i and, by abuse of notation, $\nabla W * P_i(\mathbf{M})$ is a discretization of the convolution.

Thus, after discretization, (28) becomes

$$\min_{\mathbf{M}} \quad \langle \mathbf{C}(\mathbf{M}), \mathbf{M} \rangle + \epsilon \langle \mathbf{M}, \log \mathbf{M} \rangle \quad (29a)$$

$$P_0(\mathbf{M}) = \mu, \quad P_T(\mathbf{M}) = \nu. \quad (29b)$$

Here μ and ν denote the discretized version of μ and ν respectively. This formulation (29) is akin to the MOT problem except that the unit transport cost tensor \mathbf{C} now depends on the optimization variable \mathbf{M} . This difference excludes the possibility of applying the Sinkhorn type algorithm directly to solve (29). Next we develop an algorithm to compute the solution to (29) by sequentially linearizing $\langle \mathbf{C}(\mathbf{M}), \mathbf{M} \rangle$ and then solving the resulting MOT problems.

B. Proximal Sinkhorn Belief Propagation Algorithm

Denote $\Pi(\mu, \nu)$ the set of \mathbf{M} that is consistent with the marginals μ, ν and $F(\mathbf{M}) = \langle \mathbf{C}(\mathbf{M}), \mathbf{M} \rangle$ then (29) reads

$$\min_{\mathbf{M} \in \Pi(\mu, \nu)} F(\mathbf{M}) - \epsilon \mathcal{H}(\mathbf{M}). \quad (30)$$

This is a composite optimization over the probability simplex. We can thus apply the generalized proximal gradient descent algorithm to solve it. Surprisingly, when the Bregman divergence in (21) is chosen to be the Kullback-Leibler divergence, each iteration of the algorithm on the problem (30) takes the form

$$\mathbf{M}_{k+1} = \arg\min_{\mathbf{M} \in \Pi(\mu, \nu)} \langle \nabla F(\mathbf{M}_k), \mathbf{M} \rangle + \frac{1}{\eta} \text{KL}(\mathbf{M} \parallel \mathbf{M}_k) - \epsilon \mathcal{H}(\mathbf{M}) \quad (31)$$

where $\eta > 0$ is the step size. Expanding the KL divergence term, the above becomes

$$\mathbf{M}_{k+1} = \arg\min_{\mathbf{M} \in \Pi(\mu, \nu)} \langle \nabla F(\mathbf{M}_k), \mathbf{M} \rangle - \frac{1}{\eta} \log \mathbf{M}_k, \mathbf{M} - (\epsilon + \frac{1}{\eta}) \mathcal{H}(\mathbf{M}) \quad (32)$$

which is a standard entropy regularized multi-marginal optimal transport problem (13) with cost tensor $\nabla F(\mathbf{M}_k) - \frac{1}{\eta} \log \mathbf{M}_k$.

Proposition 1. The gradient of $F(\mathbf{M}) = \langle \mathbf{C}(\mathbf{M}), \mathbf{M} \rangle$ is

$$\nabla F(\mathbf{M}) = \mathbf{C}(\mathbf{M}) + E(\mathbf{M}) \quad (33)$$

Algorithm 2 Proximal Sinkhorn Belief Propagation algorithm

Input: cost tensor \mathbf{C} , regularization ϵ , stepsize η , number of iterations K . Initialize \mathbf{M}_1 to be a uniform probability vector 1

Output: \mathbf{M}_{K+1}

for $k = 1, 2, 3, \dots, K$ **do**

 Compute $\mathbf{C}(\mathbf{M}_k) + E(\mathbf{M}_k) - \frac{1}{\eta} \log \mathbf{M}_k$

 Solve (35) using the Sinkhorn Belief Propagation algorithm (Algorithm 1) to obtain \mathbf{M}_{k+1}

end for

where $E(\mathbf{M})(x_0, x_1, \dots, x_T) = \sum_{i=0}^{T-1} E_i(x_i)$ with

$$E_i(y) = \sum_{x_i, x_{i+1}} \nabla W(x_i - y)^T [x_{i+1} - x_i + \frac{1}{T} \nabla W * P_i(\mathbf{M})] P_{i,i+1}(\mathbf{M})(x_i, x_{i+1}). \quad (34)$$

Proof. See Appendix A. \square

Plugging (33) into (32) yields the proximal gradient iteration

$$\begin{aligned} \mathbf{M}_{k+1} = \arg\min_{\mathbf{M} \in \Pi(\mu, \nu)} & \langle \mathbf{C}(\mathbf{M}_k) + E(\mathbf{M}_k) - \frac{1}{\eta} \log \mathbf{M}_k, \mathbf{M} \rangle \\ & - (\epsilon + \frac{1}{\eta}) \mathcal{H}(\mathbf{M}). \end{aligned} \quad (35)$$

Note that both $\mathbf{C}(\mathbf{M}_k)$ and $E(\mathbf{M}_k)$ have a graphical structure associated with the line graph (Figure 1). Thus, assuming \mathbf{M}_k has the same graphical structure, the solution \mathbf{M}_{k+1} to (35) also has a graphical structure corresponding to the line graph. Therefore, with proper initialization, each iteration (35) can be solved efficiently using the Sinkhorn Belief Propagation algorithm (Algorithm 1). We thus establish our Proximal Sinkhorn Belief Propagation algorithm (Algorithm 2) to solve (29). The Proximal Sinkhorn Belief Propagation algorithm inherits the convergent properties of proximal gradient algorithm and converges to a solution with sublinear rate $\mathcal{O}(1/k)$. Note that the problem (29) is in general non-convex and thus the convergence is to a local solution.

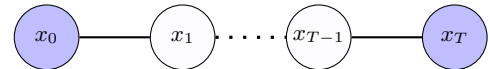


Fig. 1: Graph for the graphical OT (35)

Theorem 1. Suppose the cost tensor \mathbf{C} in (29) is bounded below. Assume every iteration is solved by the Sinkhorn Belief Propagation algorithm exactly. Then, the Proximal Sinkhorn Belief Propagation algorithm converges to a local solution to (29) with sublinear rate $\mathcal{O}(1/k)$.

Remark 2. Each iteration of our algorithm requires solving a graphical OT problem using the Sinkhorn Belief Propagation algorithm. Let D be the number of discretized grid points over space, then the complexity of the Sinkhorn Belief Propagation is $\mathcal{O}(D^2T)$.

Remark 3. In the limit case where $\epsilon = 0$, the stochastic disturbance in the dynamics vanishes and agents become deterministic. Note that Algorithm 2 applies to this deterministic setting.

C. Optimal control strategy

The optimal control policy is $\xi_t(x) = \nabla \lambda(t, x)$. Once Algorithm 2 converges, the corresponding control policy can be recovered by solving the linear equation

$$\partial_t \rho_t + \nabla \cdot (\rho_t (-\nabla W * \rho_t + \nabla \lambda(t, \cdot))) - \frac{\epsilon}{2} \Delta \rho_t = 0.$$

More specifically,

$$g_t := \partial_t \rho_t + \nabla \cdot (\rho_t (-\nabla W * \rho_t)) - \frac{\epsilon}{2} \Delta \rho_t$$

can be estimated using the solution \mathbf{M}^* to (29). It follows that λ can be recovered by solving the linear equation

$$\nabla \cdot (\rho_t \nabla \lambda(t, \cdot)) = -g_t,$$

or more precisely the least square problem

$$\min_{\lambda} \|\nabla \cdot (\rho_t \nabla \lambda(t, \cdot)) + g_t\|^2.$$

An alternative approach is based on the fact that the optimal \mathcal{P} is associated with the stochastic process

$$dX_t = -\nabla W * \mathcal{P}_t dt + \nabla \lambda(t, X_t) dt + \sqrt{\epsilon} dB_t.$$

The joint distribution of $X_{i/T}$ and $X_{(i+1)/T}$ of this process is approximately

$$\mathcal{P}_{i/T}(X_{i/T}) \mathcal{N}(X_{(i+1)/T}; X_{i/T} + [(-\nabla W * \mathcal{P}_{i/T})(X_{i/T}) + \nabla \lambda(i/T, X_{i/T})]/T, \epsilon/T),$$

where \mathcal{N} denotes a Gaussian distribution. On the other hand, it is approximated by $P_{i,i+1}(\mathbf{M}^*)$. Combining these two expressions we can solve λ and thus the optimal control policy.

D. Extension to general dynamics and cost

In the above discussions, to better illustrate our density control framework for interacting agent systems, we have restricted our attention to the simple dynamics (22). Now we extend this framework to more general dynamics⁴

$$dX_t^i = -\frac{1}{N} \sum_{j=1}^N \nabla W(X_t^i - X_t^j) dt + b(X_t) dt + \sigma(u_t^i dt + \sqrt{\epsilon} dB_t^i) \\ i = 1, \dots, N, \quad (36)$$

where $b(\cdot) \in \mathbb{R}^d$ is a continuous drift term and $\sigma \in \mathbb{R}^{d \times p}$ is the input matrix, and more general cost function

$$\int_0^1 \int_{\mathbb{R}^d} [\frac{1}{2} \|\xi_t(x)\|^2 + V(x)] \rho_t(x) dx dt.$$

In the mean field limit, the density control problem can be formulated as

$$\inf_{\rho, \xi} \int_0^1 \int_{\mathbb{R}^d} [\frac{1}{2} \|\xi_t(x)\|^2 + V(x)] \rho_t(x) dx dt \quad (37a)$$

$$\partial_t \rho_t + \nabla \cdot (\rho_t (-\nabla W * \rho_t + b + \sigma \xi_t)) - \frac{\epsilon}{2} \sum_{i,k} \frac{\partial^2 (a_{ik} \rho_t)}{\partial x_i \partial x_k} = 0 \quad (37b)$$

$$\rho_0 = \mu, \quad \rho_1 = \nu \quad (37c)$$

where $a = \sigma \sigma^T$.

The optimal strategy of (37) is

$$\xi_t(x) = \sigma^T \nabla \lambda(t, x)$$

⁴The dependence of b, σ over time is suppressed to simplify the notation.

where λ solves the PDEs

$$\partial_t \lambda + \frac{1}{2} \nabla \lambda^T a \nabla \lambda - V + \nabla \lambda^T b - \nabla \lambda^T \nabla W * \rho \\ - \int \rho(y) \nabla \lambda(y)^T \nabla W(y - x) dy + \frac{\epsilon}{2} \text{Tr}(a \nabla^2 \lambda) = 0 \quad (38a)$$

$$\partial_t \rho_t + \nabla \cdot (\rho_t (-\nabla W * \rho_t + b + a \nabla \lambda)) - \frac{\epsilon}{2} \sum_{i,k} \frac{\partial^2 (a_{ik} \rho_t)}{\partial x_i \partial x_k} = 0 \quad (38b)$$

$$\rho_0 = \mu, \quad \rho_1 = \nu. \quad (38c)$$

Following similar arguments as before we obtain an alternative formulation

$$\min_{\mathcal{P}} \quad \epsilon \text{KL}(\mathcal{P} \| \mathcal{Q}(\mathcal{P})) + \int V d\mathcal{P} \quad (39a)$$

$$(X_0)_\# \mathcal{P} = \mu, \quad (X_1)_\# \mathcal{P} = \nu, \quad (39b)$$

where $\mathcal{Q}(\mathcal{P})$ is the distribution over the path space associated with the diffusion process

$$dX_t = -[\nabla W * \mathcal{P}_t](X_t) dt + b(X_t) dt + \sqrt{\epsilon} dB_t. \quad (40)$$

The same as (29), after discretization over space and time, the problem can be written as

$$\min_{\mathbf{M} \in \Pi(\mu, \nu)} \langle \mathbf{C}(\mathbf{M}), \mathbf{M} \rangle + \epsilon \langle \mathbf{M}, \log \mathbf{M} \rangle, \quad (41)$$

but with a slightly different cost tensor

$$C(\mathbf{M})(x_0, x_1, \dots, x_T) = \frac{1}{T} \sum_{i=0}^{T-1} V(x_i) \\ + \sum_{i=0}^{T-1} \frac{T}{2} \|x_{i+1} - x_i + \frac{1}{T} \nabla W * P_i(\mathbf{M}) - \frac{1}{T} b(x_i)\|^2. \quad (42)$$

Let $F(\mathbf{M}) = \langle \mathbf{C}(\mathbf{M}), \mathbf{M} \rangle$, then the above becomes a composite optimization (30) and can be solved using the proximal gradient algorithm. The derivation is similar to that of Proposition 1 and is omitted.

Proposition 2. The gradient of $F(\mathbf{M}) = \langle \mathbf{C}(\mathbf{M}), \mathbf{M} \rangle$ with \mathbf{C} in (42) is

$$\nabla F(\mathbf{M}) = \mathbf{C}(\mathbf{M}) + E(\mathbf{M})$$

where $E(\mathbf{M})(x_0, x_1, \dots, x_T) = \sum_{i=0}^{T-1} E_i(x_i)$ with

$$E_i(y) = \sum_{x_i, x_{i+1}} \nabla W(x_i - y)^T [x_{i+1} - x_i + \frac{1}{T} \nabla W * P_i(\mathbf{M}) \\ - \frac{1}{T} b(x_i)] P_{i,i+1}(\mathbf{M})(x_i, x_{i+1}).$$

The proximal Sinkhorn belief propagation algorithm can be applied directly to solve (41) with a small modification on the expression of $E(\mathbf{M})$ as in Proposition 2.

IV. DENSITY CONTROL WITH MULTIPLE SPECIES

In this section, we extend our density control framework to account for the collective dynamics with multiple species. Consider a group of individuals comprised of L species and each has N agents. The dynamics of the i -th agents in the ℓ -th species is

$$dX_{\ell,t}^i = -\frac{1}{N} \sum_{m=1}^L \sum_{j=1}^N \nabla W_{\ell m}(X_{\ell,t}^i - X_{m,t}^j) dt + b_{\ell}(X_{\ell,t}^i) dt \\ + \sigma(u_{\ell,t}^i dt + \sqrt{\epsilon} dB_{\ell,t}^i), \quad i = 1, \dots, N,$$

where $X_{\ell,t}^i$ and $u_{\ell,t}^i$ denote the state and control of the i -th agents in the ℓ -th species respectively. The interaction potential between species ℓ and species m is assumed to be continuously differentiable and symmetric in the sense $W_{\ell m}(x) = W_{\ell m}(-x) = W_{m\ell}(x) = W_{m\ell}(-x)$. We seek L feedback policies, one for each species, such that, when they are adopted by the individuals, the group would be transformed from one configuration to another.

In the mean field region, denoting the initial distribution/configuration of species ℓ by μ_ℓ , and its target distribution by ν_ℓ , this density control problem can be formulated as

$$\inf_{\rho_1, \dots, \rho_L} \sum_{\ell=1}^L \int_0^1 \int_{\mathbb{R}^d} \left[\frac{1}{2} \|\xi_{\ell,t}(x)\|^2 + V_\ell(x) \right] \rho_{\ell,t}(x) dx dt \quad (43a)$$

$$\begin{aligned} & \partial_t \rho_{\ell,t} + \nabla \cdot (\rho_{\ell,t} (-\sum_m \nabla W_{\ell m} * \rho_{m,t} + b_\ell + \sigma \xi_{\ell,t})) \\ & - \frac{\epsilon}{2} \sum_{i,k} \frac{\partial^2 (a_{ik} \rho_{\ell,t})}{\partial x_i \partial x_k} = 0, \quad \ell = 1, 2, \dots, L \end{aligned} \quad (43b)$$

$$\rho_{\ell,0} = \mu_\ell, \quad \rho_{\ell,1} = \nu_\ell, \quad \ell = 1, 2, \dots, L, \quad (43c)$$

where $\rho_{\ell,t}$ denotes the distribution of the ℓ -th species. The constraints (43b) is a generalization of (37b) to the multi-species setting describing the distribution evolution of each species. The optimal solution to (43) can be characterized by the PDEs

$$\partial_t \lambda_\ell + \frac{1}{2} \nabla \lambda_\ell^T a \nabla \lambda_\ell - V_\ell + \nabla \lambda_\ell^T b_\ell - \nabla \lambda_\ell^T \sum_m \nabla W_{\ell m} * \rho_{m,t} \quad (44a)$$

$$- \sum_m \int \rho_{m,t}(y) \nabla \lambda_m(y)^T \nabla W_{m\ell}(y-x) dy + \frac{\epsilon}{2} \text{Tr}(a \nabla^2 \lambda_\ell) = 0$$

$$\begin{aligned} & \partial_t \rho_{\ell,t} + \nabla \cdot (\rho_{\ell,t} (-\sum_m \nabla W_{\ell m} * \rho_m + b_\ell + a \nabla \lambda_\ell)) \\ & - \frac{\epsilon}{2} \sum_{i,k} \frac{\partial^2 (a_{ik} \rho_{\ell,t})}{\partial x_i \partial x_k} = 0 \end{aligned} \quad (44b)$$

$$\rho_{\ell,0} = \mu_\ell, \quad \rho_{\ell,1} = \nu_\ell \quad (44c)$$

for all $\ell = 1, 2, \dots, L$. Here $\lambda_1, \lambda_2, \dots, \lambda_L$ are Lagrange multipliers associated with the constraints (43b) for $\ell = 1, 2, \dots, L$. The corresponding optimal control policy for the ℓ -th species is

$$u_{\ell,t} = \sigma^T \nabla \lambda_\ell(t, X_{\ell,t}).$$

To develop an efficient algorithm for (43), we reformulate it as an optimization over the path measures. More specifically, denote by \mathcal{P}^ℓ the distribution on the path space induced by species ℓ , then following a similar argument as before, we obtain the following reformulation

$$\min_{\mathcal{P}^1, \dots, \mathcal{P}^L} \sum_{\ell=1}^L \left\{ \epsilon \text{KL}(\mathcal{P}^\ell \| \mathcal{Q}^\ell(\mathcal{P}^1, \dots, \mathcal{P}^L)) + \int V_\ell d\mathcal{P}^\ell \right\} \quad (45a)$$

$$(X_0)_\# \mathcal{P}^\ell = \mu_\ell, (X_1)_\# \mathcal{P}^\ell = \nu_\ell, \ell = 1, 2, \dots, L. \quad (45b)$$

Here the distribution \mathcal{Q}^ℓ is induced by the diffusion process

$$dX_t = - \sum_m [\nabla W_{\ell m} * \mathcal{P}_t^m](X_t) dt + b_\ell(X_t) dt + \sqrt{\epsilon} \sigma dB_t,$$

which clearly depends on $\mathcal{P}^1, \mathcal{P}^2, \dots, \mathcal{P}^L$.

Discretizing the problem over space and time, the optimization variables become L tensors $\mathbf{M}^\ell, \ell = 1, 2, \dots, L$ and the optimization

problem becomes

$$\min_{\mathbf{M}^1, \dots, \mathbf{M}^L} \sum_{\ell=1}^L \{ \langle \mathbf{C}^\ell(\mathbf{M}^1, \dots, \mathbf{M}^L), \mathbf{M}^\ell \rangle + \epsilon \langle \mathbf{M}^\ell, \log \mathbf{M}^\ell \rangle \} \quad (46a)$$

$$P_0(\mathbf{M}^\ell) = \mu_\ell, \quad P_T(\mathbf{M}^\ell) = \nu_\ell, \quad \ell = 1, 2, \dots, L, \quad (46b)$$

where the cost tensor for the ℓ -th species is

$$\begin{aligned} \mathbf{C}^\ell(\mathbf{M}^1, \dots, \mathbf{M}^L)(x_0, x_1, \dots, x_T) &= \sum_{i=0}^{T-1} \frac{1}{T} V_\ell(x_i) \\ &+ \sum_{i=0}^{T-1} \frac{T}{2} \|x_{i+1} - x_i + \frac{1}{T} \sum_m \nabla W_{\ell m} * P_i(\mathbf{M}^m) - \frac{1}{T} b_\ell(x_i)\|^2. \end{aligned}$$

A more compact form of the above problem can be obtained by combining the L optimization variables $\mathbf{M}^1, \mathbf{M}^2, \dots, \mathbf{M}^L$ into a single variable \mathbf{M} . More precisely, we denote by \mathbf{M} the $T+2$ dimensional tensor where the index for the first dimension is ℓ . Similarly, we combine $\mathbf{C}^1, \mathbf{C}^2, \dots, \mathbf{C}^L$ into \mathbf{C} where

$$\mathbf{C}(\mathbf{M})(\ell, x_0, x_1, \dots, x_T) = \sum_{i=0}^{T-1} \frac{1}{T} V_\ell(x_i) + \quad (47)$$

$$\sum_{i=0}^{T-1} \frac{T}{2} \|x_{i+1} - x_i + \frac{1}{T} \sum_m [\nabla W_{\ell m} * P_{-1,i}(\mathbf{M})](\ell, x_i) - \frac{1}{T} b_\ell(x_i)\|^2.$$

In the above, we adopt an unconventional notation $P_{-1,i}(\mathbf{M})$ to denote the marginal of \mathbf{M} over (ℓ, x_i) . In terms of \mathbf{M}, \mathbf{C} , the above optimization (46) can be rewritten as

$$\min_{\mathbf{M}} \langle \mathbf{C}(\mathbf{M}), \mathbf{M} \rangle + \epsilon \langle \mathbf{M}, \log \mathbf{M} \rangle \quad (48a)$$

$$P_{-1,0}(\mathbf{M}) = \mu, \quad P_{-1,T}(\mathbf{M}) = \nu \quad (48b)$$

with $\mu = [\mu_1, \dots, \mu_L]^T$ and $\nu = [\nu_1, \dots, \nu_L]^T$.

Clearly, (48) is akin to (29). We now utilize the proximal gradient descent to solve it. Denote the set of \mathbf{M} satisfying the constraints (48b) by $\Pi(\mu, \nu)$ and $F(\mathbf{M}) = \langle \mathbf{C}(\mathbf{M}), \mathbf{M} \rangle$, then each iteration of the proximal gradient descent reads

$$\begin{aligned} \mathbf{M}_{k+1} &= \arg\min_{\mathbf{M} \in \Pi(\mu, \nu)} \langle \nabla F(\mathbf{M}_k) - \frac{1}{\eta} \log \mathbf{M}_k, \mathbf{M} \rangle \\ &- (\epsilon + \frac{1}{\eta}) \mathcal{H}(\mathbf{M}). \end{aligned} \quad (49)$$

Proposition 3. The gradient of $F(\mathbf{M}) = \langle \mathbf{C}(\mathbf{M}), \mathbf{M} \rangle$ with \mathbf{C} in (47) is

$$\nabla F(\mathbf{M}) = \mathbf{C}(\mathbf{M}) + E(\mathbf{M}) \quad (50)$$

where $E(\mathbf{M})(\ell, x_0, x_1, \dots, x_T) = \sum_{i=0}^{T-1} E_i(\ell, x_i)$ with

$$E_i(\ell, y) = \sum_{m, x_i, x_{i+1}} \nabla W_{\ell m}(x_i - y)^T [x_{i+1} - x_i + \quad (51)$$

$$\frac{1}{T} \sum_n \nabla W_{\ell n} * P_{-1,i}(\mathbf{M}) - \frac{1}{T} b_\ell(x_i)] P_{-1,i,i+1}(\mathbf{M})(\ell, x_i, x_{i+1}).$$

Proof. See Appendix B. \square

Plugging (50) into (49) yields

$$\begin{aligned} \mathbf{M}_{k+1} &= \arg\min_{\mathbf{M} \in \Pi(\mu, \nu)} \langle \mathbf{C}(\mathbf{M}_k) + E(\mathbf{M}_k) - \frac{1}{\eta} \log \mathbf{M}_k, \mathbf{M} \rangle \\ &- (\epsilon + \frac{1}{\eta}) \mathcal{H}(\mathbf{M}). \end{aligned} \quad (52)$$

Apparently, both $\mathbf{C}(\mathbf{M})$ in (47) and $E(\mathbf{M})$ in (50) have a graphical structure associated with the graph shown in Figure 2. Assume \mathbf{M}_k shares the same graphical structure, then the solution \mathbf{M}_{k+1} to (52) also has this graphical structure. Thus, with proper initialization, the graphical structure (Figure 2) is preserved through the iteration (52). Each iteration (52) is a graphical OT problem and can be solved using a (generalized) SBP algorithm. Thus, the Proximal Sinkhorn Belief Propagation algorithm (Algorithm 2) is applicable to the density control problem with multiple species as long as the SBP subroutine is tailored for the graphical structure in Figure 2.

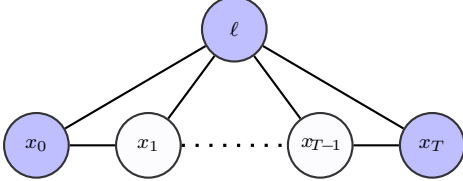


Fig. 2: Graph for graphical OT (52)

Remark 4. In the multi-species setting, the SBP algorithm is applied to the graphical OT in Figure 2. The computation complexity for each outer iteration becomes $\mathcal{O}(D^2LT)$.

Remark 5. Even though in (43) the object cost is decoupled into L separate terms, each corresponds to one species, it is straightforward to generalize the method to include cost such as $\int_0^1 \int V(x) \rho_t dx dt$ that depends on the group behavior of all the individuals.

V. LINEAR QUADRATIC CASES

A special case of particular interest is the linear quadratic density control problem where the dynamics of the individuals are linear and the costs are quadratic. That is, in the linear quadratic setting, $b_\ell, W_{m\ell}, V_\ell$ are of the form

$$b_\ell(x) = A_\ell x, \quad (53a)$$

$$W_{m\ell}(x) = \frac{1}{2} x^T \bar{A}_{m\ell} x, \text{ with } \bar{A}_{m\ell} = \bar{A}_{m\ell}^T = \bar{A}_{\ell m}, \quad (53b)$$

and

$$V_\ell(x) = \frac{1}{2} x^T Q_\ell x. \quad (53c)$$

rendering linear dynamics for each individual

$$\begin{aligned} dX_{\ell,t}^i &= -\frac{1}{N} \sum_m \sum_j \bar{A}_{\ell m} (X_{\ell,t}^i - X_{m,t}^j) dt + A_\ell X_{\ell,t}^i dt \\ &\quad + \sigma(u_{\ell,t}^i dt + \sqrt{\epsilon} dB_{\ell,t}^i), \quad i = 1, \dots, N, \end{aligned} \quad (54)$$

and quadratic cost in the mean field limit

$$\sum_{\ell=1}^L \int_0^1 \int_{\mathbb{R}^d} \left[\frac{1}{2} \|\xi_{\ell,t}(x)\|^2 + \frac{1}{2} x^T Q_\ell x \right] \rho_{\ell,t}(x) dx dt. \quad (55)$$

When the feedback strategies $\xi_{\ell,t}$ are linear and the initial distributions are Gaussian, the distribution $\rho_{\ell,t}, \ell = 1, \dots, L$ of the populations remain Gaussian all the time. Thus, we assume the marginal distributions are Gaussian, denoted by

$$\mu_\ell = \mathcal{N}(m_\ell^0, \Sigma_\ell^0), \quad \nu_\ell = \mathcal{N}(m_\ell^1, \Sigma_\ell^1), \quad \ell = 1, \dots, L.$$

When there is no interaction between the individuals, the problem reduces to the covariance control [14]. The coupling of the agents introduces extra complexities. Recently, the linear quadratic density control problem for one species ($L = 1$) has been addressed in [63].

We next present the solution when multiple species are involved. To this end, we parametrize the Gaussian distributions ρ_ℓ by

$$\rho_{\ell,t} = \mathcal{N}(m_\ell(t), \Sigma_\ell(t)). \quad (56)$$

Just as standard linear quadratic optimal control, the Lagrange multipliers $\lambda_1, \lambda_2, \dots, \lambda_L$ in (44) are quadratic, denoted by

$$\lambda_\ell(t, x) = -\frac{1}{2} x^T \Pi_\ell(t) x + n_\ell(t)^T x + c_\ell(t). \quad (57)$$

Plugging (53), (56) and (57) into the optimality condition (44) yields a coupled equation system (for all $\ell = 1, 2, \dots, L$)

$$\begin{aligned} \dot{\Pi}_\ell - \Pi_\ell \sigma \sigma^T \Pi_\ell + Q_\ell + (A_\ell - \sum_m \bar{A}_{\ell m})^T \Pi_\ell \\ + \Pi_\ell (A_\ell - \sum_m \bar{A}_{\ell m}) = 0 \end{aligned} \quad (58a)$$

$$\begin{aligned} \dot{\Sigma}_\ell - (A_\ell - \sum_m \bar{A}_{\ell m} - \sigma \sigma^T \Pi_\ell) \Sigma_\ell - \Sigma_\ell (A_\ell - \sum_m \bar{A}_{\ell m} - \sigma \sigma^T \Pi_\ell)^T \\ - \epsilon \sigma \sigma^T = 0 \end{aligned} \quad (58b)$$

$$\Sigma_\ell(0) = \Sigma_\ell^0, \quad \Sigma_\ell(1) = \Sigma_\ell^1, \quad (58c)$$

$$\begin{aligned} \dot{n}_\ell + (A_\ell - \sum_m \bar{A}_{\ell m} - \sigma \sigma^T \Pi_\ell)^T n_\ell + \sum_m \bar{A}_{\ell m} n_m \\ - \sum_m (\Pi_\ell \bar{A}_{\ell m} + \bar{A}_{\ell m} \Pi_m) m_m = 0 \end{aligned} \quad (58d)$$

$$\dot{m}_\ell - (A_\ell - \sum_m \bar{A}_{\ell m} - \sigma \sigma^T \Pi_\ell) m_\ell - \sum_m \bar{A}_{\ell m} m_m - \sigma \sigma^T n_\ell = 0 \quad (58e)$$

$$m_\ell(0) = m_\ell^0, \quad m_\ell(1) = m_\ell^1. \quad (58f)$$

In the above, (58b) and (58e) are associated with the Fokker-Planck equation (44b). To see this, note that in the mean field limit each individual in the ℓ -th species, under control policy $\sigma^T \nabla \lambda_\ell$, follows the dynamics

$$\begin{aligned} dX_t &= (A_\ell - \sum_m \bar{A}_{\ell m} - \sigma \sigma^T \Pi_\ell) X_t dt + \sum_m \bar{A}_{\ell m} m_m dt \\ &\quad + \sigma \sigma^T n_\ell dt + \sqrt{\epsilon} \sigma dB_t. \end{aligned}$$

For this linear dynamics, the Fokker-Planck equation (44b) reduces to the Lyapunov equation (58b) for the covariance and a differential equation (58e) for the mean dynamics. The PDE (44a) becomes (58a) and (58d). In particular, (58a) is a Riccati equation.

It turns out that (58) has a closed-form solution. First we observe that (58) is that Π_ℓ, Σ_ℓ can be solved from (58a)-(58c) and are independent of the value of n_ℓ, m_ℓ . Moreover, the equations for Π_ℓ, Σ_ℓ are independent to each other for different species ℓ . Thus, each pair Π_ℓ, Σ_ℓ can be computed separately. Note that the boundary conditions in (58c) for the coupled differential equations (58a)-(58c) are not conventional; the boundary values of Σ_ℓ are given on both end while no boundary value for Π_ℓ is provided. Nevertheless, closed-form solutions to (58a)-(58c) can be obtained. Let

$$H_\ell(t) = \epsilon \Sigma_\ell(t)^{-1} - \Pi_\ell(t),$$

then (58a)-(58c) become a coupled Riccati equation system

$$\dot{\Pi}_\ell - \Pi_\ell \sigma \sigma^T \Pi_\ell + Q_\ell + (A_\ell - \sum_m \bar{A}_{\ell m})^T \Pi_\ell + \Pi_\ell (A_\ell - \sum_m \bar{A}_{\ell m}) = 0$$

$$\dot{H}_\ell + H_\ell \sigma \sigma^T H_\ell - Q_\ell + (A_\ell - \sum_m \bar{A}_{\ell m})^T H_\ell + H_\ell (A_\ell - \sum_m \bar{A}_{\ell m}) = 0$$

$$\Pi_\ell(0) + H_\ell(0) = \epsilon (\Sigma_\ell^0)^{-1}, \quad \Pi_\ell(1) + H_\ell(1) = \epsilon (\Sigma_\ell^1)^{-1}.$$

This is exactly the characterization of the covariance control problem [14] for the dynamics

$$dX_t = (A_\ell - \sum_m \bar{A}_{\ell m})X_t dt + \sigma(u_t dt + \sqrt{\epsilon} dB_t). \quad (59)$$

Assume (59) is controllable, then the above Riccati equation system allows a unique solution in closed-form. We refer the reader to [14] for the exact expression for the closed-form solution.

Once Π_ℓ, Σ_ℓ are computed, we can plug them into (58d)-(58f) to solve for n_ℓ, m_ℓ . These are standard linear equations and can be solved efficiently. Once the solution to (58a)-(58f) is obtained, we can recover the optimal control as

$$\xi_{\ell,t}(x) = -\sigma^T \Pi_\ell(t)x + \sigma^T n_\ell(t), \quad (60)$$

which is a linear state feedback.

VI. NUMERICAL EXAMPLES

In this section we provide several numerical examples to illustrate the proposed framework on density control of interacting agent systems. In the first example, we demonstrate the solution to the density control problem in the linear quadratic setting for both single species and multiple species can steer the agents to target distributions. In the second example, we illustrate how a general interacting particle system evolve from an initial configuration to a target configuration.

A. Linear quadratic density control

We first consider N agents of the same species interacting with each other through the dynamics (54) with

$$A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \bar{A}_{11} = \begin{bmatrix} 0 & 0 \\ 0 & 0.5 \end{bmatrix},$$

and $\sigma = [0 \ 1]^T$. Each agent alone has a trivial second order dynamics. The choice of \bar{A}_{11} corresponds to an interaction potential that synchronizes the velocities of the agents. The noise intensity ϵ is set to be 1 and $Q_1 = I$. Our goal is to find a global feedback policy so that the distribution of the agents is transformed from

$$\mu_1 = \mathcal{N}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0.25 & 0 \\ 0 & 0.25 \end{bmatrix}\right)$$

to

$$\nu_1 = \mathcal{N}\left(\begin{bmatrix} 1.5 \\ 0.8 \end{bmatrix}, \begin{bmatrix} 0.5 & 0 \\ 0 & 0.1 \end{bmatrix}\right).$$

Figure 3 depicts the 3- σ confidence level of the Gaussian distribution of the agents. The states of the agents should be inside this envelope with probability 99.73%. We also plot several typical trajectories of the agents, which stay inside the envelop as expected.

We next add another species to it. Let

$$A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \bar{A}_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 0.5 \end{bmatrix},$$

and the interaction matrix between the two species be

$$\bar{A}_{12} = \bar{A}_{21} = \begin{bmatrix} -0.5 & 0 \\ 0 & 0 \end{bmatrix}.$$

The choice of \bar{A}_{12} imposes a repulsive potential between the agents of the two species in position. Set $Q_2 = I$. The two marginal distributions of the second species is set to be

$$\mu_2 = \mathcal{N}\left(\begin{bmatrix} -2 \\ -2 \end{bmatrix}, \begin{bmatrix} 0.25 & 0 \\ 0 & 0.25 \end{bmatrix}\right)$$

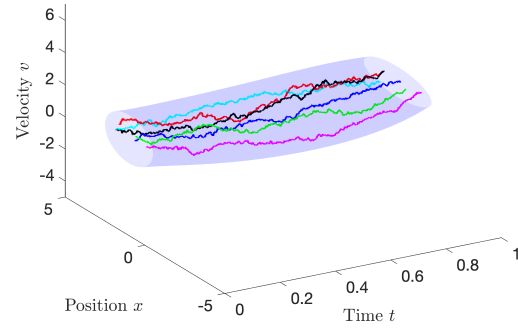


Fig. 3: Covariance evolution of interacting agents under optimal control

and

$$\nu_2 = \mathcal{N}\left(\begin{bmatrix} -1 \\ -0.8 \end{bmatrix}, \begin{bmatrix} 0.25 & 0 \\ 0 & 0.1 \end{bmatrix}\right).$$

As before, we show the 3- σ confidence envelope of the distributions of both species in Figure 4, together with some typical trajectories of the agents. It is clear from Figure 4 that the behavior of the first species is affected by the second one with the tendency to stay away from the second species.

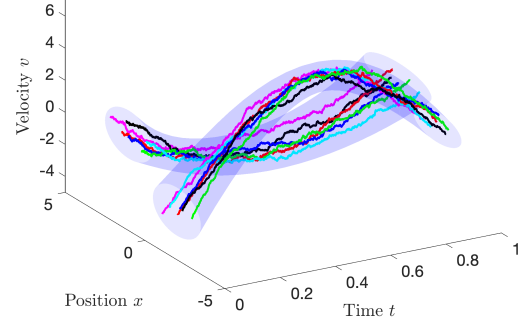


Fig. 4: Covariance evolution of interacting agents from two species under optimal control

B. Density control for general dynamics

In this example we consider N agents in one dimensional space whose dynamics are described by (36) with

$$W(x) = \frac{\beta}{|x|^\alpha}, \quad b(x) = 0, \quad \sigma = 1, \quad \epsilon = 0.1.$$

That is, each agent alone is a first order integrator but they are affected by each other through the repulsive potential W . Our goal is to steer the distribution of the agents from a bimodal initial distribution to a unimodal target distribution. For simplicity, we set $V(\cdot) \equiv 0$. Thus the objective function is the total control effort.

The evolution of the agent distribution under optimal control policy is depicted in Figure 5-8 for different values of α and β . When there is no interaction ($\beta = 0$) among the agents, the solution (Figure 5) corresponds to a standard Schrödinger bridge problem as

expected. As we increase the repulsive potential (increase α or β), the agents spread out more quickly in between due to the repulsive force. For all choices of parameters, the agents are able to transit from the bimodal initial distribution to the unimodal target distribution.

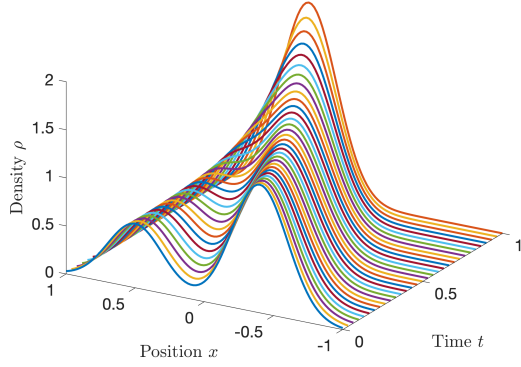


Fig. 5: Density evolution under optimal control with $\beta = 0$

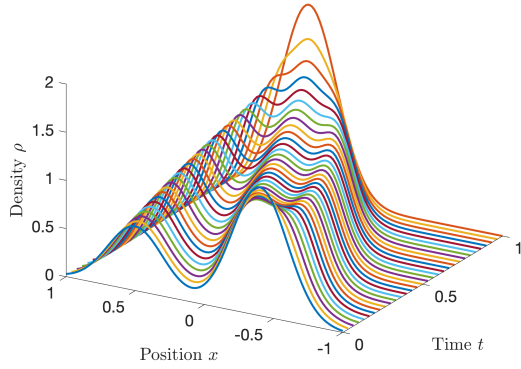


Fig. 6: Density evolution under optimal control with $\alpha = 0.15, \beta = 2$

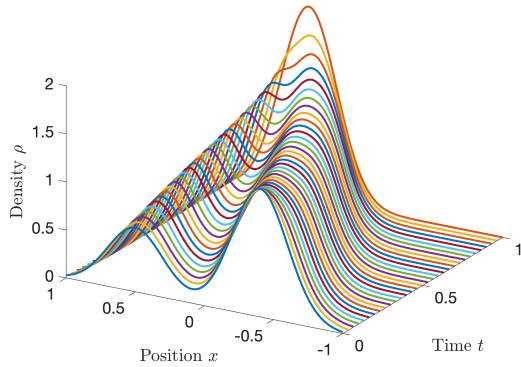


Fig. 7: Density evolution under optimal control with $\alpha = 0.2, \beta = 1$

VII. CONCLUSION

We studied a swarm control problem for a large group of agents that are constantly interacting with each other. We considered the problem in the mean field limit and formulate it as a density control

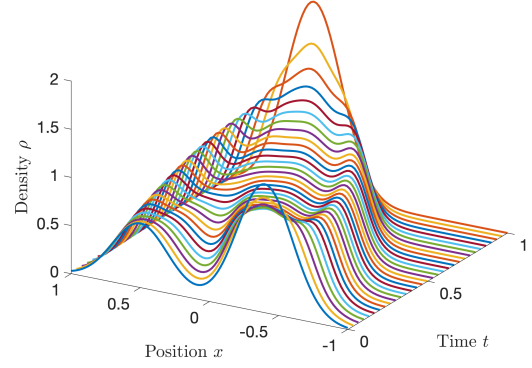


Fig. 8: Density evolution under optimal control with $\alpha = 0.2, \beta = 2$

problem. We further reformulated it as a nonlinear multi-marginal optimal transport problem with proper discretization. Leveraging the proximal gradient framework, we were able to compute the solution via iteratively linearizing the problem and solving the linearized problems with the Sinkhorn belief propagation algorithm. We extended our framework to account for swarm control with multiple species. Finally, we provided a closed-form solution to the density control problem in the linear quadratic setting. One limitation of the proposed algorithm is that it requires discretizing the state space, making it impractical for high-dimensional problems. In the future, we plan to extend our algorithm to particle-based discretization so that it has better scalability in state dimensions. Another limitation of the present work is that it assumes no boundary condition in the state space. How to incorporate explicit boundary conditions to density control is also an interesting problem to investigate in the future.

APPENDIX

A. Proof of Proposition 1

By Taylor expansion,

$$F(\mathbf{M} + \delta\mathbf{M}) - F(\mathbf{M}) = \langle \nabla F(\mathbf{M}), \delta\mathbf{M} \rangle + H.O.T.,$$

where $H.O.T.$ includes all the higher order terms. It follows that

$$\begin{aligned} \langle \nabla F(\mathbf{M}), \delta\mathbf{M} \rangle &= \langle \mathbf{C}(\mathbf{M}), \delta\mathbf{M} \rangle \\ &+ \sum_{i=0}^{T-1} T \langle \frac{1}{T} (\nabla W * P_i(\delta\mathbf{M}))^T (x_{i+1} - x_i + \frac{1}{T} \nabla W * P_i(\mathbf{M})), \mathbf{M} \rangle. \end{aligned}$$

The second term on the right hand side equals

$$\begin{aligned} &\sum_{i=0}^{T-1} \langle (\nabla W * P_i(\delta\mathbf{M}))^T (x_{i+1} - x_i + \frac{1}{T} \nabla W * P_i(\mathbf{M})), P_{i,i+1}(\mathbf{M}) \rangle \\ &= \sum_{i=0}^{T-1} \langle \delta\mathbf{M}, E_i(x_i) \rangle \end{aligned}$$

where E_i is as in (34). Hence,

$$\langle \nabla F(\mathbf{M}), \delta\mathbf{M} \rangle = \langle \mathbf{C}(\mathbf{M}), \delta\mathbf{M} \rangle + \langle E(\mathbf{M}), \delta\mathbf{M} \rangle,$$

with $E(\mathbf{M})(x_0, x_1, \dots, x_T) = \sum_{i=0}^{T-1} E_i(x_i)$, and therefore

$$\nabla F(\mathbf{M}) = \mathbf{C}(\mathbf{M}) + E(\mathbf{M}).$$

B. Proof of Proposition 3

By definition,

$$\begin{aligned} \langle \nabla F(\mathbf{M}), \delta \mathbf{M} \rangle &= \langle \mathbf{C}(\mathbf{M}), \delta \mathbf{M} \rangle + \sum_{i=0}^{T-1} T \left\langle \frac{1}{T} \left(\sum_m \nabla W_{\ell m} * P_{-1,i}(\delta \mathbf{M}) \right)^T \right. \\ &\quad \left. (x_{i+1} - x_i + \frac{1}{T} \sum_n \nabla W_{\ell n} * P_{-1,i}(\mathbf{M})) - \frac{1}{T} b_{\ell}(x_i), \mathbf{M} \right\rangle. \end{aligned}$$

The second term on the right hand side equals

$$\begin{aligned} &\sum_{i=0}^{T-1} \left\langle \left(\sum_m \nabla W_{\ell m} * P_{-1,i}(\delta \mathbf{M}) \right)^T (x_{i+1} - x_i \right. \\ &\quad \left. + \frac{1}{T} \sum_n \nabla W_{\ell n} * P_{-1,i}(\mathbf{M})) - \frac{1}{T} b_{\ell}(x_i), P_{-1,i+1}(\mathbf{M}) \right\rangle \\ &= \sum_{i=0}^{T-1} \langle \delta \mathbf{M}, E_i(\ell, x_i) \rangle \end{aligned}$$

where E_i is as in (51). Hence,

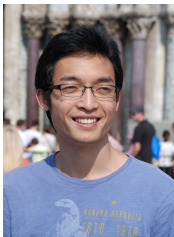
$$\langle \nabla F(\mathbf{M}), \delta \mathbf{M} \rangle = \langle \mathbf{C}(\mathbf{M}), \delta \mathbf{M} \rangle + \langle E(\mathbf{M}), \delta \mathbf{M} \rangle,$$

with $E(\mathbf{M})(\ell, x_0, x_1, \dots, x_T) = \sum_{i=0}^{T-1} E_i(\ell, x_i)$.

REFERENCES

- [1] M. Brambilla, E. Ferrante, M. Birattari, and M. Dorigo, "Swarm robotics: a review from the swarm engineering perspective," *Swarm Intelligence*, vol. 7, no. 1, pp. 1–41, 2013.
- [2] S.-J. Chung, A. A. Paranjape, P. Dames, S. Shen, and V. Kumar, "A survey on aerial swarm robotics," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 837–855, 2018.
- [3] H. P. McKean Jr, "A class of markov processes associated with nonlinear parabolic equations," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 56, no. 6, p. 1907, 1966.
- [4] M. Huang, R. P. Malhamé, P. E. Caines, et al., "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.
- [5] G. Zhu, K. M. Grigoriadis, and R. E. Skelton, "Covariance control design for hubble space telescope," *Journal of Guidance, Control, and Dynamics*, vol. 18, no. 2, pp. 230–236, 1995.
- [6] J. Ridderhof and P. Tsiotras, "Uncertainty quantification and control during mars powered descent and landing using covariance steering," in *2018 AIAA Guidance, Navigation, and Control Conference*, 2018, p. 0611.
- [7] K. Okamoto and P. Tsiotras, "Optimal stochastic vehicle path planning using covariance steering," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2276–2281, 2019.
- [8] M. Kalandros, "Covariance control for multisensor systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 4, pp. 1138–1157, 2002.
- [9] C. Sinigaglia, A. Manzoni, and F. Bagnin, "Density control of large-scale particles swarm through pde-constrained optimization," *arXiv preprint arXiv:2104.06373*, 2021.
- [10] C. Villani, *Topics in Optimal Transportation*. American Mathematical Soc., 2003, no. 58.
- [11] C. Léonard, "A survey of the Schrödinger problem and some of its connections with optimal transport," *Discrete Contin. Dyn. Syst. A*, vol. 34, no. 4, pp. 1533–1574, 2014.
- [12] Y. Chen, T. T. Georgiou, and M. Pavon, "Stochastic control liaisons: Richard Sinkhorn meets Gaspard Monge on a Schrödinger bridge," *SIAM Review*, vol. 63, no. 2, pp. 249–313, 2021.
- [13] A. Hotz and R. E. Skelton, "Covariance control theory," *International Journal of Control*, vol. 46, no. 1, pp. 13–32, 1987.
- [14] 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.
- [15] Y. Chen, T. T. Georgiou, and M. Pavon, "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.
- [16] A. Halder and E. D. B. Wendel, "Finite horizon linear quadratic Gaussian density regulator with Wasserstein terminal cost," in *Proc. American Control Conf.*, 2016.
- [17] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal steering of a linear stochastic system to a final probability distribution, Part III," *IEEE Transactions on Automatic Control*, vol. 63, no. 9, pp. 3112–3118, 2018.
- [18] E. Bakolas, "Finite-horizon covariance control for discrete-time stochastic linear systems subject to input constraints," *Automatica*, vol. 91, pp. 61–68, 2018.
- [19] V. Krishnan and S. Martínez, "Distributed optimal transport for the deployment of swarms," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 4583–4588.
- [20] K. Elamvazhuthi and S. Berman, "Mean-field models in swarm robotics: A survey," *Bioinspiration & Biomimetics*, vol. 15, no. 1, p. 015001, 2019.
- [21] D. Inoue, Y. Ito, and H. Yoshida, "Optimal transport-based coverage control for swarm robot systems: Generalization of the Voronoi tessellation-based method," *IEEE Control Systems Letters*, vol. 5, no. 4, pp. 1483–1488, 2020.
- [22] S. Biswal, K. Elamvazhuthi, and S. Berman, "Stabilization of nonlinear discrete-time systems to target measures using stochastic feedback laws," *IEEE Transactions on Automatic Control*, vol. 66, no. 5, pp. 1957–1972, 2020.
- [23] T. Zheng, Q. Han, and H. Lin, "Deployment of robotic swarms via density feedback control," *arXiv e-prints*, pp. arXiv–2006, 2020.
- [24] J. Backhoff, G. Conforti, I. Gentil, and C. Léonard, "The mean field Schrödinger problem: ergodic behavior, entropy estimates and functional inequalities," *Probability Theory and Related Fields*, pp. 1–56, 2020.
- [25] B. Pass, "Multi-marginal optimal transport: theory and applications," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 49, no. 6, pp. 1771–1790, 2015.
- [26] I. Haasler, A. Ringh, Y. Chen, and J. Karlsson, "Multimarginal optimal transport with a tree-structured cost and the schrödinger bridge problem," *SIAM Journal on Control and Optimization*, vol. 59, no. 4, pp. 2428–2453, 2021.
- [27] I. Haasler, R. Singh, Q. Zhang, J. Karlsson, and Y. Chen, "Multi-marginal optimal transport and probabilistic graphical models," *arXiv preprint arXiv:2006.14113*, 2020.
- [28] A. Beck and M. Teboulle, "Mirror descent and nonlinear projected subgradient methods for convex optimization," *Operations Research Letters*, vol. 31, no. 3, pp. 167–175, 2003.
- [29] A. Beck, *First-order methods in optimization*. SIAM, 2017.
- [30] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.
- [31] A. Bensoussan, J. Frehse, and P. Yam, *Mean field games and mean field type control theory*. Springer, 2013, vol. 101.
- [32] M. Burger, R. Pinnau, C. Totzeck, and O. Tse, "Mean-field optimal control and optimality conditions in the space of probability measures," *SIAM Journal on Control and Optimization*, vol. 59, no. 2, pp. 977–1006, 2021.
- [33] Y. Chen, T. Georgiou, and M. Pavon, "Optimal transport over a linear dynamical system," *arXiv preprint arXiv:1502.01265*, *IEEE Trans. on Automatic Control*, to appear, 2017.
- [34] 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.
- [35] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows: in metric spaces and in the space of probability measures*. Springer, 2006.
- [36] E. Schrödinger, "Über die Umkehrung der Naturgesetze," *Sitzungsberichte der Preuss Akad. Wissen. Phys. Math. Klasse, Sonderausgabe*, vol. IX, pp. 144–153, 1931.
- [37] E. Schrödinger, "Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique," in *Annales de l'institut Henri Poincaré*, vol. 2, no. 4. Presses universitaires de France, 1932, pp. 269–310.
- [38] H. Föllmer, "Random fields and diffusion processes," in *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*. Springer, 1988, pp. 101–203.
- [39] R. Fortet, "Résolution d'un système d'équations de M. Schrödinger," *J. Math. Pures Appl.*, vol. 83, no. 9, 1940.
- [40] A. Beurling, "An automorphism of product measures," *The Annals of Mathematics*, vol. 72, no. 1, pp. 189–200, 1960.
- [41] B. Jamison, "Reciprocal processes," *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, vol. 30, pp. 65–86, 1974.
- [42] 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.
- [43] P. Dai Pra, "A stochastic control approach to reciprocal diffusion processes," *Applied mathematics and Optimization*, vol. 23, no. 1, pp. 313–329, 1991.
- [44] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*.

- Springer, 1988.
- [45] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal steering of inertial particles diffusing anisotropically with losses," in *Proc. American Control Conf.*, 2015, pp. 1252–1257.
 - [46] I. Gentil, C. Léonard, and L. Ripani, "About the analogy between optimal transport and minimal entropy," *arXiv:1510.08230*, 2015.
 - [47] T. Mikami, "Monge's problem with a quadratic cost by the zero-noise limit of h-path processes," *Probability theory and related fields*, vol. 129, no. 2, pp. 245–260, 2004.
 - [48] T. Mikami and M. Thieullen, "Optimal transportation problem by stochastic optimal control," *SIAM Journal on Control and Optimization*, vol. 47, no. 3, pp. 1127–1139, 2008.
 - [49] C. Léonard, "From the Schrödinger problem to the Monge–Kantorovich problem," *Journal of Functional Analysis*, vol. 262, no. 4, pp. 1879–1920, 2012.
 - [50] R. Sinkhorn, "A relationship between arbitrary positive matrices and doubly stochastic matrices," *The annals of mathematical statistics*, vol. 35, no. 2, pp. 876–879, 1964.
 - [51] M. Cuturi, "Sinkhorn distances: Lightspeed computation of optimal transport," in *Advances in Neural Information Processing Systems*, 2013, pp. 2292–2300.
 - [52] J. Franklin and J. Lorenz, "On the scaling of multidimensional matrices," *Linear Algebra and its applications*, vol. 114, pp. 717–735, 1989.
 - [53] H. H. Bauschke and A. S. Lewis, "Dykstras algorithm with Bregman projections: A convergence proof," *Optimization*, vol. 48, no. 4, pp. 409–427, 2000.
 - [54] Z.-Q. Luo and P. Tseng, "On the convergence rate of dual ascent methods for linearly constrained convex minimization," *Mathematics of Operations Research*, vol. 18, no. 4, pp. 846–867, 1993.
 - [55] Z.-Q. Luo and P. Tseng, "On the convergence of the coordinate descent method for convex differentiable minimization," *Journal of Optimization Theory and Applications*, vol. 72, no. 1, pp. 7–35, 1992.
 - [56] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*. MIT press, 2009.
 - [57] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Understanding belief propagation and its generalizations," *Exploring artificial intelligence in the new millennium*, vol. 8, pp. 236–239, 2003.
 - [58] Q. Li, Y. Zhou, Y. Liang, and P. K. Varshney, "Convergence analysis of proximal gradient with momentum for nonconvex optimization," in *International Conference on Machine Learning*. PMLR, 2017, pp. 2111–2119.
 - [59] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal transport in systems and control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, pp. 89–113, 2021.
 - [60] E. B. Lee and L. Markus, *Foundations of optimal control theory*. Wiley, 1967.
 - [61] W. Fleming and R. Rishel, *Deterministic and Stochastic Optimal Control*. Springer, 1975.
 - [62] S. Méléard and S. Roelly-Coppoletta, "A propagation of chaos result for a system of particles with moderate interaction," *Stochastic processes and their applications*, vol. 26, pp. 317–332, 1987.
 - [63] Y. Chen, T. T. Georgiou, and M. Pavon, "Steering the distribution of agents in mean-field games system," *Journal of Optimization Theory and Applications*, vol. 179, no. 1, pp. 332–357, 2018.



Yongxin Chen (S'13–M'17–SM'22) received his BSc from Shanghai Jiao Tong University in 2011 and Ph.D. from University of Minnesota in 2016, both in Mechanical Engineering. He is currently an Assistant Professor in the School of Aerospace Engineering at Georgia Institute of Technology. He has served on the faculty at Iowa State University (2017–2018). He received the George S. Axelby Best Paper Award in 2017, the NSF Faculty Early Career Development Program (CAREER) Award in 2020, the A.V. 'Bal' Balakrishnan Award for Excellence in

Scientific Research in the Mathematics of Systems in 2021, and the Donald P. Eckman Award for Outstanding Young Engineer in the Field of Automatic Control in 2022. His current research focuses on the intersection of control theory, machine learning, robotics and optimization.