Optimal Mass Transport and Kernel Density Estimation for State-Dependent Networked Dynamic Systems

Mathias Hudoba de Badyn, Utku Eren, Behçet Açıkmeşe and Mehran Mesbahi

Abstract— **State-dependent networked dynamic systems are ones where the interconnections between agents change as a function of the states of the agents. Such systems are highly nonlinear, and a cohesive strategy for their control is lacking in the literature. In this paper, we present two techniques pertaining to the density control of such systems. Agent states are initially distributed according to some density, and a feedback law is designed to move the agents to a target density profile. We use optimal mass transport to design a feedforward control law propelling the agents towards this target density. Kernel density estimation, with constraints imposed by the state-dependent dynamics, is then used to allow each agent to estimate the local density of the agents.**

I. INTRODUCTION

Networked dynamic systems arise in many synthetic and natural systems in science and engineering [1]; in particular multi-agent systems offer an interesting control paradigm. Each agent augments the system with an additional (local) computational resource, motivating the concept of *distributed controllers & estimators* [2], [3], [4], leading to a notion of *local* control versus *global* control. The latter seeks to design control laws to guide groups of agents to a desired objective, and the former seeks to design on-board controllers for each agent to facilitate their role in the global control law.

A class of the more widely used local control laws is called *consensus*, where each agent averages data from their neighbours to compute a parameter related to their objective – for example, heading, position or a formation center [1], [5], [6]. The attractive feature of consensus is how the interconnections between agents – the so-called *network* or *graph topology* – affect the performance and agreement characteristics of the algorithm [7], [8]. Graph-theoretic characteristics, such as symmetry, structural balance and graph spectra provide additional insights into the control-theoretic behaviour of consensus [9], [10], [11], [12], [13].

Global controllers for multi-agent systems take many forms [14], eg., potential field approaches [15], smoothed particle hydrodynamics [16] and density control [17]. In [18], [19], density control with only relative measurements of position between agents is considered, and mean-field control is used to tackle multi-agent interactions by considering a mass flow in the large-N limit [20], [21], [22].

The main interest area of the current paper is networked dynamic systems in which the underlying network is time-varying. Examples of such systems are switching and proximity-based consensus [1], [6], [23] and the Vicsek flocking model [24]. State-dependency refers to networks in which the underlying graph varies based on the state of the nodes. State-dependent networks have been considered in [7], [20], [22], [25], [26], [27], [28], but few underlying principles for designing controllers to account for this difficult nonlinearity has been proposed. A wide range of real-life networked systems are state-dependent.

To tackle this problem, we propose a twofold extension of the work in [17]. First, we consider state-dependent networked dynamic systems, instead of single-integrator dynamics. Second, we propose a control method for these systems by using a feedback-based density control law that utilizes *optimal mass transport* (OMT). OMT was considered for linear systems in [29], non-linear systems in [30] and for density tracking of non-interacting agents in [31]. Our contribution considers OMT for multi-agent systems, in particular ones with state-dependent dynamics.

In the OMT problem, the initial and final densities ρ_0 & ρ_1 of the agents are specified. The solution to the problem yields a time-dependent density profile with boundary conditions imposed by ρ_0, ρ_1, and a velocity field that together satisfy the continuity equation.

We aim to use this velocity field as a feedforward control input to the state-dependent multi-agent system, coupled with a density control feedback law. Using a modified form of kernel density estimation (KDE) that takes into account the state-dependent dynamics, we will show that the combination of the two control techniques will allow us to propose a physically realizable control strategy for state-dependent networked dynamic systems.

This paper is organized as follows. The mathematical preliminaries, including notation, optimal mass transport, and density control with the KDE procedure are reviewed in §II. The problem statement and paper contributions are outlined in §III. We present the feedforward controller based on OMT in §IV, and the state-dependent KDE in §V. Examples are provided in §VI, and the paper is concluded in §VII.

II. MATHEMATICAL PRELIMINARIES AND BACKGROUND

A. Mathematical Notation

We follow the standard graph theory notation listed in [1]. A measure space $(\Sigma, \mathcal{A}, \mu)$ is a triple containing a *sample space* $\Sigma \subset \mathbb{R}^n$, a σ-algebra of subsets of Σ, and a *measure* μ that assigns the 'size' $\mu(A) \in \mathbb{R}_+$ to a set $A \in \mathcal{A}$. The *Borel σ-algebra* $\mathcal{B}$ is generated from the countable unions, intersections and complements of open subsets of $\mathbb{R}^n$. The *Lebesgue measure* λ assigns to a closed interval $[a, b] \subset \mathbb{R}$ the 'size' $b - a$; this can be extended to $\mathbb{R}^n$ by considering products of measures. A measure μ is called *absolutely continuous* with respect to a measure ν if $\nu(A) = 0 \implies \mu(A) = 0$ for $A \in \mathcal{A}$; μ is called *Lebesgue absolutely continuous* if $\nu = \lambda$. This is denoted $\mu \ll \nu$. If $\mu \ll \nu$,

there exists a function $f : \mathbb{R} \to \mathbb{R}_+$ called a *Radon-Nikodym derivative*, or *density*, of μ with respect to ν. It is denoted $f := \frac{d\mu}{d\nu}$, and satisfies the property that for all $A \in \mathcal{A}$, $\mu(A) = \int_A f d\nu$. Let $\pi(x, y)$ be a joint measure on $\mathcal{X} \times \mathcal{Y}$. We denote the set of such joint measures π by $P(\mathcal{X}, \mathcal{Y})$. The *marginal* π_x of π on $\mathcal{X}$ is defined as the push-forward under the projection map X on $\mathcal{X}$: $\pi_x = X_\# \pi$, where $X(x, y) = x$. Similarly, the marginal π_y of π on $\mathcal{Y}$ is given by $\pi_y = Y_\# \pi$, where $Y(x, y) = y$. We denote the convolution of two functions f, g as $f \star g$, or the convolution of a function f and measure μ as $f \star \mu$.

B. Optimal Mass Transport

Informally speaking, the *optimal mass transport* problem is to find a mapping between two densities that minimizes some cost. Formally speaking, we consider two measures[1] μ_0, μ_1 on $\mathbb{R}^n$ with equal mass: $\int_{\mathbb{R}^n} d\mu_0 = \int_{\mathbb{R}^n} d\mu_1$. The *optimal mass transport* problem [32], [33] is to find a measurable map $T : \mathbb{R}^n \to \mathbb{R}^n$ taking μ_0 to μ_1 via the following optimization problem:

$$
\begin{aligned}
&\text{minimize} \quad \int_{\mathbb{R}^n} c\,(x, T(x))\, d\mu_0(x) \\
&\text{subject to } \int_{x \in A} d\mu_1(x) = \int_{T(x) \in A} d\mu_0(x), \forall A \subset \mathbb{R}^n,
\end{aligned} \quad (1)
$$

where c is a cost function depending on the initial and transported masses. The constraint in Problem (1) means that μ_1 is the *push-forward* measure of μ_0 under the map T, in that for each Borel set $B \in \mathcal{B} := \sigma(\mathbb{R}^n)$, we have that $\mu_1(B) = \mu_0\left(T^{-1}(B)\right)$. This is denoted as $T_\# \mu_0 = \mu_1$.

A generalization of Problem (1) by Kantorovich is able to pick out the optimal map T, if it exists, for a certain class of costs c under the assumption of absolute continuity of the measures [34]. Here, we consider a joint distribution $\pi(x, y)$ on $\mathbb{R}^n \times \mathbb{R}^n$ and solve for the optimal admissible measure π given some cost $c(x, y)$.

The set of admissible measures π are those whose marginals are μ_0 and μ_1: $X_\# \pi = \mu_0$, $Y_\# \pi = \mu_1$. This is equivalent to requiring

$$
\pi(A \times \mathbb{R}^n) = \mu_0(A), \ \pi(\mathbb{R}^n \times B) = \mu_1(B) \quad (2)
$$

for all measurable $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^n$. The *Kantorovich relaxed optimal mass transport problem* [34] is given by

$$
\begin{aligned}
&\text{minimize} \quad \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x, y) d\pi(x, y) \\
&\text{subject to} \quad \pi \in \{\pi \in P(\mathbb{R}^n, \mathbb{R}^n) \text{ s.t. (2) holds}\}.
\end{aligned} \quad (3)
$$

Proposition 1 ([32], [33]): For quadratic costs $c(x, y) = \|x - y\|^2$, the support of the optimal joint measure $\pi^*(x, y)$ of Problem (3) is exactly the graph of the optimal map $T^*(x)$ minimizing Problem (1).

For quadratic costs, Benamou and Brenier formulated an equivalent problem in terms of a constrained fluid mechanics model.

Proposition 2 ([35]): Given Lebesgue absolutely continuous μ_0, μ_1 with Radon-Nikodym densities ρ_0, ρ_1 respectively, Problem (3) with quadratic costs is equivalent to

$$
\begin{aligned}
&\inf_{\rho, v} \quad \int_{\mathbb{R}^n} \int_0^1 \tfrac{1}{2} \|v(t, x)\|^2 \rho(t, x) dt dx \\
&\text{subject to } \tfrac{\partial \rho}{\partial t} + \nabla \cdot (v\rho) = 0 \\
&\qquad \rho(0, x) = \rho_0(x), \ \rho(1, y) = \rho_1(y).
\end{aligned} \quad (4)
$$

---

[1] If the measures are Lebesgue absolutely continuous, one can equivalently consider densities ρ_0, ρ_1.

Furthermore, the solution to Problem (4) is of the form $v(t, x) = \nabla \varphi(t, x)$, where $\varphi(t, x)$ is the Lagrange multiplier of the constraints and the solution to the Hamilton-Jacobi equation $\partial_t \phi + \frac{1}{2} |\nabla \phi|^2 = 0$.

Remark 1: The optimal map T^* of Problem (3) in the case of quadratic costs can be reconstructed from the variable $v(t, x)$ from the solution of Problem (4). This formally establishes the equivalence stated in Proposition 2 [35].

C. Density Control and Kernel Density Estimation

In [17], a feedback control law to drive a group of single-integrator agents to a desired density profile $\rho_1(x, t)$ was analyzed. The following feedback law is proposed to compute the velocity field as a function of the error in density $\Phi(x, t) := \rho(x, t) - \rho_1(x)$ as

$$
v(x, t) = -\alpha \frac{\nabla \Phi(x, t)}{\rho(x, t)}. \quad (5)
$$

Density control of multi-agent systems is impacted by the ability of individual agents to discern the local density profile from measurements of their neighbours. Since the number of agents is finite, the local density profile must be approximated from finitely many samples $r_i(t)$.

This can be accomplished using *kernel density estimation* [36], [37]. The kernel density estimate $\hat{\rho}(t, x)$ at any point $x \in \mathbb{R}^n$ and time $t \in \mathbb{R}_+$ is given by

$$
\hat{\rho}(t, x) = \int_{\mathbb{R}^n} \left[\prod_{k=1}^d \frac{1}{h_k} K\left(\frac{x^{[k]} - \xi^{[k]}}{h_k} \right) \right] dP_N(t, \xi). \quad (6)
$$

Here, $K : \mathbb{R} \to \mathbb{R}$ is called the *smoothing kernel*, h_k is called the *smoothing parameter*, and $dP_N(t, \xi)$ is a sum of Dirac measures at sample points:

$$
dP_N(t, \xi) = \frac{1}{N} \sum_{r(t) \in S(t)} \delta\,(\xi - r(t))\, d\xi.
$$

Since $\delta(\cdot)$ is the Dirac delta functional, Equation (6) can be written as

$$
\hat{\rho}(t, x) = \frac{1}{N} \sum_{i=1}^N \left[\prod_{k=1}^d \frac{1}{h_k} K\left(\frac{x^{[k]} - r_i^{[k]}}{h_k} \right) \right].
$$

The control law (5) then uses the estimate $\hat{\rho}(x, t)$ in place of knowledge of the true density $\rho(x, t)$, where the sample points $r(t)$ are taken to be the agent states:

$$
v(x, t) = -\alpha \frac{\nabla(\hat{\rho}(x, t) - \rho_1(x))}{\hat{\rho}(x, t)}.
$$

In [17], Gaussian kernels were used – this induces an all-to-all communication; every agent is able to sample the position every other agent. The control law (5) has a convergence guarantee listed in Theorem 6 of [17].

III. PROBLEM STATEMENT AND CONTRIBUTIONS

In this paper, we consider state-dependent networked dynamic systems with N agents on a bounded region $\mathcal{R} \subset \mathbb{R}^n$, where the ith agent's state evolves according to the dynamics

$$
\dot{x}_i = f(\mathcal{G}(\mathbf{x}), \mathbf{x}) + B u_i(\mathbf{x}, t), \ \mathbf{x} := (x_1, \ldots, x_n).
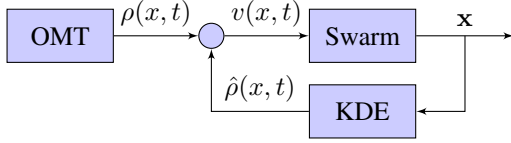$$

Fig. 1: Block diagram of density control scheme

A prototypical example of such a system is *state-dependent consensus*

$$\dot{x}_i = \sum_{j \neq i} A(x_i, x_j) \cdot (x_i - x_j) + u_i, \tag{7}$$

where the edge weight $w_{ij} := A(x_i, x_j)$ changes depending on the state of the agent i and its neighbour j. In general, one can consider an *interaction kernel* $H(x)$ that generates a consensus-like dynamics by convolution with the Dirac measure supported at agent states [20], [22]:

$$\mu_N(x) = \frac{1}{N} \sum_{j=1}^{N} \mathbf{1}_{\{x_j\}}(x) = \frac{1}{N} \sum_{j=1}^{N} \delta(x - x_j). \tag{8}$$

Using Equation (8), we can write a general multi-agent system as

$$\dot{x}_i = (H \star \mu_N)(x_i) + u_i.$$

A simple example motivated by robotics is proximity-based edge switching, where $A(x_i, x_j) = 1$ if $\|x_i - x_j\| \leq r$ and 0 otherwise, where r is some communication radius. This corresponds to an interaction kernel $H(x) = x\mathbf{1}_{\|x\| \leq r}(x)$.

As the number of agents N grows sufficiently large, one can consider the time-dependent *density* of agents $\rho(x, t)$ over a region of the state space. In the context of mean-field control, the formal large-N limit of the dynamics (7) produces the *mean field dynamics* [20], [22],

$$\frac{\partial \rho}{\partial t} + \nabla_x \cdot [(\mathcal{P}(\rho(x,t), t) + u)\rho] = 0 \tag{9}$$
$$\mathcal{P}(\rho(x,t), x) = \int A(x,y)(y-x)\rho(y,t)dy.$$

We now state the contributions of this paper, namely the feedforward OMT scheme with kernel density estimation shown in Figure 1.

Contribution 1: Feed-Forward Density Control With Optimal Mass Transport.

We consider a generalization of the Brenier-Benamou OMT problem (4) with the continuity equation constraint replaced by the mean-field dynamics (9):

$$\begin{aligned} \inf_{\rho, v} \quad & \int \int_0^1 \frac{1}{2} \|v(t, x)\|^2 \rho(t, x) dt dx \\ \text{subject to} \quad & \partial_t \rho + \nabla_x \cdot [(\mathcal{P}(\rho(x,t), x) + v)\rho] = 0 \\ & \mathcal{P}(\rho(x,t), x) = \int A(x,y)(y-x)\rho(y,t)dy \\ & \rho(0, x) = \rho_0(x), \ \rho(1, y) = \rho_1(y). \end{aligned} \tag{10}$$

The solution to Problem (10) yields two variables with important physical interpretations. The time-varying density $\rho(t, x)$ represents the mass of agents with dynamics (7) with inputs $u_i(x, t) := v(x_i, t)$; the velocity field $v(x, t)$ is precisely the input u_i given to agent i at position x at time t.

Problem (10) assumes that the initial and final masses of agents are distributed in the mean-field limit according to

densities ρ_0 and ρ_1, respectively. When considering finitely many agents, any initial density will take the form of Equation (8) - namely, it will be a Dirac measure supported at the agent states, also called the *empirical density*.

Hence, in general, the boundary conditions on the density in Problem (10) can be either *deterministic* (in the case of Dirac measures supported at the agent states), or *probabalistic* (in the sense that the initial/final agent states $x_i(0)$ and $x_i(1)$ are randomly distributed according to the densities ρ_0 and ρ_1). In the latter case, as $N \to \infty$, the Dirac measure supported at the agent states at time t converges in a formal sense to the density $\rho(\cdot, t)$ [20], [22].

In either case, we consider the velocity field $v(x, t)$ as a feed-forward input to the dynamics (7). Since the number of agents is finite, the empirical density at time t will only approximate the density $\rho(x, t)$ from the solution of Problem (10).

Contribution 2: Feedback Density Control with Kernel Density Estimation and State-Dependent Constraints.

In a state-dependent networked dynamic system, eg., Equation (7), the existence of an edge indicates some notion of information transfer between agents. Hence, a physical estimation scheme and density control law can only allow i to sample those agents j such that $A(x_i, x_j) \neq 0$. The second contribution of this paper is to extend the KDE procedure in [17] by solving a quadratic program for an optimal kernel that takes into account the state-dependent communication constraints.

IV. FEEDBACK CONTROL OF STATE-DEPENDENT NETWORKED DYNAMIC SYSTEMS

Consider the state-dependent consensus dynamics (7). Let $v_1(x, t)$ denote the velocity field from the solution to Problem (10), and let $v_2(x, t)$ denote the velocity field from the control law (5). Our proposed control law is then given by the velocity field (with $\alpha > 0$),

$$u(x, t) = \begin{cases} v_1(x, t) - \alpha \dfrac{\nabla(\rho(x,t) - \rho_1(x))}{\rho(x,t)} & 0 \leq t \leq 1 \\ v_2(x, t) - \mathcal{P}(\mu(x,t), x) & t \geq 1 \end{cases} \tag{11}$$

$$\mathcal{P}(\mu(x,t), x) = \int A(x,y)(y-x)\mu(y,t)dy.$$

Of course, the switch at $t = 1$ is completely arbitrary, and can be altered by changing the time horizon of Problem (10).

The main result of this section is the following theorem, an extension of Theorem 6 in [17]. Informally, it states that as the number of agents in the system tends as $N \to \infty$, the velocities of the agents performing the state-dependent control law (11) will vanish asymptotically.

Theorem 1: Consider a system of N agents $\mathcal{S}(t)$ on a bounded region $\mathcal{R} \subset \mathbb{R}^n$ with individual dynamics given by (7) and with control law (11). Further suppose that the initial swarm density $\rho(x, t)$ and target density $\rho_1(x)$ satisfy the boundary condition $\nabla\Phi(x, t) = 0$ on $\partial\mathcal{R}$. As $t \to \infty$, for sufficiently large N the error density $\Phi(x, t)$ converges to zero: $\lim_{t \to \infty} \Phi(x, t) = 0$, for $x \in \mathcal{R}$ and so $\hat{\rho}(x, t) \to \rho_1(x)$. Furthermore, the velocities of all agents vanish asymptotically: $\lim_{t \to \infty} \dot{x}(t) = 0$, for $x \in \mathcal{R}$.
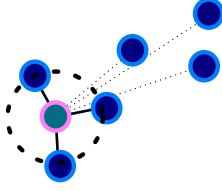The proof is discussed in the Appendix.

Fig. 2: Illustration of (proximity-based) state-dependent constraints on the KDE procedure. Dotted lines indicate samples the center agent cannot measure.

V. DENSITY ESTIMATION FOR KERNELS WITH COMPACT SUPPORT

Consider a state-dependent consensus dynamics as in Equation (7) where $A(x_i, x_j)$ is a state-dependent edge weight. In order to implement the density control law, each agent must be able to estimate the density of nearby agents to generate the correct velocity field. In this section, we discuss optimal kernels designed to achieve this task that are subject to the state-dependent constraints imposed by $A(x_i, x_j)$.

The state-dependent constraints in some (informal) sense denote 'information transfer' between agents i and j. If $A_{ij} = 0$, then agents i and j cannot detect each other, and the KDE procedure should reflect this. To illustrate this notion, consider Figure 2.

In 1D kernel density estimation, there are two parameters selected *a priori* that influence the quality of the estimated probability density function, namely the kernel K and the smoothing parameter h. We consider an optimal selection of K subject to the state-dependent constraints; we leave the task of selecting h for future work. For now, we just need the following assumption on h as a function of the number of samples: $\lim_{N \to \infty} Nh(N) = \infty$. The standard metric for measuring the quality of the estimated probability density function is given by the *mean integrated square error* $E_{\text{MISE}} := \mathbb{E}\left[\int \left(\hat{\rho}(x) - \rho(x)\right) dx\right].$

By extracting out the dependence on the number of samples N, and the choice of smoothing parameter h, one can obtain the *asymptotic mean integrated square error* (AMISE) [36]: $E_{\text{MISE}} := E_{\text{AMISE}} + o\left((hn)^{-1} + h^4\right)$. One can factor the AMISE into a product of two terms, one depending on h and one depending on K: $E_{\text{AMISE}} = C_1(K)C_2(h)$, where

$$C_1(K) := \left[\left(\int K(x)^2 dx\right)^4 \left(\int x^2 K(x) dx\right)^2\right]^{1/5}.$$

Hence, by fixing $a := \int x^2 K(x) dx$ depending on the length of the boundary of our estimation horizon, the only parameter left to optimize is the *roughness* of $K(x)$: $\int K(x)^2 dx$. We can write an optimization problem as follows:

$$
\begin{array}{ll}
\text{minimize} & \int K(x)^2 dx \\
\text{subject to} & \int K(x) = 1, \ \int x K(x) = 0 \\
& \int x^2 K(x) = a^2 < \infty, \ K(x) \geq 0.
\end{array} \quad (12)
$$

In one dimension, the solution to Problem (12) is given by [36],

$$K^a(x) = \frac{3}{4} \frac{1}{a\sqrt{5}} \left[1 - \left(\frac{x}{a\sqrt{5}}\right)^2\right] \mathbf{1}_{\{|x| \leq a\sqrt{5}\}}. \quad (13)$$

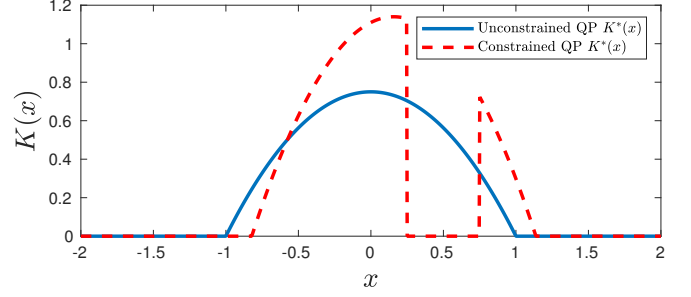Fig. 3: Optimal kernels with unconstrained support, and support constrained to $[-2, 2] \setminus [1/4, 3/4]$, with second moment $a = 5^{-1/2}$. The unconstrained kernel solution is exactly given by Equation (13).

We wish to find the solution to a modified version of this problem where we enforce a compact support constraint of the form $\{K(x) = 0, \ x \in \mathcal{A}, \ \mathcal{A}^c \text{ compact}\}$. To this end, notice that Problem (12) can be numerically solved by discretizing it as follows.

Denote the region of the problem as $X = \{x : |x| \leq B\}$. Discretize X into N points spaced dx apart. Let the vector $x := \{x_i\}_{i=1}^N \in [-B, B]^N$ consist of these points, and let $k := \{k_i\}_{i=1}^N$ be the vector of the kernel K evaluated at these points, i.e. $k_i = K(x_i)$. Discretizing the integrals yields a quadratic program of the form:

$$
\begin{array}{ll}
\text{minimize} & k^T k \\
\text{subject to} & \sum_{i=1}^N k_i dx = 1, \ \sum_{i=1}^N x_i k_i = 0 \\
& \sum_{i=1}^N x_i^2 k_i dx = a^2, \ k_i \geq 0, \ 1 \leq i \leq N.
\end{array} \quad (14)
$$

As discussed before, an agent's state-dependent density estimate will depend on sampling points from agents that have an edge between them. Hence, in the density estimate for agent i, the kernel K will only depend upon the state of agent i and its neighbours $\mathcal{N}_i$. The density estimate of agent i is written as

$$\hat{\rho}_i(t, x(t)) = \frac{1}{Nh^d} \sum_{j \in \mathcal{N}_i} \left[\prod_{k=1}^d K_i \left(\frac{x_i(t) - x_j(t)}{h}, x_j\right)\right]$$

where the support of the kernel is restricted to the support of the state-dependent edge weight $A(x_i, x_j)$:

$$A(x_i, x_j) = 0 \implies K(h^{-1}(x_i - x_j), x_j) = 0.$$

To extend Problem (14) to multi-dimensional systems, we consider *multiplicative kernels* for $x_i \in \mathbb{R}^n$, where each dimension is estimated independently: $K(\mathbf{x}) = \prod_{k=1}^N K_k(x_k)$. This yields the final optimal kernel problem with compact support constraint. For brevity, we show the explicit form for the 2D problem, as it is clear (yet notationally cumbersome) how to write the general ND problem:

$$
\begin{array}{ll}
\text{minimize} & \sum_i^{N_x} \sum_j^{N_y} k_{ij}^2 \\
\text{subject to} & \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} k_{ij} dx dy = 1, \ k_{ij} \geq 0, \ \forall i,j \\
& \sum_{i=1}^{N_x} x_i k_{ij} = 0, \ \sum_{i=1}^{N_x} x_i^2 k_{ij} dx = a^2, \ \forall j \\
& \sum_{j=1}^{N_y} y_j k_{ij} = 0, \ \sum_{j=1}^{N_y} y_j^2 k_{ij} dx = a^2, \ \forall i \\
& k_{ij} = 0 \text{ if } (x_i, y_j) \in \mathcal{A}, \ \mathcal{A}^c \text{ compact.}
\end{array}
$$

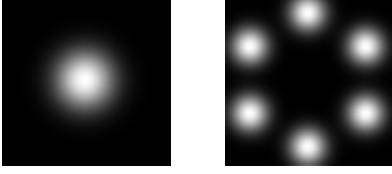It is important to note that removing a compact interval from the kernel may bias the density estimate - this

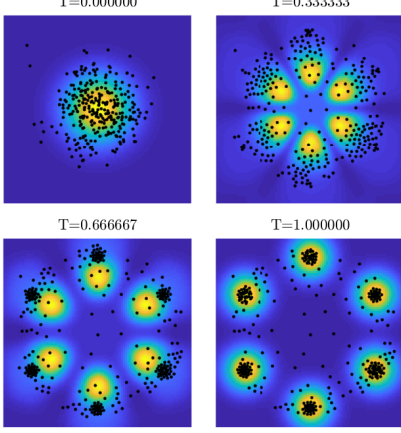Fig. 4: Left: Initial density ρ_0. Right: Target density ρ_1.

Fig. 5: Optimal density profiles over time, and superimposed agent states using the feedback density control law.

is unavoidable. The compact support constraint defines a *selection-biased distribution (SBD)*; each agent samples the distribution $g(x) = w(x)\rho(x)/\mu$, with $w(x) = \mathbf{1}_{\mathcal{A}^c}(x), \mu = \int w(x)\rho(x)dx$. The standard *unbiased* kernel density estimate of a SBD involves multiplying the kernel by a factor of $\mu/w(x)$ [38], [39], which is unbounded for our choice of $w(x)$. Techniques for unbiasing $\hat{\rho}(x)$ will be left for future work.

VI. EXAMPLES

We numerically simulate $N = 200$ agents with interaction kernel $H(x) = x\mathbf{1}_{\|x\|\le 0.01}(x)$. The velocity field $\mathcal{P}[\mu](x)$ is evaluated with the Dirac measure (8), effectively yielding N single-integrator agents that are able to only sample agents a short distance away from each other.

The optimal mass transport problem was solved using open-source code, utilizing a primal-dual algorithm [40], [41]. The density profile and velocity field was calculated over a $100 \times 100 \times 100$ grid in x, y and t space. The initial density $\rho_0(x)$ was a 2D Gaussian at the center of a $[0,1]^2$ grid, and the target density was a ring of 2D Gaussians, as shown in Figure 4. The superimposed optimal density profile over time, and the states of the agents over time integrating the feedback and feedforward control law are shown in Figure 5. As one can see, the agents are more organized around the final density distributions when using the feedback law as opposed to just integrating the feedforward law, as shown in Figure 6.

VII. CONCLUSION

In this paper, we examined density control of state-dependent networked dynamic systems. We utilized the optimal mass transport problem to design a feed-forward velocity field propelling agents with initial conditions sampled from a density profile ρ_0 to some target density ρ_1. We then
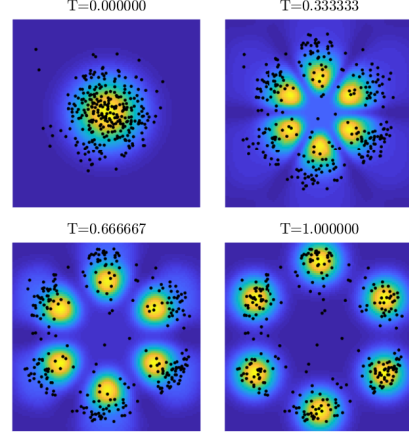

Fig. 6: Optimal density profiles over time, and superimposed agent states with only the feedforward control.

tackled the problem of using a density feedback control law with sparse measurements dictated by the state-dependent edge switching constraints of the agents. We utilized kernel density estimation to convert measurements of neighboring agents into a local estimate of the swarm density, which was then used to calculate a feedback density control law. In particular, a quadratic program was designed to find the optimal kernel subject to the state-dependent edge switching.

There are many open problems remaining, here we discuss several. First, the selection of an optimal interaction distance $r = h$ for proximity-based edge switching. This will depend on, for example, ρ_0, ρ_1 and N. If h is large, this will require more on-board computation and sensing capability; if h is small, agents will be isolated. Second, one can consider the task of determining a state-dependent kernel yielding an unbiased estimate of ρ.

APPENDIX

We now state the proof of Theorem 1.

Proof: First, recall the following theorem about consistency of the estimate $\hat{\rho}(x,t)$.

Theorem 2 ([42]): Consider a kernel density estimation scheme for the target density $\rho(x)$. Suppose the smoothing parameter h is chosen as a function of the number of samples N: $\lim_{N\to\infty} Nh(N) = \infty$. Then, at each point of continuity x of ρ, the estimator $\hat{\rho}_N(x)$ is weakly consistent in that for all $\epsilon > 0$, $\lim_{N\to\infty} \mathbb{P}(|\rho_N(x) - \rho(x)| > \epsilon) = 0$.

By Theorem 2, under the assumption on the smoothing parameter h, we have that as $N \to \infty$, $\hat{\rho}(t,x) \to \rho(t,x)$ with probability 1 for any finite t.

Consider the following Lyapunov function:

$$V(t) = \int_{\mathcal{R}} \left(\frac{\rho(x,t)}{\alpha}\right)^2 \dot{x}^T \dot{x}\ dx. \qquad (15)$$

As $N \to \infty$, the velocity field $\dot{x}$ approaches the mean-field limit [20]: $\dot{x} = \mathcal{P}(\mu,t) + u(x,t)$, where $\mathcal{P}(\mu,x) = \int A(x,y)(y-x)\mu(y,t)dy$, and $\mu := \mu(y,t)$ is the measure satisfying $\partial t + \nabla \cdot ((\mathcal{P}(\mu,t) + u(x,t))\mu) = 0$. Under the control law (11), it follows that for sufficiently large t, the Lyapunov function (15) can be written as

$$V(t) = \int_{\mathcal{R}} \nabla\Phi(x,t)^T \nabla\Phi(x,t)dx.$$

The time derivative of $V(t)$ is then given by

$$\dot{V}(t) = \alpha \int_{\mathcal{R}} \xi(x,t)^T \Delta \xi(x,t) dx,$$

where $\xi(x,t) := \nabla \Phi(x,t)$. Since $\nabla \Phi(x,t) = 0$ on $\partial \mathcal{R}$, we have that $\xi(x,t) = 0$ on $\partial \mathcal{R}$ which is a Dirichlet boundary condition. It follows that $\dot{V}(t) < 0$ since the Dirichlet problem for the Laplace operator has strictly negative eigenvalues [43].

Therefore, by LaSalle's Invariance Principle, we can conclude that $\lim_{t\to\infty} \nabla \Phi(x,t) = 0$, and so $\lim_{t\to\infty} \Phi(x,t) =$ constant. However, since $\nabla \Phi(x,t) = 0$ on $\partial \mathcal{R}$, the mass $\int_{\mathcal{R}} \Phi(x,t)dx$ is conserved for all $t > 0$ (in that $\int_{\mathcal{R}} \Phi(x,t)dx = 0$) and so we have that $\int_{\mathcal{R}} \hat{\rho}(x,t)dx = \int_{\mathcal{R}} \rho(x,t)dx$. Consequently, it follows that $\lim_{t\to\infty} \Phi(x,t) = 0$, and hence $\lim_{t\to\infty} \rho(x,t) = \lim_{t\to\infty} \hat{\rho}(x,t)$ which completes the proof. ∎

REFERENCES

[1] M. Mesbahi and M. Egerstedt, *Graph-Theoretic Methods in Multiagent Networks*. Princeton University Press, 2010.

[2] B. Açıkmeşe, M. Mandić, and J. L. Speyer, "Decentralized observers with consensus filters for distributed discrete-time linear systems," *Automatica*, vol. 50, no. 4, pp. 1037–1052, 2014.

[3] M. Hudoba de Badyn and M. Mesbahi, "Large-scale distributed Kalman filtering via an optimization approach," in *Proc. IFAC World Congress*, Toulouse, France, 2017, pp. 11 229–11 234.

[4] H. Li, C. Feng, H. Ehrhard, Y. Shen, B. Cobos, F. Zhang, K. Elamvazhuthi, S. Berman, M. Haberland, and A. L. Bertozzi, "Decentralized stochastic control of robotic swarm density: Theory, simulation, and experiment," in *Proc. IEEE International Conference on Intelligent Robots and Systems*, Vancouver, Canada, 2017, pp. 4341–4347.

[5] M. Hudoba de Badyn and M. Mesbahi, "Growing controllable networks via whiskering and submodular optimization," in *Proc. 55th IEEE Conference on Decision and Control*, Las Vegas, USA, 2016, pp. 867–872.

[6] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[7] Y. Kim and M. Mesbahi, "On maximizing the second smallest eigenvalue of a state-dependent graph Laplacian," *IEEE Transactions on Automatic Control*, vol. 51, no. 1, pp. 116–120, 2006.

[8] V. Krishnan and S. Martínez, "Identification of critical node clusters for consensus in large-scale spatial networks," in *IFAC-PapersOnLine*, vol. 50-1, 2017, pp. 14 721–14 726.

[9] A. Rahmani, M. Ji, M. Mesbahi, and M. Egerstedt, "Controllability of multi-agent systems from a graph-theoretic perspective," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 162–186, 2009.

[10] S. Alemzadeh, M. Hudoba de Badyn, and M. Mesbahi, "Controllability and stabilizability analysis of signed consensus networks," in *Proc. IEEE Conference on Control Technology and Applications*, Kohala Coast, USA, 2017, pp. 55–60.

[11] M. Hudoba de Badyn, S. Alemzadeh, and M. Mesbahi, "Controllability and data-driven identification of bipartite consensus on nonlinear signed networks," in *Proc. 56th IEEE Conference on Decision and Control*, Melbourne, Australia, 2017.

[12] C. Altafini, "Consensus problems on networks with antagonistic interactions," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 935–946, 2013.

[13] M. Hudoba de Badyn, A. Chapman, and M. Mesbahi, "Network entropy: A system-theoretic perspective," in *Proc. 54th IEEE Conference on Decision and Control*, Osaka, Japan, 2015, pp. 5512–5517.

[14] K. R. Harrison, A. P. Engelbrecht, and B. M. Ombuki-Berman, "Optimal parameter regions and the time-dependence of control parameter values for the particle swarm optimization algorithm," *Swarm and Evolutionary Computation*, pp. 1–40, 2018.

[15] L. Chaimowicz, N. Michael, and V. Kumar, "Controlling swarms of robots using interpolated implicit functions," in *Proc. 2005 IEEE International Conference on Robotics and Automation*, Barcelona, Spain, 2005, pp. 2487–2492.

[16] L. C. Pimenta, G. A. Pereira, N. Michael, R. C. Mesquita, M. M. Bosque, L. Chaimowicz, and V. Kumar, "Swarm coordination based on smoothed particle hydrodynamics technique," *IEEE Transactions on Robotics*, vol. 29, no. 2, pp. 383–399, 2013.

[17] U. Eren and B. Açıkmeşe, "Velocity field generation for density control of swarms using heat equation and smoothing kernels," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 9405–9411, 2017.

[18] V. Krishnan and S. Martínez, "Distributed control for spatial self-organization of multi-agent swarms," *ArXiv preprint arXiv:1705.03109*, pp. 1–33, 2017.

[19] ——, "Self-organization in multi-agent swarms via distributed computation of diffeomorphisms," in *Mathematical Theory of Networks and Systems*, Minneapolis, USA, 2016.

[20] G. Albi, Y.-P. Choi, M. Fornasier, and D. Kalise, "Mean field control hierarchy," *Applied Mathematics & Optimization*, vol. 76, no. 1, pp. 95–135, 2017.

[21] P. E. Caines, "Mean field games," in *Encyclopedia of Systems and Control*. Springer, 2013, pp. 1–6.

[22] M. Fornasier and F. Solombrino, "Mean-field optimal control," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 20, no. 4, pp. 1123–1152, 2014.

[23] M. Nabi-Abdolyousefi and M. Mesbahi, "System theory over random consensus networks: Controllability and optimality properties," in *Proc. 50th IEEE Conference on Decision and Control*, Orlando, USA, 2011, pp. 2323–2328.

[24] A. Tahbaz-Salehi and A. Jadbabaie, "On recurrence of graph connectivity in Vicsek's model of motion coordination for mobile autonomous agents," *Proceedings of the American Control Conference*, pp. 699–704, 2007.

[25] M. Mesbahi, "On a dynamic extension of the theory of graphs," in *Proceedings of the American Control Conference*, vol. 2, Anchorage, USA, 2002, pp. 1234–1239.

[26] ——, "On state-dependent dynamic graphs and their controllability properties," *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 387–392, 2005.

[27] A. Awad, A. Chapman, E. Schoof, A. Narang-Siddarth, and M. Mesbahi, "Time-scale separation on networks: Consensus, tracking, and state-dependent interactions," in *Proc. 54th IEEE Conference on Decision and Control*, Osaka, Japan, 2015, pp. 6172–6177.

[28] ——, "Time-scale separation in networks: State-dependent graphs and consensus tracking," *IEEE Transactions on Control of Network Systems*, vol. 5870, no. c, pp. 1–1, 2018.

[29] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal transport over a linear dynamical system," *IEEE Transactions on Automatic Control*, vol. 62, no. 5, pp. 2137–2152, 2017.

[30] K. Elamvazhuthi and P. Grover, "Optimal transport over nonlinear systems via infinitesimal generators on graphs," *arXiv:1612.01193*, 2016.

[31] Y. Chen and J. Karlsson, "State tracking of linear ensembles via optimal mass transport," *arXiv preprint arXiv: 1803.02296*, pp. 1–6, 2018.

[32] C. Villani, *Topics in Optimal Transportation*. American Mathematical Society, 2003.

[33] ——, *Optimal Transport: Old and New*. Springer Science & Business Media, 2009, vol. 338.

[34] L. V. Kantorovich, "On a problem of Monge," *Uspekhi Matematicheskikh Nauk*, vol. 3, pp. 225–226, 1948.

[35] 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.

[36] M. P. Wand and M. C. Jones, *Kernel Smoothing*. CRC Press, 1994.

[37] E. Parzen, "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, 1962.

[38] R. D. Gill, Y. Vardi, and J. A. Wellner, "Large sample theory of empirical distributions in biased sampling models," *The Annals of Statistics*, vol. 16, no. 3, pp. 1069–1112, 1988.

[39] C. O. Wu and A. Q. Mao, "Minimax kernels for density estimation with biased data," *Annals of the Institute of Statistical Mathematics*, vol. 48, no. 3, pp. 451–467, 1996.

[40] N. Papadakis, G. Peyré, and E. Oudet, "Optimal transport with proximal splitting," *SIAM Journal on Imaging Sciences*, vol. 7, no. 1, pp. 212–238, 2014.

[41] G. Peyré, "The numerical tours of signal processing," *Advanced Computational Signal and Image Processing IEEE Computing in Science and Engineering*, vol. 13, no. 4, pp. 94–97, 2011.

[42] D. Wied and R. Weißbach, "Consistency of the kernel density estimator: A survey," *Statistical Papers*, vol. 53, no. 1, pp. 1–21, 2012.

[43] P. Li and S.-T. Yau, "On the Schrodinger equation and the eigenvalue problem," *Commun. Math. Phys. Mathematical Physics*, vol. 88, no. 81, pp. 309–318, 1983.