Mean-Field Controllability and Decentralized Stabilization of Markov Chains

Karthik Elamvazhuthi, Matthias Kawski, Shiba Biswal, Vaibhav Deshmukh, and Spring Berman

Abstract—In this paper, we present several novel results on controllability and stabilizability properties of the Kolmogorov forward equation of a continuous time Markov chain (CTMC) evolving on a finite state space, using the transition rates as the control parameters. First, we characterize all the stationary distributions that are stabilizable using time-independent control parameters. We then present a result on small-time local and global controllability of the system from and to strictly positive equilibrium distributions when the underlying graph is strongly connected. Additionally, we show that any target distribution can be reached asymptotically using timevarying control parameters. For bidirected graphs, we construct rational and polynomial density feedback laws that stabilize strictly positive stationary distributions while satisfying the additional constraint that that the feedback law takes zero value at equilibrium. This last result enables the construction of decentralized density feedback controllers, using tools from linear systems theory and sum-of-squares based polynomial optimization, that stabilize a swarm of agents modeled as a CTMC to a target state distribution with no state-switching at equilibrium. We validate the effectiveness of the constructed feedback laws with stochastic simulations of the CTMC for finite numbers of agents and numerical solutions of the corresponding mean-field models.

I. INTRODUCTION

In this paper, we address the problem of redistributing a large number of homogeneous agents among a set of states, such as tasks to be performed or spatial locations to occupy. This problem has many applications in swarm robotics, for example, such as environmental monitoring, surveillance, disaster response, and autonomous construction. In recent years, approaches to this problem have been developed in which the agents are programmed to switch stochastically between states at tunable transition rates. In some of these approaches [5], [19], [20], the agents' states evolve according to a continuous time Markov chain (CTMC), and their state distribution is controlled using the corresponding mean-field model, given by the Kolmogorov forward equation. Similar approaches in [1], [2] specify that the agents' dynamics evolve in discrete time, in which case their state evolution is described by a discrete time Markov chain (DTMC). These methods enable the scalable design of robot controllers due

to the independence of the control methodology from the number of agents.

Various methods for control synthesis in this framework have been proposed not only for applications in swarm robotics, but also in the context of mean-field games and optimal transport theory. The works [5] and [8] address an open-loop optimal control problem for the Kolmogorov forward equation, in which the control parameters, which are the transition probabilities or rates of the associated CTMC, are constrained to be time-invariant. These approaches have also been extended to the case of time-varying control parameters [3], [22], [15]. In [2], [11], [20], feedback controllers are designed to drive a Markov chain to a target distribution. In contrast to traditional control approaches for Markov chains that use only the agent states as feedback [21], these works use the densities of agents in different states as feedback and continuously recompute the control parameters such that the target distribution is stabilized. To avoid requiring agents to have global information on these densities, *decentralized* control approaches were developed either by a priori restricting the controller to have a decentralized structure [20] or by designing a centralized controller and then using estimation algorithms to estimate the global density of the swarm in a decentralized manner [11]. There has also been some related recent work on mean-field games, in which Hamilton-Jacobi-Bellman (HJB) based methods are used for control synthesis [17]. However, unlike HJB based methods in classical control theory, approaches based on mean-field games do not result in density feedback controllers unless analytical solutions can be derived. The synthesized control inputs are open-loop in nature and yield the desired behavior only for predefined fixed initial conditions of the mean-field model [4].

In this paper, we present novel results on the controllability and stabilizability of the mean-field control problem for CTMCs. We study local and global controllability properties of the forward equation when the control inputs are required to be zero at equilibrium. The case when control inputs are not constrained to be zero at equilibrium is comparatively much easier, since local controllability follows directly from linearization-based arguments, so we do not consider this case here. We also demonstrate that it is possible to compute density-independent transition rates of a CTMC that make any probability distribution with a strongly connected support (to be defined later) invariant and globally stable. Similar work in [1] has characterized the class of stabilizable stationary distributions for DTMCs with control parameters that are time- and density-invariant; we characterize this class of distributions for CTMCs with the same type of control

This work was supported by ONR Young Investigator Award N00014-16-1-2605 and by the Arizona State University Global Security Initiative.

Karthik Elamvazhuthi, Shiba Biswal, Vaibhav Deshmukh, and Spring Berman are with the School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ, 85281 USA {karthikevaz, sbiswal, vdeshmuk, Spring.Berman}@asu.edu.

Matthias Kawski is with the School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ, 85281 USA {kawski@asu.edu}.

parameters (see Theorem IV.5). We show that this result can be further strengthened by employing time-varying control parameters that make the system asymptotically controllable to any feasible probability distribution.

In addition, we address the stabilization of mean-field models using decentralized density feedback laws under the constraint that the transition rates are required to be zero at equilibrium. Such a constraint is needed in swarm robotic applications to prevent robots from constantly switching between states at equilibrium. The problem of unnecessary state-switching was previously addressed for CTMCs in [20] as a variance control problem, and for DTMCs in [3] using a decentralized density estimation strategy that implements centralized feedback laws and ensures that the transition matrix is the identity matrix at equilibrium. In this paper, we investigate the CTMC case in more detail. In contrast to [20], we explicitly show that any (strictly positive) distribution is stabilizable using a decentralized feedback law, and we impose the additional constraint that transition rates must be zero at equilibrium. Moreover, the controller in [20] was proved to be stabilizing with the assumption that negative transition rates are admissible, and was then implemented with a saturation condition in order to avoid negative rates, in which case the stability guarantees are lost. We show how this issue can be resolved with a linear controller by interpreting a negative flow from one state to another as a positive flow of appropriate magnitude in the opposite direction. While the algorithmic construction of linear controllers has low computational complexity, these controllers violate positivity constraints on the control inputs. To realize linear controllers in practice for our problem, we show that for bidirected graphs, we can implement linear controllers with rational feedback laws that mimic their behavior. However, since this approach results in unbounded controls, we also propose a feedback law that is a positive polynomial function of the densities and therefore guarantees the global boundedness of the controls.

The proofs of the results presented in this paper will be made available in a forthcoming manuscript [13]. Preliminary versions of these proofs are given in [14], [6].

II. NOTATION

We first define the notation that will be used to formulate the problems addressed in this paper. We denote by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ a directed graph with a set of M vertices, $\mathcal{V} = \{1, 2, ..., M\}$, and a set of $N_{\mathcal{E}}$ edges, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$. An edge from vertex $i \in \mathcal{V}$ to vertex $j \in \mathcal{V}$ is denoted by $e = (i, j) \in \mathcal{E}$. We define a source map $S : \mathcal{E} \to \mathcal{V}$ and a target map $T : \mathcal{E} \to \mathcal{V}$ for which S(e) = i and T(e) = jwhenever $e = (i, j) \in \mathcal{E}$. There is a *directed path* of length s from node $i \in \mathcal{V}$ to node $j \in \mathcal{V}$ if there exists a sequence of edges $\{e_i\}_{i=1}^s$ in \mathcal{E} such that $S(e_1) = i$, $T(e_s) = j$, and $S(e_k) = T(e_{k-1})$ for all $1 \leq k < s - 1$. A directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is called *strongly connected* if for every pair of distinct vertices $v_0, v_T \in \mathcal{V}$, there exists a *directed path* of edges in \mathcal{E} connecting v_0 to v_T . We assume that $(i, i) \notin \mathcal{E}$ for all $i \in \mathcal{V}$. The graph \mathcal{G} is said to be *bidrected* if $e \in \mathcal{E}$ implies that $\tilde{e} = (T(e), S(e))$ also lies in \mathcal{E} .

We denote the *M*-dimensional Euclidean space by \mathbb{R}^{M} . $\mathbb{R}^{M \times N}$ will refer to the space of $M \times N$ matrices, and \mathbb{R}_{+} will refer to the set of positive real numbers. Given a vector $\mathbf{x} \in \mathbb{R}^{M}$, x_{i} will refer to the i^{th} coordinate value of \mathbf{x} . The 2-norm of the vector $\mathbf{x} \in \mathbb{R}^{M}$ is denoted by $\|\mathbf{x}\|_{2} = \sqrt{\sum_{i} x_{i}^{2}}$. For a matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$, A^{ij} will refer to the element in the i^{th} row and j^{th} column of \mathbf{A} . For a subset $B \subset \mathbb{R}^{M}$, $\operatorname{int}(B)$ will refer to the interior of the set B.

III. PROBLEM FORMULATION

We consider a swarm of N autonomous agents whose states evolve in continuous time according to a Markov chain with finite state space $\mathcal{V} = \{1, ..., M\}$, the vertex set of a given graph \mathcal{G} . As an example application of interest, \mathcal{V} can represent a set of spatial locations that are obtained by partitioning the agents' environment. The agents must reallocate among the states to achieve a target population distribution at equilibrium. The edge set \mathcal{E} of \mathcal{G} defines the pairs of vertices (states) between which the agents can transition. Denoting the set of admissible control inputs by $U \subset \mathbb{R}$, the agents' transition rules are determined by the control parameters $u_e: [0,\infty) \to U$ for each $e \in \mathcal{E}$, and are known as the transition rates of the associated CTMC. An agent in state v_1 at time t decides to switch to state v_2 at probability per unit time $u_e(t)$, $e = (v_1, v_2)$. We will mostly focus on the case where $U \subset \mathbb{R}_+$, i.e., the $u_e(t)$ obey positivity constraints, since transition rates must always be positive for a CTMC. However, we will also consider the case where transition rates may be negative, in order to facilitate analysis of the case with positivity constraints. The state of each agent $i \in \{1, ..., N\}$ is defined by a stochastic process $X_i(t)$ that evolves on the state space \mathcal{V} according to the conditional probabilities

$$\mathbb{P}(X_i(t+h) = T(e)|X_i(t) = S(e)) = u_e(t)h + o(h) \quad (1)$$

for each $e \in \mathcal{E}$. Here, o(h) is the little-oh symbol and \mathbb{P} is the underlying probability measure induced on the the space of events Ω (which will be left undefined, as is common) by the stochastic processes $\{X_i(t)\}_{i=1}^N$. Let $\mathcal{P}(\mathcal{V}) = \{\mathbf{y} \in \mathbb{R}^M_+; \sum_v y_v = 1\}$ be the simplex of probability densities on \mathcal{V} . Corresponding to the CTMC is a set of ordinary differential equations (ODEs) which determines the evolution of the probability densities $\mathbb{P}(X_i(t) = v) = x_v(t) \in \mathbb{R}_+$. Since $\{X_i\}_{i=1}^N$ is a set of independent and identically distributed random variables, the *Kolmogorov forward equation* can be represented by a single linear system of ODEs,

$$\dot{\mathbf{x}}(t) = \sum_{e \in \mathcal{E}} u_e(t) \mathbf{B}_e \mathbf{x}(t), \quad t \in [0, \infty), \qquad (2)$$
$$\mathbf{x}(0) = \mathbf{x}^0 \in \mathcal{P}(\mathcal{V}),$$

where \mathbf{B}_e are control matrices whose entries are given by

$$B_{e}^{ij} = \begin{cases} -1 & \text{if } i = j = S(e), \\ 1 & \text{if } i = T(e), \ j = S(e), \\ 0 & \text{otherwise.} \end{cases}$$
(3)

The focus of this paper is to study controllability and stabilizability properties of the control system (2). To describe the controllability problem of interest, we first recall some controllability notions from nonlinear control theory [7].

Definition III.1. Given U and $\mathbf{x}^0 \in \mathcal{P}(\mathcal{V})$, we define $R^U(\mathbf{x}^0, t)$ to be the set of all $\mathbf{y} \in \mathcal{P}(\mathcal{V})$ for which there exists an admissible control, $\mathbf{u} = \{u_e\}_{e \in \mathcal{E}}$, taking values in U such that there exists a trajectory of system (2) with $\mathbf{x}(0) = \mathbf{x}^0, \mathbf{x}(t) = \mathbf{y}$. The reachable set from \mathbf{x}^0 at time T is defined to be

$$R_T^U(\mathbf{x}^0) = \bigcup_{0 \le t \le T} R^U(\mathbf{x}^0, t).$$
(4)

Definition III.2. The system (2) is said to be **small-time locally controllable (STLC)** from an equilibrium distribution $\mathbf{x}^d \in \mathcal{P}(\mathcal{V})$ if the set of reachable states $R_T^U(\mathbf{x}^d)$ contains a neighborhood of $\mathbf{x}^d \in \mathcal{P}(\mathcal{V})$ in the subspace topology of $\mathcal{P}(\mathcal{V})$ (as a subset of \mathbb{R}^M) for any T > 0.

Here, we have defined local controllability in terms of the subspace topology of $\mathcal{P}(\mathcal{V})$. This is because the set $\mathcal{P}(\mathcal{V})$ is invariant for the system (2) of controlled ODEs, and hence one cannot expect controllability to a full neighborhood of \mathbf{x}^d . Informally, this just means that, due to conservation of mass, one cannot create or destroy agents by manipulating their rates of transitioning from one vertex to another.

Our first problem of interest can be framed as follows:

Problem III.3. Given $\mathbf{x}^d \in \mathcal{P}(\mathcal{V})$, determine if the system (2) is STLC from $\mathbf{x}^d \in \mathcal{P}(\mathcal{V})$.

The above problem allows the control inputs to be timevarying. We now pose a problem in which the control inputs are constrained to be time-independent.

Problem III.4. Given $\mathbf{x}^d \in \mathcal{P}(\mathcal{V})$, determine if there exist positive control parameters $\{u_e\}_{e \in \mathcal{E}}$ for the system (2) such that $\lim_{t\to\infty} \|\mathbf{x}(t) - \mathbf{x}^d\| = 0$ for all $\mathbf{x}^0 \in \mathcal{P}(\mathcal{V})$.

We provide a complete characterization of the stationary distributions that are stabilizable for this case. Although density- and time-independent transition rates of CTMCs have been previously computed in an optimization framework [5], the question of which equilibrium distributions are feasible has remained unresolved for the case where the target distribution is not strictly positive on all vertices. While only strictly positive target distributions have been considered in previous work on control of swarms governed by CTMCs [5], we address the more general case in which the target densities of some states can be zero. This question was addressed in [1] for swarms governed by DTMCs. The problem has also been investigated in the context of consensus protocols [9] for strictly positive distributions, where what is referred to as "advection on graphs" is in fact the forward equation corresponding to a CTMC. In our controller synthesis, we will relax the assumption of strict positivity for desired target distributions.

Next, we consider the feedback stabilization problem for

system (2). Consider the following system:

$$\dot{\mathbf{x}}(t) = \sum_{e \in \mathcal{E}} k_e(\mathbf{x}(t)) \mathbf{B}_e \mathbf{x}(t), \quad t \in [0, \infty), \quad (5)$$
$$\mathbf{x}(0) = \mathbf{x}^0 \in \mathcal{P}(\mathcal{V}).$$

Problem III.5. Given $\mathbf{x}^d \in \mathcal{P}(\mathcal{V})$, determine whether there exists a decentralized feedback law, defined as a collection of maps $\tilde{k}_e : \mathbb{R}^2 \to \mathbb{R}_+$ where $k_e(\mathbf{y}) = \tilde{k}_e(y_{S(e)}, y_{T(e)})$ for each $e \in \mathcal{E}$ and $\mathbf{y} \in \mathbb{R}^M$, such that for the closed-loop system (5), \mathbf{x}^d is asymptotically stable and $k_e(\mathbf{x}^d) = 0$ for each $e \in \mathcal{E}$.

We note that we were able to describe the state evolution of the agents by system (2) when the transition rates were density-independent because the agents' states were independent and identically distributed (i.i.d.) random variables in that case. However, when the density feedback control law $\{u_e(\mathbf{x})\}_{e \in \mathcal{E}}$ is used, the independence of the stochastic processes $X_i(t)$ is lost. This implies that the evolution of the probability distribution cannot be described by system (2). However, if we invoke the mean-field hypothesis and take the limit $N \to \infty$, then we can model the evolution of the probability distribution according to a nonlinear Markov *chain.* In this limit, the discrete number of agents $N_v(t,\omega)$ in state $v \in \mathcal{V}$ at time $t \in [0, T]$, where ω is used to emphasize that N_v is a measurable function of the sample path Ω , converges to the continuous agent population $x_v(t)$ in an appropriate sense, provided that solutions of system (5) are defined until a given final time T > 0. A rigorous process for taking this limit in a stochastic process setting is described in [16], [18].

IV. ANALYSIS

A. Controllability

In this section, we state several results on controllability that address Problems III.3 and III.4.

Intuitively, if the graph G is not strongly connected, then we should not have local controllability. This is formalized in the following proposition.

Proposition IV.1. If the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is not strongly connected, then the system (2) is not locally controllable.

When the graph is strongly connected, we can transport a small amount of mass from one vertex to another using a sequence of control inputs that are associated with the edges of the graph. This observation enables us to state the following STLC property of system (2):

Proposition IV.2. Let $U = [0, \epsilon]$ for some $\epsilon > 0$. If the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is strongly connected, then the system (2) is STLC from every point in $int(\mathcal{P}(\mathcal{V}))$.

Note that Proposition IV.2 does not immediately follow from classical conditions for controllability such as the Kalman rank condition in linear systems theory or the Lie Bracket conditions in geometric control theory, due to the positivity constraints on the control inputs and the fact that 0 does not lie in the interior of the set of control inputs U. The constructive nature of the proof of Proposition IV.2 is largely due to the sparse structure of the matrix \mathbf{B}_e defined in (3), which enables an explicit representation of the matrix exponential $\exp(t\mathbf{B}_e)$. For any edge $e = (i, j) \in \mathcal{E}$, the exponential of the control matrix \mathbf{B}_e is a stochastic matrix with entries given by

$$(\exp t\mathbf{B}_e)^{k\ell} = \begin{cases} 1 & \text{if } k = \ell \neq S(e) \\ e^{-t} & \text{if } k = \ell = S(e) \\ 1 - e^{-t} & \text{if } k = T(e) \text{ and } \ell = S(e) \\ 0 & \text{otherwise.} \end{cases}$$
(6)

Rather than deriving a general formula for the corresponding product on the group for arbitrary edges $(i, j), (j, \ell) \in \mathcal{E}$, we state the product for the special case of $\mathcal{V} = \{1, 2, 3\}$ and edges e = (1, 2) and e' = (2, 3) as an illustration:

$$e^{t\mathbf{B}_{e'}}e^{s\mathbf{B}_{e}} = \begin{pmatrix} e^{-s} & 0 & 0\\ e^{-t}(1-e^{-s}) & e^{-t} & 0\\ (1-e^{-t})(1-e^{-s}) & 1-e^{-t} & 1 \end{pmatrix}.$$
 (7)

Despite the assumption that \mathcal{G} is strongly connected, there are some limitations on the controllability of system (2) due to the nature of its control vector fields. In particular, global controllability can only be guaranteed for target densities that lie in the interior of the domain of the simplex $\mathcal{P}(\mathcal{V})$, as stated in the theorem below.

Theorem IV.3. If the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is strongly connected, then the system (2) is small-time globally controllable from every point in the interior of the simplex $\mathcal{P}(\mathcal{V})$.

In fact, we can state the following broader result. If \mathcal{G} is strongly connected, then the system (2) is also *path* controllable: given any trajectory $\gamma(t)$ in $\mathcal{P}(\mathcal{V})$ that is defined over a finite time interval [0,T] and is once differentiable with respect to the time variable t, there exists a control law $\mathbf{u} : [0,T] \rightarrow [0,\infty)^{N_{\mathcal{E}}}$ such that the solution of the control system (2) satisfies $\mathbf{x}(t) = \gamma(t)$ for all $t \in [0,T]$. This is true because conical combinations of the collection of vectors $\{\mathbf{B}_{e}\mathbf{y}\}_{e\in\mathcal{E}}$ span the tangent space of $\mathcal{P}(\mathcal{V})$ whenever \mathbf{y} lies in the interior of $\mathcal{P}(\mathcal{V})$. For example, given a strongly connected graph \mathcal{G} , if $(i, j) \in \mathcal{E}$ and there exists a directed path μ from j to i, then $-\mathbf{B}_{(i,j)}\mathbf{1} = \sum_{e\in\mu} \mathbf{B}_{e}\mathbf{1}$.

It would be desirable to extend the above result to include target distributions that lie on the boundary of $\mathcal{P}(\mathcal{V})$. However, as stated previously, one cannot expect to reach target distributions on the boundary in finite time. The boundary of $\mathcal{P}(\mathcal{V})$ is unreachable if the initial condition of system (2) starts from the interior of $\mathcal{P}(\mathcal{V})$, even if one uses possibly unbounded but measurable inputs with finite Lebesgue integral. The following counterexample clarifies this point.

Example IV.4. Consider system (2) for a bidirected graph G with two vertices:

$$\begin{aligned} \dot{x}_1(t) &= -u_{(1,2)}(t)x_1(t) + u_{(2,1)}x_2(t), \quad (8) \\ \dot{x}_2(t) &= u_{(1,2)}(t)x_1(t) - u_{(2,1)}x_2(t), \\ x_1(0) &= x_1^0, \quad x_2(0) = x_2^0. \end{aligned}$$

Let $u_{(1,2)}, u_{(2,1)} \in L^1_+(0,1)$, the set of positive-valued measurable inputs with finite integrals over the time interval (0,1). Then the solution, $\mathbf{x}(t) = [x_1(t) \ x_2(t)]^T$, satisfies:

$$x_{1}(t) = x_{1}^{0} - \int_{0}^{t} (u_{(1,2)}(\tau)x_{1}(\tau) - u_{(2,1)}(\tau)x_{2}(\tau))d\tau, \quad (9)$$

 $x_{2}(t) = x_{2}^{0} + \int_{0}^{0} (u_{(1,2)}(\tau)x_{1}(\tau) - u_{(2,1)}(\tau)x_{2}(\tau))d\tau, (10)$ such that $x_{1}^{0} \in (0, 1)$ and $x_{2}^{0} = 1 - x_{1}^{0}$. We assume, without

such that $x_1^0 \in (0, 1)$ and $x_2^0 = 1 - x_1^0$. We assume, without loss of generality, that $x_1(t) > 0$ for all $t \in [0, 1)$. Then for each $T \in [0, 1)$, Equations (9) and (10) imply that:

$$x_1(T) = x_1^0 - \int_0^T \left(u_{(1,2)}(\tau) + u_{(2,1)}(\tau) - \frac{u_{(2,1)}(\tau)}{x_1(\tau)} \right) x_1(\tau) d\tau.$$

From this equation, we can conclude that

$$x_{1}(1) \geq x_{1}^{0} - \int_{0}^{1} (u_{(1,2)}(\tau) + u_{(2,1)}(\tau)\tilde{x}_{1}(\tau))d\tau \quad (11)$$

$$= \exp\left(-\int_{0}^{1} (u_{(1,2)}(\tau) + u_{(2,1)}(\tau))d\tau\right)x_{1}^{0},$$

where \tilde{x}_1 is the solution of the differential equation

$$\dot{\tilde{x}}_1(t) = -(u_{(1,2)}(t) + u_{(2,1)}(t))\tilde{x}_1(t),$$
 (12)
$$\tilde{x}_1(0) = x_1^0.$$

Therefore, it must be true that $\exp\left(-\int_{0}^{1}(u_{(1,2)}(\tau)+u_{(2,1)}(\tau))d\tau\right)x_{1}^{0}\leq 0$, which yields a contradiction since $x_{1}^{0}\neq 0$.

The above observation is not a significant disadvantage, since each point in $\mathcal{P}(\mathcal{V})$ is at least *asymptotically controllable*. Before demonstrating why this is true, we address Problem III.4 and hence give a complete characterization of the class of equilibrium stationary distributions that are stabilizable using time-independent control inputs.

Theorem IV.5. Let \mathcal{G} be a strongly connected graph. Suppose that $\mathbf{x}^0 \in \mathcal{P}(\mathcal{V})$ is an initial distribution and $\mathbf{x}^d \in \mathcal{P}(\mathcal{V})$ is a desired distribution. Additionally, assume that \mathbf{x}^d has strongly connected support. Then there is a set of parameters, $a_e \in [0, \infty)$ for each $e \in \mathcal{E}$, such that if $u_e(t) = a_e$ for all $t \in [0, \infty)$ and for each $e \in \mathcal{E}$ in system (2), then the solution $\mathbf{x}(t)$ of this system satisfies $\|\mathbf{x}(t) - \mathbf{x}^d\| \leq Me^{-\lambda t}$ for all $t \in [0, \infty)$ and for some positive parameters M and λ that are independent of \mathbf{x}^0 .

Now we are ready to state the asymptotic controllability result referred to previously.

Theorem IV.6. Let \mathcal{G} be a strongly connected graph. Suppose that $\mathbf{x}^0 \in \mathcal{P}(\mathcal{V})$ is the initial distribution, and $\mathbf{x}^d \in \mathcal{P}(\mathcal{V})$ is the desired distribution. Then for each $e \in \mathcal{E}$, there exists a set of time-dependent control parameters $u_e : \mathbb{R}_+ \to \mathbb{R}_+$, $e \in \mathcal{E}$, such that the solution $\mathbf{x}(t)$ of the controlled ODE (2) satisfies $\lim_{t\to\infty} \mathbf{x}(t) = \mathbf{x}^d$.

The above result can be proved for control inputs that are globally bounded in time. Hence, it follows that any point in $\mathcal{P}(\mathcal{V})$, including those lying on the boundary of $\mathcal{P}(\mathcal{V})$, can be stabilized using a full-state feedback controller [10].

B. Stabilization

Now we investigate the stabilizability properties of the system (2). Note that stabilizability using centralized feedback follows from the controllability result in Theorem IV.3. Hence, our focus in this section is to establish stabilizability using decentralized control laws.

Lemma IV.7. Let \mathcal{G} be a strongly connected graph. Suppose that $\mathbf{x}^d \in \operatorname{int}(\mathcal{P}(\mathcal{V}))$. Let $k_e : \mathbb{R}^M \to (-\infty, \infty)$ be given by $k_e(\mathbf{y}) = x_{T(e)}^d y_{S(e)} - x_{S(e)}^d y_{T(e)}$ in system (5), for each $e \in \mathcal{E}$ and each $\mathbf{y} \in \mathbb{R}^M$. Then for this system, the equilibrium point \mathbf{x}^d is locally exponentially stable on the space $\mathcal{P}(\mathcal{V})$, *i.e.*, there exists r > 0 such that $\|\mathbf{x}_0 - \mathbf{x}^d\|_2 < r$ and $\mathbf{x}_0 \in \mathcal{P}(\mathcal{V})$ imply that the solution $\mathbf{x}(t)$ of the system satisfies the inequality

$$\|\mathbf{x}(t) - \mathbf{x}^d\|_2 \le M_0 e^{-\lambda t} \tag{13}$$

for all $t \in [0,\infty)$ and for some parameters $M_0 > 0$ and $\lambda > 0$ that depend only on r.

The above lemma implies that if negative transition rates are admissible, then there exists a linear feedback law, $\{k_e\}_{e \in \mathcal{E}}$, such that $k_e(\mathbf{x}^d) = 0$ for each $e \in \mathcal{E}$ and the desired equilibrium point is locally exponentially stable.

A desirable property of the control system (2) is that stabilization of the desired equilibrium can be achieved using a linear feedback law that satisfies positivity constraints away from equilibrium and is zero at equilibrium. However, any stabilizing linear control law that is zero at equilibrium must in fact be zero everywhere. On the other hand, in the next theorem we show that whenever \mathcal{G} is bidirected, any feedback control law that violates positivity constraints can be implemented using a rational feedback law of the form $k(\mathbf{x}) = a(\mathbf{x}) + b(\mathbf{x}) \frac{f(\mathbf{x})}{g(\mathbf{x})}$, such that $k(\mathbf{x})$ satisfies the positivity constraints and is zero at equilibrium.

Theorem IV.8. Let \mathcal{G} be a bidirected graph. Let $k_e : \mathbb{R}^M \to (-\infty, \infty)$ be a map for each $e \in \mathcal{E}$ such that there exists a unique global solution of the system (5). Additionally, assume that $\mathbf{x}(t) \in \operatorname{int}(\mathcal{P}(\mathcal{V}))$ for each $t \in [0, \infty)$. Consider the functions $m_e^p : \mathbb{R}^M \to \{0, 1\}$ and $m_e^n : \mathbb{R}^M \to \{0, 1\}$, defined as follows for each $e \in \mathcal{E}$:

$$m_e^p(\mathbf{y}) = 1 \quad if \quad k_e(\mathbf{y}) \ge 0, \quad 0 \quad otherwise;$$
(14)

 $m_e^n(\mathbf{y}) = 1 \quad if \quad k_e(\mathbf{y}) \le 0, \quad 0 \quad otherwise.$ (15)

Let $c_e : \mathbb{R}^M \to [0,\infty)$ be given by

$$c_e(\mathbf{y}) = m_e^p(\mathbf{y})k_e(\mathbf{y}) - m_{\tilde{e}}^n(\mathbf{y})k_{\tilde{e}}(\mathbf{y})\frac{y_{T(e)}}{y_{S(e)}}.$$
 (16)

Then the solution $\tilde{\mathbf{x}}(t)$ of the following system,

$$\dot{\tilde{\mathbf{x}}} = \sum_{e \in \mathcal{E}} c_e(\tilde{\mathbf{x}}(t)) \mathbf{B}_e \tilde{\mathbf{x}}(t), \quad t \in [0, \infty), \quad (17)$$
$$\tilde{\mathbf{x}}(0) = \mathbf{x}^0 \in \operatorname{int}(\mathcal{P}(\mathcal{V})),$$

is unique, defined globally, and satisfies $\tilde{\mathbf{x}}(t) = \mathbf{x}(t)$ for all $t \in [0, \infty)$.

In the above theorem, it is required that $\mathbf{x}(t) \in int(\mathcal{P}(\mathcal{V}))$ for all $t \ge 0$. This assumption on the initial distribution can be avoided if one uses polynomial feedback instead, as in the following theorem.

Theorem IV.9. Let \mathcal{G} be a bidirected graph. Suppose that $\mathbf{x}^d \in \operatorname{int}(\mathcal{P}(\mathcal{V}))$. Let $k_e : \mathbb{R}^M \to [0, \infty)$ be given by $k_e(\mathbf{y}) = (y_{S(e)} - x_{S(e)}^d)^2 + (y_{T(e)} - x_{T(e)}^d)^2$ in system (5), for each $e \in \mathcal{E}$ and each $\mathbf{y} \in \mathbb{R}^M$. Then for this system, the equilibrium point \mathbf{x}^d is globally asymptotically stable.

V. NUMERICAL SIMULATIONS

In this section, we numerically verify the effectiveness of the decentralized feedback controllers that are defined in Lemma IV.7 (the linear controller) and Theorem IV.9 (the nonlinear controller). The controllers were constructed to redistribute populations of N = 80 and N = 1200agents on the six-vertex bidirected graph shown in Fig. 1. In all cases, the initial distribution of agents was set to $\mathbf{x}^0 = [0.2 \ 0.1 \ 0.2 \ 0.15 \ 0.2 \ 0.15]^T$, and the desired distribution was $\mathbf{x}^d = [0.1 \ 0.2 \ 0.05 \ 0.25 \ 0.15 \ 0.25]^T$. For both feedback controllers, the numerical solution of the mean-field model (5) was compared to stochastic simulations of the CTMC characterized by expression (1). This CTMC was simulated using an approximating DTMC that evolves in discrete time. The probability that an agent *i* in state (vertex) $S(e), e \in \mathcal{E}$, at time t transitions to state T(e) at time $t + \Delta t$ was set to:

$$\mathbb{P}(X_i(t+\Delta t) = T(e)|X_i(t) = S(e)) = k_e\left(\frac{1}{N}\mathbf{N}^p(t)\right)\Delta t.$$

Here, $\{k_e\}_{e \in \mathcal{E}}$ is the set of feedback laws and $\mathbf{N}^p(t) = [N_1^p(t) \ N_2^p(t) \ \dots \ N_M^p(t)]^T$, where $N_v^p(t)$ is the number of agents in state $v \in \mathcal{V}$ at time t. We assume that each agent can measure the agent populations in its current state and in adjacent states.

In Figs. 2 and 3, we compare simulations of the closedloop system (5) with the feedback controllers to simulations of the open-loop system (2). The controller for the openloop system was constructed by setting the right-hand side of system (2) equal to $\mathbf{Gx} = -\mathbf{L}(\mathcal{G})\mathbf{Dx}$, where $\mathbf{L}(\mathcal{G})$ is the Laplacian matrix of the graph \mathcal{G} and \mathbf{D} is a diagonal matrix with entries $D^{ij} = 1/x_i^d$ if i = j, $D^{ij} = 0$ otherwise. This makes the desired distribution \mathbf{x}^d invariant for the corresponding CTMC. The transition rates (control inputs) for this controller were defined as $u_e(t) = G^{T(e)S(e)}$ for all $t \in [0,\infty)$, $e \in \mathcal{E}$. Fig. 2 shows that the open-loop controller produces large variances in the agent populations at steady-state. As an expected consequence of the law of large numbers, these variances are smaller for N = 1200agents than for N = 80 agents. In comparison, the variances are much smaller when the feedback controllers are used, as shown in Fig. 3. This is due to the property of the feedback controllers that as the agent densities approach their desired equilibrium values, the transition rates tend to



Fig. 1: Six-vertex bidirected graph.

zero. This property reduces the number of unnecessary agent state transitions at equilibrium. Using open-loop control, the agents' states keep switching and never reach steady-state values. In contrast, using the feedback controllers, the agents' states remain constant after a certain time.

As discussed in Section III, the underlying assumption of using the mean-field models (2) and (5) is that the swarm behaves like a continuum. That is, the ODEs (2) and (5) are valid as number of agents $N \rightarrow \infty$. Hence, it is imperative to check the performance of the feedback controllers for different agent populations. We observe in Figs. 3b and 3d that in the case of N = 1200 agents, the stochastic simulation follows the mean-field model solution quite closely for both feedback controllers. In addition, in all simulations, the numbers of agents in each state remain constant after some time; in the case of N = 80 agents, the fluctuations stop earlier than in the case of N = 1200 agents.

Lastly, we observe that the nonlinear controller produces a significantly slower convergence rate to equilibrium than the linear controller. Figs. 3c and 3d show that the mean-field model with the nonlinear controller converges asymptotically to the desired equilibrium, and hence still exhibits a small discrepancy from the stochastic simulations at time t = 500 s. This discrepancy is absent when the linear controller is used. We could increase the convergence rate for the nonlinear controller by encoding it as a constraint in our algorithmic procedure, described in [12], for constructing the linear and nonlinear feedback controllers.

VI. CONCLUSION

In this paper, we have presented several fundamental results on controllability properties of forward equations of CTMCs that are associated with strongly connected graphs. We showed asymptotic controllability of distributions that are not strictly positive, with target densities equal to zero for some states. In addition, we constructed decentralized, density-dependent rational and polynomial feedback laws that stabilize the corresponding mean-field model, with control inputs that equal zero at equilibrium. In future work, we plan to investigate the design of polynomial feedback laws that improve the stability of the closed-loop system beyond asymptotic stability and optimize the convergence rate to equilibrium. Another direction of future work is to characterize the effect of noise in estimates of the agent densities on the convergence properties of the proposed control laws.

REFERENCES

- Behçet Açıkmeşe and David S Bayard. A Markov chain approach to probabilistic swarm guidance. In *American Control Conference* (ACC), pages 6300–6307. IEEE, 2012.
- [2] Saptarshi Bandyopadhyay, Soon-Jo Chung, and Fred Y Hadaegh. Inhomogeneous Markov chain approach to probabilistic swarm guidance algorithm. In 5th Int'l. Conf. on Spacecraft Formation Flying Missions and Technologies, Munich, Germany, 2013.
- [3] Saptarshi Bandyopadhyay, Soon-Jo Chung, and Fred Y Hadaegh. A probabilistic Eulerian approach for motion planning of a large-scale swarm of robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3822–3829, 2016.
- [4] Alain Bensoussan, Jens Frehse, Phillip Yam, et al. Mean Field Games and Mean Field Type Control Theory, volume 101. Springer, 2013.
- [5] Spring Berman, Adám Halász, M. Ani Hsieh, and Vijay Kumar. Optimized stochastic policies for task allocation in swarms of robots. *IEEE Transactions on Robotics*, 25(4):927–937, 2009.
- [6] Shiba Biswal, Karthik Elamvazhuthi, and Spring Berman. Meanfield controllability and decentralized stabilization of Markov chains, part II: Asymptotic controllability and polynomial feedbacks. arXiv preprint arXiv:1703.08515, 2017.
- [7] Anthony M Bloch. Nonholonomic Mechanics and Control, volume 24. Springer, 2015.
- [8] Stephen Boyd, Persi Diaconis, and Lin Xiao. Fastest mixing Markov chain on a graph. SIAM Review, 46(4):667–689, 2004.
- [9] Airlie Chapman. Advection on graphs. In Semi-Autonomous Networks, pages 3–16. Springer, 2015.
- [10] Francis H Clarke, Yuri S Ledyaev, Eduardo D Sontag, and Andrei I Subbotin. Asymptotic controllability implies feedback stabilization. *IEEE Transactions on Automatic Control*, 42(10):1394–1407, 1997.
- [11] Nazlı Demir, Utku Eren, and Behçet Açıkmeşe. Decentralized probabilistic density control of autonomous swarms with safety constraints. *Autonomous Robots*, 39(4):537–554, 2015.
- [12] Vaibhav Deshmukh, Karthik Elamvazhuthi, Shiba Biswal, Zahi Kakish, and Spring Berman. Mean-field stabilization of Markov chain models for robotic swarms: Computational approaches and experimental results. Submitted to *IEEE Robotics and Automation Letters* (*RA-L*), 2017.
- [13] Karthik Elamvazhuthi, Shiba Biswal, Vaibhav Deshmukh, and Spring Berman. Mean-field controllability and decentralized stabilization of Markov chains: Elaboration on theoretical results. In preparation, 2017.
- [14] Karthik Elamvazhuthi, Vaibhav Deshmukh, Matthias Kawski, and Spring Berman. Mean-field controllability and decentralized stabilization of Markov chains, part I: Global controllability and rational feedbacks. arXiv preprint arXiv:1703.08243, 2017.
- [15] Karthik Elamvazhuthi and Piyush Grover. Optimal transport over nonlinear systems via infinitesimal generators on graphs. arXiv preprint arXiv:1612.01193, 2016.
- [16] Stewart N Ethier and Thomas G Kurtz. Markov Processes: Characterization and Convergence, volume 282. John Wiley & Sons, 2009.
- [17] Diogo A. Gomes, Joana Mohr, and Rafael Rigão Souza. Continuous time finite state mean field games. *Applied Mathematics & Optimization*, 68(1):99–143, 2013.
- [18] Vassili N. Kolokoltsov. *Nonlinear Markov Processes and Kinetic Equations*, volume 182. Cambridge University Press, 2010.
- [19] Alcherio Martinoli, Kjerstin Easton, and William Agassounon. Modeling swarm robotic systems: A case study in collaborative distributed manipulation. *The International Journal of Robotics Research*, 23(4-5):415–436, 2004.
- [20] T. William Mather and M. Ani Hsieh. Synthesis and analysis of distributed ensemble control strategies for allocation to multiple tasks. *Robotica*, 32(02):177–192, 2014.
- [21] Martin L Puterman. Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley & Sons, 2014.
- [22] Justin Solomon, Raif Rustamov, Leonidas Guibas, and Adrian Butscher. Continuous-flow graph transportation distances. arXiv preprint arXiv:1603.06927, 2016.



Fig. 2: Trajectories of the mean-field model (thick lines) and the corresponding stochastic simulations (thin lines).



Fig. 3: Trajectories of the mean-field model (*thick lines*) and the corresponding stochastic simulations (*thin lines*).