Mean-Field Models in Swarm Robotics: A Survey

Karthik Elamvazhuthi and Spring Berman

School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ, 85287

E-mail: {karthikevaz, spring.berman}@asu.edu

Abstract. We present a survey on the application of fluid approximations, in the form of mean-field models, to the design of control strategies in swarm robotics. Mean-field models that consist of ordinary differential equations, partial differential equations, and difference equations have been used in the swarm robotics literature, depending on whether the state of each agent and the time variable take values from a discrete or continuous set. These *macroscopic* models are independent of the number of agents in the swarm, and hence can be used to synthesize robot control strategies in a scalable manner, in contrast to individual-based *microscopic* models of swarms that represent finite numbers of discrete agents. Moreover, mean-field models are amenable to rigorous investigation using tools from dynamical systems theory, control theory, stochastic processes, and analysis of partial differential equations, enabling new insights and provable guarantees on the dynamics of collective behaviors. In this paper, we survey the applications of these models to problems in swarm robotics that include coverage, task allocation, self-assembly, consensus, and environmental mapping.

1. Introduction

There has been a significant amount of work on swarm robotic systems over the last two decades. A major challenge is to develop modeling and control techniques for these large-scale multi-robot systems that are scalable with the swarm population size [1]. One approach to address this issue, inspired by modeling methodologies used in the natural sciences such as fluid dynamics [2], statistical mechanics [3], and biology [4, 5], is to treat the swarm as a continuum. The starting point of this approach is the *Kolmogorov forward equation* of a stochastic process, which describes the spatio-temporal evolution of the probability density associated with the process. For a finite number of agents that are each modeled using such a stochastic process, the state space of the forward equation, a linear dynamical system, is dependent on the number of agents N. On the other hand, in the limit as the number of agents tends to infinity, one can approximate the N-agent linear forward equation with a single, possibly nonlinear, forward equation with parameters that can be functions of the probability density. The resulting equation, known as the *mean-field model*, is defined on the set of probability densities that determine the probability of an agent being in a given state at a specific time. When the number of agents in the swarm is large, this approximation is valid if all agents follow the same control laws (i.e., the swarm is homogeneous) and the control laws of each agent are not dependent on other agents' identities, but only on the agent's own state or the local density of the swarm. This *identity-invariance* of the control laws implies that the dimension of the state space of the mean-field model depends on the dimension of the state space of a single agent, and hence is independent of the actual number of agents in the swarm. Therefore, the scalability of any controller design methodology that is based on mean-field models is dependent on the number of admissible states of a single agent, rather than on the total number of agents in the swarm.

In this paper, we survey the application of mean-field models to different problems in swarm robotics such as coverage, task allocation, self-assembly, consensus, and mapping. Many of these problems can be framed as problems of feedback stabilization or parameter identification for the corresponding mean-field model. Feedback is known to play an important role in natural and synthetic collective systems that demonstrate *self-organizing behavior* [6]. When studying self-organizing systems, a natural scientist is interested in understanding the feedback mechanisms that result in the observed self-organizing behavior, such as a particular division of labor or a consensus on a new nest site. On the other hand, a swarm roboticist is interested in determining the feedback mechanisms that should be programmed into a swarm of robots in order to achieve a target self-organizing behavior, such as a desired distribution of robots among tasks or physical sites. The formal design and analysis of feedback laws is a major focus of the field of control theory. This field studies solutions to *control problems*, which entail the selection of suitable values for tunable parameters in a given system that cause the system to exhibit a target dynamical behavior.

Many existing works on robotic swarms approach the task of programming a swarm to achieve desired self-organizing behaviors as a control problem, formulated in terms of a mean-field description of the swarm. In this survey, we classify these works according to the type of mean-field model that they use within a control-theoretic framework. In Section 2, we describe swarm control problems that use finite-dimensional mean-field models in the form of ordinary differential equations and difference equations, in which case each agent has a finite number of states and the time variable is continuous or discrete. In Section 3, we discuss much more challenging swarm control problems in terms of infinite-dimensional mean-field models in the form of partial differential equations, for which the agents' state space is continuous and the time variable is continuous. Due to the limited amount of work on swarms in which the agents have a continuous state space and time is a discrete variable, we only briefly introduce the mean-field model for this case in Section 4, where we also describe some potential future research directions. Throughout the paper, we organize the works in terms of the complexity of their control problem, as determined by the amount of information that is available to each robot in the swarm to execute its controller. Types of information that each robot could possibly access include the robot's own state or the state distribution of other robots within its local sensing range.

While there have been several surveys on swarm robotics [1, 7–9], multi-robot systems [10–12] and the broader field of multi-agent systems [13], and networked computational devices [14], the contribution of this paper is to provide a review of works that specifically use mean-field models to predict and control collective behaviors in robotic swarms. We note that the use of mean-field models in robotic swarm control has been previously discussed in the literature under different terminology, including macroscopic models [15], Rate Equation models [16], and probabilistic swarm guidance [17]. To our knowledge, the 2004 paper [18] is the earliest survey of works on the application of mean-field models (referred to as macroscopic models) to the analysis and control of swarm robotic collective behaviors. The more recent works [1, 8, 9] devote part of their surveys to reviewing some prior works on mean-field models of robotic swarms. Our use of the term *mean-field model* connects the swarm control problems that we discuss in this paper to the broader literature on mean-field control theory [19], which has been a subject of intensive research within the engineering and applied mathematics communities in recent years.

2. Finite-Dimensional Mean-Field Models

In this section, we introduce finite-dimensional mean-field models in which the time variable is continuous (Section 2.1) or discrete (Section 2.2).

2.1. Continuous-time models

There are N autonomous agents whose states evolve in continuous time according to a Markov chain with a finite state space defined as the vertex set $\mathcal{V} = \{1, ..., M\}$. For example, the vertices in $\mathcal{V}$ can represent a set of tasks that the agents must perform, or a set of spatial locations obtained by partitioning the agents' environment. The edge set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ defines the pairs of vertices between which the agents can transition. The directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is assumed to be strongly connected. The agents' transition rules are determined by the control parameters $u_e : [0, \infty) \to \mathbb{R}_{\geq 0}$ for each $e \in \mathcal{E}$, and are known as the *transition rates* of the associated continuous-time Markov chain (CTMC). The state of each agent $i \in \{1, ..., N\}$ at time t 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 $e = (S(e), T(e)) \in \mathcal{E}$, where $S(e)$ and $T(e)$ denote the source and target vertices of the edge e, respectively. Here, $\mathbb{P}$ is the underlying probability measure induced on the space of events Ω by the stochastic processes $\{X_i(t)\}_{i=1}^N$. Informally, $u_e(t)h$ is the probability of an agent jumping from state i to

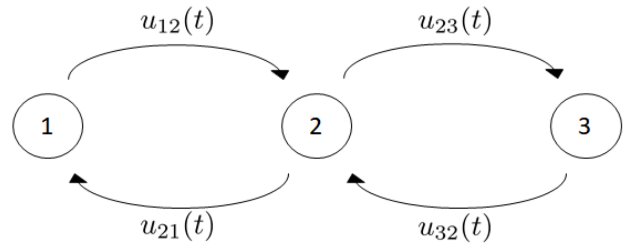

Figure 1. Bidirected graph with 3 vertices, representing agent states.

state j within an infinitesimally small time interval h. The little-o notation $o(h)$ indicates that as h tends to 0, the term $\frac{o(h)}{h}$ converges to 0. Let $\mathcal{P}(\mathcal{V}) = \{\mathbf{y} \in \mathbb{R}^M_{\geq 0};\ \sum_v y_v = 1\}$ be the set of probability densities on $\mathcal{V}$, and let int $\mathcal{P}(\mathcal{V}) := \{\mathbf{y} \in \mathcal{P}(\mathcal{V});\ y_v > 0$ for all $v \in \mathcal{V}\}$ be the interior of the set $\mathcal{P}(\mathcal{V})$, that is, the set of probability densities $\mathbf{y}$ for which all entries y_v are positive. Corresponding to the CTMC is a system of ordinary differential equations (ODEs) that determines the time evolution of the probability densities $\mathbb{P}(X_i(t) = v) = x_v(t) \in \mathbb{R}_{\geq 0}$. If $X_i(0)$ are independent and identically distributed (IID), then the processes $\{X_i(t)\}_{i=1}^N$ are also IID, and the *Kolmogorov forward equation* can be represented by a single linear system of ODEs,

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

where $\mathbf{x}^0$ represents the initial distribution of the random variables $X_i(0)$ and $\mathbf{B}_e \in \mathbb{R}^{M \times M}$ are control matrices whose entries at row i and column j are given by

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

For example, consider a 3-state Markov chain, for which the corresponding graph $\mathcal{G}$ is illustrated in Fig. 1. The system of ODEs (2) in this case is given by:

$$\begin{aligned} \dot{x}_1(t) &= -u_{12}(t)x_1(t) + u_{21}(t)x_2(t) \qquad (3) \\ \dot{x}_2(t) &= -(u_{21}(t) + u_{23}(t))x_2(t) + u_{12}(t)x_1(t) \\ &\quad + u_{32}(t)x_3(t) \\ \dot{x}_3(t) &= -u_{32}(t)x_3(t) + u_{23}(t)x_2(t) \\ x_1(0) &= x_1^0, \quad x_2(0) = x_2^0, \quad x_3(0) = x_3^0. \end{aligned}$$

From this perspective, it can be seen that the structure of the matrix $\mathbf{B}_e$ enforces the conservation of agents in the swarm: for each edge $e \in \mathcal{E}$, the rate of decrease $u_e(t)x_{S(e)}(t)$ in the population fraction $x_{S(e)}(t)$ of agents at vertex $S(e)$ equals the rate of increase in the population fraction $x_{T(e)}(t)$ of agents at the adjacent vertex $T(e)$. Note that the control parameter $u_e(t)$ determines this rate of change in agent population at the source and target vertices of edge e.

Let $\chi_v : \mathcal{V} \to \{0,1\}$ represent the indicator function of the vertex v. As $N \to \infty$, the population fraction of agents at a vertex v, given by $\frac{1}{N}\sum_{i=1}^{N} \chi_v(X_i(t))$, converges to $x_v(t)$ for each $t \in [0,\infty)$. This follows from the law of large numbers due to the random variables $X_i(t)$ being IID. Thus, by modeling a swarm using the approximation $N \to \infty$, swarm control problems can be posed in terms of the deterministic quantity $\mathbf{x}(t)$ rather than the random variables X_i, enabling control of the *mean-field* behavior of the swarm. Therefore, control or estimation problems where the objectives are functions of the population fractions $\frac{1}{N}\sum_{i=1}^{N} \chi_v(X_i(t))$ can be replaced by problems where the objectives are functions of the probability distribution or *population density* $\mathbf{x}(t)$. An instance of this *mean-field control problem* [19] is when the goal is to design the control inputs $u_e(t)$ such that $\mathbf{x}(T) = \mathbf{x}^d$ for a target distribution $\mathbf{x}^d \in \mathcal{P}(\mathcal{V})$ and time $T > 0$. Another example of this type of control problem is the *mean-field stabilization problem*, where the goal is to design non-negative, possibly time-varying parameters k_e such that $u_e(t) = k_e$ for all $t \geq 0$ and a given $\mathbf{x}^d \in \mathcal{P}(\mathcal{V})$ is an asymptotically stable equilibrium point of system (2). When the control inputs $u_e(t)$ are independent of time and the population density $\mathbf{x}(t)$, we will say that they are in *state-feedback form*. Here, the term *state-feedback* refers to the fact that agent i requires only knowledge of its current state $X_i(t)$ to execute the control action, and not the mean-field term $\mathbf{x}(t)$.

The regulation of division of labor in biological swarms such as social insect colonies [20], as well as problems of task allocation [16] and spatial coverage [21] in swarm robotics, are all instances of the mean-field stabilization problem. Social insects are hypothesized to maintain a division of labor in their colonies through cooperation [20] and heterogeneity of colony members [22, 23]. However, in robotic swarms, it can be challenging to control multiple robots to perform cooperative tasks and to coordinate heterogeneous robots with different capabilities. Therefore, it is of interest to determine whether the mean-field stabilization problem can be solved by a swarm of non-interacting homogeneous agents that are each cognizant only of its own current state. Phrased another way, is mean-field stabilization possible using state-feedback laws? This section surveys some works that have addressed this question.

The following result is fundamental in analyzing the long-time behavior of Markov chains and has therefore also played an important role in the solvability of the mean-field control and stabilization problems. It follows from the *Perron-Frobenius* theorem [24] and plays an important role in the stabilization of the mean-field model (2) using time-independent state-feedback laws.

Theorem 2.1 *Consider the mean-field model* (2)*, for which the corresponding graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ *is strongly connected. Suppose that* $u_e(t) = k_e$ *is a (time-independent) state-feedback law and is positive for each* $e \in \mathcal{E}$*. Then* 0 *is an eigenvalue of the matrix* $\sum_{e\in\mathcal{E}} k_e \mathbf{B}_e$*, and it has the largest real part of all the eigenvalues of this matrix. Moreover, this eigenvalue is simple. Hence, the solution* $\mathbf{x}(t)$ *of system* (2) *exponentially converges to a unique limit* $\mathbf{x}^\infty \in \operatorname{int} \mathcal{P}(\mathcal{V})$*, which is a vector with all elements positive.*

In the above theorem, *exponential convergence* refers to the property that the distance between the solution $\mathbf{x}(t)$ and the limiting distribution $\mathbf{x}^\infty$ is bounded by $Me^{-\lambda t}$ for some constants $M, \lambda > 0$ that are independent of the initial distribution $\mathbf{x}^0$. Using this theorem, the problem of designing state-feedback laws with the goal of achieving exponential stabilization with maximal decay rate λ is considered in [25] for a multi-robot stochastic task allocation scenario. It was shown that this problem can be framed as a convex optimization problem. A drawback of using state-feedback laws is that the control inputs $u_e(t)$ remain non-zero at equilibrium, and hence agents might continue switching between states at equilibrium; i.e., **the system being in** *macroscopic equilibrium* **does not imply that it is in** *microscopic equilibrium*. To reduce the frequency of switching at equilibrium, [26] introduced biologically-inspired control laws that are functions of the population density $\mathbf{x}(t)$. We will refer to such control laws as *mean-field feedback laws*. In particular, a mean-field feedback law is a family of functions $k_e : \mathcal{P}(\mathcal{V}) \to [0,\infty)$ such that the control inputs are defined as $u_e(t) = k_e(\mathbf{x}(t))$ for all $t \geq 0$ and all $e \in \mathcal{E}$. Mean-field feedback laws can describe the mechanism of *quorum sensing* in biological swarms such as bacterial colonies [27] and house-hunting ants [28], whereby individuals' assessment of population density triggers a change in their behavior and gives rise to coordinated collective phenomena. In [26], the following mean-field feedback law k_e, referred to as "ensemble feedback," is considered:

$$k_e(\mathbf{x}) = k_e^* + \sigma_{S(e)}(x_{S(e)}, q_{S(e)})(\alpha - 1)k_e^*, \qquad (4)$$

where k_e^* is a baseline transition rate, $q_{S(e)}$ is a predefined quorum fraction of the desired agent population at vertex $S(e)$, $\sigma_{S(e)} = (1 + \exp[\gamma(q_{S(e)} - \frac{x_{S(e)}}{x^d_{S(e)}})])^{-1} \in [0,1]$ is an analytic switching function, and γ and α are suitably chosen constants. If the quorum at vertex $S(e)$ is exceeded, then the control law (4) increases the transition rate to adjacent vertex $T(e)$ up to αk_e^*;

i.e., overpopulation at $S(e)$ triggers agents to leave the vertex more quickly. It was shown in [26] that for $\mathbf{x}^d \in \text{int } \mathcal{P}(\mathcal{V})$, i.e. the set of probability distributions that are positive everywhere on $\mathcal{V}$, the solutions of system (2) with mean-field feedback law (4) converge to $\mathbf{x}^d$ as $t \to \infty$.

Note that, when the control inputs $u_e(t)$ are functions of the agent population fractions, which converge to the mean-field distribution $\mathbf{x}$ in the limit $N \to \infty$, the random variables X_i are not IID. Therefore, the validity of the limit $\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \chi_v(X_i(t)) = x_v(t)$ does not follow from the law of large numbers. Instead, one can apply the *dynamic law of large numbers*, which is proved in [29].

Theorem 2.2 ***(Mean-field/Fluidic Limit) [29]*** *Suppose that the transition rates $u_e(t)$ of each agent are given by*

$$u_e(t) = v_e\left(\frac{1}{N}\sum_{i=1}^{N}\chi_1(X_i(t)), ..., \frac{1}{N}\sum_{i=1}^{N}\chi_M(X_i(t))\right), \tag{5}$$

where $v_e : \mathcal{P}(\mathcal{V}) \to [0,\infty)$ is a differentiable function for each $e \in \mathcal{E}$. Consider the solution $\mathbf{x}(t)$ of the following system of ordinary differential equations,

$$\dot{\mathbf{x}}(t) = \sum_{e\in\mathcal{E}} v_e(x_1, ..., x_M)\mathbf{B}_e\mathbf{x}(t), \quad t \in [0,\infty), \tag{6}$$

$$\mathbf{x}(0) = \mathbf{x}^0 \in \mathcal{P}(\mathcal{V}).$$

Then for every $t \geq 0$,

$$\lim_{N\to\infty} \sup_{s\geq t} |\mathbf{Y}_N(s) - \mathbf{x}(s)| = 0 \quad \textit{almost surely} \tag{7}$$

where for each $s \geq 0$, the random variable $\mathbf{Y}_N(s)$ is given by

$$\mathbf{Y}_N(s) = \left[\frac{1}{N}\sum_{i=1}^{N}\chi_1(X_i(t)) \; ... \; \frac{1}{N}\sum_{i=1}^{N}\chi_M(X_i(t))\right]^T$$

and for each $\mathbf{y} \in \mathbb{R}^M$, $|\mathbf{y}| := \sum_{i=1}^{M}|y_i|$.

There has been an extensive amount of work on generalizing the above result to cases where the functions v_e are possibly discontinuous [30, 31] or where the mean-field model is a hybrid system with continuous as well as discrete states [32].

In their most general form, mean-field feedback laws require that agents can measure the population density vector $\mathbf{x}(t)$. Given the typical sensing and communication constraints on agents in a swarm, it is desirable that the mean-field feedback laws are *local*; that is, the control inputs u_e are functions of the population density only at the source vertex $S(e)$ (i.e., the agent's current state), the target vertex $T(e)$, or both. The problem of reducing agent fluctuations at equilibrium is framed as a variance control problem in [33], using local mean-field feedback laws of the form $u_e(\mathbf{x}) = \alpha_e + \beta_e\frac{x_{S(e)}}{x_{T(e)}}$ for suitable choices of the parameters α_e and β_e. The works [26, 33] emphasize the fact that while state-feedback laws are sufficient to solve the mean-field stabilization problem, mean-field feedback laws can produce improved performance in terms of reducing the variance of the swarm distribution about the target equilibrium distribution.

Before one proceeds to design control laws, it is important to know which distributions are stabilizable. The works [25, 26, 33] require the assumption that $\mathbf{x}^d \in \text{int } \mathcal{P}(\mathcal{V})$. When $\mathcal{G}$ is bidirected, it follows by construction from [34] that, if $\mathbf{x}^d \in \text{int } \mathcal{P}(\mathcal{V})$, then there exists a state-feedback law that asymptotically stabilizes $\mathbf{x}^d$. From Theorem 2.1, it can be seen that the assumption that $\mathcal{G}$ is bidirected can be relaxed in order for the stabilization result to still hold. Suppose that $\mathcal{G}$ is strongly connected, the parameters k_e are positive, and $\mathbf{x}^\infty$ is the unique (up to a scaling factor) eigenvector of the matrix $\sum_{e\in\mathcal{E}} k_e\mathbf{B}_e$ corresponding to 0. Then for the state-feedback law $\tilde{k}_e = k_e\frac{x^\infty_{S(e)}}{x^d_{S(e)}}$, we have that $\mathbf{x}^d$ is the unique eigenvector of the matrix $\sum_{e\in\mathcal{E}} \tilde{k}_e\mathbf{B}_e = \sum_{e\in\mathcal{E}} k_e\mathbf{B}_e\mathbf{D}$, where $\mathbf{D}$ is the diagonal matrix $diag(\frac{x_1^\infty}{x_1^d}, \frac{x_2^\infty}{x_2^d}, ..., \frac{x_M^\infty}{x_M^d})$. Thus, $\mathbf{x}^d$ is the globally asymptotically stable equilibrium point of system (2). In order to relax the assumption that $\mathbf{x}^d \in \text{int } \mathcal{P}(\mathcal{V})$, [35, 36] considered the problem of which elements of $\mathcal{P}(\mathcal{V})$ are reachable asymptotically or in finite time, and can be stabilized using state-feedback laws and local mean-field feedback laws. In particular, it was shown that any element of int $\mathcal{P}(\mathcal{V})$ can be reached in finite time, and any element of $\mathcal{P}(\mathcal{V})$ can be reached asymptotically in infinite time, using time-varying feedback laws. It was additionally shown that probability distributions with *strongly connected supports* could be stabilized using state-feedback laws and local mean-field feedback laws. We say that a probability distribution has a *strongly connected support* if the sub-graph induced by the vertices on which the probability distribution is positive is strongly connected. For example, for the graph shown in Fig. 1, the probability distribution $\mathbf{x} = [0.5\ 0.5\ 0]^T$ has a strongly connected support, whereas the probability distribution $\mathbf{x} = [0.5\ 0\ 0.5]^T$ has a support that is not strongly connected.

The assumption that the probability distribution must have a strongly connected support is relaxed in [37]. State-feedback laws cannot be used to stabilize arbitrary probability distributions, since the graph associated with the corresponding CTMC is disconnected, resulting in the mean-field model having multiple equilibrium distributions. In [37], it is proved that *mean-field feedback laws* are needed to stabilize arbitrary probability distributions. In particular, it is shown that if $\mathbf{x}^d \in \mathcal{P}(\mathcal{V})$ is a target equilibrium distribution, then by defining the control inputs $u_e(t)$ as

$$u_e(t) = f_e(x_{S(e)}(t)), \tag{8}$$

where f_e is any non-decreasing differentiable function such that $f_e(y) = 0$ if and only if $y = x^d_{S(e)}$, the system (2) is asymptotically stable about the equilibrium $\mathbf{x}^d$. This control law is based on the principle, "Leave the current vertex with a positive probability if the agent population at the vertex is above the target population." This control law has the important property that agents do not switch between vertices at equilibrium.

A method for computing optimal time-varying state-feedback laws in order to achieve a target distribution in finite time is shown in the work [38] on computational optimal transport. For certain cost functions, this optimal control problem can be treated in a convex optimization framework. For example, for a given $T > 0$ and $\mathbf{x}^d \in \mathcal{P}(\mathcal{V})$, consider the following optimization problem:

$$\inf_{u_e(t)\geq 0, x_v \geq 0} \sum_{e\in\mathcal{E}} \int_0^T u_e^2(t) x_{S(e)}(t) dt \tag{9}$$

subject to the bilinear constraints defined by system (2), with

$$\mathbf{x}(T) = \mathbf{x}^d. \tag{10}$$

This optimization problem is non-convex. However, it can be transformed into the following equivalent convex optimization problem:

$$\inf_{r_e(t)\geq 0, x_v(t) \geq 0} \sum_{e\in\mathcal{E}} \int_0^T \frac{r_e^2(t)}{x_{S(e)}(t)} dt \tag{11}$$

subject to the linear constraints

$$\dot{\mathbf{x}}(t) = \sum_{e\in\mathcal{E}} r_e(t)\mathbf{B}_e\mathbf{1}, \quad t \in [0, \infty), \tag{12}$$

$$\mathbf{x}(0) = \mathbf{x}^0, \quad \mathbf{x}(T) = \mathbf{x}^d,$$

where $\mathbf{1} \in \mathbb{R}^M$ is the vector with all elements equal to 1. This approach of convexifying optimization problems with objective functions such as the one in (9) and constraints (2), (10) was introduced in [38] in order to adapt the fluid-dynamic version of the optimal transport problem [39], where the state space is continuous, to the case of discrete state spaces. See Section 3 for more details. We note that the cost function in (9) has a simpler structure than the one considered in [38].

Numerical construction of mean-field feedback laws is a much more computationally challenging task, in comparison with the synthesis of state-feedback laws. Computational approaches based on Linear Matrix Inequalities [40] and Sum-of-squares methods [41] are used to numerically construct decentralized mean-field feedback laws in [42]. Execution of mean-field feedback strategies requires knowledge of the distribution of robots in each state. One approach to estimate the robot distribution is to use a centralized observer, such as an overhead camera [42]. An alternative approach, which does not rely on a centralized authority to observe the swarm, is to use encounter rates between agents to estimate population densities, as observed in natural swarms such as ant colonies [28, 43]. A model for estimating population densities of swarms as a function of inter-agent encounter rates is proposed and experimentally validated in [44]. This encounter rate model was used in [45] to implement mean-field feedback laws for a swarm task allocation problem.

The work [46] considers the effect of heterogeneity in the robot populations on the optimal robot control policies. In this work, $\mathcal{V}$ denotes not only the states that robots can occupy, but also the types of different robots. The problem of identifying the minimum number of robots of each type in order to achieve a given goal is framed as an optimization problem. There have also been efforts to develop mean-field models of heterogeneous collectives comprised of both autonomous robots and living organisms [47, 48]. For instance, the work [49] experimentally validates a mean-field model of a mixed group of robots and cockroaches that influence each other during a collective decision-making process to choose a common shelter.

In some scenarios, it is useful to consider mean-field models where different types of agents or agents in different states interact at particular probability rates and then physically bond or change their states. Such models are commonly used to describe the dynamics of chemical reaction networks (CRNs) [50, 51], and have been adopted in several works in swarm robotics. A CRN model of a swarm represents agents of different types or in different states as distinct *species* that are analogous to chemical species. A *reaction* occurs when a combination of *reactant* species converts into a combination of *product* species at a certain *reaction rate constant*. Suppose that a reaction r in a CRN has reactants $a_i \in \mathbb{R}_{>0}$, $i = 1, ..., n$, that combine with probability $k_r(t)\Delta T$ in an infinitesimally small amount of time ΔT to form products $b_j \in \mathbb{R}_{>0}$, $j = 1, ..., m$. Here, $k_r(t)$ is the reaction rate constant. We denote this reaction by $r = [(a_1, ..., a_n), (b_1, ..., b_m)]$. Let M be the total number of reactant and product species in the entire CRN; then the vector of agent population densities in each species is given by $\mathbf{x} \in \mathbb{R}^M$. Define a vector field $f_r : \mathbb{R}^M \to \mathbb{R}^M$ associated with reaction r that has entries $(f_r(\mathbf{x}))_{a_i} = -\prod_{i=1}^n x_{a_i}$ for $i \in \{1, ..., n\}$, $(f_r(\mathbf{x}))_{b_j} = \prod_{i=1}^n x_{a_i}$ for $j \in \{1, ..., m\}$, and 0 otherwise. Then the resulting mean-field model can be written as follows, where $\mathcal{R}$ is the set of all reactions in the CRN:

$$\dot{\mathbf{x}}(t) = \sum_{r\in\mathcal{R}} k_r(t) f_r(\mathbf{x}(t)), \quad t \in [0,\infty), \tag{13}$$
$$\mathbf{x}(0) = \mathbf{x}^0 \in \mathbb{R}^M_{\geq 0},$$

The system of equations (13) simplifies to the form of system (2) when only *unimolecular reactions* are admissible; i.e, all reactions in the CRN are of the form $r = [a, b]$, where $a, b \in \mathcal{V}$.

The first application of this type of mean-field model to simulating the behavior of a robotic swarm was in [52], which introduced a CRN model for a stick-pulling experiment performed by a swarm of robots that do not explicitly communicate or coordinate with one another. Using the mean-field model, the authors identify optimal state-dependent control parameters to improve the system's performance. In [18], the authors study the application of these types of models to a number of tasks performed by a swarm of robots, including collaborative pulling, foraging, and aggregation. In [53], the authors use a mean-field model to study the effect of spatial interference on the performance of robots in a collective foraging task. A mean-field model based on a CRN is used in [54] for a task in which a swarm of robots must assemble a collection of parts into target amounts of final products using stochastic control policies determined by the reaction rate constants. The authors optimize the reaction rate constants to improve the system's rate of convergence to the target numbers of products. In [55], the authors use a CRN-based mean-field model to design stochastic robot attachment-detachment policies that drive a swarm to specified spatial distributions around multiple payloads for a collective transport task. A CRN is used to model a stochastic self-assembly task in [56], and methods are developed to estimate the reaction rates in the CRN model using high-fidelity physics-based simulations. In [57], the authors present an optimization-based method to maximize the yield of a stochastic self-assembly process by finding the optimal reaction rates, and validate the method using a CRN model. These authors also introduce an integral feedback controller in [58] to stabilize a CRN model of another stochastic self-assembly process. In [59], the authors develop a CRN model of a scenario in which robots collaboratively screen an environment for undesirable agents, and use this model to find the optimal parameters to achieve the goal.

CRN models have also been used extensively to model collective decision-making problems in swarm robotics, where a group of robots must collectively decide among a number of available options using limited information and interactions. Collective decision-making leads to stabilization problems that differ from classical formulations: the target probability distribution to which the agents should stabilize is not predefined by a centralized authority, and this distribution is a non-local function of the states or the agent populations in the states, while the robot control laws are constrained to be local. In [60], the authors consider a modified form of the majority rule opinion dynamics, studied in the literature on opinion dynamics [61], for a scenario where a swarm of robots must decide between two different actions with different execution times, but without any prior knowledge of the execution times. Similarly, CRN models that have been used to describe honeybee nest site selection strategies [62] have found applications in swarm robotics [63]. In [64], the authors add stochastic disturbances to a mean-field model of collective decision making, introduced in [65], and compare simulations of the model to data obtained from an experiment in which a swarm of robots collectively decide between two spatially distributed features. This work showed how spatial effects due to inter-agent collisions and inhomogeneity in the distribution of the swarm can reduce the predictive power of the mean-field model. In [66], CRN models of honeybee nest site selection are used to design unmanned aerial vehicle control policies for non-uniform spatial coverage. In this work, the states represent spatial sites as well as tasks. See [67, 68] for extensive surveys on the topic of collective decision-making problems in swarm robotics with some applications of mean-field models.

Other recent work that uses a CRN-based mean-field framework for swarm applications considers the problem of keeping individual robot types private [69]. A *privacy model* that uses notions from differential privacy is developed to understand the privacy preservation capabilities of the swarm as a function of the reaction parameters.

2.2. Discrete-time models

In discrete-time mean-field models, the state of each agent $i \in \{1, ..., N\}$ is defined by a discrete-time Markov chain (DTMC) $X_i(n)$, $n \in \mathbb{Z}_+$, that evolves on the state space $\mathcal{V}$ according to the conditional probabilities

$$\mathbb{P}(X_i(n+1) = T(e) | X_i(n) = S(e)) = u_e(n) \tag{14}$$

with control parameters $u_e(n) \in [0, 1]$ that satisfy the constraint

$$\sum_{e\in\mathcal{E}, S(e)=v} u_e(n) = 1 \tag{15}$$

for all $v \in \mathcal{V}$ and all $n \in \mathbb{Z}_+$. The parameters $u_e(n)$ are the *transition probabilities* that are associated with each edge $\mathcal{E}$. The probability distribution $\mathbf{x}(n) \in \mathbb{R}^M$ of the DTMC $X_i(n)$, given by $\mathbb{P}(X_i(n) = v) = x_v(n) \in$

$\mathbb{R}_{\geq 0}$ for all $v \in \mathcal{V}$, evolves according to the mean-field model

$$\begin{aligned} \mathbf{x}(n+1) &= \sum_{e \in \mathcal{E}} u_e(n) \mathbf{B}_e \mathbf{x}(n), \quad n \in \mathbb{Z}_+, \\ \mathbf{x}(0) &= \mathbf{x}^0 \in \mathcal{P}(\mathcal{V}), \end{aligned} \tag{16}$$

where the entries of $\mathbf{B}_e \in \mathbb{R}^{M \times M}$ are given by

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

The above model is the discrete-time analogue of model (2). The problem of stabilizing the solution $\mathbf{x}(n)$ of the system (16) for swarm models was first considered in [70]. In this work, the authors develop an iterative scheme to construct a (time-independent) state feedback law u_e such that $\lim_{n\to\infty} \mathbf{x}(n) = \mathbf{x}^{eq}$, where $\mathbf{x}^{eq} \in \text{int } \mathcal{P}(\mathcal{V})$ is a target stationary probability distribution. In [71], the authors use a discrete-time mean-field model of the form (16), with experimentally identified parameters, to model a swarm of robots that inspect a turbine system. This model is used to optimize the performance of the robotic swarm for the inspection task in [72].

In [17], the authors investigate general conditions on the graph $\mathcal{G}$ under which time-independent state feedback laws $u_e \geq 0$ can be designed such that the solution of the system (16) converges to a given stationary distribution $\mathbf{x}^{eq}$. The authors construct a DTMC using a variant of the Metropolis-Hastings algorithm [73] and show that if the vector $\mathbf{x}^{eq}$ has a strongly connected support and the graph $\mathcal{G}$ is symmetric, then one can find parameters $u_e \geq 0$ such that this stabilization problem can be solved. The authors also provide a Linear Matrix Inequality based method for computing the parameters u_e such that a target $\mathbf{x}^{eq}$ is exponentially stable with a given decay rate. The following theorem is the discrete-time version of Theorem 2.1, and it provides a theoretical foundation for the results proved in [17].

Theorem 2.3 *Consider the mean-field model* (16)*, for which the corresponding graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is strongly connected. Suppose that the transition probabilities u_e are positive and constant. Additionally, suppose that there exists a time $n \in \mathbb{Z}^+$ such that, for each $v, w \in \mathcal{V}$, there exists a directed path of length n from v to w. Then* 1 *is the eigenvalue of the matrix $\sum_{e \in \mathcal{E}} u_e \mathbf{B}_e$ with the largest modulus. Moreover, this eigenvalue is simple. Hence, the solution $\mathbf{x}(t)$ of system* (16) *exponentially converges to a unique limit $\mathbf{x}^\infty \in \text{int } \mathcal{P}(\mathcal{V})$ for which all the elements are positive.*

A drawback of using time-independent state-feedback laws is that, as for the case of CTMCs, the agents do not stop transitioning between states once the mean-field model (16) reaches equilibrium. In order to resolve this issue, the authors in [74] consider the problem of constructing time-varying parameters $u_e(n)$ such that $\lim_{n\to\infty} u_e(n) = 1$ for all $e = (v,v) \in \mathcal{E}$, $v \in \mathcal{V}$. This problem is framed as a linear programming problem that each agent i must solve in order to compute its own optimal transition probabilities $u_e^i(n)$ at each time n so that the swarm reaches the target distribution while minimizing a particular objective functional. Strictly speaking, this linear programming approach is not a mean-field approach, since the problem is formulated for a finite number of agents and it is not clear whether the transition probabilities $u_e(n)$ have well-defined limits as $N \to \infty$. The state-feedback laws constructed in [74] depend on the distance of the swarm from the target distribution, and hence require global knowledge of the swarm distribution at each time n. This requirement is then relaxed by implementing a filtering algorithm that each agent uses to estimate the distribution of the swarm over all the states through local measurements (also known as local sampling) of the agent distribution in its current state.

To our knowledge, a similar approach for stabilizing a swarm modeled using DTMCs to a target distribution with local mean-field feedback laws, such as those designed for CTMCs in [37], has not been considered in the literature so far. However, the result for CTMCs that is given in Theorem IV.8 in [37] can be extended to DTMCs, as stated in the following theorem. The proof of this theorem is similar to the proof of Theorem IV.8 in [37] and is omitted here for brevity.

Theorem 2.4 *Suppose that the graph $\mathcal{G}$ is strongly connected and has self-edges at every vertex. Additionally, suppose that $\mathbf{x}^{eq} \in \mathcal{P}(\mathcal{V})$. For each $e \in \mathcal{E}$, let $k_e : [0,1] \to [0,1]$ be a continuous function such that k_e is non-zero over the interval $(x^{eq}_{S(e)}, 1]$, and*

$$\sum_{e \in \mathcal{E}; S(e)=v, T(e)\neq v} k_e(z) \leq \frac{z - x_v^{eq}}{z}, \quad \sum_{e \in \mathcal{E}; S(e)=v} k_e(y) = 1 \tag{17}$$

for each $v \in \mathcal{V}$ and for all $z < x_v^{eq}$ and all $y \in [0,1]$. Additionally, if e is not a self-edge, then the function satisfies $k_e(z) = 0$ for all $z < x^{eq}_{S(e)}$. Then $\mathbf{x}^{eq} \in \mathcal{P}(\mathcal{V})$ is a globally asymptotically stable equilibrium point for the system

$$\begin{aligned} \mathbf{x}(n+1) &= \sum_{e \in \mathcal{E}} k_e(x_{S(e)}(n)) \mathbf{B}_e \mathbf{x}(n), \quad n \in \mathbb{Z}_+, \\ \mathbf{x}(0) &= \mathbf{x}^0 \in \mathcal{P}(\mathcal{V}). \end{aligned} \tag{18}$$

In [75], the authors address a swarm stabilization problem in which the control laws must satisfy certain density constraints on the solution of the mean-field model. The authors adapt classical Markov decision process (MDP) theory [76] to construct stochastic or *randomized* state-feedback laws with constraints on the probability distribution of the stochastic process that models agent motion, such as constraints on robot densities.

3. Infinite-Dimensional Mean-Field Models

In this section, we describe infinite-dimensional mean-field models in which the time variable is continuous. We start with the case where the state space Ω of each agent, indexed by $i \in \{1, 2, ..., N\}$, is a subset of the Euclidean space $\mathbb{R}^n$. The position of each agent i evolves according to a stochastic process $\mathbf{Z}_i(t) \in \Omega$, where t denotes time. We initially assume that the agents are non-interacting. Therefore, the random variables $\mathbf{Z}_i(t)$ are independent and identically distributed, and we can drop the subscript i and define the problem in terms of a single stochastic process $\mathbf{Z}(t) \in \Omega$. The deterministic motion of each agent is defined by a velocity vector field $\mathbf{v}(\mathbf{x}, t) \in \mathbb{R}^n$, where $\mathbf{x} \in \Omega$. This motion is perturbed by an n-dimensional *Wiener process* $\mathbf{W}(t)$, which models noise. This process can be a model for stochasticity arising from inherent sensor and actuator noise. Alternatively, noise could be actively programmed into the agents' motion [77] to implement more exploratory agent behaviors [78] and to take advantage of the smoothening effect of the process on the agents' probability densities. Given the velocity field $\mathbf{v}(\mathbf{x}, t)$ and a diffusion coefficient $D > 0$, the position of each agent evolves according to a *diffusion process* $\mathbf{Z}(t)$ that satisfies the following stochastic differential equation (SDE) [79]:

$$\begin{aligned} d\mathbf{Z}(t) &= \mathbf{v}(\mathbf{Z}(t), t)dt + \sqrt{2D}d\mathbf{W}(t), \\ \mathbf{Z}(0) &= \mathbf{Z}_0. \end{aligned} \tag{19}$$

Given a final time $T > 0$, the Kolmogorov forward equation or *Fokker-Planck equation* [80] corresponding to the SDE (19) is given by:

$$\begin{aligned} & y_t = D\Delta y - \nabla \cdot (\mathbf{v}(\mathbf{x}(t), t)y) \quad in \;\; \Omega \times [0, T], \\ & y(\cdot, 0) = y^0 \quad in \;\; \Omega. \end{aligned} \tag{20}$$

The solution $y(\mathbf{x}, t)$ of this equation represents the probability density of a single agent occupying position $\mathbf{x} \in \Omega$ at time t, or alternatively, the density of a population of agents at this position and time. The PDE (20) is related to the SDE (19) through the relation $\mathbb{P}(\mathbf{Z}(t) \in \Gamma) = \int_\Gamma y(\mathbf{x}, t)d\mathbf{x}$ for all $t \in [0, T]$ and all measurable $\Gamma \subset \Omega$. In [81], the authors use the model (20) with a constant velocity field $\mathbf{v}$ to simulate a swarm of miniature robots performing an inspection task and validate the model experimentally. In [82], the authors design swarm strategies mimicking fluid flow behavior [2] by constructing state-feedback laws $\mathbf{v}$ that are piecewise constant with respect to space for the model (20) with $D = 0$, using the Helmholtz-Hodge decomposition of a vector field.

The work [83] considers a PDE model of the form

$$\begin{aligned} y_t(\mathbf{x}, \mathbf{v}) = & -\mathbf{v} \cdot \nabla_{\mathbf{x}} \cdot (y(\mathbf{x}, \mathbf{v})) - \lambda(\mathbf{x}, \mathbf{v})y((\mathbf{x}, \mathbf{v})) \\ & + \int T_{\mathbf{v}'}(v, v')\lambda(\mathbf{x}, \mathbf{v})y(\mathbf{x}, \mathbf{v}', t)d\mathbf{v}', \end{aligned} \tag{21}$$

where $\mathbf{x}$ denotes the position coordinates and $\mathbf{v}$ denotes the velocity coordinates. The parameter λ denotes the rate at which a robot jumps to a random value of $\mathbf{v}$ according to the parameter $T_{\mathbf{v}'}$, a function known as the *jump pdf.* Inspired by chemotaxis behavior observed in bacterial colonies [84], the authors design suitable λ and $T_{\mathbf{v}'}$ such that the robots converge to a target probability density that is positive everywhere. This result is generalized to a larger class of controllable nonlinear systems in [85].

There have been a number of works on numerical construction of state-feedback laws for a swarm of agents that follow the dynamics (19). In [86], the authors consider the problem of designing a time-varying, state-dependent velocity $u_1(\mathbf{x}, t)$ and turning rate $u_2(\mathbf{x}, t)$ with the vector field $\mathbf{v}$ in (20) given by

$$\mathbf{v}(\mathbf{x}, t) = \begin{bmatrix} u_1(\mathbf{x}, t)\cos(x_1) \\ u_1(\mathbf{x}, t)\sin(x_2) \\ u_2(\mathbf{x}, t) \end{bmatrix}.$$

The authors use optimal control to compute the control inputs $u_1(\mathbf{x}, t)$ and $u_2(\mathbf{x}, t)$ that transport a swarm from an initial probability density to a target density. The optimal control of PDEs that govern stochastic processes has received considerable attention in the mathematics literature [87–90]. Similar optimal control problems have also been investigated in the mathematics and control theory literature on *mean-field games* [19, 91–94]. The application of mean-field games to swarm robotics problems has begun only recently [95]. A promising approach to numerically constructing state-feedback laws comes from optimal transport theory. While this approach has thus far not been applied to control swarms of robots, we mention it here due to its applicability in this domain. Consider the following optimization problem:

$$\inf_{\mathbf{v}} \int_0^T \int_\Omega |\mathbf{v}(\mathbf{x}, t)|^2 y(\mathbf{x}, t) d\mathbf{x} dt \tag{22}$$

subject to the constraints

$$\begin{aligned} & y_t = -\nabla \cdot (\mathbf{v}(\mathbf{x}, t)y), \\ & y(0) = y^0, \quad y(T) = y^d, \end{aligned} \tag{23}$$

where y^0 and y^d are the initial and target probability densities, respectively. The optimization problem (22)-(23) was introduced to develop a computationally tractable approach to calculating the 2-Wasserstein distance [96]. In swarm robotics applications, this can be viewed as an optimal control problem that computes a state-feedback law $\mathbf{v}(\mathbf{x}, t)$ which drives a swarm from an initial probability density y^0 to a target probability density y^d in time T. However, this optimization problem is non-convex in the decision variables $\mathbf{v}$ and ρ. If we perform the change of variable $\mathbf{m} = \frac{\mathbf{v}}{\rho}$, we can instead consider the equivalent convex optimization problem,

$$\inf_{\mathbf{m},\rho\geq 0} \int_0^1 \int_\Omega \frac{|\mathbf{m}(\mathbf{x},t)|^2}{y(\mathbf{x},t)} d\mathbf{x}dt \tag{24}$$

subject to the constraints

$$\begin{aligned} &y_t = -\nabla \cdot (\mathbf{m}(\mathbf{x},t)), \\ &y(0) = y^0, \quad y(1) = y^d. \end{aligned} \tag{25}$$

Due to this convexification, one can guarantee that any locally optimal solution of the optimization problem (24)-(25) is also globally optimal. This offers an advantage over objective functionals that are more commonly used in optimal control of PDEs [97], for which global optimality of locally optimal solutions is much more difficult to guarantee.

The work [98] considers the problem of stabilizing the PDE (20) to a target probability density y_∞. It is shown that if the diffusion coefficient is defined as the spatially-dependent function $c/\sqrt{y_\infty}$ for any positive constant c, then the solution of the PDE converges to ρ_∞. The effectiveness of this control law is experimentally verified with robot experiments in [99]. This strategy is extended to the case where agents evolve on compact manifolds in [100]. An alternative approach to stabilize a swarm to a target distribution is to set D to a positive constant and $\mathbf{v} = D\frac{\nabla \rho_\infty}{\rho_\infty}$, which also results in the solution converging to ρ_∞ [101–105]. The long-time behavior of SDEs with gradient drift has been extensively treated in the mathematics and physics literature. In applications beyond swarm robotics, the problem of controlling the PDE (20) to a target probability density using a time-dependent state-feedback law $\mathbf{v}(\mathbf{x},t)$ has been investigated in optimal transport theory [39] and stochastic control [106] for the case where $\Omega = \mathbb{R}^n$, in work on control of PDEs [105] when Ω is a bounded subset of $\mathbb{R}^n$, and in the theory of mean-field games [107] when Ω is a torus.

While models of the form (20), with control parameters that are functions of the swarm density, have been extensively analyzed in the mathematics literature [108–112], there has been very little work on using such models to construct mean-field feedback laws for stabilization of robotic swarms. In [113], the authors design mean-field feedback laws where the vector field $\mathbf{v}$ in (20) is set to a suitable integral functional of the density so that the agents achieve consensus. A similar approach for the analysis of consensus in swarms is also considered in [114]. In [115], the authors construct a mean-field feedback law by interpreting the linear heat equation as a nonlinear advection equation with a density-dependent velocity field as follows. The diffusion coefficient D is set to zero, and the control law is defined as $\mathbf{v}(\mathbf{x},t) = -\frac{\nabla e(\mathbf{x},t)}{y(\mathbf{x},t)}$ for all $\mathbf{x} \in \Omega$ and all $t \geq 0$, where $e(\mathbf{x},t) = y(\mathbf{x},t) - y^d(\mathbf{x})$ and y^d is the target probability density. Then model (20) becomes

$$\begin{aligned} &e_t = \Delta e \qquad in \;\; \Omega \times [0,T], \\ &e(\cdot,0) = e^0 \quad in \;\; \Omega. \end{aligned} \tag{26}$$

Using the relation between models (20) and (26), one can show that the swarm density $y(\cdot,t)$ converges to the target probability density y^d as $t \to \infty$. Similarly, one can see that if $\mathbf{v}(\mathbf{x},t) = \frac{\nabla(y(\mathbf{x},t)/y^d(\mathbf{x}))}{y(\mathbf{x},t)}$ for all $\mathbf{x} \in \Omega$ and all $t \geq 0$, then (20) becomes

$$\begin{aligned} &y_t = \Delta\left(\frac{y}{y^d}\right) \quad in \;\; \Omega \times [0,T], \\ &y(\cdot,0) = y^0 \qquad in \;\; \Omega. \end{aligned} \tag{27}$$

While the analysis of the closed-loop systems (26)-(27) is straightforward due to their linearity, the solutions of these PDEs make sense only for initial conditions that are positive everywhere on Ω; otherwise, the control law $\mathbf{v}$ is unbounded. An alternative is to set $\mathbf{v}(\mathbf{x},t) = -b(\mathbf{x})\frac{\nabla y(\mathbf{x},t)}{y^d(\mathbf{x},t)}$, where $b(\mathbf{x})$ is a positive function. The resulting closed-loop system is a weighted variation of a well-known nonlinear PDE called the porous media equation [116]. According to results established in the mathematics literature [117], it is known that under particular technical assumptions on $b(\mathbf{x})$ and $y^d(\mathbf{x})$, the swarm density $y(\cdot,t)$ converges to the target probability density y^d as $t \to \infty$. These types of control laws are used for stabilizing swarms to target probability densities in the recent works [100], for robots evolving on compact manifolds without boundary, and [118], for robots evolving on a subset of a Euclidean space with boundary.

In models of robotic swarms, it is useful to consider *hybrid* variants of the SDE (19) to account for the fact that each robot, in addition to a continuous spatial state $\mathbf{Z}(t)$, can be associated with a discrete state $Y(t) \in \mathcal{V}$ at each time t. For such scenarios, we can define a hybrid switching diffusion process $(\mathbf{Z}(t), Y(t))$ as a system of SDEs of the form

$$\begin{aligned} &d\mathbf{Z}(t) = \mathbf{v}(Y(t),\mathbf{Z}(t),t)dt + \sqrt{2\mathbf{D}} \cdot d\mathbf{W}(t), \\ &\mathbf{Z}(0) \;= \mathbf{Z}_0, \end{aligned} \tag{28}$$

where $\mathbf{v} : \mathcal{V} \times \Omega \times [0,T] \to \mathbb{R}^n$ is the state- and time-dependent velocity vector field, and $\mathbf{D} \in \mathbb{R}_+^M$ is a vector of positive elements D_k, the diffusion coefficient associated with discrete state $k \in \mathcal{V}$. Let $\mathbf{v}_k$ denote the velocity field associated with discrete state $k \in \mathcal{V}$ and $\mathbf{y} = [y_1 \, ... \, y_M]^T$ denote the the probability density of the process $(\mathbf{Z}(t), Y(t))$. Then the forward equation for this system of SDEs is given by the system of PDEs

$$(y_k)_t = D_k \Delta y_k - \nabla \cdot (\mathbf{v}_k(\mathbf{x},t) y_k) + \mathcal{F}_k \quad in \quad \Omega \times [0, t_f],$$
$$y_k(\cdot, 0) = y_k^0 \quad in \quad \Omega, \tag{29}$$

where $k \in \mathcal{V}$ and $\mathcal{F}_k = \sum_{e \in \mathcal{E}} \sum_{j \in \mathcal{V}} u_e(t) B_e^{kj} y_j$, with $\mathbf{B}_e$ defined as in Subsection 2.1. The PDE (29) is related to the SDE (28), for each $k \in \mathcal{V}$, through the relation $\mathbb{P}(Y(t) = k, \mathbf{Z}(t) \in \Gamma) = \int_\Gamma y_k(\mathbf{x}, t) d\mathbf{x}$ for all $t \in [0, T]$ and all measurable $\Gamma \subset \Omega$.

The class of models (29) is used in [119] to model microscopic robots that reside in a fluid. In this work, some components of the vector $\mathbf{y}$ are used to model robot densities, and some model them densities of chemicals that the robots follow. In [120, 121], the authors consider a 3-state model for a coverage task, with diffusion coefficients equal to 0, in which the time-dependent transition rates are optimized using infinite-dimensional optimal control theory [122]. Each state is associated with an uncontrolled velocity vector field, corresponding to left-translation, right-translation, and remaining stationary. In [123, 124], these models are applied to study collective migration and collective perception tasks in swarms. To simulate the phenomenon of emergent taxis, the authors construct *mean-field feedback laws* in the sense that the velocity fields and diffusion coefficients are functions of the population densities, as in biological models of chemotaxis.

In [125], the authors use model (29) to simulate the coverage activity of a swarm of robotic bees in a commercial pollination problem. The framework presented in [125] is used in [126] to optimize time-dependent (and state-independent) robot velocities and state transition rates using optimal control theory of PDEs [97]. Additionally, [126] considers the problem of identifying the spatial distribution of resources in the environment from temporal robot data and frames this as a problem of identifying coefficients in model (29) using PDE-constrained optimization. Following a similar approach, [127] addresses the problem of mapping the boundaries of regions of interest in an environment from temporal robot data. In [105], the authors analytically construct control laws $\mathbf{v}_k(\mathbf{x}, t)$ and $u_e(t)$ to transport a swarm modeled by (29) from an initial probability density to a target density, thus establishing the controllability of the system (29).

When the parameters $\mathbf{v}_k(\mathbf{x}, t)$ and $u_e(t)$ are independent of the density $\mathbf{y}$, the convergence of the solution of the mean-field model (29) to the density of a finite-size swarm with agents that follow the SDEs (28) can be concluded from the law of large numbers. However, such convergence results thus far have been mostly qualitative. A more quantitative convergence analysis of the model presented in [126] is performed in [128], where the density of the finite-agent model is shown to converge to the solution of the mean-field model as the number of agents tends to infinity. Using this convergence result, performance bounds are derived in [128] for the optimal control strategies constructed in [126] as a function of the approximation error due to the finiteness of the agent population.

Finally, another type of infinite-dimensional mean-field model is a delay differential equation (DDE) that describes the time evolution of a finite number of moments of the density of a swarm. In [129], the authors derive this type of model to describe the acceleration of the center of mass of a swarm of agents that interact through pairwise potentials. The model shows that for inter-agent coupling strength and communication time delays above certain thresholds, a sufficiently high noise intensity causes the swarm to transition from a misaligned state to an aligned state.

4. Future Opportunities

The previous sections discussed a wide range of applications of mean-field models in swarm robotics. However, there remain various challenges that must be overcome in order for this modeling and control methodology to be developed into a broadly applicable tool for designing swarm robotic control strategies.

One of the main challenges is the issue of developing efficient computational methods for designing robot control laws. While it is true, as stated in the Introduction, that the scalability of computational methods using mean-field models is a function of the dimensionality of a single agent's state space and not the number of agents, this does not necessarily mean that these methods are not computationally intensive. For example, consider the case where the mean-field model is a PDE. Since the state space is a continuum and analytical solutions of PDEs are available only for a small number of special cases, the state space needs to be approximated as a finite set using a numerical discretization process such as gridding. To obtain an approximation with acceptable numerical accuracy, the number of states in the approximation could be very large, even though this number does not depend on the population size of the swarm.

Even when the mean-field model is finite-dimensional, efficient numerical computation of control laws can be a challenge if the goal is to construct mean-field feedback laws, since the resulting closed-loop system is nonlinear. When the mean-field model is a PDE, this computation becomes an even greater challenge due to the large dimensionality of the set of possible mean-field feedback laws.

In addition, mean-field models assume that the number of agents is infinite, whereas swarms have a finite number of agents in practice. Since mean-field formulations do not account for correlations and fluctuations that are captured by stochastic formulations [130], there needs to be further work on quantitative, rather than just qualitative, error analysis of mean-field models, such as bounds on the variance of agent population densities about the target distribution as a function of the number of agents. Moreover, it is not clear how mean-field models can capture microscopic effects, such as inter-robot collisions, without introducing additional nonlinearities into the model. The quantitative characterization of error is especially relevant for many applications in swarm robotics, in which the number of robots in the swarm is much smaller than the number of discrete entities (e.g., molecules, particles) in other types of large-scale collective systems with dynamics that are typically approximated by mean-field models. Mean-field models of robotic swarms should also account for the effects of disturbances and uncertainty in system parameters on the swarm dynamics. The modeling of disturbances may be especially relevant in cases where rare events significantly affect the performance or stability of the swarm [131]. Mean-field models can also have limited predictive power when they have multiple locally stable equilibria [132]. In this case, the mean-field model can fail to capture the long-time behavior of a swarm with a small number of agents. Accounting for such inaccuracies in mean-field models during the controller design process is another challenging avenue of future research.

There are also various control-theoretical challenges associated with mean-field models, such as establishing the controllability of mean-field models with state-feedback laws and the stabilizability of these models with mean-field feedback laws. While a number of works have initiated some research on this topic, numerous open problems remain, especially for PDE mean-field models, which can require highly complicated techniques for analysis. Even fundamental properties such as existence and uniqueness of solutions of PDEs, which are necessary to prove before developing a full-fledged analysis of the stability properties of the system, are difficult to guarantee. An alternative approach is to assume the existence of solutions that are suitably differentiable in time and space, and then proceed with a formal stability analysis that is not necessarily mathematically rigorous. This is a suitable alternative only if the end goal is purely computational in nature, such as the numerical construction of control laws. However, such assumptions on the existence and uniqueness of solutions can lead to erroneous conclusions, such as establishing the asymptotic stability of a system that is not in fact asymptotically stable about an equilibrium point. This is a well-known issue in mathematics. See Example IV.7 in [37] for an instance where this assumption can lead to such incorrect results when applying an ODE mean-field model to a swarm coverage problem.

Another interesting direction of future work is the quantitative investigation of relative benefits of mean-field feedback laws versus state-feedback laws. One advantage of mean-field feedback laws is that they can achieve microscopic equilibrium at the same time as macroscopic equilibrium. On the other hand, state-feedback laws can be more robust to robot failures than mean-field feedback laws: if robots cannot distinguish between functional and non-functional robots, then their estimates of the swarm population density will be too high, and this error will affect mean-field feedback laws (and the resulting collective performance of the swarm) but not state-feedback laws.

In addition, there is potential for using discrete-time mean-field models of agents evolving on a continuous state space for swarm robotic applications. Like PDEs, these models are more accurate descriptions of swarm behavior than their discrete-state space counterparts, due to the continuity of their state space. On the other hand, since time is discrete-valued in these models, they are much easier to treat mathematically than PDEs. A DTMC evolving on a continuous state space Ω can be constructed using a kernel function $k : \Omega \times \Omega \to \mathbb{R}_{\geq 0}$. The time evolution of $\rho_n(\mathbf{y})$, the probability density of the DTMC at time n, is then given by

$$\rho_{n+1}(\mathbf{y}) = \int_\Omega k(\mathbf{x}, \mathbf{y}) \rho_n(\mathbf{x}) d\mathbf{x}, \quad n \in \mathbb{Z}_+ \tag{30}$$

for all $\mathbf{y} \in \Omega$. The function k represents a *stochastic state-feedback law* that must be designed so that ρ_n reaches a target density either at a given $n \in \mathbb{Z}_+$, or as $n \to \infty$. Here, $k(\mathbf{x}, \mathbf{y}) d\mathbf{y}$ denotes the probability of an agent jumping to the point $\mathbf{y}$, given that the agent is at $\mathbf{x}$. More precisely, the probability of an agent jumping to a set $A \subseteq \Omega$, given that the agent is initially at $\mathbf{x}$, is $\int_A k(\mathbf{x}, \mathbf{y}) d\mathbf{y}$. Such models and their generalizations have been studied extensively in the mathematics literature on the long-time behavior of Markov chains on continuous state spaces [133,134]. Inverse problems that are of interest in swarm robotics, like the design of robot control laws from macroscopic specifications on the swarm collective behavior, have been considered only very recently in the control theory literature [135,136].

For discrete-time, continuous-state space mean-field models of swarms in which the agents interact with one another, it is necessary to define a density-dependent kernel or *stochastic mean-field feedback law*, $k : \mathcal{P}(\Omega) \times \Omega \times \Omega \to \mathbb{R}_{\geq 0}$, where $\mathcal{P}(\Omega)$ denotes the set of probability densities on Ω. In this case, the time evolution of the probability density $\rho_n(\mathbf{y})$ is given by

$$\rho_{n+1}(\mathbf{y}) = \int_\Omega k(\rho_n, \mathbf{x}, \mathbf{y}) \rho_n(\mathbf{x}) d\mathbf{x}, \quad n \in \mathbb{Z}_+ \tag{31}$$

for all $\mathbf{y} \in \Omega$. This type of model has been used to describe opinion dynamics [137], and it has many potential applications in swarm robotics, as is evident from the related literature cited in Sections 2 and 3.

5. Conclusion

In summary, this survey has presented an overview of the large number of works on the use of mean-field models to solve control problems in swarm robotics. As we discuss in the previous section, there remain significant open challenges in extending the current use of these models in designing robot control laws to a general controller synthesis methodology for achieving a variety of swarm collective behaviors. Many of these challenges will likely require solutions that are highly interdisciplinary in nature, cutting across disciplines such as robotics, control theory, mathematics, and physics. We hope that this survey paper will foster further research in these directions.

Acknowledgments

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

References

[1] Brambilla M, Ferrante E, Birattari M and Dorigo M 2013 *Swarm Intelligence* **7** 1–41
[2] Young D F, Munson B R, Okiishi T H and Huebsch W W 2010 *A brief introduction to fluid mechanics* (John Wiley & Sons)
[3] Van Kampen N G 2007 *Stochastic processes in physics and chemistry* 3rd ed North-Holland Personal Library (Elsevier)
[4] Brauer F, Castillo-Chavez C and Castillo-Chavez C 2001 *Mathematical models in population biology and epidemiology* vol 40 (Springer)
[5] Bressloff P C 2014 *Stochastic processes in cell biology* vol 41 (Springer)
[6] Camazine S, Deneubourg J L, Franks N R, Sneyd J, Bonabeau E and Theraula G 2003 *Self-organization in biological systems* (Princeton university press)
[7] Francesca G and Birattari M 2016 *Frontiers in Robotics and AI* **3** 29
[8] Hamann H 2018 *Swarm robotics: A formal approach* (Springer)
[9] Nedjah N and Junior L S 2019 *Swarm and Evolutionary Computation* 100565
[10] Khamis A, Hussein A and Elmogy A 2015 Multi-robot task allocation: A review of the state-of-the-art *Cooperative Robots and Sensor Networks 2015* (Springer) pp 31–51
[11] Robin C and Lacroix S 2016 *Autonomous Robots* **40** 729–760
[12] Parker L E 2008 *Springer Handbook of Robotics* 921–941
[13] Oh K K, Park M C and Ahn H S 2015 *Automatica* **53** 424–440
[14] Beal J and Viroli M 2015 *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **373** 20140220
[15] Agassounon W, Martinoli A and Easton K 2004 *Autonomous Robots* **17** 163–192
[16] Lerman K, Jones C, Galstyan A and Matarić M J 2006 *The International Journal of Robotics Research* **25** 225–241
[17] Açıkmeşe B and Bayard D S 2015 *Asian Journal of Control* **17** 1105–1124
[18] Lerman K, Martinoli A and Galstyan A 2004 A review of probabilistic macroscopic models for swarm robotic systems *International Workshop on Swarm Robotics* (Springer) pp 143–152
[19] Bensoussan A, Frehse J, Yam P *et al.* 2013 *Mean field games and mean field type control theory* vol 101 (Springer)
[20] Beshers S N and Fewell J H 2001 *Annual review of entomology* **46** 413–440
[21] Correll N and Martinoli A 2007 Robust distributed coverage using a swarm of miniature robots *Proceedings 2007 IEEE International Conference on Robotics and Automation* (IEEE) pp 379–384
[22] Wilson E O 1978 *Journal of the Kansas Entomological Society* 615–636
[23] Muscedere M L and Traniello J F 2012 *PLoS One* **7** e31618
[24] Berman A and Plemmons R J 1994 *Nonnegative matrices in the mathematical sciences* vol 9 (SIAM)
[25] Berman S, Halász Á, Hsieh M A and Kumar V 2009 *IEEE Transactions on Robotics* **25** 927–937
[26] Hsieh M A, Halász Á, Berman S and Kumar V 2008 *Swarm Intelligence* **2** 121–141
[27] Miller M B and Bassler B L 2001 *Annual Reviews in Microbiology* **55** 165–199
[28] Pratt S C 2005 *Behavioral Ecology* **16** 488–496
[29] Ethier S N and Kurtz T G 2009 *Markov processes: characterization and convergence* vol 282 (John Wiley & Sons)
[30] Gast N and Gaujal B 2012 *Performance Evaluation* **69** 623–642
[31] Roth G and Sandholm W H 2013 *SIAM Journal on Control and Optimization* **51** 525–555
[32] Bortolussi L, Hillston J, Latella D and Massink M 2013 *Performance Evaluation* **70** 317–349
[33] Mather T W and Hsieh M A 2014 *Robotica* **32** 177–192
[34] Halász A, Hsieh M A, Berman S and Kumar V 2007 Dynamic redistribution of a swarm of robots among multiple sites *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE) pp 2320–2325
[35] Elamvazhuthi K, Kawski M, Biswal S, Deshmukh V and Berman S 2017 Mean-field controllability and decentralized stabilization of markov chains *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on* (IEEE) pp 3131–3137
[36] Elamvazhuthi K and Grover P 2018 *Journal of Computational Dynamics* **5** 1–32
[37] Elamvazhuthi K, Biswal S and Berman S 2018 Mean-field stabilization of robotic swarms to probability distributions with disconnected supports *American Control Conference (ACC), 2018* (IEEE) pp 885–892
[38] Solomon J, Rustamov R, Guibas L and Butscher A 2016 *arXiv preprint arXiv:1603.06927*
[39] Benamou J D and Brenier Y 2000 *Numerische Mathematik* **84** 375–393
[40] Boyd S, El Ghaoui L, Feron E and Balakrishnan V 1994 *Linear matrix inequalities in system and control theory* vol 15 (SIAM)
[41] Chesi G 2011 *Domain of attraction: analysis and control via SOS programming* vol 415 (Springer Science & Business Media)
[42] Deshmukh V, Elamvazhuthi K, Biswal S, Kakish Z and Berman S 2018 *IEEE Robotics and Automation Letters* **3** 1985–1992

[43] Gordon D M and Mehdiabadi N J 1999 *Behavioral Ecology and Sociobiology* **45** 370–377

[44] Mayya S, Pierpaoli P, Nair G and Egerstedt M 2019 *IEEE Transactions on Robotics* **35** 21–34

[45] Mayya S, Wilson S and Egerstedt M 2019 *Swarm Intelligence* 1–29

[46] Prorok A, Hsieh M A and Kumar V 2017 *IEEE Transactions on Robotics* **33** 346–358

[47] Romano D, Donati E, Benelli G and Stefanini C 2019 *Biological cybernetics* **113** 201–225

[48] Garnier S 2011 From ants to robots and back: How robotics can contribute to the study of collective animal behavior *Bio-inspired self-organizing robotic systems* (Springer) pp 105–120

[49] Halloy J, Sempo G, Caprari G, Rivault C, Asadpour M, Tâche F, Saïd I, Durier V, Canonge S, Amé J M *et al.* 2007 *Science* **318** 1155–1158

[50] Feinberg M 2019 *Foundations of Chemical Reaction Network Theory* vol 202 (Springer)

[51] Gillespie D T, Hellander A and Petzold L R 2013 *The Journal of Chemical Physics* **138** 05B201_1

[52] Lerman K, Galstyan A, Martinoli A and Ijspeert A 2001 *Artificial Life* **7** 375–393

[53] Lerman K and Galstyan A 2002 *Autonomous Robots* **13** 127–141

[54] Matthey L, Berman S and Kumar V 2009 Stochastic strategies for a swarm robotic assembly system *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on* (IEEE) pp 1953–1958

[55] Wilson S, Pavlic T P, Kumar G P, Buffin A, Pratt S C and Berman S 2014 *Swarm Intelligence* **8** 303–327

[56] Haghighat B, Thandiackal R, Mordig M and Martinoli A 2017 Probabilistic modeling of programmable stochastic self-assembly of robotic modules *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE) pp 4656–5663

[57] Klavins E, Burden S and Napp N 2006 Optimal rules for programmed stochastic self-assembly. *Robotics: Science and Systems* (Philadelphia, PA)

[58] Napp N, Burden S and Klavins E 2011 *Autonomous Robots* **30** 57–71

[59] Mermoud G, Matthey L, Evans W C and Martinoli A 2010 Aggregation-mediated collective perception and action in a group of miniature robots *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 2-Volume 2* (International Foundation for Autonomous Agents and Multiagent Systems) pp 599–606

[60] de Oca M A M, Ferrante E, Scheidler A, Pinciroli C, Birattari M and Dorigo M 2011 *Swarm Intelligence* **5** 305–327

[61] Krapivsky P L and Redner S 2003 *Physical Review Letters* **90** 238701

[62] Seeley T D, Visscher P K, Schlegel T, Hogan P M, Franks N R and Marshall J A 2012 *Science* **335** 108–111

[63] Reina A, Valentini G, Fernández-Oto C, Dorigo M and Trianni V 2015 *PLoS One* **10** e0140950

[64] Reina A, Bose T, Trianni V and Marshall J A 2018 Effects of spatiality on value-sensitive decisions made by robot swarms *Distributed Autonomous Robotic Systems* (Springer) pp 461–473

[65] Pais D, Hogan P M, Schlegel T, Franks N R, Leonard N E and Marshall J A 2013 *PloS one* **8** e73216

[66] Albani D, Manoni T, Nardi D and Trianni V 2018 Dynamic UAV swarm deployment for non-uniform coverage *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems* (International Foundation for Autonomous Agents and Multiagent Systems) pp 523–531

[67] Valentini G 2017 *Achieving Consensus in Robot Swarms: Design and Analysis of Strategies for the best-of-n Problem* vol 706 (Springer)

[68] Valentini G, Ferrante E and Dorigo M 2017 *Frontiers in Robotics and AI* **4** 9

[69] Prorok A and Kumar V 2016 A macroscopic privacy model for heterogeneous robot swarms *International Conference on Swarm Intelligence* (Springer) pp 15–27

[70] Chattopadhyay I and Ray A 2009 *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **39** 1505–1515

[71] Correll N and Martinoli A 2006 System identification of self-organizing robotic swarms *Distributed Autonomous Robotic Systems 7* (Springer) pp 31–40

[72] Correll N 2008 Parameter estimation and optimal control of swarm-robotic systems: A case study in distributed task allocation *2008 IEEE International Conference on Robotics and Automation* (IEEE) pp 3302–3307

[73] Chib S and Greenberg E 1995 *The American Statistician* **49** 327–335

[74] Bandyopadhyay S, Chung S J and Hadaegh F Y 2017 *IEEE Transactions on Robotics* **33** 1103–1123

[75] El Chamie M, Yu Y, Açıkmeşe B and Ono M 2019 *IEEE Transactions on Automatic Control* **64** 1003–1018

[76] Puterman M L 2014 *Markov decision processes: discrete stochastic dynamic programming* (John Wiley & Sons)

[77] Dimidov C, Oriolo G and Trianni V 2016 Random walks in swarm robotics: an experiment with kilobots *International Conference on Swarm Intelligence* (Springer) pp 185–196

[78] Viswanathan G M, Da Luz M G, Raposo E P and Stanley H E 2011 *The physics of foraging: an introduction to random searches and biological encounters* (Cambridge University Press)

[79] Gardiner C 2009 *Stochastic methods* vol 4 (springer Berlin)

[80] Risken H 1996 Fokker-planck equation *The Fokker-Planck Equation* (Springer) pp 63–95

[81] Prorok A, Correll N and Martinoli A 2011 *International Journal of Robotics Research* **30** 574–589

[82] Kingston P and Egerstedt M 2011 Distributed-infrastructure multi-robot routing using a Helmholtz-Hodge decomposition *2011 50th IEEE Conference on Decision and Control and European Control Conference* (IEEE) pp 5281–5286

[83] Mesquita A R, Hespanha J P and Åström K 2008 Optimotaxis: A stochastic multi-agent optimization procedure with point measurements *International Workshop on Hybrid Systems: Computation and control* (Springer) pp 358–371

[84] Wu M, Roberts J W, Kim S, Koch D L and DeLisa M P 2006 *Appl. Environ. Microbiol.* **72** 4987–4994

[85] Mesquita A R and Hespanha J P 2012 *IEEE Transactions on Automatic Control* **57** 2588–2598

[86] Foderaro G, Ferrari S and Wettergren T A 2014 *Automatica* **50** 149–154

[87] Annunziato M and Borzì A 2010 *Mathematical Modelling and Analysis* **15** 393–407

[88] Annunziato M and Borzì A 2013 *Journal of Computational and Applied Mathematics* **237** 487–507

[89] Annunziato M and Borzì A 2018 *EMS Surveys in Mathematical Sciences* **5** 65–98

[90] Fleig A and Guglielmi R 2017 *Journal of Optimization Theory and Applications* **174** 408–427

[91] Lasry J M and Lions P L 2007 *Japanese Journal of Mathematics* **2** 229–260

[92] Huang M, Caines P E and Malhamé R P 2007 *IEEE Transactions on Automatic Control* **52** 1560–1571

[93] Caines P E, Huang M and Malhamé R P 2017 Mean field games *Handbook of Dynamic Game Theory* ed T Basar G Z e (Springer) pp 1–28

[94] Carmona R and Delarue F 2018 *Probabilistic Theory of*

Mean Field Games with Applications I-II (Springer)

[95] Liu Z, Wu B and Lin H 2018 A mean field game approach to swarming robots control *2018 Annual American Control Conference (ACC)* (IEEE) pp 4293–4298

[96] Villani C 2008 *Optimal transport: old and new* vol 338 (Springer Science & Business Media)

[97] Tröltzsch F 2010 *Optimal control of partial differential equations: theory, methods, and applications* vol 112 (American Mathematical Society)

[98] Elamvazhuthi K, Adams C and Berman S 2016 Coverage and field estimation on bounded domains by diffusive swarms *Decision and Control (CDC), 2016 IEEE 55th Conference on* (IEEE) pp 2867–2874

[99] Li H, Feng C, Ehrhard H, Shen Y, Cobos B, Zhang F, Elamvazhuthi K, Berman S, Haberland M and Bertozzi A L 2017 Decentralized stochastic control of robotic swarm density: Theory, simulation, and experiment *Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on* (IEEE) pp 4341–4347

[100] Elamvazhuthi K and Berman S 2018 Nonlinear generalizations of diffusion-based coverage by robotic swarms *2018 IEEE Conference on Decision and Control (CDC)* (IEEE) pp 1341–1346

[101] Stroock D W 1993 Logarithmic Sobolev inequalities for Gibbs states *Dirichlet forms* (Springer) pp 194–228

[102] Markowich P A and Villani C 1999 On the trend to equilibrium for the Fokker-Planck equation: an interplay between physics and functional analysis *Physics and Functional Analysis, Matematica Contemporanea (SBM) 19*

[103] Ambrosio L, Savaré G and Zambotti L 2009 *Probability Theory and Related Fields* **145** 517–564

[104] Breiten T, Kunisch K and Pfeiffer L 2018 *ESAIM: Control, Optimisation and Calculus of Variations* **24** 741–763

[105] Elamvazhuthi K, Kuiper H, Kawski M and Berman S 2019 *IEEE Transactions on Automatic Control* **64** 2282–2297

[106] Blaquiere A 1992 *Dynamics and Control* **2** 235–253

[107] Porretta A 2014 *Dynamic Games and Applications* **4** 231–256

[108] Bodnar M and Velazquez J J L 2005 *Mathematical Methods in the Applied Sciences* **28** 1757–1779

[109] Topaz C M, Bertozzi A L and Lewis M A 2006 *Bulletin of Mathematical Biology* **68** 1601

[110] Bertozzi A L, Laurent T and Rosado J 2011 *Communications on Pure and Applied Mathematics* **64** 45–83

[111] Carrillo J A, Choi Y P and Hauray M 2014 The derivation of swarming models: mean-field limit and Wasserstein distances *Collective dynamics from bacteria to crowds* (Springer) pp 1–46

[112] Carrillo J A, Fornasier M, Toscani G and Vecil F 2010 Particle, kinetic, and hydrodynamic models of swarming *Mathematical modeling of collective behavior in socio-economic and life sciences* (Springer) pp 297–336

[113] Kingston P and Egerstedt M 2010 *IFAC Proceedings Volumes* **43** 215–220

[114] Canuto C, Fagnani F and Tilli P 2008 A Eulerian approach to the analysis of rendez-vous algorithms *Proceedings of the 17th IFAC world congress (IFAC08)* pp 9039–9044

[115] Eren U and Açıkmeşe B 2017 *IFAC-PapersOnLine* **50** 9405–9411

[116] Vázquez J L 2007 *The porous medium equation: mathematical theory* (Oxford University Press)

[117] Grillo G, Muratori M and Porzio M M 2013 *Discrete & Continuous Dynamical Systems - A* **33** 3599–3640

[118] Krishnan V and Martínez S 2018 Distributed optimal transport for the deployment of swarms *2018 IEEE Conference on Decision and Control (CDC)* (IEEE) pp 4583–4588

[119] Galstyan A, Hogg T and Lerman K 2005 Modeling and mathematical analysis of swarms of microscopic robots *Proceedings 2005 IEEE Swarm Intelligence Symposium, 2005. SIS 2005.* (IEEE) pp 201–208

[120] Milutinovic D and Lima P 2006 *IEEE Transactions on Robotics* **22** 1280–1285

[121] Milutinovic D L and Lima P U 2007 *Cells and robots: modeling and control of large-size agent populations* (New York)

[122] Fattorini H O 1999 *Infinite dimensional optimization and control theory* vol 54 (Cambridge University Press)

[123] Hamann H and Wörn H 2008 *Swarm Intelligence* **2** 209–239

[124] Hamann H 2010 *Space-time continuous models of swarm robotic systems: Supporting global-to-local programming* vol 9 (Springer Science & Business Media)

[125] Berman S, Kumar V and Nagpal R 2011 Design of control policies for spatially inhomogeneous robot swarms with application to commercial pollination *Robotics and Automation (ICRA), 2011 IEEE International Conference on* (IEEE) pp 378–385

[126] Elamvazhuthi K, Kuiper H and Berman S 2018 *Automatica* **95** 356–367

[127] Ramachandran R K, Elamvazhuthi K and Berman S 2018 An optimal control approach to mapping GPS-denied environments using a stochastic robotic swarm *Robotics Research* (Springer) pp 477–493

[128] Zhang F, Bertozzi A L, Elamvazhuthi K and Berman S 2018 *IEEE Transactions on Automatic Control* **63** 1563–1578

[129] Mier-y-Teran-Romero L, Forgoston E and Schwartz I B 2012 *IEEE Transactions on Robotics* **28** 1034–1044

[130] Gillespie D T 1976 *Journal of Computational Physics* **22** 403–434

[131] Goutsias J and Jenkinson G 2013 *Physics reports* **529** 199–264

[132] El Samad H, Khammash M, Petzold L and Gillespie D 2005 *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal* **15** 691–711

[133] Meyn S P and Tweedie R L 2012 *Markov chains and stochastic stability* (Springer Science & Business Media)

[134] Diaconis P and Freedman D 1997 On Markov chains with continuous state space Tech. rep. Technical Report

[135] Saldi N, Basar T and Raginsky M 2018 *SIAM Journal on Control and Optimization* **56** 4256–4287

[136] Elamvazhuthi K, Grover P and Berman S 2018 *IEEE Control Systems Letters* **3** 168–173

[137] Mirtabatabaei A, Jia P and Bullo F 2014 *SIAM Journal on Applied Dynamical Systems* **13** 425–446