

# Supervised Self-Organization of Homogeneous Swarms Using Ergodic Projections of Markov Chains

Ishanu Chattopadhyay and Asok Ray

**Abstract**—This paper formulates a self-organization algorithm to address the problem of global behavior supervision in engineered swarms of arbitrarily large population sizes. The swarms considered in this paper are assumed to be homogeneous collections of independent identical finite-state agents, each of which is modeled by an irreducible finite Markov chain. The proposed algorithm computes the necessary perturbations in the local agents' behavior, which guarantees convergence to the desired observed state of the swarm. The ergodicity property of the swarm, which is induced as a result of the irreducibility of the agent models, implies that while the local behavior of the agents converges to the desired behavior only in the time average, the overall swarm behavior converges to the specification and stays there at all times. A simulation example illustrates the underlying concept.

**Index Terms**—Discrete event systems, ergodic projections, finite-state irreducible Markov chains, swarms.

## I. INTRODUCTION AND MOTIVATION

WITH the recent advances in sensor technology and the affordable miniaturization of mobile computing platforms, swarms of simple agents are now capable of performing a variety of complex coordinated tasks. Potential applications of such human-engineered swarms range from self-organizing sensor fields for military surveillance and civilian search and rescue operations to coordinated handling and transportation of large objects. Motion coordination among teams of autonomous agents has extensively been studied [1], [2], with special emphasis on the formation control of large groups [3].

The essential distinction between large groups of autonomous agents and a *swarm* needs to be clarified. An example from nature is pertinent here. Wolves hunt in packs; however, a pack is not a swarm. Members of a pack play special roles in the highly coordinated process in the sense that removal of a few members renders the pack ineffective until the missing members are reinstated. In contrast, a colony of honeybees functions as a swarm, where removal of a few hundred workers has little impact on the colony as a whole. This distinction is not merely due to the size of the group; it arises from a difference

in the operational philosophy. In a swarm, an individual has no importance, and attaining the group objective is all that matters.

For human-engineered systems, the preceding philosophy translates to having little or no performance requirements on individual agents, and control of global behavior is of *sole* importance. Thus, the problem in swarm control is more than to just come up with decentralized control policies; it is one of engineering a framework that embodies a fundamental survival philosophy observed in nature. Largely, the reported work in this field addresses coordination of groups that are not large enough to qualify as swarms [1], [2], and often makes application-specific modeling assumptions [4], [5]. Recently, Belta and Kumar [6] have proposed an abstract framework for computing decentralized controllers to operate on locally available sensor information that realizes coordinated task execution for finite autonomous teams. However, there are performance criteria to be satisfied at the agent level, and the analysis does not necessarily scale to arbitrarily large population sizes. Recently, bio-inspired approaches to swarm control have been reported [7], in which the authors use a self-organized communication method inspired from males of chorusing insects, which are known for the rapid synchronization of their acoustic signals in a chorus. Attempts have been made to combine behavior-based and systems-theoretic approaches to give better and qualitatively more functional controllers for swarms [8]. This method controls the mean (average position) and variance (spread) of a swarm of robots to direct movement in underwater mine-hunting applications and makes use of “artificial potential fields” as the key tool for developing underlying system behaviors. A more applied approach to the problem is reported in [9], which uses “digital pheromones” to bias the movements of individual units within a swarm toward areas that are attractive and away from areas that are dangerous or unattractive and presents efficient methods for performing pheromone field calculations in the graphics processing unit (GPU). Other analytical concepts have been investigated as well, such as the Hill Function Construction [10] to guide swarms to acoustic sources. However, the reported approaches to swarm control suffer from lack of generality. While working well for the particular applications cited, the problem of formulating a general framework for controlling swarms is apparently still open. Furthermore, almost all of the reported works address the problem of coordinating swarm movement in the physical space; the formulation of a control approach that is capable of handling more general tasks has received little attention. In this paper, we lay the initial framework to address these deficiencies for homogeneous swarms.

Manuscript received May 21, 2008; revised October 31, 2008 and March 10, 2009. This work was supported in part by the U.S. Army Research Laboratory and the U.S. Army Research Office (ARO) under Grant W911NF-07-1-0376, by the U.S. Office of Naval Research under Grant N00014-08-1-380, and by the National Aeronautics and Space Administration (NASA) under Cooperative Agreement NNX07AK49A. This paper was recommended by Associate Editor R. Selmic.

The authors are with the Pennsylvania State University, University Park, PA 16802 USA (e-mail: icx128@psu.edu; axr2@spu.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSMCB.2009.2020173

We model homogeneous swarms as collections of independent identical finite-state Markov chains that capture the operational behavior of individual agents. In particular, we associate states with distinct behaviors that the agents can execute. Thus, any behavior that satisfies the rather weak requirements stated in the sequel can be represented by an agent state. The key concepts pertaining to Markov chain modeling used in this paper are that of irreducible chains and their corresponding ergodic projection [11]–[13], and the latter is related to the stationary distribution attained on the model states. For the sake of brevity, well-known results are presented without proof, which can be found in standard texts on the subject [13]–[16]. The transition probabilities, i.e., the probabilities of behavior switching, are assumed to be controllable in a certain probabilistic sense [17], [18] to the agents themselves, which forms the basis of the proposed supervision algorithm. The details are presented in the sequel.

The exact configuration or the “microstate” of the swarm at any given time is given by a detailed enumeration of the local states of the individual independent agents. Drawing an analogy with statistical mechanics [19], the problem here is to control the macroscopically measurable behavior of the swarm without having access to the microstates, i.e., the detailed enumeration previously described. The assumption of independent operation of the agents is similar to the ideal gas concept in statistical thermodynamics, where the individual gas molecules are noninteracting, i.e., the energy of intermolecular interaction is assumed to be insignificant with respect to the kinetic energy of the individual molecules. In the present framework, the operational independence of the agents implies that they are able to execute transitions in the local states, i.e., switch behaviors, independent of the neighbors. Thus, we assume that the neighbors have little effect in addition to what is already modeled in the local Markov chains. This requirement is not too restrictive. For example, if in a given scenario the local Markov chains do not model collisions, then we would have to guarantee that interagent collisions are limited. This can be achieved rather simply by executing local reactive collision avoidance routines in each agent in addition to the main control algorithm. The proposed algorithm provides a control methodology under which every agent in the swarm receives the same control broadcast, thus eliminating the need to talk to specific agents or individually control them, but the observed global behavior converges to the desired specification. Furthermore, this control broadcast is required only once when the desired global behavior needs to be updated; the swarm self-organizes after the initial broadcast to reflect the desired global configuration. Practical issues such as interagent communication, noise corruption, and actuation delays, although critical, are not addressed in this paper and would be investigated in future work on actual system implementations.

This paper is organized in four sections including the present section. Section II states the necessary definitions and concepts and presents the formal statement of the control problem. Section III presents the main results along with an illustrative application example. This paper is summarized and concluded in Section IV with recommendations for future work.

## II. PRELIMINARIES AND NOTATIONS

This section provides preliminary concepts and notations that facilitate understanding of the concepts presented in the sequel.

*Notation 1:* In the sequel, we denote the cardinality of any set  $A$  as  $\text{CARD}(A)$ .

*Definition 1:* A finite-state homogeneous Markov chain is a triple  $G = (Q, \Pi, \varphi^0)$ , where  $Q$  is a set of states with  $\text{CARD}(Q) = n \in \mathbb{N}$ ,  $\Pi$  is the  $(n \times n)$  stationary transition probability matrix such that  $\forall i, j \in \{1, \dots, n\}$ ,  $\Pi_{ij} \geq 0$  with  $\sum_i \Pi_{ij} = 1$ , and  $\varphi^0 \in [0, 1]^{\text{CARD}(Q)}$ ,  $\sum_i \varphi^0 = 1$ , is the initial occupancy distribution over the model states.

*Remark 1:* The transition matrix  $\Pi$  of a finite Markov chain is always a stochastic matrix [13], [16]. Every stochastic matrix  $\Pi$  has at least one unity eigenvalue, and all the eigenvalues of  $\Pi$  are located within or on the unit disk.

*Definition 2:* A finite-state homogeneous Markov chain  $G = (Q, \Pi, \varphi^0)$ , with  $\text{CARD}(Q) = n \geq 2$ , is called irreducible if, for any pair  $(i, j)$ ,  $1 \leq i, j \leq n$ , there exists a positive integer  $k(i, j) \leq n$  such that the  $ij$ th element of the  $k$ th power of  $\Pi$  is strictly positive, i.e.,  $\Pi_{ij}^k > 0$ . In this case, the stochastic transition matrix  $\Pi$  is also an irreducible matrix [13], [16].

*Remark 2:* The following standard results are crucial for the development in the sequel. Proofs can be found in any standard text on Markov chain modeling or stochastic matrices [11], [13]–[16].

For any  $n \times n$  irreducible stochastic matrix  $\Pi$  with  $n > 1$ , the diagonal terms are strictly less than unity, i.e.,  $\Pi_{ii} < 1 \forall i$ . Upon unity sum normalization, the left eigenvector vector  $\varphi = [\varphi_1 \ \varphi_2 \ \dots \ \varphi_n]$  corresponding to the unique unity eigenvalue of  $\Pi$  is called the stationary probability vector, where  $\sum_i \varphi_i = 1$ , and  $\varphi_j > 0 \forall j$ . The stationary probability vector  $\varphi$  satisfies the following condition:

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=0}^k \Pi^j = \begin{bmatrix} \dots & \varphi & \dots \\ \vdots & \vdots & \vdots \\ \dots & \varphi & \dots \end{bmatrix}. \quad (1)$$

The (unity rank) matrix on the right-hand side of (1) is called the ergodic projection and is denoted as  $\mathcal{C}(\Pi)$ . The rows of  $\mathcal{C}(\Pi)$  define the stationary probability distribution for the irreducible chain  $G$  in the sense that  $\varphi \Pi = \varphi$ , and, in general,  $\mathcal{C}(\Pi) \Pi = \Pi \mathcal{C}(\Pi) = \mathcal{C}(\Pi)$ .

*Definition 3 (Probabilistic Disabling and Enabling):* Let  $G = (Q, \Pi, \varphi^0)$  be an irreducible Markov chain, and let the  $ij$ th probability be nonzero, i.e.,  $\Pi_{ij} > 0$ . Referring to Fig. 1, the probabilistic disabling and enabling [17], [18] from state  $i$  to state  $j$  are, respectively, defined in terms of perturbations in the state transition matrix  $\Pi$  as

$$\text{DISABLING :} \quad \Pi_{ij} \mapsto (1 - \gamma) \Pi_{ij}, \quad \gamma \in [0, 1] \quad (2a)$$

$$\Pi_{ii} \mapsto \Pi_{ii} + \gamma \Pi_{ij} \quad (2b)$$

$$\text{ENABLING :} \quad \Pi_{ii} \mapsto \Pi_{ii} - \gamma \Pi_{ij} \quad (2c)$$

$$\text{where } \gamma \in [0, 1] \wedge \Pi_{ii} \geq \gamma \Pi_{ij}$$

$$\Pi_{ij} \mapsto (1 + \gamma) \Pi_{ij}. \quad (2d)$$

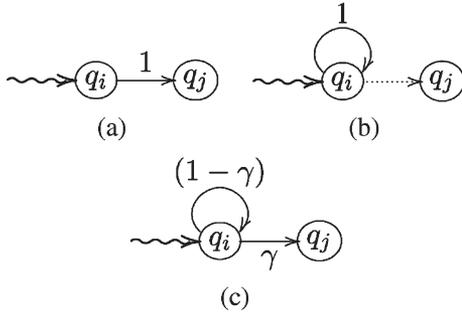


Fig. 1. Probabilistic disabling of finite-state Markov chains. (a) Enabled transition. (b) Disabled transition. (c) Probabilistically disabled transition.

*Remark 3:* The concept of probabilistic “disabling” was first reported in [17] and has widely been used in the supervisory control of discrete event systems [18], [20]. However, the use of such formalism to control-engineered swarms has not been reported yet.

In the sequel, the analysis is restricted to finite-state Markov chains for which every transition with nonzero occurrence probability could be controlled in the sense of Definition 3 for an arbitrary choice of the parameter  $\gamma$  in  $[0, 1]$ .

For a physical agent modeled by an irreducible Markov chain, the preceding assumption implies that the agent is capable of deciding to execute a particular transition with a specific probability, e.g., seven out of ten times it happens to reach the state at which the particular transition is defined. The self-loop arising from disabling simply ensures that state changes via such disabled transitions are prevented; so if the agent attempts to execute a “disabled” transition, then it comes back to the originating state. In the sequel, we associate states with operational behaviors. The preceding notion of controllability does not pose any serious restriction on the behavior types that can be defined in this framework; however, the agent must be free to choose which behavior it will execute next; this behavioral transition must not be uncontrollable. This is equivalent to assuming that there are no uncontrollable edges in the agent graph defined in the sequel.

*Definition 4 (Controlled Descendant):* Let the Markov chain  $G = (Q, \Pi, \wp^0)$  be derived from a given finite-state Markov chain  $\tilde{G} = (Q, \tilde{\Pi}, \wp^0)$ . Then,  $G$  is defined to be a controlled descendant of  $\tilde{G}$  if the following condition is satisfied:

$$\forall i \neq j, \quad \tilde{\Pi}_{ij} = 0 \implies \Pi_{ij} = 0. \quad (3)$$

A descendant  $G$  is obtained by applying probabilistic disabling or enabling to one or more transitions of the Markov chain  $\tilde{G}$ .

*Definition 5 (Agent in Swarm Modeling):* An agent is a connected digraph  $\mathcal{A} = (Q, \Delta)$ , where each state  $i \in Q$  represents a distinct predefined behavior, and  $\Delta \in \{0, 1\}^{\text{CARD}(Q) \times \text{CARD}(Q)}$  is a matrix such that  $\Delta_{ij} = 1$  implies that there exists a controllable transition from state  $i$  to state  $j$ .

The matrix  $\Delta$  in Definition 5 specifies state transitions (i.e., behavior switching) of the agent  $\mathcal{A}$ 's state (i.e., behavior) in the sense that  $\mathcal{A}$  decides to continue in the current state or make a transition to another state. The probabilities of state transitions constitute a (finite-state) irreducible Markov chain, with the irreducibility property following from the connected-

ness of the agent graph. Without any control, it is assumed that the state transition probabilities are uniformly distributed over the defined transitions at each state, and it follows from the connectedness of the agent graph that the uncontrolled agent corresponds to an *irreducible* Markov chain. Therefore, the specifications of switching probabilities must obey the constraint that no transition, which represents a behavior switch, with nonzero probability is defined at a state if the corresponding edge does not exist in the graph of the agent. The notion is formalized next.

*Definition 6:* Let the uncontrolled behavior of an agent  $A = (Q, \Delta)$  be represented by an irreducible Markov chain  $G^0 = (Q, \Pi^0, \wp^0)$ , where each row of  $\Pi^0$  reflects the uniform probability distribution over the defined set of transitions from the state that the particular row corresponds to (see preceding discussion). Let the agent be controlled by modifying the state-based transition probabilities, i.e., specifying an irreducible Markov chain  $G = (Q, \Pi, \wp^0)$ . Then, following Definition 4, a control policy is defined to be admissible if  $G$  is a controlled descendant of  $G^0$ .

*Definition 7:* In the sense of Definition 5, a homogeneous swarm  $\mathcal{S}$  is defined to be a collection of independent identical agents  $A = (Q, \Delta)$ , each of which is represented by the same (finite-state) irreducible Markov chain  $G = (Q, \Pi, \wp^0)$ . Formally

$$\mathcal{S} = \{G^\alpha : \alpha \in \mathbb{X}\} \quad (4)$$

such that  $G^\alpha = G$ , and  $\mathbb{X}$  is an index set.

Note that no restriction is imposed on the cardinality of the index set  $\mathbb{X}$ ; hence, the swarm can be finite, countably infinite, or uncountable.

*Notation 2:* Denoting the cardinality of the state set  $Q$  as  $\text{CARD}(Q)$ , the following notation is used for the collection of orthonormal basis vectors for the space  $\mathbb{R}^{\text{CARD}(Q)}$ :

$$\mathcal{B} = \left\{ \nu^i \in \mathbb{R}^{\text{CARD}(Q)} : \nu_j^i = \delta_{ij} \right\}. \quad (5)$$

Note that if the cardinality of the state set is  $n$ , then  $\nu^i$  is a vector in  $\mathbb{R}^n$ , whereas  $\nu_j^i$  denotes the  $j$ th entry in the vector  $\nu^i$ .

*Definition 8 (Local State):* The local or intensive state  $q^\alpha(t) \in \mathcal{B}$  for the swarm  $\mathcal{S} = \{G^\alpha : \alpha \in \mathbb{X}\}$  at time  $t \in [0, \infty)$  is defined as

$$q^\alpha(t)|_i = \begin{cases} 1, & \text{if } q_i \text{ is the current state for agent } G^\alpha \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

*Definition 9 (Observed State):* The observed or extensive state  $q^{\mathcal{S}}(t)$  of a swarm  $\mathcal{S}$  at time  $t$  is defined as

$$q^{\mathcal{S}}(t) = \lim_{\mathbb{Y} \rightarrow \mathbb{X}} \frac{1}{\mu(\mathbb{Y})} \int_{\alpha \in \mathbb{Y} \subseteq \mathbb{X}} q^\alpha(t) d\mu(\alpha) \quad (7)$$

where  $q^\alpha(t)$  is the local swarm state at time  $t$ , and  $\mu$  is the appropriate measure for the index set  $\mathbb{X}$ . Note that  $\mu$  is the counting measure if  $\mathbb{X}$  is finite or countable; if  $\mathbb{X}$  is a continuum, then  $\mu$  is the appropriate Lebesgue measure.

If the behavior of the swarm is ergodic, then the time average of the local state for a given  $\alpha \in \mathbb{X}$  converges to the ensemble average at equilibrium, i.e.,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t q^\alpha(\tau) d\tau = \wp \quad (8a)$$

$$\lim_{\mathbb{Y} \rightarrow \mathbb{X}} \frac{1}{\mu(\mathbb{Y})} \int_{\alpha \in \mathbb{Y} \subseteq \mathbb{X}} q^\alpha(t) d\mu = \wp \quad \forall t. \quad (8b)$$

### A. Analogy With Statistical Mechanics

In the terminology of statistical mechanics, a microstate [19] of the swarm constitutes a detailed enumeration of the local states of each individual agent.

Specifically, if the total number of (identical) agents in the swarm is  $N$ , then the total number of possible microstates ( $\Omega$ ) is given by

$$\Omega = \text{CARD}(Q)^N. \quad (9)$$

For example, even for two-state agents, the number of available microstates for a swarm of 100 such agents is  $2^{100}$ , which is ominously large. The observed measurable state of the swarm is the average effect of the local states of the member agents. Thus, for a given observed state, there exists a large number of possible configurations or microstates that realize the observed measurable state. In this paper, the swarm microstates are not considered to be important; the observed measurable state is what needs to be controlled. Such an approach embodies the fundamental philosophy that the individual agents do not have to locally meet the performance requirements; the global objective is what is solely important.

In the statistical mechanical framework, a homogeneous swarm is visualized as an ensemble of identical microsystems that are the individual agents. As previously stated, a microstate of the system is an enumeration of the local states of each agent in the swarm, which is not externally observable and, more importantly, cannot individually be controlled. The state of the swarm in the sense of Definition 9 is the average effect of all the microstates [19] and is, thus, an observable macrostate. The analogy is largely similar to the case of an ideal gas, where the microstates correspond to the detailed enumeration of position and momentum of each gas molecule, whereas the microcanonical ensemble average [19], [21] simply enumerates a fraction of particle population in each energy level without considering which particular molecules are at what level at any point in time. Thus, this ensemble average can be identified with the observed swarm state. It is generally impossible to control the energy levels of individual gas molecules; but the gas temperature, which is a function of the state of the microcanonical ensemble, can be controlled at ease. The driving philosophy presented in this paper is to formulate an implementable policy that controls the observed swarm state without accessing the agent microstates. In the proposed algorithm, an external supervisor computes the state transition (i.e., behavior switching)

probabilities of the agents subject to the defined constraints. However, control communications are assumed to occur as general broadcasts, and no individual agents are individually controlled. The decision and control problem is formally stated in the sequel.

### B. Statement of the Control Problem

Following Definition 7, let  $\mathcal{S} = \{G^\alpha : \alpha \in \mathbb{X}\}$  be a homogeneous swarm, where  $G^\alpha = (Q, \Pi, \wp^0)$  is the irreducible Markov chain corresponding to the uncontrolled agent  $G^0 = (Q, \Pi^0, \wp^0)$  and a target observed state  $\wp^*$  for the swarm, where  $\sum_{i=1}^{\text{CARD}(Q)} \wp_i^* = 1$ , and  $\wp_j^* > 0 \quad \forall j$ . The problem is to synthesize a controlled descendant  $G^*$  of  $G^0$  such that the perturbed transition matrix  $\Pi^*$  satisfies the following conditions.

- 1)  $G^*$  is a finite-state ergodic Markov chain, i.e., the associated state transition matrix  $\Pi^*$  is irreducible and acyclic [13].
- 2) Ergodic projection matrix

$$\mathcal{C}(\Pi^*) = \begin{bmatrix} \cdots & \wp^* & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \wp^* & \cdots \end{bmatrix}.$$

The preceding conditions have the following implications.

- 1) Since  $G^*$  is a disabled descendant of  $G^0$ ,  $G^*$  has a nonzero transition probability of switching states via a particular event if the corresponding transition is defined in the underlying agent graph.
- 2) It follows from (8a) and (8b) that the desired swarm state is achieved at equilibrium under the assumption of ergodicity.
- 3) The irreducibility of  $G^*$  (i.e., irreducibility of  $\Pi^*$ ) implies that the limiting time average of the local state of each agent is independent of the initial conditions. The fact that we demand  $G^*$  to be ergodic guarantees that the time average of the local states coincides with the observed swarm state in the limit for a sufficiently large swarm. This follows from the fact that for an ergodic chain with a transition probability matrix  $\Pi^*$  and for a sequence  $\{x^{[m]}\}$  of unity-sum-normalized vectors in  $\mathbb{R}^n$  satisfying  $\forall m \in \mathbb{N}, x^{[m]} = x^{[m-1]}\Pi^*$ , we have

$$\lim_{m \rightarrow \infty} x^{[m]} = \wp^*. \quad (10)$$

The necessity of a large agent population for convergence of the observed swarm state will be clarified in the sequel.

## III. MAIN RESULTS

This section presents an analytical formulation of the supervised self-organization algorithm and addresses the associated issues of computational complexity. In the sequel, unless otherwise mentioned, the cardinality of the state set is assumed to be  $n$ .

### A. Derivation of the Self-Organization Algorithm

This section formulates a recursive algorithm to solve the control problem stated in Section II-B. To this end, two supporting lemma and two theorems are presented.

*Definition 10:* Let  $\varphi^* \in \mathbb{R}^n$  be a unity-sum-normalized non-negative vector, and let  $\Pi \in \mathbb{R}^{n \times n}$  be an irreducible stochastic matrix with the stationary probability vector  $\varphi$ . A perturbation  $\bar{\Pi} \in \mathbb{R}^{n \times n}$  of the irreducible stochastic matrix  $\Pi$  is defined as follows:

$$\bar{\Pi} = \Pi + \mathcal{K}\mathcal{E}[\Pi - \mathbb{I}] \Rightarrow \bar{\Pi} - \mathbb{I} = [\mathbb{I} + \mathcal{K}\mathcal{E}][\Pi - \mathbb{I}] \quad (11)$$

where

$$\mathcal{E}_{ij} \triangleq \delta_{ij}(\varphi_i - \varphi_i^*) \quad (12a)$$

$$\mathcal{K}_{ij} \triangleq \begin{cases} \frac{\delta_{ij}}{\varphi_i^*}, & \text{if } \mathcal{E}_{ii} < 0 \\ 0, & \text{otherwise} \end{cases}. \quad (12b)$$

*Remark 4:* The following properties hold based on (11), (12a), and (12b) in Definition 10 and the facts stated in Remark 1.

- 1)  $\mathcal{K}_{ii}\mathcal{E}_{ii} \in (-1, 0]$ , and, hence,  $\mathcal{K}_{ii}\mathcal{E}_{ii} \leq \Pi_{ii}/1 - \Pi_{ii}$  because  $\Pi_{ii} \in [0, 1)$ .
- 2)  $\Pi_{ij} \geq 0 \Rightarrow \bar{\Pi} \geq 0$ , i.e.,  $\bar{\Pi}$  is a nonnegative matrix.

*Lemma 1:* The perturbation  $\bar{\Pi}$  in Definition 10 is an irreducible stochastic matrix.

*Proof:* Since both  $\mathcal{K}$  and  $\mathcal{E}$  are diagonal matrices, and  $\Pi$  is a stochastic matrix, it follows that

$$\sum_j \bar{\Pi}_{ij} = \sum_j \Pi_{ij} + \mathcal{K}_{ii}\mathcal{E}_{ii} \left( \sum_j \Pi_{ij} - 1 \right) = 1 \forall i. \quad (13)$$

The stochasticity of  $\bar{\Pi}$  is established by combining (13) with Property 3 in Remark 4. The irreducibility of  $\bar{\Pi}$  is proved next. Let  $\hat{\varphi}$  be an elementwise nonnegative vector representing a direction in the eigenspace of  $\bar{\Pi}$  corresponding to its unity eigenvalue, i.e.,  $\hat{\varphi}[\bar{\Pi} - \mathbb{I}] = 0$ . Such a  $\hat{\varphi}$  is guaranteed to exist for all stochastic matrices [22]. Equation (11) yields

$$\hat{\varphi}[\mathbb{I} + \mathcal{K}\mathcal{E}][\Pi - \mathbb{I}] = 0.$$

Since  $\varphi$  is unique,  $\hat{\varphi}[\mathbb{I} + \mathcal{K}\mathcal{E}] = \varphi$  with a scalar multiplicity of 1. Then, it follows from Definition 10 and Property 2 of Remark 4 that the stationary probability vector of  $\bar{\Pi}$  is obtained as

$$\bar{\varphi} \triangleq \frac{\hat{\varphi}}{\|\hat{\varphi}\|_1} = \frac{1}{\sum_i \hat{\varphi}_i} \hat{\varphi} \quad (14)$$

where  $\hat{\varphi} \equiv \varphi[\mathbb{I} + \mathcal{K}\mathcal{E}]^{-1}$ .

Since  $\bar{\Pi}$  has a unique stationary probability vector  $\bar{\varphi}$  that is positive elementwise, it is irreducible [13], [16]. ■

*Lemma 2:* The stationary probability vector  $\bar{\varphi}$  of the stochastic matrix  $\bar{\Pi}$  satisfies the following strict inequality:

$$\|\bar{\varphi} - \varphi_*\|_\infty < \|\varphi - \varphi_*\|_\infty \quad (15)$$

where  $\|x\|_\infty$  is the max norm of the finite-dimensional vector  $x$ .

*Proof:* It follows from (14) and Property 2 of Remark 4 that

$$\|\hat{\varphi}\|_1 = \sum_i \frac{\varphi_i}{1 + \mathcal{K}_{ii}\mathcal{E}_{ii}} \quad (16)$$

$$= \sum_{i:\mathcal{E}_{ii}<0} \frac{\varphi_i}{1 + \mathcal{K}_{ii}\mathcal{E}_{ii}} + \sum_{i:\mathcal{E}_{ii}\geq 0} \varphi_i. \quad (17)$$

An application of the bounds on  $\mathcal{K}_{ii}$  [see (12a) and (12b)] in (17) results in

$$\|\hat{\varphi}\|_1 > \sum_{i:\mathcal{E}_{ii}<0} \varphi_i + \sum_{i:\mathcal{E}_{ii}\geq 0} \varphi_i = \sum_i \varphi_i = 1. \quad (18)$$

Usage of the identity  $[\mathbb{I} + \mathcal{K}\mathcal{E}]^{-1} = \mathbb{I} - \mathcal{K}\mathcal{E}[\mathbb{I} + \mathcal{K}\mathcal{E}]^{-1}$  yields

$$\begin{aligned} \tilde{\varphi} - \varphi^* &= \varphi - \varphi^* - \varphi\mathcal{K}\mathcal{E}[\mathbb{I} + \mathcal{K}\mathcal{E}]^{-1} \\ \implies \tilde{\varphi} - \varphi^* &= (\varphi - \varphi^*)[\mathbb{I} - A] \end{aligned} \quad (19)$$

where it follows from (12a) and (12b) that

$$\begin{aligned} A_{ij} &= \frac{\delta_{ij}\mathcal{K}_{ii}\varphi_i}{1 + \mathcal{K}_{ii}(\varphi_i - \varphi_i^*)} \\ \implies A_{ii} &= \frac{\mathcal{K}_{ii}\varphi_i}{1 + \mathcal{K}_{ii}\mathcal{E}_{ii}} \in \{0, 1\}. \end{aligned} \quad (20)$$

It is noted from (16) that

$$\begin{aligned} 1 - \|\hat{\varphi}\|_1 &= 1 - \sum_i \frac{\varphi_i}{1 + \mathcal{K}_{ii}\mathcal{E}_{ii}} \\ &= \sum_i \left( \varphi_i - \frac{\varphi_i}{1 + \mathcal{K}_{ii}\mathcal{E}_{ii}} \right) = \sum_i \left( \frac{\varphi_i\mathcal{K}_{ii}\mathcal{E}_{ii}}{1 + \mathcal{K}_{ii}\mathcal{E}_{ii}} \right) \\ &= \begin{bmatrix} \vdots \\ \mathcal{E}_{ii} \\ \vdots \end{bmatrix}^T \begin{bmatrix} \ddots & & 0 \\ & A_{ii} & \\ 0 & & \ddots \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \\ &= (\varphi - \varphi^*)A \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}. \end{aligned}$$

Then, it follows from (14) and (19) that

$$\begin{aligned} \bar{\varphi} - \varphi^* &= \frac{\tilde{\varphi}}{\|\hat{\varphi}\|_1} - \varphi^* \\ &= \frac{\varphi - \varphi^*}{\|\hat{\varphi}\|_1} + \left( \frac{1 - \|\hat{\varphi}\|_1}{\|\hat{\varphi}\|_1} \right) \varphi^* \\ &= \left( \frac{\varphi - \varphi^*}{\|\hat{\varphi}\|_1} \right) [\mathbb{I} - A] + \frac{1}{\|\hat{\varphi}\|_1} (\varphi - \varphi^*)A\boldsymbol{\eta}\varphi^* \\ &= \left( \frac{\varphi - \varphi^*}{\|\hat{\varphi}\|_1} \right) [\mathbb{I} - A + A\boldsymbol{\eta}\varphi^*] \\ &= \left( \frac{\varphi - \varphi^*}{\|\hat{\varphi}\|_1} \right) W \end{aligned} \quad (21)$$

where  $\boldsymbol{\eta} \triangleq [1 \ \cdots \ 1]^T$ , and  $W \triangleq [\mathbb{I} - A + A\boldsymbol{\eta}\varphi^*]$ .

Next, it is shown that  $W$  in (21) is a stochastic matrix. It is obvious from (20) that the off-diagonal elements of  $W$  are nonnegative, and the diagonal elements are given as

$$W_{ii} = 1 - A_{ii} + A_{ii}\varphi_i^* = \begin{cases} 1, & \text{if } A_{ii} = 0 \\ \varphi_i^*, & \text{if } A_{ii} = 1. \end{cases}$$

Furthermore, since  $\eta\varphi^*$  is a stochastic matrix of rank 1 with all rows identically equal to  $\varphi^*$ , it follows that

$$\sum_j W_{ij} = 1 - A_{ii} + A_{ii} \sum_j \varphi_j^* = 1. \quad (22)$$

Therefore,  $W$  is a stochastic matrix, and, hence, the induced norm  $\|W\|_\infty = 1$ . Using the inequality  $\|\widehat{\varphi}\|_1 > 1$  from (18), (21) yields

$$\|\widehat{\varphi} - \varphi^*\|_\infty \leq \left( \frac{\|\varphi - \varphi^*\|_\infty}{\|\widehat{\varphi}\|_1} \right) \|W\|_\infty < \|\varphi - \varphi^*\|_\infty. \quad (23)$$

The proof is now complete.  $\blacksquare$

*Theorem 1:* Let  $\varphi^* \in \mathbb{R}^n$  be a unity-sum-normalized non-negative vector (for  $n > 1$ ), and let  $\Pi \in \mathbb{R}^{n \times n}$  be an irreducible stochastic matrix. Then, the recursive procedure

$$\Pi^{[r+1]} = \Pi^{[r]} + \mathcal{K}^{[r]} \mathcal{E}^{[r]} [\Pi^{[r]} - \Pi], \quad \Pi^{[0]} = \Pi \quad (24)$$

where  $\mathcal{E}^{[r]}$  and  $\mathcal{K}^{[r]}$  satisfy the conditions specified in (12a) and (12b), has the following.

- 1) Iteratively estimates an irreducible stochastic matrix  $\Pi^*$  with  $\varphi^* \Pi^* = \varphi^*$ , i.e., we have

$$\lim_{r \rightarrow \infty} \varphi^{[r]} \Pi^{[r]} = \lim_{r \rightarrow \infty} \varphi^{[r]} \lim_{r \rightarrow \infty} \Pi^{[r]} = \varphi^* \Pi^* = \varphi^*.$$

- 2)  $\|\varphi^{[r]} - \varphi^*\|_\infty$  strictly monotonically converges to zero.
- 3)  $\forall i \neq j, \Pi_{ij}^0 = 0 \implies \Pi_{ij}^* = 0$ .

*Proof:* It follows from Lemma 2 that  $\{\|\varphi^{[r]} - \varphi^*\|_\infty\}_{r \in \mathbb{N}}$  is a strictly monotonically decreasing sequence. The nonnegativity of norm implies that this sequence necessarily converges, which in turn implies that  $\{\varphi^{[r]}\}_{r \in \mathbb{N}}$  is a Cauchy sequence in  $\mathbb{R}$ . Recalling (21), it is noted that

$$\begin{aligned} \varphi^{[r+1]} - \varphi^{[r]} &= (\varphi^{[r]} - \varphi^*) \left( \frac{W}{\|\widehat{\varphi}^{[r]}\|_1} - \mathbb{I} \right) \\ \implies \|\varphi^{[r+1]} - \varphi^{[r]}\|_\infty &\left\| \left[ \frac{W}{\|\widehat{\varphi}^{[r]}\|_1} - \mathbb{I} \right]^{-1} \right\|_\infty \geq \|\varphi^{[r]} - \varphi^*\|_\infty \\ \implies \|\varphi^{[r+1]} - \varphi^{[r]}\|_\infty &\left( \frac{\|\widehat{\varphi}^{[r]}\|_1}{\|\widehat{\varphi}^{[r]}\|_1 - 1} \right) \geq \|\varphi^{[r]} - \varphi^*\|_\infty \end{aligned} \quad (25)$$

where (25) follows from the stochasticity of  $W$ , and the inequality  $\|\widehat{\varphi}^{[r]}\|_1 > 1$ , which implies that  $(\|\widehat{\varphi}^{[r]}\|_1 - 1/\|\widehat{\varphi}^{[r]}\|_1)[\mathbb{I} - (W/\|\widehat{\varphi}^{[r]}\|_1)]^{-1}$  is a stochastic matrix with unity infinity norm. Then, we have

$$\|\varphi^{[r+1]} - \varphi^{[r]}\|_\infty \geq \|\varphi^{[r]} - \varphi^*\|_\infty \left( 1 - \frac{1}{\|\widehat{\varphi}^{[r]}\|_1} \right).$$

For proof of statement 2 in the theorem by contradiction, let us assume that

$$\lim_{r \rightarrow \infty} \|\varphi^{[r]} - \varphi^*\|_\infty = \epsilon > 0. \quad (26)$$

The strict monotonicity of the sequence  $\{\|\varphi^{[r]} - \varphi^*\|_\infty\}_{r \in \mathbb{N}}$  implies that

$$\|\varphi^{[r]} - \varphi^*\|_\infty > \epsilon \quad \forall r \in \mathbb{N}. \quad (27)$$

It is claimed that

$$\|\widehat{\varphi}^{[r]}\|_1 > 1 + \epsilon \quad (28)$$

which follows from (12a), (12b), and (17) that  $\|\widehat{\varphi}^{[r]}\|_1 = \sum_i \max(\varphi_i^{[r]}, \varphi_i^*)$ . Then, (28) yields

$$\begin{aligned} \|\varphi^{[r+1]} - \varphi^{[r]}\|_\infty &\geq \epsilon \left( 1 - \frac{1}{1 + \epsilon} \right) \quad \forall r \in \mathbb{N} \\ \implies \|\varphi^{[r+1]} - \varphi^{[r]}\|_\infty &\geq \frac{\epsilon^2}{1 + \epsilon} \quad \forall r \in \mathbb{N} \end{aligned} \quad (29)$$

which contradicts the fact that  $\{\varphi^{[r]}\}_{r \in \mathbb{N}}$  forms a Cauchy sequence. Hence, it is concluded that  $\epsilon = 0$ , which implies that  $\|\varphi^{[r]} - \varphi^*\|_\infty$  monotonically converges to zero. The proof of statements 1 and 2 is now complete.

Statement 3 follows from noting that the recursive procedure guarantees

$$\Pi_{ij}^{[r]} = 0 \implies \Pi_{ij}^{[r+1]} = 0 \quad \forall r \in \mathbb{N} \quad \text{and} \quad \forall i \neq j. \quad \blacksquare$$

Theorem 1 [see (24)] computes the stochastic matrix  $\Pi^{[r^*]}$  after  $r_*$  iterations such that the resulting stationary distribution for  $\Pi^{[r^*]}$  coincides with the target distribution  $\varphi^*$  in the limit. Starting from an arbitrary initial distribution  $x^{[0]}$ , a necessary condition for the global swarm behavior to converge to the target distribution is as follows:

$$\forall m \in \mathbb{N} \cup \{0\}, \quad x^{[m+1]} = x^{[m]} \Pi^{[r^*]} \implies \lim_{m \rightarrow \infty} x^{[m]} = \varphi^*. \quad (30)$$

The preceding condition is guaranteed only if the matrix  $\Pi^{[r^*]}$  is irreducible, acyclic, and stochastic. However, the stochastic irreducible matrix  $\Pi^{[r^*]}$  is not guaranteed to be acyclic. This property is achieved via modification of  $\Pi^{[r^*]}$  via the following lemma.

*Lemma 3:* Given an  $(n \times n)$  irreducible stochastic matrix  $\Pi$ , the modified matrix  $P \triangleq (1 - \theta)\Pi + \theta\mathbb{I}_n$  is stochastic, irreducible, and acyclic, where the scalar parameter  $\theta \in (0, 1)$ , and  $\mathbb{I}_n$  is the  $(n \times n)$  identity matrix.

*Proof:* The stochasticity property of  $P$  is retained because both  $\Pi$  and  $\mathbb{I}_n$  are nonnegative matrices and  $\theta > 0$ . Therefore,  $P$  is a nonnegative matrix with each row sum being unity, which implies that  $P$  is stochastic. The irreducibility property of  $P$  is retained because every eigenvector of  $P$  is an eigenvector of  $\Pi$ . Since  $\Pi$  is irreducible, no nonnegative left eigenvector of  $\Pi$  has a zero coordinate. Therefore, no nonnegative left eigenvector of  $P$  has a zero coordinate, which implies that

$P$  is irreducible [13]. Finally, since  $\text{Trace}(P) \geq n\theta > 0$ , the irreducible stochastic matrix  $P$  must be acyclic [13] regardless of whether  $\Pi$  is cyclic or acyclic. ■

Algorithm 1, which is called Supervised Self-organization of Swarms ( $\mathcal{S}^3$ ), is formulated based on Theorem 1 after modifying the irreducible matrices  $\Pi^{[r]}$  to the respective irreducible and acyclic (i.e., ergodic) matrices based on Lemma 3 to solve the control problem stated in Section II-B. The  $\mathcal{S}^3$  algorithm is discussed in Section III-B; the simulation results and the associated computational aspects are presented in Section III-C. The particular choice of  $\theta$  in applications of Lemma 3 is related to the convergence rate of the solution, which is a topic of future research.

**Algorithm 1:** Supervised Self-organization of Swarms ( $\mathcal{S}^3$ )

**input:**  $\Pi^0, \varphi^*, \text{Tolerance}$   
**output:**  $\Pi^*$

```

1 begin
2   Set  $r = 0, \text{Error} = 1;$  /* Initialize */
3   while ( $\text{Error} > \text{Tolerance}$ ) do
4     Compute  $\mathcal{E}^{[r]}$ ;
5     Compute  $\mathcal{K}^{[r]}$ ;
6      $\Pi^{[r+1]} = \Pi^{[r]} + \mathcal{K}^{[r]}\mathcal{E}^{[r]}[\Pi^{[r]} - \mathbb{I}];$ 
7     Compute  $\bar{\varphi}^{[r+1]}$ ; /* stationary
      distribution [12]*/
8      $\text{Error} = \|\bar{\varphi}^{[r+1]} - \varphi^*\|_{\infty};$ 
9     Set  $r = r + 1;$ 
10  endw
11  Choose  $\theta \in (0, 1);$ 
12   $\Pi^* = (1 - \theta)\Pi^{[r-1]} + \theta\mathbb{I};$  /* Modification of
      the transition probability matrix via
      Lemma 3 */
13 end

```

### B. Key Observations on Applicability of Algorithm $\mathcal{S}^3$

The following key observations are made regarding the applicability of  $\mathcal{S}^3$  in practical scenarios.

There are no restrictions on the choice of the target probability vector  $\varphi^*$ , except for the requirement that it needs to be a bona fide probability vector, i.e.,  $\forall i \varphi_i^* \geq 0$ , and  $\sum_i \varphi_i^* = 1$ . It is important to consider the situation where the chosen  $\varphi^*$  has one or more zero entries. It is noted that, for any irreducible stochastic matrix, the stationary vector is elementwise strictly positive [13], [16], and Algorithm 1 yields an irreducible stochastic matrix at every iteration (see Lemma 1). In such a situation, the target may not exactly be achieved for any finite number of iterations of Algorithm 1; however, one can approach the target with arbitrary precision as guaranteed by Theorem 1.

The perturbed transition matrix  $\Pi^*$ , which is computed by Algorithm 1, is not the only possible solution to the control problem stated in Section II-B. Algorithm 1 is merely an efficient procedure for computing an admissible solution to the problem. An additional solution criteria may be useful. An example is the perturbed stochastic matrix for which the magnitude of the second largest eigenvalue is the least among all the admissible solutions. The rationale is that the magnitude of the second largest eigenvalue of the transition matrix is directly

related to the settling time of the corresponding Markov chain. The swarm population controlled with such a transition matrix is likely to converge to its stationary distribution faster than any other admissible solution [12], [23], [24]. Thus, while the present procedure computes an admissible solution, a procedure that yields a solution under the previously stated constraint can be regarded as the optimal solution with respect to the settling time (also known in the literature as “mixing time”). Future work will address the integration of such additional constraints with the algorithm presented in this paper.

For a given swarm  $\mathcal{S} = \{G^\alpha : \alpha \in \mathbb{X}\}$ , the proposed algorithm causes the observed or the extensive state  $q^{\mathcal{S}}(t)$  to converge to the specified target  $\varphi^*$ . However, the local or intensive state  $q^\alpha(t)$ , which corresponds to the state of the particular agent  $\alpha \in \mathbb{X}$ , does not “converge” at any time  $t \in [0, \infty)$ . Each agent can be in one and only one of the local states at any time, i.e.,  $q^\alpha(t) \in \mathcal{B}, \forall t \in [0, \infty)$ . In fact,  $q^\alpha(t)$  keeps switching such that the time average of the local states, for any agent, converges to the target distribution  $\varphi^*$ . However, in spite of these local fluctuations, if the computed transition matrix  $\Pi^*$  is irreducible and acyclic, then the observed swarm state does not change after convergence and remains at or close to  $\varphi^*$  for all times. For example, in a swarm of two-state agents, if the target is given by [0.2 0.8], then it is guaranteed that at any point in time, after convergence, about 20% of the agents can be found in state 1 and 80% in state 2, but it is not certain which particular agents will be in the specific states at a specific instant. In addition, each agent will be either in [1 0] or in [0 1] at any point in time and will keep visiting both states as dictated by the computed  $\Pi^*$ . Thus, the swarm behavior emerges as a global effect of the local agent actions. This is a direct consequence of the homogeneity of the swarm, the connectedness of the agent graph (see Definition 5), the ergodicity of the transition matrix computed by  $\mathcal{S}^3$ , and the implicit assumption of a *sufficiently large swarm population*. The requirement of a large population arises from the fact that each individual agent can be at one and only one state at any time. Thus, if there are only two agents in the previously described two-state “swarm,” then the only swarm states that can possibly be observed are [1 0], [0 1] and [0.5 0.5], and this set of possible observed swarm states increases with the swarm size. Then, the maximum error between a particular coordinate  $i$  of the target vector, which corresponds to the target stationary probability at state  $i$ , and the value achieved at equilibrium is bounded above by  $1/N$ , where  $N$  is the population size. This follows from the fact that the increase (decrease) in the state probability by adding (subtracting) one agent from a given state, averaged over time, is given by  $1/N$ . Therefore, the infinity or max norm of the error between an arbitrary target vector and the distribution achieved at equilibrium should approximately be inversely proportional to the swarm size. This approximate hyperbolic relationship, namely, (Max Norm of Error  $\times$  Number of Agents  $\simeq$  constant), is validated by simulation, as shown in Fig. 2. Fig. 2(a) shows a four-state connected agent graph to simulate the effects of increasing swarm size on the convergence rate. The fitted curve in Fig. 2(b) validates the preceding hyperbolic relationship with a small root mean square error (RMSE) of 0.031. Fig. 2(c) illustrates the observed max norm of error as a function of

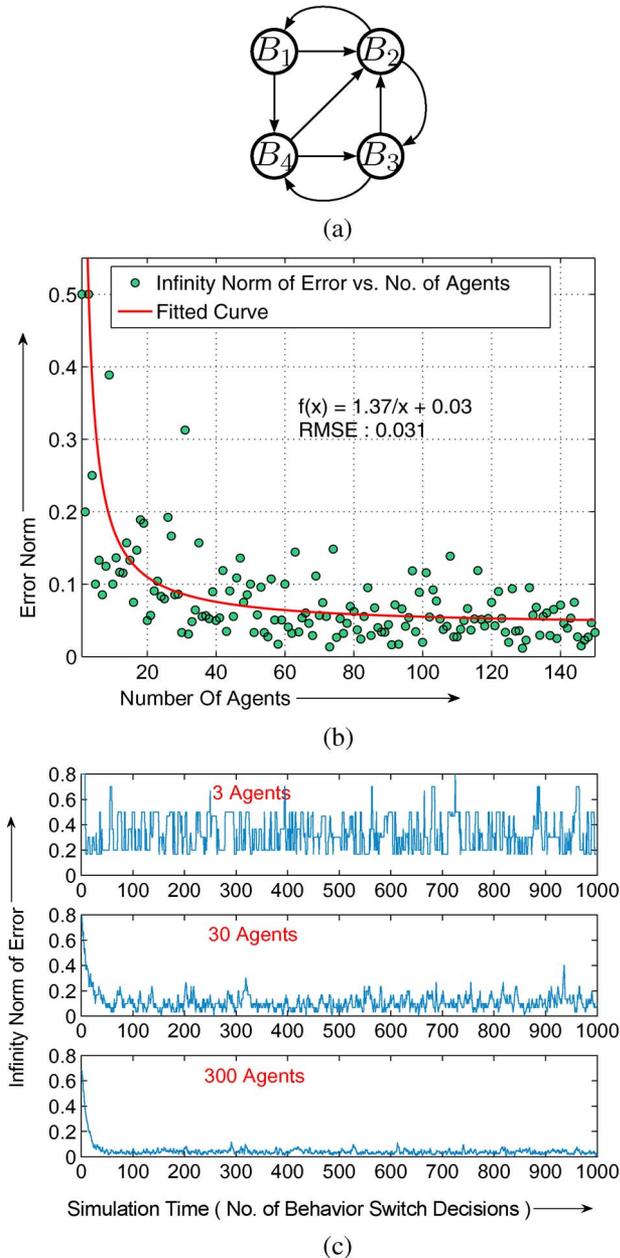


Fig. 2. Effect of population size. (a) Agent graph. (b) Hyperbolic relationship between the infinity norm of error between the observed state and the target distribution as a function of the swarm size. (c) Behavior convergence for different swarm sizes as a function of simulation time.

simulation steps (or decision ticks) for the cases of three, 30, and 300 agents, respectively. Section III-D further illustrates this concept with a simulated example of a mobile sensor field, where the swarm rapidly converges to the detected activity hotspots in the field. However, this apparent coordination is a pseudoemergent phenomenon, since under the current model there is little coordination at the local level with the agents operating independent of their neighbors. The case where the agent behaviors are allowed to be statistically dependent is a topic of future work.

The above-described dependence of convergence on swarm size is related to the concept of the thermodynamic limit [19], [21] in the sense of statistical mechanics—the larger the

system, the smaller the probability of thermal microfluctuations resulting in more uniform macroscopic properties. In light of the preceding discussion, the addition of new agents would increase the size of the swarm and leads to better convergence due to a closer approximation of the thermodynamic limit. Removal of agents, either intentionally or due to failures, has the reverse effect. The pertinent question is that how many agents can be deleted before reaching an unacceptable performance degradation. This issue is related to the size of the remaining number of functional agents in the swarm. The control policy will have an acceptable implementation as long as the swarm population is large enough to spatially reflect the locally time-averaged behavior with acceptable precision, i.e., as long as the swarm population is large enough for the validity of the ergodicity assumption stated in (8a) and (8b).

*Example 1:* Let us consider a four-state connected agent graph [see Fig. 2(a)] and a target probability vector  $\varphi^*$  as enumerated in (3.1), i.e.,

$$\varphi^* = \begin{bmatrix} 0.2 \\ 0.0 \\ 0.5 \\ 0.3 \end{bmatrix}$$

$$\Pi^{[50]} = \begin{bmatrix} 0.9 & 0.04 & 0 & 0.06 \\ 0.73 & 0.01 & 0.26 & 0 \\ 0 & 0.04 & 0.9 & 0.06 \\ 0 & 0 & 0.13 & 0.87 \end{bmatrix}$$

$$\varphi^{[50]} = [0.20 \quad 0.02 \quad 0.49 \quad 0.29]. \quad (31)$$

Application of  $\mathcal{S}^3$  leads to the transition matrix  $\Pi^{[50]}$  after 50 iterations, which is sufficiently close to the target distribution. Note that  $\Pi^{[50]}$  is irreducible and acyclic as verified by computing the eigenvalues ( $\lambda_1 = 1, \lambda_2 = 0.9, \lambda_3 = 0.81, \lambda_4 = -0.03$ ), which reveals that there is only one eigenvalue on the unit circle. This implies that the Markov chain corresponding to  $\Pi^{[50]}$  is ergodic. In particular, any initial distribution is guaranteed to converge via repeated right multiplication by  $\Pi^{[50]}$  to  $\varphi^{[50]}$ . This is verified by assuming a uniform initial distribution  $[0.25 \ 0.25 \ 0.25 \ 0.25]$ , and the subsequent convergence is shown in Fig. 3. The abscissa in Fig. 3 is annotated as “Number of Decision Steps,” since each right multiplication step can be assumed to be a local behavior switching decision by the agents. Note that the convergence shown in Fig. 3 is not sufficient for the convergence of the observed swarm state. The latter requires a sufficiently large population size, as explained in detail above.

### C. Computational Complexity of $\mathcal{S}^3$

The simulation results indicate the high computational efficiency of the proposed control algorithm. A possible bottleneck in Algorithm 1 is the computation of the stable probability vector  $\varphi^{[r]}$  in each step of the iterative refinement process. The computation of the stationary distributions of irreducible Markov chains is well studied [12], and efficient algorithms have been reported. Numerical results are illustrated in Fig. 4, which suggest a quadratic bound on the asymptotic run-time complexity. Fig. 4 was generated by considering

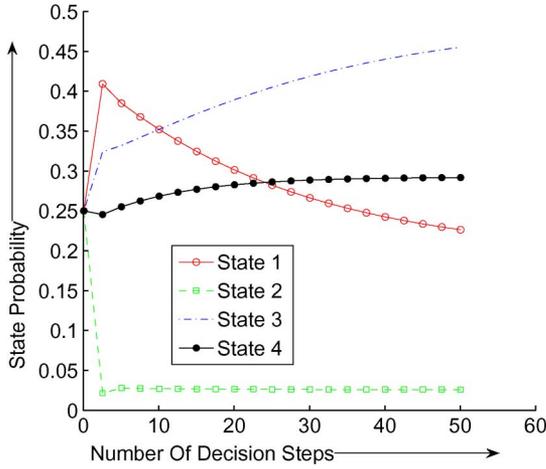


Fig. 3. Convergence of a uniform initial distribution to the target vector as a function of the number of right multiplication by  $\Pi^{[50]}$ .

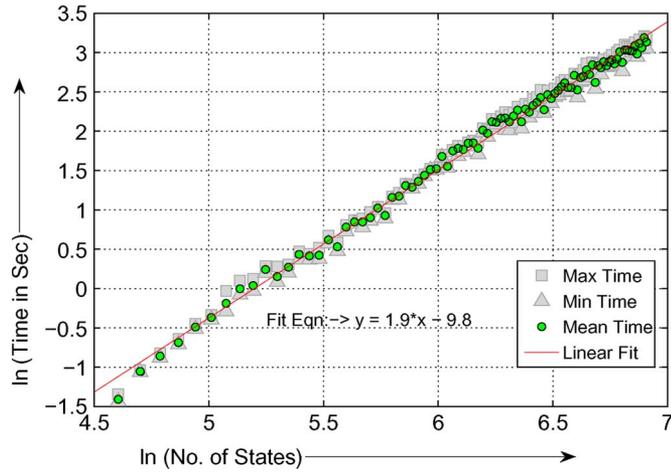


Fig. 4. Algorithm execution time versus number of states.

1000 randomly chosen input stochastic matrices for each  $N$  in the range  $[100, 1000]$ . The “Mean Time” is the average runtime over the input samples considered for each  $N$ .

Fig. 5 exhibits a rapid decrease of the error norm with respect to the number of iterations in a sample case (note the logarithmic scale of the ordinate). The convergence rate is determined by the second largest eigenvalue of the computed stochastic matrix  $\Pi^*$ . The rapid convergence observed in the simulation results from the fact that we initiate  $\mathcal{S}^3$  Algorithm 1 with  $\Pi^0$  for which the transition probabilities are defined to be uniform over the transition set at each state. Estimation of a rigorous upper bound on the second largest eigenvalue of the computed transition matrix  $\Pi^*$  is a topic of future research.

#### D. Supervised Self-Organization of a Simulated Mobile Sensor Field

This section presents the simulated control of a mobile sensor field as an application of the proposed control algorithm. The swarm is assumed to consist of a large number of identical mobile sensors in a sufficiently large rectangular 2-D grid of dimension  $N \times M$ . In this scenario, the parameters are chosen as

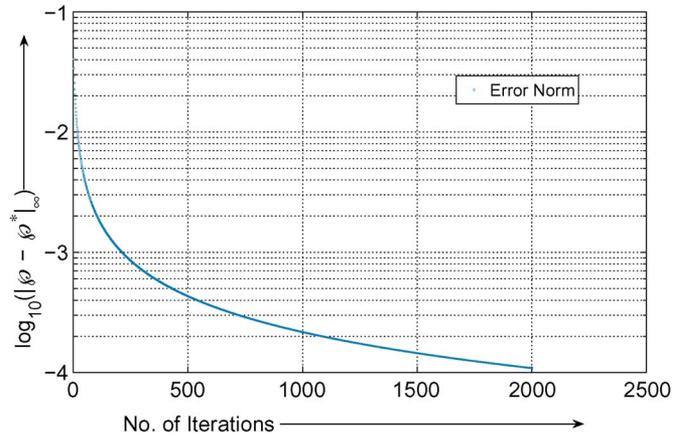


Fig. 5. Illustrative case: error norm versus number of iterations.

$N = M = 1000$ . The graph for each agent, therefore, consists of  $N \times M$  nodes that represent the spatial grid locations. Each agent may decide to move to any of the adjacent grid locations in one step, which implies that each agent state has eight defined edges. The uncontrolled agents correspond to the irreducible finite-state Markov chain  $G^0$  with the transition probabilities defined to be uniform over the outgoing edges at each state. The simulation experiment is conducted as follows.

At the beginning, the majority of the sensors are concentrated at the top-left corner of the grid [see Fig. 6(a)] with the exception of a few that are randomly distributed over the remaining grid locations. Detection of activity by the latter at three different locations is communicated to an external supervisor that determines that the sensor field density needs to peak at the corresponding “hotspots.” The supervision policy is computed via the  $\mathcal{S}^3$  algorithm and is communicated to the sensors via a general broadcast. The eight plates in Fig. 6 exhibit the progressive effects of the swarm control algorithm to achieve the goal of locating the “hotspots.” The field density gradually moves out from the top-left corner [see Fig. 6, plates (a) and (b)] and self-organizes to peak at the desired locations [see Fig. 6, plates (g) and (h)].

Invoking the ideal gas analogy from statistical mechanics, this control is achieved without considering the system microstates, i.e., without referring or accessing the detailed enumeration of the local states of each agent in the swarm.

It is important to note that the control broadcast occurs only once each time the tactical scenario (i.e., location of the hotspots) changes. No communication with the supervisor is necessary for the subsequent self-organization process. Each agent needs to know its current location, which can be obtained from onboard GPS, and communication with neighbors is unnecessary in this example.

#### IV. SUMMARY AND FUTURE WORK

This paper has presented an algorithm called ( $\mathcal{S}^3$ ) for the global behavior supervision of homogeneous engineered swarms of potentially unbounded population sizes. The swarm is modeled as an arbitrarily large collection of independent identical finite-state agents. The algorithm is for computing the necessary perturbations in the switching probabilities for the

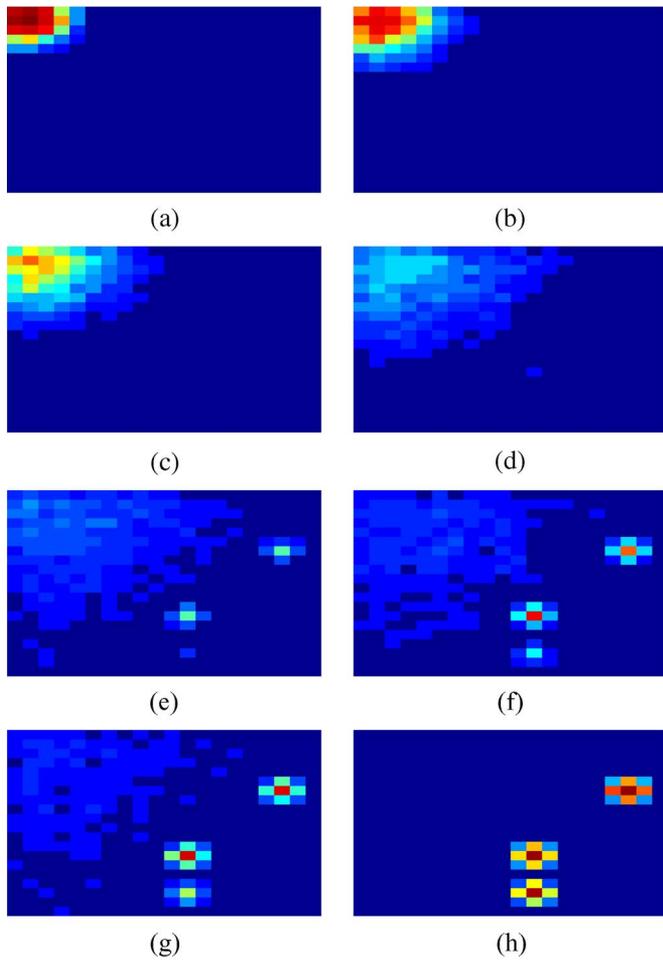


Fig. 6. Self-organization of the sensor field subsequent to control broadcast. The initial field is concentrated at the top-left corner as seen in (a), which subsequently organizes to peak at the activity hotspots. The spatial convergence is illustrated by plates (b)–(h).

individual agents that guarantee convergence of the observed swarm state to a desired distribution. A simulation example is presented to illustrate the concept.

Future research is planned to pursue the following areas.

- 1) Choice of the scalar parameter  $\theta$  in applications of Lemma 3: This issue is related to the convergence rate of the solution and will be investigated for obtaining optimal solutions to the control problem stated in this paper.
- 2) Estimation of a rigorous upper bound on the magnitude of the second largest eigenvalue of the computed transition matrix: The convergence rate (e.g., settling time, also known as “mixing time”) of the overall swarm is faster if the bound is smaller and slows down as it approaches unity from below.
- 3) Generalization to swarms of interacting agents: This issue is analogous to extending the ideal gas formulation in basic thermodynamics to that of real gases and is of enormous importance from the implementation standpoint.
- 4) Execution of the  $S^3$  in a distributed manner: The control algorithm presented in this paper requires a general broadcast when the desired target distribution changes, after which the organization proceeds in a distributed

manner without central supervision. Execution of the  $S^3$  algorithm itself in a distributed manner would eliminate the need for general broadcast. This task involves estimation of the stationary distribution of a Markov chain via only performing local computations.

- 5) Implementation in real-world systems: It requires resolution of several issues such as observation delays, uncertainties (e.g., due to interagent communication, noise corruption and actuation), and broadcast bandwidth limitations prior to deployment.

## REFERENCES

- [1] F. Zhang, M. Goldgeier, and P. S. Krishnaprasad, “Control of small formations using shape coordinates,” in *ICRA*, 2003, pp. 2510–2515.
- [2] M. Egerstedt and X. Hu, “Formation constrained multi-agent control,” *IEEE Trans. Robot. Autom.*, vol. 17, no. 6, pp. 947–951, Dec. 2001.
- [3] L. Chaimowicz, N. Michael, and V. Kumar, “Controlling swarms of robots using interpolated implicit functions,” in *ICRA*, 2005, pp. 2487–2492.
- [4] J. Cortes, S. Martinez, T. Karatas, and F. Bullo, “Coverage control for mobile sensing networks,” *IEEE Trans. Robot. Autom.*, vol. 20, no. 2, pp. 243–255, Apr. 2004.
- [5] S. Kalantar and U. Zimmer, “Distributed shape control of homogeneous swarms of autonomous underwater vehicles,” *Auton. Robot.*, vol. 22, no. 1, pp. 37–53, Jan. 2007.
- [6] C. Belta and V. Kumar, “Abstraction and control for groups of robots,” *IEEE Trans. Robot.*, vol. 20, no. 5, pp. 865–875, Oct. 2004.
- [7] M. Hartbauer and H. Römer, “A novel distributed swarm control strategy based on coupled signal oscillators,” *Bioinspiration Biomimetics*, vol. 2, no. 3, pp. 42–56, Sep. 2007.
- [8] Y. C. Tan and B. Bishop, “Evaluation of robot swarm control methods for underwater mine countermeasures,” in *Proc. 36th Southeastern Symp. Syst. Theory*, 2004, pp. 294–298.
- [9] B. Walter, A. Sannier, D. Reiners, and J. Oliver, “UAV swarm control: Calculating digital pheromone fields with the GPU,” *JDMS*, vol. 3, no. 3, pp. 167–176, 2006.
- [10] J. L. Dohner, G. R. Eisler, B. J. Driessen, and J. Hurtado, “Cooperative control of vehicle swarms for acoustic target localization by energy flows,” *Trans. ASME, J. Dyn. Syst. Meas. Control*, vol. 126, no. 4, pp. 891–895, Dec. 2004.
- [11] W. J. Stewart, *Introduction to the Numerical Solution of Markov Chains*. Princeton, NJ: Princeton Univ. Press, 1994.
- [12] W. Stewart, *Computational Probability: Numerical Methods for Computing Stationary Distribution of Finite Irreducible Markov Chains*. New York: Springer-Verlag, 1999.
- [13] A. Berman and R. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. Philadelphia, PA: SIAM, 1994, corrected republication, with supplement, of work first published in 1979 by Academic Press.
- [14] M. Iosifescu, *Finite Markov Processes and Their Applications*. New York: Wiley, 1980.
- [15] V. Romanovsky, *Discrete Markov Chains*. Groningen, The Netherlands: Noordhoff, 1978.
- [16] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*, 2nd ed. New York: Springer-Verlag, 1960.
- [17] M. Lawford and W. Wonham, “Supervisory control of probabilistic discrete event systems,” in *Proc. 36th Midwest Symp. Circuits Syst.*, 1993, pp. 327–331.
- [18] I. Chattopadhyay and A. Ray, “Language-measure-theoretic optimal control of probabilistic finite-state systems,” *Int. J. Control*, vol. 80, no. 8, pp. 1271–1290, Aug. 2007.
- [19] R. K. Pathria, *Statistical Mechanics*, 2nd ed. London, U.K.: Butterworth, 1996.
- [20] I. Chattopadhyay and A. Ray, “Renormalized measure of regular languages,” *Int. J. Control*, vol. 79, no. 9, pp. 1107–1117, 2006.
- [21] D. F. Styer, “What good is the thermodynamic limit?” *Amer. J. Phys.*, vol. 72, no. 1, pp. 25–29, Jan. 2004.
- [22] R. B. Bapat and T. E. S. Raghavan, *Nonnegative Matrices and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 1997.
- [23] M. Jerrum and A. Sinclair, “Approximating the permanent,” *SIAM J. Comput.*, vol. 18, no. 6, pp. 1149–1178, 1989.
- [24] L. Lovasz and M. Simonovits, “The mixing rate of Markov chains, an isoperimetric inequality, and computing the volume,” in *Proc. 31st Annu. Symp. Foundations Comput. Sci.*, vol. 1, Oct. 1990, pp. 346–354.



**Ishanu Chattopadhyay** received the Masters and Doctoral degrees in mechanical engineering from the Pennsylvania State University (Pennstate), University Park, in 2005 and 2006, respectively.

He is currently a Research Associate with the Networked Robotics and Systems Laboratory, Department of Mechanical Engineering, Pennstate. He is actively involved in several DoD projects funded by the U.S. Army and the U.S. Navy. His principal research interests include autonomous systems, distributed control of large-scale systems, robotics, sensor networks, and application of biological principles in control of human-engineered complex systems.



**Asok Ray** received the Ph.D. degree in mechanical engineering and the graduate degrees in each discipline of electrical engineering, mathematics, and computer science from Northeastern University, Boston, MA.

He previously held research and academic positions with the Massachusetts Institute of Technology, Cambridge, and Carnegie-Mellon University, Pittsburgh, PA, as well as research and management positions with GTE Strategic Systems Division, Charles Stark Draper Laboratory, and MITRE Corporation. Since July 1985, he has been with the Pennsylvania State University (Pennstate), University Park, where he is currently a Distinguished Professor of mechanical engineering. He has been a Senior Research Fellow with the National Aeronautics and Space Administration (NASA) Glenn Research Center under a National Academy of Sciences award. He has authored or coauthored over 450 research publications including 225 scholarly articles in refereed journals such as *TRANSACTIONS OF IEEE*, *American Society of Mechanical Engineers (ASME)*, and *American Institute of Aeronautics and Astronautics (AIAA)*, and research monographs. His research experience and interests include control and optimization of continuously varying and discrete-event dynamical systems, intelligent instrumentation for real-time distributed systems, and modeling and analysis of complex dynamical systems from thermodynamic perspectives in both deterministic and stochastic settings, as applied to aeronautics and astronautics, undersea vehicles and surface ships, power and processing plants, and robotics.

Dr. Ray is a Fellow of ASME and of the World Innovative Foundation (WIF).