2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

Supervised Self-organization of Large Homogeneous Swarms Using Ergodic Projections of Markov Chains★

Ishanu Chattopadhyay‡ Asok Ray ‡
ixc128@psu.edu axr2@psu.edu

Mechanical Engineering Department
The Pennsylvania State University
University Park, PA 16802, USA

***Abstract*— This paper formulates a self-organization algorithm to addresses the problem of emergent behavior supervision in engineered swarms of arbitrary population size. Based on collections of independent identical finite-state agents, the algorithm is derived to compute necessary perturbations in the local agents' behavior, which guarantees convergence to the desired observed state of the swarm. A simulation example illustrates the underlying concept.**

***Index Terms*— Swarms; Finite State Ergodic Markov Chains; Ergodic Projections; Discrete Event Systems;**

1. Introduction & Motivation

With recent advances in sensor technology and 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 & rescue operations to coordinated handling and transportation of large objects. Motion coordination among teams of autonomous agents has been studied extensively [1], [2] with special emphasis on formation control of large groups [3].

The major distinction between a large group of autonomous agents and a *swarm* is first clarified, and an example from biology 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 honey bees functions as a swarm, where removal of a few hundred workers have 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 above philosophy translates to having little or no performance requirements on individual agents and control of emergent 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 [2], [1] 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 carry over to an unbounded population size. Apparently, formulation of algorithms that reflect the above philosophy has not yet been reported in open literature.

This paper addresses mathematical modeling of engineered swarms as arbitrary collections of independent finite-state Markov chains and formulates an algorithm to realize emergent behavior supervision in sufficiently large non-interacting populations of agents. The analysis, presented here, does not address inter-agent communication, noise corruption, actuation delays, and the associated practical implementation issues, because the objective is to formulate a relatively simple and unambiguous approach that lays the framework rather than having one that attempts to offer a complete solution to the complex overall problem, which is a topic of future research.

The paper is organized in four sections including the present one. Section 2 states the necessary definitions and concepts and presents the formal statement of the control problem. Section 3 presents the main results along with an illustrative application example. The paper is summarized and concluded in Section 4 with recommendations for future work.

2. Preliminaries & Notations

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

Definition 2.1: *A finite-state homogeneous Markov chain is a pair $G = (Q, \Pi)$ where Q is a set of states with* $\textsc{Card}(Q) = n \in \mathbb{N}$*, and Π is the $(n \times n)$ stationary transition probability matrix such that $\forall i, j \in \{1, \cdots, n\}$, $\Pi_{ij} \geqq 0$ with $\sum_i \Pi_{ij} = 1$; and Π is called a stochastic matrix.*

Remark 2.1: *Every stochastic matrix Π has at least one unity eigenvalue; and all eigenvalues of Π are located within or on the unit disk.*

Definition 2.2: *A finite-state homogeneous Markov chain $G = (Q, \Pi)$, with $\textsc{Card}(Q) = n \geq 2$, 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 Π is strictly positive, i.e., $\Pi^{(k)}_{ij} > 0$; and Π is called an irreducible matrix [7]. An irreducible finite-state homogeneous Markov chain G is also called ergodic [7].*

★This work has been supported in part by the U.S. Army Research Laboratory and the U.S. Army Research Office (ARO) under Grant No. W911NF-07-1-0376, by the U.S. Office of Naval Research under Grant No. N00014-08-1-380, and by NASA under Cooperative Agreement No. NNX07AK49A. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

Remark 2.2: *For any $n \times n$ irreducible stochastic matrix Π 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 $\wp = [\wp_1\ \wp_2\ \cdots\ \wp_n]$ corresponding to the unique unity eigenvalue of Π is called the stationary probability vector, where $\sum_i \wp_i = 1$ and $\wp_j > 0\ \forall j$. The stationary probability vector $\wp$ has the following property.*

$$\lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k} \Pi^j = \begin{bmatrix} \cdots & \wp & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \wp & \cdots \end{bmatrix} \quad (1)$$

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

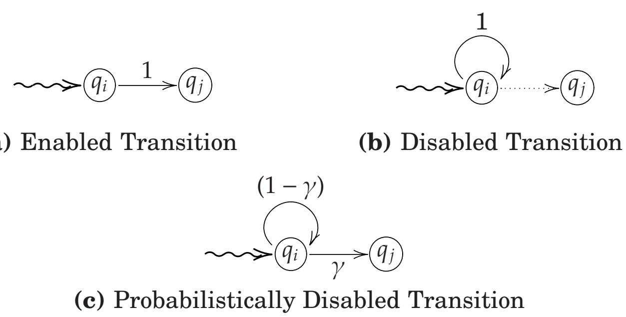

Fig. 1. Probabilistic Disabling of Finite-State Markov Chains

Definition 2.3: (Probabilistic Disabling and Enabling) *[8] Let $G = (Q, \Pi)$ be an ergodic Markov chain and let the ij^{th} probability be non-zero, i.e., $\Pi_{ij} > 0$. Referring to Fig. 1, the probabilistic disabling and enabling from the state i to the state j are respectively defined in terms of perturbations in the state transition matrix Π as*

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

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

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

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

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

In the sequel, the analysis is restricted to finite state Markov chains for which every transition with non-zero occurrence probability could be controlled in the sense of Definition 2.3 for an arbitrary choice of the parameter γ in $[0,1]$.

Definition 2.4: (Controlled Descendant) *Let the Markov chain $G = (Q, \Pi)$ be derived from a given finite-state Markov chain $\widetilde{G} = (Q, \widetilde{\Pi})$. Then, G is defined to be a controlled descendant of $\widetilde{G}$ if the following condition is satisfied.*

$$\forall i \neq j,\ \widetilde{\Pi}_{ij} = 0 \Longrightarrow \Pi_{ij} = 0 \quad (3)$$

A descendant G is obtained by applying probabilistic disablings or enablings to one or more transitions of the Markov chain $\widetilde{G}$.

Definition 2.5: (Agent in Swarm Modeling) *An agent is a connected graph $\mathcal{A} = (Q, \Delta)$, where each state $i \in Q$ represents a distinct predefined behavior and each transition $(i, j) \in \Delta$ represents a controllable transition from state i to state j.*

The matrix Δ in Definition 2.5 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) ergodic Markov chain. 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 *ergodic* Markov chain. Therefore, specifications of switching probabilities must obey the constraint that no transition, representing a behavior switch, with non-zero 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 2.6: *Let the uncontrolled behavior of an agent $A = (Q, \Delta)$ be represented by an ergodic Markov chain $G^0 = (Q, \Pi^0)$. Let the agent be controlled by specifying transition probabilities in an ergodic Markov chain $G = (Q, \Pi)$. Then, following Definition 2.4, a control policy is defined to be admissible if G is a controlled descendant of G^0.*

Definition 2.7: *In the sense of Definition 2.5, a homogeneous swarm $\mathscr{S}$ is defined to be a collection of independent identical agents $A = (Q, \Delta)$, each of which is represented by the same (finite-state) ergodic Markov chain $G = (Q, \Pi)$. Formally,*

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

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 denumerable.

Notation 2.1: *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 $\mathbb{R}^{\text{CARD}(Q)}$.*

$$\mathcal{B} = \{\boldsymbol{\nu}^i \in \mathbb{R}^{\text{CARD}(Q)} : \nu^i_j = \delta_{ij}\} \quad (5)$$

Definition 2.8: (Local State) *The local or the intensive state $q^\alpha(t) \in \mathcal{B}$ for the swarm $\mathscr{S} = \{\mathscr{G}^\alpha : \alpha \in \mathbb{X}\}$ at time $t \in [0, \infty)$ is defined as:*

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

Definition 2.9: (Observed State) *The observed or the extensive state $q^{\mathscr{S}}(t)$ of a swarm $\mathscr{S}$ at time t is defined as*

$$q^{\mathscr{S}}(t) = \lim_{\Upsilon\to\mathbb{X}} \frac{1}{\mu(\Upsilon)} \int_{\alpha\in\Upsilon\subseteq\mathbb{X}} q^\alpha(t)\ \mathrm{d}\mu(\alpha) \quad (7)$$

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

Under the assumption of ergodicity for a sufficiently large swarm, the time average of the local state for a given $\alpha \in \mathbb{X}$ converges to the ensemble average at an equilibrium point, *i.e.*,

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

$$\lim_{\Upsilon \to \mathbb{X}} \frac{1}{\mu(\Upsilon)} \int_{\alpha \in \Upsilon \subseteq \mathbb{X}} q^{\alpha}(t)\, \mathrm{d}\mu = \wp \ \forall t \tag{8b}$$

Remark 2.3: ***(A Statistical Mechanical Analogy)*** *In the statistical mechanical framework, a homogeneous swarm can be visualized as an ensemble of identical microsystems that are the individual agents. The agent states correspond to the microstates of the overall system, which are not observable externally. The state of the swarm in the sense of Definition 2.9 is the average effect of all the microstates and is thus an observable macrostate. The analogy is largely similar to the case of an ideal gas where the microstates correspond to the kinetic energies of the individual non-interacting molecules; and the observed macrostate is temperature that 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 can be controlled at ease. Hence, the driving philosophy, presented in this paper, is to formulate an implementable policy that controls the observed swarm state without accessing the agent microstates.*

An external supervisor manipulates state transition (i.e., behavior switching) probabilities of the agents. However, control communications are assumed to occur as general broadcasts and no individual agent can be controlled individually. Under this constraint, the decision & control problem is formally stated as follows.

A. Statement of the Control Problem

Following Definition 2.7, let $\mathscr{S} = \{G^{\alpha} : \alpha \in \mathbb{X}\}$ be a homogeneous swarm, where $G^{\alpha} = (Q, \Pi)$ is the ergodic Markov chain corresponding to the uncontrolled agent $G^0 = (Q, \Pi^0)$ and a target observed state $\wp^{\star}$ for the swarm, where $\sum_{i=1}^{\textsc{Card}(Q)} \wp_i^{\star} = 1$ and $\wp_j^{\star} > 0 \ \forall j$. The problem is to synthesize a controlled descendant $G^{\star}$ of G^0 such that the perturbed transition matrix $\Pi^{\star}$ satisfies the following conditions.

1) *$G^{\star}$ is a finite state ergodic Markov chain i.e., the associated state transition matrix $\Pi^{\star}$ is irreducible; and*
2) *Ergodic projection matrix* $C(\Pi^{\star}) = \begin{bmatrix} \cdots & \wp^{\star} & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \wp^{\star} & \cdots \end{bmatrix}$

The above conditions have the following implications.

1) Since $G^{\star}$ is a disabled descendant of G^0, i.e., $G^{\star}$ has a non-zero transition probability of switching states via a particular event if the corresponding transition is defined in the underlying agent graph.
2) It follows from Eqn. (8a) and Eqn. (8b) that the desired swarm state is achieved at equilibrium.
3) Ergodicity of $G^{\star}$ (*i.e.*, irreducibility of $\Pi^{\star}$) implies that the initial state of the agents is unimportant for convergence to the desired swarm state.

3. Main Results

This section presents analytical formulation of the supervised self-organization algorithm and addresses the associated issues of computational complexity.

A. Derivation of the Supervised Self-organization Algorithm

This subsection formulates a recursive algorithm to solve the control problem stated in Section 2-A. To this end, two supporting lemma and a theorem are presented.

Definition 3.1: *Let $\wp^{\star} \in \mathbb{R}^n$ be a ℓ_1 (i.e., sum)-normalized non-negative vector and let $\Pi \in \mathbb{R}^{n \times n}$ be an irreducible stochastic matrix with the stationary probability vector $\wp$. A perturbation $\overline{\Pi} \in \mathbb{R}^{n\times n}$ of the irreducible stochastic matrix Π is defined as follows.*

$$\overline{\Pi} = \Pi + \mathscr{K}\mathscr{E}\Big[\Pi - \mathbb{I}\Big] \ \Rightarrow \ \overline{\Pi} - \mathbb{I} = \Big[\mathbb{I} + \mathscr{K}\mathscr{E}\Big]\Big[\Pi - \mathbb{I}\Big] \tag{9}$$

$$\text{where } \mathscr{E}_{ij} \triangleq \delta_{ij}(\wp_i - \wp_i^{\star}) \tag{10a}$$

$$\text{and } \mathscr{K}_{ij} \triangleq \begin{cases} \frac{\delta_{ij}}{\wp_i^{\star}} & \text{if } \mathscr{E}_{ii} < 0 \\ 0 & \text{otherwise} \end{cases} \tag{10b}$$

Remark 3.1: *The following properties hold based on Eqns. (9), (10a) and (10b) in Definition 3.1 and the facts stated in Remark 2.1.*

1) *$\mathscr{K}_{ii}\mathscr{E}_{ii} \leq 0 \Rightarrow \mathscr{K}_{ii}\mathscr{E}_{ii} \leq \frac{\Pi_{ii}}{1-\Pi_{ii}}$ because $\Pi_{ii} \in [0,1)$.*
2) *$1 + \mathscr{K}_{ii}\mathscr{E}_{ii} \in (0,1]$ because $\mathscr{K}_{ii}\mathscr{E}_{ii} \in (-1,0]$.*
3) *$\overline{\Pi}_{ij} \geq 0 \Rightarrow \overline{\Pi}$ is a non-negative matrix.*

Lemma 3.1: *The perturbation $\overline{\Pi}$ in Definition 3.1 is an irreducible stochastic matrix.*

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

$$\sum_j \overline{\Pi}_{ij} = \sum_j \Pi_{ij} + \mathscr{K}_{ii}\mathscr{E}_{ii}(\sum_j \Pi_{ij} - 1) = 1 \ \forall i \tag{11}$$

Stochasticity of $\overline{\Pi}$ is established by combining Eqn. (11) with Property 3 in Remark 3.1. Irreducibility of $\overline{\Pi}$ is proved next. Let $\widehat{\wp}$ be an elementwise non-negative vector representing a direction in the eigenspace of $\overline{\Pi}$ corresponding to its unity eigenvalue, i.e., $\widehat{\wp}\Big[\overline{\Pi} - \mathbb{I}\Big] = 0$. Such a $\widehat{\wp}$ is guaranteed to exist for all stochastic matrices [9]. Eqn. (9) yields

$$\widehat{\wp}\Big[\mathbb{I} + \mathscr{K}\mathscr{E}\Big]\Big[\Pi - \mathbb{I}\Big] = 0$$

Since $\wp$ is unique, $\widehat{\wp}\Big[\mathbb{I} + \mathscr{K}\mathscr{E}\Big] = \wp$ with scalar multiplicity of 1. Then, it follows from Definition 3.1 and Property 2 of Remark 3.1 that the stationary probability vector of $\overline{\Pi}$ is obtained as

$$\overline{\wp} \triangleq \frac{\widehat{\wp}}{||\widehat{\wp}||_1} = \frac{1}{\sum_i \widehat{\wp}_i}\widehat{\wp} \ \text{ where } \ \widehat{\wp} = \wp\Big[\mathbb{I} + \mathscr{K}\mathscr{E}\Big]^{-1} \tag{12}$$

Since $\overline{\Pi}$ has a unique stationary probability vector $\overline{\wp}$ that is positive elementwise, it is irreducible [7]. ❐

Lemma 3.2: *The stationary probability vector $\overline{\wp}$ of the stochastic matrix $\overline{\Pi}$ satisfies the following strict inequality.*

$$||\overline{\wp} - \wp_{\star}||_{\infty} < ||\wp - \wp_{\star}||_{\infty} \tag{13}$$

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

Proof: It follows from Eqn. 12 and Property 2 of Remark 3.1 that

$$||\widehat{\wp}||_1 = \sum_i \frac{\wp_i}{1 + \mathscr{K}_{ii}\mathscr{E}_{ii}} \tag{14}$$

$$= \sum_{i:\mathscr{E}_{ii}<0} \frac{\wp_i}{1 + \mathscr{K}_{ii}\mathscr{E}_{ii}} + \sum_{i:\mathscr{E}_{ii}\geq 0} \wp_i \tag{15}$$

An application of the bounds on $\mathscr{K}_{ii}$ (See Eqns. (10a) and (10b)) in Eqn. (15) results

$$\|\widehat{\wp}\|_1 > \sum_{i:\mathscr{E}_{ii}<0} \wp_i + \sum_{i:\mathscr{E}_{ii}\geqq 0} \wp_i = \sum_i \wp_i = 1 \tag{16}$$

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

$$\widehat{\wp} - \wp^\star = \wp - \wp^\star - \wp\mathscr{K}\mathscr{E}\Big[\mathbb{I}+\mathscr{K}\mathscr{E}\Big]^{-1}$$
$$\Longrightarrow \widehat{\wp} - \wp^\star = \Big(\wp - \wp^\star\Big)\Big[\mathbb{I} - A\Big] \tag{17}$$

where it follows from Eqns. (10a) and (10b) that

$$A_{ij} = \frac{\delta_{ij}\mathscr{K}_{ii}\wp_i}{1+\mathscr{K}_{ii}(\wp_i - \wp_i^\star)} \;\Rightarrow\; A_{ii} = \frac{\mathscr{K}_{ii}\wp_i}{1+\mathscr{K}_{ii}\mathscr{E}_{ii}} \in \{0,1\} \tag{18}$$

It is noted from Eqn. 14 that

$$1 - \|\widehat{\wp}\|_1 = 1 - \sum_i \frac{\wp_i}{1+\mathscr{K}_{ii}\mathscr{E}_{ii}}$$
$$= \sum_i \Big(\wp_i - \frac{\wp_i}{1+\mathscr{K}_{ii}\mathscr{E}_{ii}}\Big) = \sum_i \Big(\frac{\wp_i\mathscr{K}_{ii}\mathscr{E}_{ii}}{1+\mathscr{K}_{ii}\mathscr{E}_{ii}}\Big)$$
$$= \begin{bmatrix} \vdots \\ \mathscr{E}_{ii} \\ \vdots \end{bmatrix}^T \begin{bmatrix} \ddots & & 0 \\ & A_{ii} & \\ 0 & & \ddots \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = (\wp - \wp^\star)A \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Then, it follows from Eqns.(12) and (17) that

$$\overline{\wp} - \wp^\star = \frac{\widehat{\wp}}{\|\widehat{\wp}\|_1} - \wp^\star = \frac{\wp - \wp^\star}{\|\widehat{\wp}\|_1} + \Big(\frac{1-\|\widehat{\wp}\|_1}{\|\widehat{\wp}\|_1}\Big)\wp^\star$$
$$= \Big(\frac{\wp - \wp^\star}{\|\widehat{\wp}\|_1}\Big)\Big[\mathbb{I} - A\Big] + \frac{1}{\|\widehat{\wp}\|_1}(\wp - \wp^\star)A\boldsymbol{\eta}\wp^\star$$
$$= \Big(\frac{\wp - \wp^\star}{\|\widehat{\wp}\|_1}\Big)[\mathbb{I} - A + A\boldsymbol{\eta}\wp^\star] = \Big(\frac{\wp - \wp^\star}{\|\widehat{\wp}\|_1}\Big)W \tag{19}$$

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

Next it is shown that W in Eqn. (19) is a stochastic matrix. It is obvious from Eqn. (18) that off-diagonal elements of W are non-negative and the diagonal elements are given as

$$W_{ii} = 1 - A_{ii} + A_{ii}\wp_i^\star = \begin{cases} 1 & \text{if } A_{ii} = 0 \\ \wp_i^\star & \text{if } A_{ii} = 1 \end{cases}$$

Furthermore, since $\boldsymbol{\eta}\wp^\star$ is a stochastic matrix of rank 1 with all rows identically equal to $\wp^\star$, it follows that

$$\sum_j W_{ij} = 1 - A_{ii} + A_{ii}\sum_j \wp_j^\star = 1 \tag{20}$$

Therefore, W is a stochastic matrix and hence the induced norm $\|W\|_\infty = 1$. Using the inequality $\|\widehat{\wp}\|_1 > 1$ from Eqn. (16), Eqn. (19) yields

$$\Big\|\overline{\wp} - \wp^\star\Big\|_\infty \leqq \Big(\frac{\|\wp - \wp^\star\|_\infty}{\|\widehat{\wp}\|_1}\Big)\|W\|_\infty < \|\wp - \wp^\star\|_\infty \tag{21}$$

The proof is now complete. ❐

Theorem 3.1: *Let $\wp^\star \in \mathbb{R}^n$ be a ℓ_1-normalized nonnegative vector (for n>1) and $\Pi \in \mathbb{R}^{n\times n}$ be an irreducible stochastic matrix. Then, the recursive procedure*

$$\Pi^{[r+1]} = \Pi^{[r]} + \mathscr{K}^{[r]}\mathscr{E}^{[r]}\Big[\Pi^{[r]} - \mathbb{I}\Big],\; \Pi^{[0]} = \Pi \tag{22}$$

where $\mathscr{E}^{[r]}$ and $\mathscr{K}^{[r]}$ satisfy the conditions specified in Eqns (10a) and (10b),

1) *iteratively estimates an irreducible stochastic matrix $\Pi^\star$ with $\wp^\star\Pi^\star = \wp^\star$ i.e. we have*

$$\lim_{r\to\infty}\wp^{[r]}\Pi^{[r]} = \lim_{r\to\infty}\wp^{[r]}\lim_{r\to\infty}\Pi^{[r]} = \wp^\star\Pi^\star = \wp^\star$$

2) *$\|\wp^{[r]} - \wp^\star\|_\infty$ strictly monotonically converges to zero.*

3) *$\forall i \neq j,\; \Pi^0_{ij} = 0 \Longrightarrow \Pi^\star_{ij} = 0$*

Proof: It follows from Lemma 3.2 that $\{\|\wp^{[r]} - \wp^\star\|_\infty\}_{r\in\mathbb{N}}$ is a strictly monotonically decreasing sequence . Non-negativitity of norm implies that this sequence necessarily converges, which in turn implies that $\{\wp^{[r]}\}_{r\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. Recalling Eq.(19), it is noted that

$$\wp^{[r+1]} - \wp^{[r]} = \Big(\wp^{[r]} - \wp^\star\Big)\Big(\frac{W}{\|\widehat{\wp}^{[r]}\|_1} - \mathbb{I}\Big)$$
$$\Rightarrow \Big\|\wp^{[r+1]} - \wp^{[r]}\Big\|_\infty \Big\|\Big[\frac{W}{\|\widehat{\wp}^{[r]}\|_1} - \mathbb{I}\Big]^{-1}\Big\|_\infty \geqq \Big\|\wp^{[r]} - \wp^\star\Big\|_\infty$$
$$\Rightarrow \Big\|\wp^{[r+1]} - \wp^{[r]}\Big\|_\infty \left(\frac{\Big\|\widehat{\wp}^{[r]}\Big\|_1}{\Big\|\widehat{\wp}^{[r]}\Big\|_1 - 1}\right) \geqq \Big\|\wp^{[r]} - \wp^\star\Big\|_\infty \tag{23}$$

where Eq. (23) follows from stochasticity of W and the inequality $\Big\|\widehat{\wp}^{[r]}\Big\|_1 > 1$, which implies that $\frac{\|\widehat{\wp}^{[r]}\|_1 - 1}{\|\widehat{\wp}^{[r]}\|_1}\Big[\mathbb{I} - \frac{W}{\|\widehat{\wp}^{[r]}\|_1}\Big]^{-1}$ is a stochastic matrix with unity infinity norm. Then we have,

$$\Big\|\wp^{[r+1]} - \wp^{[r]}\Big\|_\infty \geqq \Big\|\wp^{[r]} - \wp^\star\Big\|_\infty \left(1 - \frac{1}{\Big\|\widehat{\wp}^{[r]}\Big\|_1}\right) \tag{24}$$

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

$$\lim_{r\to\infty}\|\wp^{[r]} - \wp^\star\|_\infty = \epsilon > 0 \tag{25}$$

Strict monotonicity of the sequence $\{\|\wp^{[r]} - \wp^\star\|_\infty\}_{r\in\mathbb{N}}$ implies that

$$\Big\|\wp^{[r]} - \wp^\star\Big\|_\infty > \epsilon \quad \forall r \in \mathbb{N} \tag{26}$$

It is claimed that

$$\Big\|\widehat{\wp}^{[r]}\Big\|_1 > 1 + \epsilon \tag{27}$$

which follows from Eqns. (10a), (10b), and (15)) that $\Big\|\widehat{\wp}^{[r]}\Big\|_1 = \sum_i \max(\wp_i^{[r]}, \wp_i^\star)$. Then, Eqn. (27) yields

$$\Big\|\wp^{[r+1]} - \wp^{[r]}\Big\|_\infty \geqq \epsilon\Big(1 - \frac{1}{1+\epsilon}\Big) \quad \forall r \in \mathbb{N}$$
$$\Rightarrow \Big\|\wp^{[r+1]} - \wp^{[r]}\Big\|_\infty \geqq \frac{\epsilon^2}{1+\epsilon} \quad \forall r \in \mathbb{N} \tag{28}$$

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

Statement (3) follows from noting that the recursive procedure guarantees:

$$\Pi^{[r]}_{ij} = 0 \Longrightarrow \Pi^{[r+1]}_{ij} = 0 \quad \forall r \in \mathbb{N} \text{ and } \forall i \neq j$$

❐

Algorithm 1 is formulated based on Theorem 3.1 to solve the control problem stated in Section 2-A. The computational aspects of Algorithm 1, called Supervised Self-organization of Swarms (S^3), are discussed in Section 3-B.

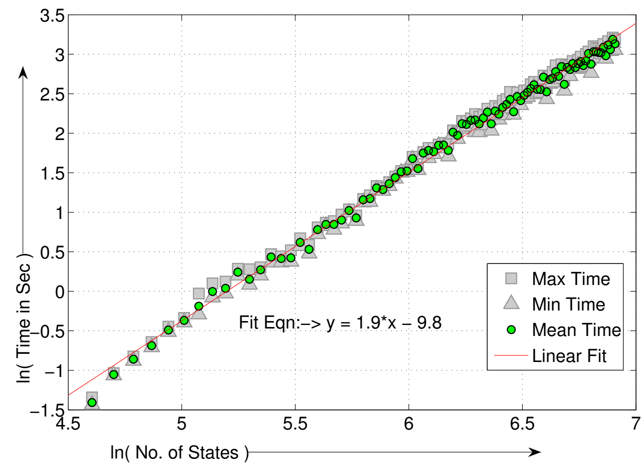

Fig. 2. Algorithm Execution Time versus Number of States

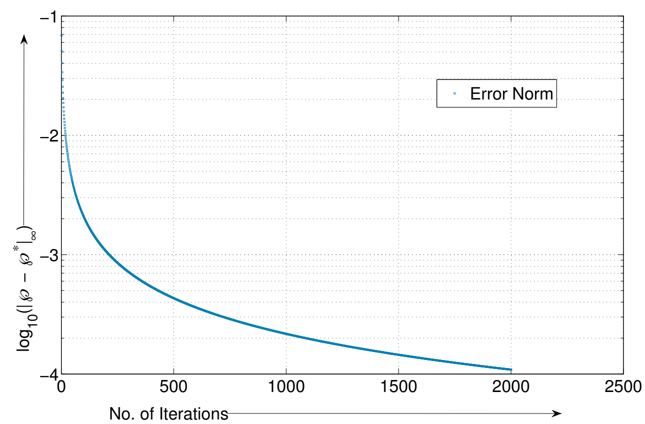

Fig. 3. Illustrative Case: Error Norm versus Number of Iterations

Algorithm 1: Supervised Self-organization of Swarms (S^3)

input : $\Pi^0, \wp^\star$, **Tolerance**
output: $\Pi^\star$
begin
 Set $r = 0$, **Error** $= 1$; /* Initialize */
 while (**Error** > **Tolerance**) **do**
  Compute $\mathscr{E}^{[r]}$;
  Compute $\mathscr{K}^{[r]}$;
  $\Pi^{[r+1]} = \Pi^{[r]} + \mathscr{K}^{[r]}\mathscr{E}^{[r]}\left[\Pi^{[r]} - \mathbb{I}\right]$;
  Compute $\overline{\wp}^{[r+1]}$; /* stationary distribution [10] */
  Error $= \|\overline{\wp}^{[r+1]} - \wp^\star\|_\infty$;
  Set $r = r + 1$;
 endw
 $\Pi^\star = \Pi^{[r-1]}$;
end

B. Computational Complexity

Simulation results indicate high computational efficiency of the proposed control algorithm. A possible bottleneck in Algorithm 1 is computation of the stable probability vector $\wp^{[r]}$ in each step of the iterative refinement process. The computation of stationary distributions of irreducible Markov chains is well-studied [10] and efficient algorithms have been reported. Numerical results are illustrated in Figure 2, which suggest a quadratic bound on the asymptotic run-time complexity. Figure 2 was generated by considering 1,000 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.

Figure 3 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 $\Pi^\star$. The rapid convergence observed in simulation results from the fact that we initiate S^3 Algorithm 1) with Π^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 $\Pi^\star$ is a topic of future research.

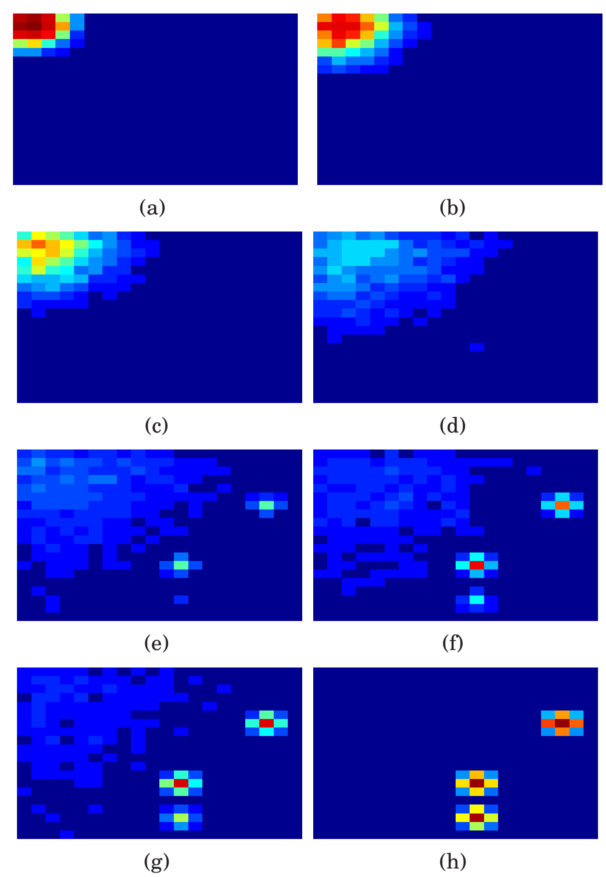

Fig. 4. Self-organization of the Sensor Field subsequent to Control Broadcast. The initial field is concentrated at the top-right corner as seen in (a) which subsequently organizes to peak at the activity hotspots. The spatial convergence is illustrated by the plates (b)-(h). A similar simulation result for a smaller problem can be viewed in the movie form by running the *"swarm.avi"* supplied as an attachment

C. Supervised Self-organization of A Simulated Mobile Sensor Field

This section presents 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 $N = M = 1000$. The graph for each agent therefore consists of $N \times M$ nodes representing the spatial grid locations. Each agent may decide to move to any of the adjacent grid locations in one step implying that each agent state has 8 defined edges. The uncontrolled agents correspond to the ergodic 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 Figure 4(a)) with the exception of a few that are distributed randomly over the remaining grid locations. Detection of activity by the the latter at three different locations is communicated to an external supervisor which determines that the sensor field density needs to peak at the corresponding "hotspots". The supervision policy is computed via the S^3 Algorithm and is communicated to the sensors via a general broadcast. The eight plates in Figure 4 exhibit 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 Figure 4, plates (a) & (b)) and self-organizes to peak at the desired locations (See Figure 4 plates (g) & (h)). A similar simulation result for a smaller problem can be viewed in the movie form by running the *"swarm.avi"* supplied as an attachment. It is seen in the attached movie that as the activity hotspots change (at a slower time scale), the supervised sensor swarm rapidly self-organizes to track the active zones. 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.

4. Summary, Conclusions, and Future Work

This paper presents an algorithm, called Supervised Self-organization of Swarms (S^3), for supervision of emergent behavior supervision of homogeneous engineered swarms of potentially unbounded population size. The swarm is modeled as an arbitrary collection of independent identical finite-state agents. The algorithm for computing the necessary perturbations in the switching probabilities for the 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) *Estimation of a rigorous upper bound on the magnitude of the second largest eigenvalue of the computed transition matrix:* The convergence rate of the overall swarm is faster if the bound is smaller and slows down as it approaches unity from below.
2) *Generalization to swarms of interacting agents:* This 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.
3) *Investigation of the possibility of executing the proposed algorithm in a distributed manner rather than on an external supervisor*
4) *Resolution of practical implementation issues concerning observation delays, broadcast bandwidth limitations etc. prior to deployments in real-world systems*

References

[1] F. Zhang, M. Goldgeier, and P. S. Krishnaprasad, "Control of small formations using shape i coordinates," in *ICRA*, pp. 2510–2515, IEEE, 2003.

[2] M. Egerstedt and X. Hu, "Formation constrained multi-agent control," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 6, pp. 947–951, 2001.

[3] L. Chaimowicz, N. Michael, and V. Kumar, "Controlling swarms of robots using interpolated implicit functions," in *ICRA*, pp. 2487–2492, IEEE, 2005.

[4] J. Cortes, S. Martinez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 2, pp. 243–255, 2004.

[5] S. Kalantar and U. R. Zimmer, "Distributed shape control of homogeneous swarms of autonomous underwater vehicles," *Autonomous Robotics*, vol. 22, no. 1, pp. 37–53, 2007.

[6] C. Belta and V. Kumar, "Abstraction and control for groups of robots," *IEEE Transactions on Robotics*, vol. 20, no. 5, pp. 865–875, 2004.

[7] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1994. Corrected republication, with supplement, of work first published in 1979 by Academic Press.

[8] I. Chattopadhyay and A. Ray, "Language-measure-theoretic optimal control of probabilistic finite-state systems," *Int. J. Control*, August,2007.

[9] R. B. Bapat and T. E. S. Raghavan, *Nonnegative Matrices and Applications*. Cambridge: Cambridge University Press, 1997.

[10] W. Stewart, *Computational Probability: Numerical methods for computing stationary distribution of finite irreducible Markov chains*. New York: Springer, 1999.