

Distributed Learning and Cooperative Control for Multi-Agent Systems [★]

Jongeun Choi ^a, Songhwai Oh ^b, Roberto Horowitz ^c

^a*Department of Mechanical Engineering, Michigan State University, East Lansing, MI 48824-1226, USA*

^b*Electrical Engineering and Computer Science, School of Engineering, University of California, Merced, CA 95344-2039, USA*

^c*Department of Mechanical Engineering, University of California at Berkeley, CA 94720-1740, USA*

Abstract

This paper presents a novel class of resource-constrained multi-agent systems for cooperatively estimating an unknown field of interest and locating peaks of the field. Each agent is resource constrained and has limited capabilities in terms of sensing, computation, and communication; hence a centralized approach is not desirable and not feasible. We propose an algorithm for distributed learning and cooperative control of a multi-agent system so that a global goal of the overall system is achieved from locally acting agents. The proposed algorithm is executed by each agent independently to estimate an unknown field of interest from noisy measurements and to coordinate multiple agents in a distributed manner to discover peaks of the unknown field. Each mobile agent maintains its own local estimate of the field and updates the estimate cooperatively using its own measurements and measurements from nearby agents. Then each agent moves towards peaks of the field using the gradient of its estimated field. Agents are coordinated using a distributed rule so that they avoid collision while maintaining communication connectivity. The proposed algorithm is based on a recursive spatial estimation of an unknown field of interest using noisy measurements. We show that the closed-loop dynamics of the proposed multi-agent system can be transformed into a form of a stochastic approximation algorithm and prove its convergence using the Ljung's ordinary differential equation (ODE) approach. Our theoretical results are also verified in simulation.

Key words: Cooperative control; Multi-agent systems; Recursive parameter estimation

1 Introduction

In recent years, significant enhancements have been made in the areas of sensor networks and mobile sensing agents. Emerging technologies have been reported on coordination of mobile sensing agents [1–5]. Mobile sensing agents form an ad-hoc wireless communication network in which each agent is resource constrained, i.e., it operates under a short communication range, limited memory, and limited computational power. To perform various tasks such as exploration, surveillance, and environmental monitoring, distributed coordination is required for mobile sensing agents to adapt to environments to achieve a global goal. Among challenging problems in distributed coordination of mobile sensing agents, gradient climbing over an unknown field of interest

has attracted much attention of environmental scientists and control engineers [6,7]. This has numerous applications including environmental monitoring and toxic-chemical plume tracing. An interesting practical application is to trace harmful algal blooms in a lake. For certain environmental conditions, rapidly reproducing harmful algal blooms in lakes and in oceans can produce cyanotoxins [8]. Exposure to water contaminated with algal cyanotoxins causes serious acute and chronic health effects to humans and adverse effects to aquatic life [8,9]. The level of chlorophyll is a measure closely related to harmful algal blooms. Hence, there have been efforts to generate the estimated fields of chlorophyll over the areas of concern (Fig. 1). Having had the aforementioned motivation, the objective of our work is to develop theoretically-sound control algorithms for multi-agent systems to trace peaks of a scalar field of interest (for example, harmful algal blooms, temperature, pH, salinity, toxins, and chemical plumes etc.). In general, these scalar parameters provide rich information about quality of environments such as the air, lakes, and public water systems.

The most common approach to this tracing problem has been

* This paper was not presented at any IFAC meeting. Corresponding author Jongeun Choi Tel. +1-517-432-3164. Fax +1-517-353-1750.

Email addresses: jchoi@egr.msu.edu (Jongeun Choi), songhwai.oh@ucmerced.edu (Songhwai Oh), horowitz@me.berkeley.edu (Roberto Horowitz).

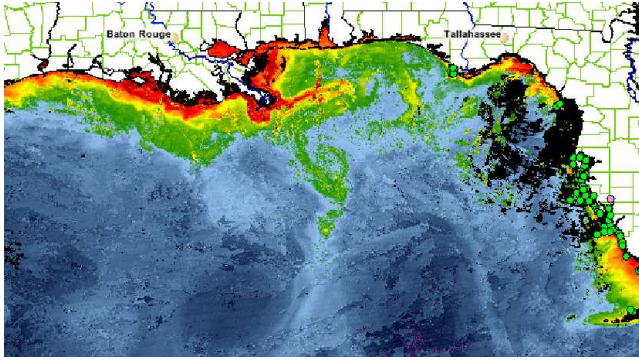


Fig. 1. The estimated field of chlorophyll a generated by the harmful algal blooms observation system [9] by NOAA. (Photo courtesy of NOAA).

biologically inspired *chemotaxis* [10,11], in which a mobile sensing agent is driven according to a local gradient of a field of interest. However, with this approach, the convergence rate can be slow and the mobile robot may get stuck in the local maxima of the field. The cooperative network of agents that performs adaptive gradient climbing in a distributed environment was presented in [6,7]. The centralized network can adapt its configuration in response to the sensed environment in order to optimize its gradient climb.

This problem of gradient climbing constantly occurs in biological species. Aquatic organisms search for favorable regions that contain abundant resources for their survival. For example, it is well-known that fish schools climb gradients of nutrients to locate the densest source of food. To locate resources, fish schools use “taxis”, a behavior in which they navigate habitats according to local gradients in uncertain environments. Grünbaum [12] showed that schooling behavior can improve the ability of performing taxis to climb gradients, since the swarming alignment tendency can average out the stochastic sampling errors of individuals.

Tanner [3] and Olfati-Saber [4] presented comprehensive analyses of the flocking algorithm by Reynolds [13]. This flocking algorithm was originally developed to simulate the movements of flocking birds in computer graphics where each artificial bird follows a set of rather simple distributed rules [13]. A bird in a flock coordinates with the movements of its neighboring flock mates and tries to stay close to its neighbors while avoiding collisions. In general, the collective swarm behaviors of birds/fish/ants/bees are known to be the outcomes of natural optimization [14,15].

In this paper, we extend the recent development in multi-agent systems [3,4] and develop novel distributed learning and cooperative control algorithms for multi-agent systems. The learning and control algorithms are performed at each agent using only local information. However, they are designed so that agents as a whole exhibit *collective intelligence*, i.e., a collection of agents achieves a global goal. In a resource-constrained multi-agent system, the communication range of each agent is limited as compared to the size

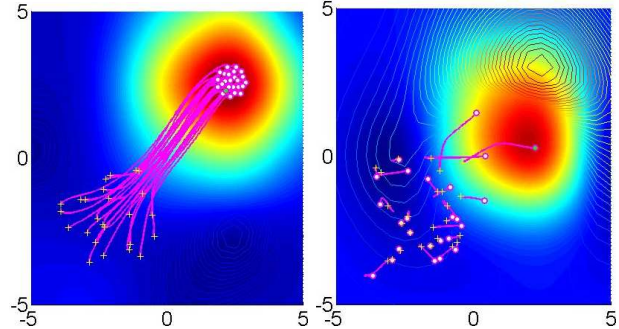


Fig. 2. Left: Trajectories of the proposed multi-agent system. Right: Trajectories of field estimating agents without communication and the swarming effort. The estimated field by agent 1 is shown as a background in colors. Agent 1 is plotted as a green dot. Thin contour lines represent the error field between the true field and the estimated field. (+) and (o) represent, respectively, initial and final locations. Solid lines represent trajectories of agents. See more details about the simulation in Section 5.

of a surveillance region. Hence, agents cannot perform the coverage control as shown in [16]. Instead, each agent receives measurements from its neighboring agents within its communication range. Upon receiving cooperative measurements, each mobile sensing agent will recursively update the estimate of an unknown static field of interest. The recursive estimation is based on the nonparametric method called *kernel regression* in order to represent a wide range of physical phenomena. To locate the maximum (or source) of the field, the sensing agent will climb the gradient of its own estimated field. The proposed cooperative control mimics the individual and social behaviors of a distributed pack of animals communicating locally to search for their densest resources in an uncertain environment. The fish school’s efficient performance of climbing nutrient gradients to search food resources and the exceptional geographical mapping capability of biological creatures have motivated the development of our multi-agent systems. Simulation results in Section 5 strongly support our idea and validate the effectiveness of the proposed multi-agent systems with cooperative control as compared to field estimating agents without cooperative control. As shown in Fig. 2, the proposed multi-agent system collectively locate the maximum of the unknown field rapidly while, without communication and the swarming effort, only a couple of agents near the maximum point can slowly estimate and climb the gradient of the field.

This paper also presents convergence properties of the proposed distributed learning and cooperative control algorithms by transforming the closed-loop dynamics of a multi-agent system into a form of a stochastic approximation algorithm. Our theoretical results are based on the ordinary differential equation (ODE) approach [17,18]. We also present a set of sufficient conditions for which the convergence is guaranteed with probability one.

This paper is organized as follows. In Section 2, we briefly introduce the mobile sensing network model, notations related to a graph, and artificial potentials to form a swarming

behavior. A recursive radial basis function learning algorithm for mapping the field of interest is presented in Section 3.1. In Section 3.2, cooperatively learning control is described with a stochastic approximation gain. Section 4 analyzes the convergence properties of the proposed coordination algorithm based on the ODE approach. In Section 5, the effectiveness of the proposed multi-agent system is demonstrated by simulation results with respect to different fields of interest and conditions.

2 Mobile Sensing Agent Network

In this section, we describe the mathematical framework for mobile sensing agent networks and explain notations used in this paper.

Let $\mathbb{R}, \mathbb{R}_{\geq 0}, \mathbb{R}_{> 0}, \mathbb{Z}, \mathbb{Z}_{\geq 0}, \mathbb{Z}_{> 0}$ denote, respectively, the set of real, non-negative real, positive real, integer, non-negative integer, and positive integer numbers. The positive definiteness (respectively, semi-definiteness) of a matrix A is denoted by $A \succ 0$ (respectively, $A \succeq 0$). $I_n \in \mathbb{R}^{n \times n}$ denotes the identity matrix of size n .

2.1 Models for Mobile Sensing Agents

Let N_s be the number of sensing agents distributed over the surveillance region $\mathcal{M} \subset \mathbb{R}^2$, which is assumed to be a convex and compact set. The identity of each agent is indexed by $\mathcal{I} := \{1, 2, \dots, N_s\}$. Let $q_i(t) \in \mathcal{M}$ be the location of the i -th sensing agent at time $t \in \mathbb{R}_{\geq 0}$ and let $q := \text{col}(q_1, q_2, \dots, q_{N_s}) \in \mathbb{R}^{2N_s}$ be the configuration of the multi-agent system. The discrete time, high-level dynamics of agent i is modeled by

$$q_i(t + \Delta_t) = q_i(t) + \Delta_t v_i(t), \quad (1)$$

where $q_i(t) \in \mathbb{R}^2$ and $v_i(t) \in \mathbb{R}^2$ are, respectively, the position and the control input of agent i at time $t \in \mathbb{R}_{\geq 0}$. $\Delta_t \in \mathbb{R}_{> 0}$ denotes the iteration step size (or sampling time). We assume that the measurement $y(q_i(t))$ of the i -th sensor includes the scalar value of the field $\mu(q_i(t))$ and sensor noise $w(t)$, at its position $q_i(t)$ and a sampled time t ,

$$y(q_i(t)) := \mu(q_i(t)) + w(t), \quad (2)$$

where $\mu : \mathcal{M} \rightarrow [0, \mu_{max}]$ is an unknown field of interest.

2.2 Graph-Theoretic Representation

The group behavior of mobile sensing agents and their complicated interactions with neighbors are best treated by a graph with edges. Let $G(q) := (\mathcal{I}, \mathcal{E}(q))$ be an undirected communication graph such that an edge $(i, j) \in \mathcal{E}(q)$ if and only if agent i can communicate with agent $j \neq i$. We assume that each agent can communicate with its neighboring agents within a limited transmission range given by a radius

of r , as depicted in Fig. 3. Therefore, $(i, j) \in \mathcal{E}(q)$ if and only if $\|q_i(t) - q_j(t)\| \leq r$. For example, the i -th agent in Fig. 3 communicates with and collects measurements from all four neighboring sensing agents in the i -th agent's communication range. We define the neighborhood of agent i with a configuration of q by $\mathcal{N}(i, q) := \{j \in \mathcal{I} \mid (i, j) \in \mathcal{E}(q)\}$. The adjacency matrix $A := [a_{ij}]$ of an undirected graph G is a symmetric matrix such that $a_{ij} = k_3 \in \mathbb{R}_{> 0}$ if vertex i and vertex j are neighbors and $a_{ij} = 0$ otherwise. Notice that an adjacency matrix A can be also defined in a smooth fashion in terms of q [4]. The scalar graph Laplacian $L = [l_{ij}] \in \mathbb{R}^{N_s \times N_s}$ is a matrix defined as $L := D^A - A$, where D^A is a diagonal matrix whose diagonal entries are row sums of A , i.e., $D^A := \text{diag}(\sum_{j=1}^{N_s} a_{ij})$. The 2-dimensional graph Laplacian is defined as $\hat{L} := L \otimes I_2$, where \otimes is the Kronecker product. For instance, the corresponding A , L and \hat{L} for a graph in Fig. 3 are:

$$A = k_3 \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad L = k_3 \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & -1 & 0 \\ -1 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\hat{L} = L \otimes I_2 = k_3 \begin{bmatrix} 2I_2 & -I_2 & -I_2 & 0 \\ -I_2 & 2I_2 & -I_2 & 0 \\ -I_2 & -I_2 & 2I_2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where 0 denotes a zero matrix with appropriate dimensions. Let a state $p_i \in \mathbb{R}^2$ be associated to agent i for all $i \in \mathcal{I}$ under a topology of an undirected graph G . Two agents i and j are said to agree whenever they have the same states, i.e., $p_i = p_j$. The quadratic disagreement function $\Psi_G : \mathbb{R}^{2N_s} \rightarrow \mathbb{R}_{\geq 0}$ evaluates the group disagreement in the network of agents:

$$\Psi_G(p) := \frac{1}{4} \sum_{(i,j) \in \mathcal{E}(q)} a_{ij} \|p_j - p_i\|^2, \quad (3)$$

where $p := \text{col}(p_1, p_2, \dots, p_{N_s}) \in \mathbb{R}^{2N_s}$. A disagreement function [4,19] can be obtained via the Laplacian \hat{L} :

$$\Psi_G(p) = \frac{1}{2} p^T \hat{L} p, \quad (4)$$

and hence the gradient of $\Psi_G(p)$ with respect to p is given by

$$\nabla \Psi_G(p) = \hat{L} p. \quad (5)$$

The properties shown in (4) and (5) will be used in the convergence analysis in Section 4.

2.3 Swarming Behavior

A group of agents are coordinated to collect (noisy) samples from a stationary field at diverse locations for the purpose

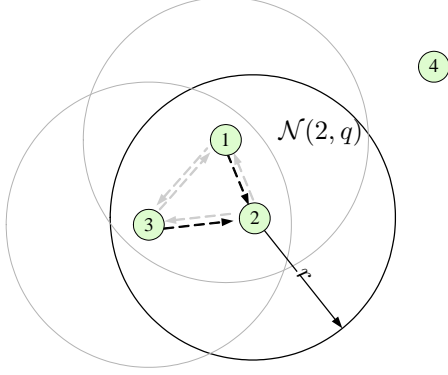


Fig. 3. The model of the mobile sensing agent network. The agent 2 gathers measurements from two neighboring sensing agents 1 and 3 in a r interactive range. Hence, the collective measurements of agent 2 will be sampled at locations denoted by agents 1 and 3.

of estimating the field of interest. A set of artificial potential functions creates a swarming behavior of agents and provides agents with obstacle avoidance capabilities. We use attractive and repulsive potential functions similar to ones used in [3,4,20] to generate a swarming behavior. To enforce a group of agents to satisfy a set of algebraic constraints $\|q_i - q_j\| = d$ for all $j \in \mathcal{N}(i, q)$, we introduce a smooth collective potential function

$$\begin{aligned} U_1(q) &:= \sum_i \sum_{j \in \mathcal{N}(i, q), j \neq i} U_{ij}(\|q_i - q_j\|^2) \\ &= \sum_i \sum_{j \in \mathcal{N}(i, q), j \neq i} U_{ij}(r_{ij}), \end{aligned} \quad (6)$$

where $r_{ij} := \|q_i - q_j\|^2$. The pair-wise attractive/repulsive potential function $U_{ij}(\cdot)$ in (6) is defined by

$$U_{ij}(r_{ij}) := \frac{1}{2} \left(\log(\alpha + r_{ij}) + \frac{\alpha + d^2}{\alpha + r_{ij}} \right), \text{ if } r_{ij} < d_0^2, \quad (7)$$

otherwise (i.e., $r_{ij} \geq d_0^2$), it is defined according to the gradient of the potential, which will be described shortly. Here $\alpha, d \in \mathbb{R}_{>0}$ and $d < d_0$. The gradient of the potential with respect to q_i for agent i is given by

$$\begin{aligned} \nabla U_1(q_i) &:= \frac{\partial U_1(q)}{\partial \tilde{q}_i} \Big|_{\tilde{q}_i = q_i} = \sum_{j \neq i} \frac{\partial U_{ij}(r)}{\partial r} \Big|_{r=r_{ij}} (q_i - q_j) \\ &= \begin{cases} \sum_{j \neq i} \frac{(r_{ij} - d^2)(q_i - q_j)}{(\alpha + r_{ij})^2} & \text{if } r_{ij} < d_0^2 \\ \sum_{j \neq i} \rho \left(\frac{\sqrt{r_{ij} - d_0^2}}{|d_1 - d_0|} \right) \frac{\|d_0^2 - d^2\|}{(\alpha + d_0^2)^2} (q_i - q_j) & \text{if } r_{ij} \geq d_0^2, \end{cases} \end{aligned} \quad (8)$$

where $\rho : \mathbb{R}_{\geq 0} \rightarrow [0, 1]$ is the bump function [4]

$$\rho(z) := \begin{cases} 1, & z \in [0, h); \\ \frac{1}{2} \left[1 + \cos \left(\pi \frac{(z-h)}{(1-h)} \right) \right], & z \in [h, 1]; \\ 0, & \text{otherwise.} \end{cases}$$

Notice that ρ varies smoothly from 1 to 0 as the scalar input increases. (6), (7), and (8) will produce a continuously differentiable (C^1) reaction potential force between any two agents as depicted in Fig. 4. Parameters α , d , d_0 , and d_1 will shape the artificial potential function. A typical way to choose those parameters are explained as follows. In equations (6), (7), and (8), a non-zero gain factor α is introduced to prevent the reaction force from diverging at $r_{ij} = \|q_i - q_j\|^2 = 0$. As illustrated in Fig. 4, this potential yields a reaction force that is attracting when the agents are too far away and repelling when a pair of two agents are too close. It has an equilibrium point at a distance of d . d_0 will be chosen at the location where the slope of the potential force first becomes zero (Fig. 4) as $\sqrt{r_{ij}}$ increases from zero. For $\sqrt{r_{ij}} > d_0$, the bump function will shape the the potential force to become zero smoothly when the relative distance reaches to d_1 which is slightly shorter than the radius of the transmission range r . Hence, in general, we configure parameters such that $d < d_0 < d_1 < r$, which will force the gradient of the potential function due to agent j in (8) to be a zero vector before the communication link to agent i is disconnected from agent j . In this way, we can construct a smooth collective potential force between any two agents in spite of the limited communication range. We also introduce a potential U_2 to model the environment. U_2 enforces each agent to stay inside the closed and connected surveillance region in \mathcal{M} and prevents collisions with obstacles in \mathcal{M} . We construct U_2 such that it is radially unbounded in q , i.e.,

$$U_2(q) \rightarrow \infty \text{ as } \|q\| \rightarrow \infty. \quad (9)$$

The condition in (9) can be used for making a Lyapunov function candidate radially unbounded. Define the total artificial potential by

$$U(q) := k_1 U_1(q) + k_2 U_2(q), \quad (10)$$

where $k_1, k_2 \in \mathbb{R}_{>0}$ are weighting factors. A swarming behavior and an obstacle avoidance capability of each agent will be developed in Section 3.2.

3 Distributed Learning and Cooperative Control

In this section, we describe distributed learning and cooperative control algorithms. The sensing agent will receive measurements from its neighboring agents within a limited transmission range. Upon receiving measurements, each mobile sensing agent will recursively update the estimate of an unknown static field of interest using the distributed learning algorithm. Based on the estimated field, each agent moves to the peak of the field using the cooperative control algorithm.

3.1 Distributed Learning

We introduce a distributed learning algorithm for each mobile sensing agent to estimate a static field of interest μ :

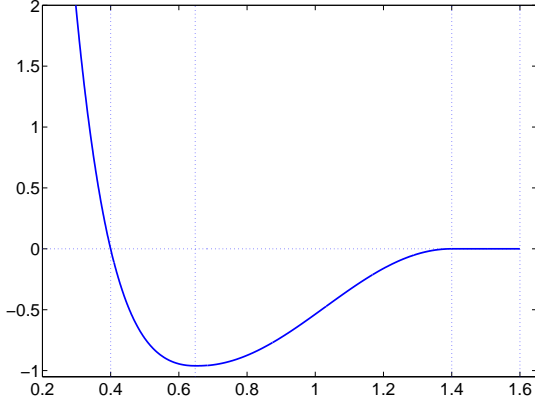


Fig. 4. The reaction force between two agents is generated by the potential function in (6), (7), and (8) with respect to $\|q_i - q_j\|$. Here parameters $d = 0.4$, $d_0 = 0.648$, $d_1 = 1.4$ and $r = 1.6$ are used.

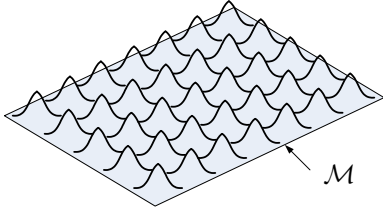


Fig. 5. Uniformly distributed Gaussian bases over a surveillance region \mathcal{M} .

$\mathcal{M} \rightarrow [0, \mu_{max}]$. Suppose that the scalar field $\mu(\nu)$ is generated by a network of radial basis functions¹:

$$\mu(\nu) := \sum_{j=1}^m \phi_j(\nu) \theta_j = \phi^T(\nu) \Theta, \quad (11)$$

where $\phi^T(\nu)$ and Θ are defined respectively by

$$\begin{aligned} \phi^T(\nu) &:= [\phi_1(\nu) \ \phi_2(\nu) \ \cdots \ \phi_m(\nu)], \\ \Theta &:= [\theta_1 \ \theta_2 \ \cdots \ \theta_m]^T \in \Theta, \end{aligned}$$

where $\Theta \subset \mathbb{R}^m$ is a compact set. Gaussian radial basis functions $\{\phi_j(\nu)\}$ are given by

$$\phi_j(\nu) := \frac{1}{\Gamma_j} \exp\left(\frac{-\|\nu - \kappa_j\|^2}{\sigma_j^2}\right), \quad (12)$$

where σ_j is the width of the Gaussian basis and Γ_j is a normalizing constant. Centers of basis functions

¹ We have considered a fairly simple parameterization for the field of interest to focus more on the design and the convergence analysis of learning agents. See more general models used for the field of interest in [21,22].

$\{\kappa_j | j \in \{1, \dots, m\}\}$ are uniformly distributed in the surveillance region \mathcal{M} as shown in Fig. 5. $\Theta \in \Theta \subset \mathbb{R}^m$ is the true parameter of the regression model in (11). From (2), we have observations through sensors at the location ν_k , $y(\nu_k) = \phi^T(\nu_k) \Theta + w(k)$, where k is a measurement sampling index. Based on the observations and regressors $\{(y(\nu_k), \phi(\nu_k))\}_{k=1}^n$, our objective is to find $\hat{\Theta}$ which minimizes the least-squares error:

$$\sum_{k=1}^n |y(\nu_k) - \phi^T(\nu_k) \hat{\Theta}|^2. \quad (13)$$

Remark 1 Our environmental model in (11) can be viewed as a nonparametric approach to model spatial phenomena. Other popular nonparametric approaches are Gaussian processes [23] [24] and kriging models [25]. A dynamical Θ in (11) was used to represent a time-varying trend in the space-time Kalman filter model [26] to model spatio-temporal random fields.

Noiseless Measurements

Let us first consider the measurement model (2) without the sensor noise $w(k)$. Similar spatial estimation algorithms with a known sensor noise level for achieving the minimum variance of the estimation error can be found in [21,22]. For a set $\{(y(\nu_k), \phi(\nu_k))\}_{k=1}^n$, the optimal least-squares estimation solution is well-known [27]:

$$\hat{\Theta}(n) = P(n, 1) \Phi^T(n, 1) Y(n, 1), \quad (14)$$

where (for simplicity, we abuse notations by letting $y(k) := y(\nu_k)$ and $\phi(k) := \phi(\nu_k)$)

$$\begin{aligned} Y(n, s) &:= [y(s) \ y(s+1) \ \cdots \ y(n)]^T \in \mathbb{R}^{n-s+1}, \\ \Phi(n, s) &:= [\phi(s) \ \cdots \ \phi(n)]^T \in \mathbb{R}^{n-s+1 \times m}, \\ P(n, s) &:= [\Phi^T(n, s) \Phi(n, s)]^{-1} \\ &= \left[\sum_{k=s}^n \phi(k) \phi^T(k) \right]^{-1} \in \mathbb{R}^{m \times m}. \end{aligned}$$

During a time interval between the coordination iteration indices t and $t + \Delta_t$ as in (1), we suppose that a sensing agent has collected s number of samples from itself and its neighbors. Assume that at the previous iteration, the agent has already updated the field $\hat{\mu}(\cdot)$ based on the previous data set $\{(y(k), \phi(k))\}_{k=1}^{n-s}$, where $n-s$ is the total number of past measurements. Now the sensing agent needs to update the field $\hat{\mu}(\cdot)$ upon receiving cooperatively measured s number of points $\{(y(k), \phi(k))\}_{k=n-s+1}^n$.

We then have the following algorithm. Assume that $\Phi^T(t) \Phi(t)$ is nonsingular for all t . For the collected s number of observations and regressors

$\{(y(k), \phi(k))\}_{k=n-s+1}^n$, consider the recursive algorithm given as

$$\begin{aligned} K(n) &= P(n-s)\Phi_*^T [I_s + \Phi_* P(n-s)\Phi_*^T]^{-1}, \\ P(n) &= [I_m - K(n)\Phi_*]P(n-s), \\ \hat{\Theta}(n) &= \hat{\Theta}(n-s) + K(n) [Y_* - \Phi_* \hat{\Theta}(n-s)], \\ \hat{\mu}(\nu) &:= \phi^T(\nu)\hat{\Theta}(n), \end{aligned} \quad (15)$$

where some abbreviations are defined: $Y_* := Y(n, n-s+1) \in \mathbb{R}^s$, $\Phi_* = \Phi(n, n-s+1) \in \mathbb{R}^{s \times m}$, $\Phi^T(n) := \Phi^T(n, 1) \in \mathbb{R}^{n \times m}$, $Y(n) := Y(n, 1) \in \mathbb{R}^n$ and $P(n) := P(n, 1) \in \mathbb{R}^{m \times m}$. Then it is straightforward to see that the recursive estimation presented in (15) is the least-squares estimation that minimizes the error function in (13).

Remark 2 $\Phi^T(n)\Phi(n)$ is always singular for $n < m$. $\Phi^T(n)\Phi(n)$ is nonsingular for $n \geq m$ except for the case where measurements are only taken at a set of measure zero, for example, a line splitting two Gaussian radial basis functions equally such that $\phi_i(\nu) = \phi_j(\nu)$. In practice, we start the recursive LSE algorithm in (15) with initial states $\hat{\Theta}(0)$ and $P(0) \succ 0$ which corresponds to the situation in which the parameters have an a priori distribution and keep running the recursive algorithm with new measurements. With these initial values, we have

$$P^{-1}(n) := P^{-1}(0) + \Phi^T(n)\Phi(n) \succ 0. \quad (16)$$

In the next subsection, we elaborate on the case of noisy observation and the resulting effects on the estimated field and its gradient.

Noisy Measurements

Consider the measurement model (2) with the sensor noise $w(k)$, which is assumed to be a white noise sequence with an unknown variance W :

$$\mathbb{E}(w(k)) = 0, \quad \mathbb{E}(w(k)w(z)) = \begin{cases} W \succ 0 & \text{if } k = z \\ 0 & \text{if } k \neq z \end{cases}, \quad (17)$$

where \mathbb{E} denotes the expectation operator. Moreover, we assume that there exists $L < \infty$ so that

$$|w(k)| < L \quad \text{with probability one (w.p.1)} \quad \forall k. \quad (18)$$

Given the measurement data set

$$\{y(\mu) \mid \mu \in S\} \quad \text{where } S = \{\nu_k \mid 1 \leq k \leq n\}$$

and the sensor noise $\{w(k) \mid k \in \{1, \dots, n\}\}$ defined in (17) and (18), an agent will estimate $\hat{\Theta}(n)$ using the recursive

LSE algorithm in (15). Let the estimation error vector be $\tilde{\Theta}(n) := \hat{\Theta}(n) - \Theta$. We also define the error of the estimated field at the location $\nu \in \mathcal{R}$ by

$$\tilde{\mu}(S, \nu) := \hat{\mu}(S, \nu) - \mu(\nu) = \phi^T(\nu)\tilde{\Theta}(|S|), \quad (19)$$

where $|S|$ is the cardinality of the set S . The error of the estimated field at $\nu \in \mathcal{M}$ is then obtained by

$$\tilde{\mu}(S, \nu) = \mathbb{E}(\tilde{\mu}(S, \nu)) + \epsilon(S, \nu), \quad (20)$$

where

$$\begin{aligned} \mathbb{E}(\tilde{\mu}(S, \nu)) &:= \phi^T(\nu) \left[P(|S|) \sum_{\nu_t \in S} \phi(\nu_t)\phi^T(\nu_t) - I_m \right] \Theta, \\ \epsilon(S, \nu) &:= \phi^T(\nu) \left[P(|S|) \sum_{t=1}^{|S|} \phi(\nu_t)w(t) \right], \end{aligned}$$

where $|S|$ is the total number of collective measurements for the associated agent. For persistent exciting coordination strategies ($\Phi_*^T \Phi_* \succ 0$), the estimator is asymptotically² unbiased

$$\lim_{|S| \rightarrow \infty} \mathbb{E}(\tilde{\mu}(S, \nu)) = 0, \quad \forall \nu \in \mathcal{M}. \quad (21)$$

For a large number of $|S|$, the variance of the estimation error is given by

$$\begin{aligned} \mathbb{E}(\epsilon(S, \nu)\epsilon^T(S, \nu)) &= \phi^T(\nu)WP(|S|)\phi(\nu), \\ &= \phi^T(\nu)\frac{W}{|S|}R^{-1}(S)\phi(\nu), \end{aligned} \quad (22)$$

where $R(S)$ is defined by

$$R(S) := \left[\frac{P^{-1}(0)}{|S|} + \frac{1}{|S|} \sum_{\nu_k \in S} \phi(\nu_k)\phi^T(\nu_k) \right]. \quad (23)$$

Remark 3 From (22), it is straightforward to see that the estimation error variance is a function of the evaluated position ν in \mathcal{M} , is proportional to the variance W , and decreases at the rate of $1/|S|$ and $R^{-1}(S)$. $R(S)$ asymptotically serves as a time average of outer products of basis functions evaluated at the measurement points in S , which implies that the error variance is smaller at places where the agent has collected more samples.

The gradient of the field of interest is denoted by

$$\nabla \mu(\nu) := \frac{\partial \mu(x)}{\partial x} \Big|_{x=\nu}. \quad (24)$$

² It is asymptotically unbiased if a priori distribution of $\Theta(0)$ and $P(0)$ is not available.

From (11), we have

$$\nabla\mu(\nu) = \frac{\partial\phi^T(x)}{\partial x}\Big|_{x=\nu} \Theta =: \phi'^T(\nu)\Theta \in \mathbb{R}^{2 \times 1}, \quad (25)$$

where $\phi'^T(\nu) \in \mathbb{R}^{2 \times m}$. Thus, the gradient of the estimated field based on observations $S := \{\nu_k\}_{k=1}^n$ and $\{y(\mu)\}_{\mu \in S}$ is given by

$$\nabla\hat{\mu}(S, \nu) := \phi'^T(\nu)\hat{\Theta}(|S|) \in \mathbb{R}^{2 \times 1}. \quad (26)$$

The error of the estimated gradient at the location $\nu \in \mathcal{M}$ is obtained by

$$\begin{aligned} \nabla\tilde{\mu}(S, \nu) &:= \phi'^T(\nu)\hat{\Theta}(|S|) - \nabla\mu(\nu) = \phi'^T(\nu)\tilde{\Theta}(|S|) \\ &= \mathbb{E}(\nabla\tilde{\mu}(S, \nu)) + \nabla\epsilon(S, \nu), \end{aligned} \quad (27)$$

where

$$\begin{aligned} \mathbb{E}(\nabla\tilde{\mu}(S, \nu)) &= \phi'^T(\nu) \left[P(|S|) \sum_{\nu_k \in S} \phi(\nu_k)\phi^T(\nu_k) - I_m \right] \Theta, \\ \nabla\epsilon(S, \nu) &:= \phi'^T(\nu) \left[P(|S|) \sum_{k=1}^{|S|} \phi(\nu_k)w(k) \right]. \end{aligned}$$

For $\Phi_*^T \Phi_* \succ 0$, the gradient estimator is asymptotically unbiased

$$\lim_{|S| \rightarrow \infty} \mathbb{E}(\nabla\tilde{\mu}(S, \nu)) = 0, \quad \forall \nu \in \mathcal{M}. \quad (28)$$

The covariance matrix $\mathbb{E}(\nabla\epsilon(S, \nu)\nabla\epsilon^T(S, \nu))$ is obtained by

$$\phi'^T(\nu) \frac{W}{|S|} R^{-1}(S) \phi'(\nu), \quad (29)$$

where $R(S)$ is defined in (23). Now we present our collaboratively learning control protocol.

3.2 Cooperative Control

Each of mobile agents receives measurements from neighbors, then it updates its gradient of the estimated field using $\hat{\Theta}$ from the recursive algorithm presented in (15). Subsequently, based on this updated gradient, the control for its coordination will be decided. Hereafter, we apply a new time notation used for the coordination, to the recursive LSE algorithm in (15). In particular, we replace $n - s \in \mathbb{Z}_{\geq 0}$ by $t \in \mathbb{Z}_{\geq 0}$ and $n \in \mathbb{Z}_{\geq 0}$ by $t + 1 \in \mathbb{Z}_{\geq 0}$ in (15) such that the resulting recursive algorithm with the new time index for

agent i at its position $q_i(t)$ is given by

$$\begin{aligned} K_i(t+1) &= P_i(t)\Phi_{*i}^T (I_s + \Phi_{*i}P_i(t)\Phi_{*i}^T)^{-1}, \\ P_i(t+1) &= (I_m - K_i(t+1)\Phi_{*i})P_i(t), \\ \hat{\Theta}_i(t+1) &= \hat{\Theta}_i(t) + K_i(t+1) \left[Y_{*i} - \Phi_{*i}\hat{\Theta}_i(t) \right], \\ \nabla\hat{\mu}_i(t, q_i(t)) &= \phi'^T(q_i(t))\hat{\Theta}_i(t+1), \end{aligned} \quad (30)$$

where $\nabla\hat{\mu}_i(t, \nu) : \mathbb{Z}_{\geq 0} \times \mathcal{M} \rightarrow \mathbb{R}^2$ denotes the gradient of the estimated field at ν based on measurements before the time $t + 1$. Y_{*i} and Φ_{*i} of agent i are defined in the same way as Y_* and Φ_* are defined in (15). Y_{*i} is the collection of collaboratively measured data. From (2), for all $j \in \mathcal{N}(i, q(t)) \cup \{i\}$, we have

$$Y_{*i} = \Phi_{*i}\Theta + \begin{bmatrix} \vdots \\ w_j(k) \\ \vdots \end{bmatrix} =: \Phi_{*i}\Theta + w_{*i}(t), \quad (31)$$

where the sampled time of the measurements can vary among sensors but we label the time index by t for any sampled time contained in a measurement period between t and $t + 1$. $w_j(k)$ is the measurement noise of sensor j , and is independently and identically distributed over $j \in \mathcal{I}$. We also define a new variable $w_{*i}(t)$ as in (31) for later use.

Based on the latest update of the gradient of the estimated field $\nabla\hat{\mu}_i(t, q_i(t))$, a distributed control $v_i(t + 1)$ in (1) for agent i is proposed by

$$v_i(t+1) := \frac{\gamma(t+1)}{\Delta_t} \left[\frac{\Delta_t}{\gamma(t)} v_i(t) + \gamma(t)u_i(t) \right], \quad (32)$$

with

$$\begin{aligned} u_i(t) &:= -\nabla U(q_i(t)) - k_{di} \frac{\Delta_t}{\gamma(t)} v_i(t) \\ &+ \sum_{j \in \mathcal{N}(i, q(t))} a_{ij}(q(t)) \left(\frac{\Delta_t(v_j(t) - v_i(t))}{\gamma(t)} \right) \\ &+ k_4 \nabla\hat{\mu}_i(t, q_i(t)), \end{aligned} \quad (33)$$

where $k_4 \in \mathbb{R}_{>0}$ is a gain factor for the estimated gradient and $k_{di} \in \mathbb{R}_{\geq 0}$ is a gain for the velocity feedback. The first term in the right-hand side of (33) is the gradient of the artificial potential defined in (10) which attracts agents while avoiding collisions among them. Also it restricts the movements of agents inside \mathcal{M} ; appropriate artificial potentials can be added to $U(q_i)$ for agents to avoid obstacles in \mathcal{M} . The second term in (33) provides damping. The third term in (33) is an effort for agent i to match its velocity with those of neighbors. This term is used for the ‘‘velocity consensus’’ and serves as a damping force among agents. The gradient ascent of the estimated field is provided as the last term.

The control for the coordination of sensing agents gradually decreases for perfect tracking of the maximum of an unknown field in spite of the estimation based on the noisy measurements. We have proposed the control protocol in (32) with a standard adaptive gain sequence $\gamma(t)$ that satisfies the following properties

$$\begin{aligned} \gamma(t) > 0, \quad \sum_{t=1}^{\infty} \gamma(t) = \infty, \quad \sum_{t=1}^{\infty} \gamma^2(t) < \infty, \\ \limsup_{t \rightarrow \infty} [1/\gamma(t) - 1/\gamma(t-1)] < \infty. \end{aligned} \quad (34)$$

This gain sequence is often used for stochastic approximation algorithms [17,18] and enables us to apply the ODE approach [17,28,29] for convergence analysis.

For the convenience of analysis, we change variables. In particular, we introduce $p_i(t)$, a scaled version of the velocity state $v_i(t)$:

$$p_i(t) := \frac{\Delta_t}{\gamma(t)} v_i(t), \quad (35)$$

where $v_i(t)$ is the control input to agent i as defined in (32). After the change of variables in (35), the resulting dynamics of agent i is given by

$$\begin{cases} q_i(t+1) = q_i(t) + \gamma(t)p_i(t), \\ p_i(t+1) = p_i(t) + \gamma(t)u_i(t), \end{cases} \quad (36)$$

where we applied new notations to (1) by replacing $\Delta_t v_i(t)$ by $\gamma(t)p_i(t)$, $t + \Delta_t \in \mathbb{R}_{\geq 0}$ by $t + 1 \in \mathbb{Z}_{\geq 0}$ and $t \in \mathbb{R}_{\geq 0}$ by $t \in \mathbb{Z}_{\geq 0}$.

Incorporating the discrete time model in (36) along with the proposed control in (32) gives

$$\begin{aligned} q_i(t+1) &= q_i(t) + \gamma(t)p_i(t), \\ p_i(t+1) &= p_i(t) + \gamma(t) \left\{ -\nabla U(q_i(t)) - k_{di}p_i(t) \right. \\ &\quad \left. - \nabla \Psi_G(p_i(t)) + k_4 \phi^T(q_i(t)) \hat{\Theta}_i(t+1) \right\}, \end{aligned} \quad (37)$$

where $\nabla \Psi_G(p_i(t))$ is the gradient of the disagreement function (defined in (3) and (5)) with respect to p_i :

$$\nabla \Psi_G(p_i(t)) = \sum_{j \in \mathcal{N}(i, q(t))} a_{ij}(q(t))(p_i(t) - p_j(t)).$$

In the next section, we will transform our multi-agent system into a recursive stochastic algorithm with states

$$x(t) := \text{col}(q_1, \dots, q_{N_s}(t), p_1(t), \dots, p_{N_s}(t)),$$

and

$$\varphi(t) := \text{col}(\tilde{\Theta}_1(t), \dots, \tilde{\Theta}_{N_s}(t)).$$

4 Convergence Analysis

In order to analyze the convergence properties of (30), (37) and (34), we utilize Ljung's ordinary differential equation (ODE) approach developed in [17,28,29]. In particular, Ljung [17,28] presented an analysis technique of general recursive stochastic algorithms in the canonical form of

$$x(t) = x(t-1) + \gamma(t)Q(t; x(t-1), \varphi(t)), \quad (38)$$

along with the observation process

$$\varphi(t) = g(t; x(t-1), \varphi(t-1), e(t)). \quad (39)$$

In order to use the ODE approach, for this nonlinear observation process in (39), the following regularity conditions in [28] need to be satisfied. Let D_R be a subset of the x space in (38), where the regularity conditions hold.

- C1:** $\|g(x, \varphi, e)\| < C$ for all φ, e for all $x \in D_R$.
- C2:** The function $Q(t, x, \varphi)$ is continuously differentiable with respect to x and φ for $x \in D_R$. The derivatives are, for fixed x and φ , bounded in t .
- C3:** $g(t; x, \varphi, e)$ is continuously differentiable with respect to $x \in D_R$.
- C4:** Define $\bar{\varphi}(t, \bar{x})$ as

$$\bar{\varphi}(t, \bar{x}) = g(t; \bar{x}, \varphi(t-1, \bar{x}), e(t)), \quad \bar{\varphi}(0, \bar{x}) = 0, \quad (40)$$

and assume that $g(\cdot)$ has the property

$$\|\bar{\varphi}(t, \bar{x}) - \varphi(t)\| < C \max_{n \leq k \leq t} \|\bar{x} - x(k)\|,$$

if $\bar{\varphi}(n, \bar{x}) = \varphi(n)$. This means that small variations in x in (39) are not amplified to a higher magnitude for the observations φ .

- C5:** Let $\bar{\varphi}_1(t, \bar{x})$ and $\bar{\varphi}_2(t, \bar{x})$ be solutions of (40) with $\bar{\varphi}_1(s, \bar{x}) := \varphi_1^0$ and $\bar{\varphi}_2(s, \bar{x}) := \varphi_2^0$. Then define D_s as the set of all \bar{x} for which the following holds:

$$\|\bar{\varphi}_1(t, \bar{x}) - \bar{\varphi}_2(t, \bar{x})\| < C(\varphi_1^0, \varphi_2^0)\lambda^{t-s}(\bar{x}),$$

where $t > s$ and $\lambda(\bar{x}) < 1$. This is the region of exponential stability of (39).

- C6:** $\lim_{t \rightarrow \infty} \mathbb{E}Q(t, \bar{x}, \bar{\varphi}(t, \bar{x}))$ exists for $\bar{x} \in D_R$ and is denoted by $f(\bar{x})$. The expectation is over $\{e(\cdot)\}$.
- C7:** $e(\cdot)$ is a sequence of independent random variables.
- C8:** $\sum_{t=1}^{\infty} \gamma(t) = \infty$.
- C9:** $\sum_{t=1}^{\infty} \gamma^p(t) < \infty$ for some p .
- C10:** $\gamma(\cdot)$ is a decreasing sequence.
- C11:** $\lim_{t \rightarrow \infty} \sup[1/\gamma(t) - 1/\gamma(t-1)] < \infty$.

For practical algorithm implementation, the projection or saturation is often introduced [17,18] to meet the boundedness condition required in the ODE approach [17]. Since

dynamics of agents are given by a single integrator, i.e., the position of the agent can be controlled

$$q_i(t+1) = q_i(t) + \gamma(t)p_i(t),$$

where $p_i(t)$ is the control, we can apply the usual saturation given by $[\cdot]_D$

$$x(t) = [\Omega(t)]_D = \begin{cases} \Omega(t), & \Omega(t) \in D \\ x(t-1), & \Omega(t) \notin D, \end{cases} \quad (41)$$

where $x(t)$ and $\Omega(t)$ denote the left- and right-hand sides of (38) respectively, i.e., the projected algorithm updates only if the updated value belongs to D otherwise it keeps the previous state. Our closed-loop system in (37) will be converted to the canonical form in (38). Throughout the paper, we assume that the projection is applied to the resulted algorithm in the form of (38). The projection disappears in the averaged updating directions. Hence, the convergence properties of the projected algorithm can be studied as if there was no projection in (38). For more details, see [29,30] and references therein.

We will then utilize the following corollary reported in [29].

Corollary 4 (Ljung [29]) *Consider the algorithm (38), (39) and (41) subject to the regularity conditions C1-C11. Let D_R be an open connected subset of D_S . Let D in (41) be the compact subset of D_R such that the trajectories of the associated ODE*

$$\frac{d}{d\tau}x(\tau) = f(x(\tau)) \quad (42)$$

where

$$f(x) := \lim_{t \rightarrow \infty} \mathbb{E}Q(t; x, \bar{\varphi}(t, x)),$$

that start in D remain in a closed subset \bar{D}_R of D_R for $\tau > 0$. Assume that the differential equation (42) has an invariant set D_c with domain of attraction $D_A \supset D$.

Then either

$$x(t) \rightarrow D_c, \text{ with probability one as } t \rightarrow \infty, \quad (43)$$

or

$$x(t) \rightarrow \partial D, \text{ with probability one as } t \rightarrow \infty, \quad (44)$$

where ∂D is the boundary of D .

The conclusion (44) is possible only if there is a trajectory of the differential equation in (42) that leaves D in (41).

Now we present our main results. The following lemma shows how to transform our coordination and estimation algorithms to the canonical forms in (38) and (39).

Lemma 5 *The algorithms (37) and (30) can be transformed into the forms of (38) and (39) respectively, using the following definitions:*

$$\begin{aligned} q(t) &:= \text{col}(q_1(t), \dots, q_{N_s}(t)) \in \mathbb{R}^{2N_s}, \\ p(t) &:= \text{col}(p_1(t), \dots, p_{N_s}(t)) \in \mathbb{R}^{2N_s}, \\ x(t) &:= [q^T(t), p^T(t)]^T \in \mathbb{R}^{4N_s}, \\ Q(t; x(t-1), \varphi(t)) & \\ &:= \begin{bmatrix} p \\ -\nabla U(q) - (\hat{L}(q) + K_d)p - \nabla \hat{C}(\varphi, q) \end{bmatrix}, \end{aligned} \quad (45)$$

where $K_d = \text{diag}(k_{d1}, \dots, k_{dN_s}) \otimes I_2 \succ 0$. The gradient of the estimated cost function $\nabla \hat{C}(\varphi(t), q(t-1)) \in \mathbb{R}^{2N_s}$ is defined by

$$\begin{aligned} &-k_4 \text{col}(\nabla \hat{\mu}_1(t-1, q_1(t-1)), \dots, \\ &\quad \nabla \hat{\mu}_{N_s}(t-1, q_{N_s}(t-1))) \\ &= -k_4 \text{col}(\phi^{T^T}(q_1(t-1))\hat{\Theta}_1(t), \dots, \\ &\quad \phi^{T^T}(q_{N_s}(t-1))\hat{\Theta}_{N_s}(t)). \end{aligned}$$

For the observation process in (39), we have:

$$\begin{aligned} \varphi(t) &= g(t; x(t-1), \varphi(t-1), e(t)) \\ &= A(t; x(t-1))\varphi(t-1) + B(t; x(t-1))e(t), \end{aligned} \quad (46)$$

where

$$\begin{aligned} \varphi(t) &:= \text{col}(\tilde{\Theta}_1(t), \dots, \tilde{\Theta}_{N_s}(t)) \in \mathbb{R}^{mN_s}, \\ A(t; x(t-1)) &:= \text{diag}(I_m - K_1(t)\Phi_{*1}, \dots, \\ &\quad I_m - K_{N_s}(t)\Phi_{*N_s}) \in \mathbb{R}^{mN_s \times mN_s}, \\ B(t; x(t-1)) &:= \text{diag}(K_1(t), \dots, K_{N_s}(t)) \in \mathbb{R}^{mN_s \times O}, \\ e(t) &:= \text{col}(w_{*1}(t-1), \dots, w_{*N_s}(t-1)) \in \mathbb{R}^O, \end{aligned}$$

where O varies according to the number of collaborative measurements at each iteration.

Proof: From (30), notice that:

$$\tilde{\Theta}_i(t) = [I_m - K_i(t)\Phi_{*i}]\tilde{\Theta}_i(t-1) + K_i(t)w_{*i}(t-1). \quad (47)$$

The rest of the proof is straightforward and so is omitted. \square

Two lemmas to validate the regularity conditions C1-C11 will be presented under the following assumptions:

M1: Each agent collects $s \geq m$ number of measurements at locations $\{\nu_k\}_{k=1}^s$ from itself and neighbors so that

$$\sum_{k=1}^s \phi(\nu_k)\phi^T(\nu_k) \succ 0,$$

where m is in (11).

M2: The artificial potential force and the adjacency matrix are continuously differentiable with respect to q and derivatives are bounded.

M3: The projection algorithm (41) is applied to the coordination algorithm (38). Let D in (41) be a convex and compact set defined by $D := \mathcal{M}^{N_s} \times \mathcal{M}_p$, where $\mathcal{M}_p := [p_{\min}, p_{\max}]^{2N_s}$.

Remark 6 M1 can be viewed as a persistent excitation condition in adaptive control [27]. M2 can be satisfied, for instance, see (8) and [4]. M3 is used to satisfy the boundedness condition for the ODE approach and it is also very useful to model the realistic control saturations for mobile vehicles.

Lemma 7 Let $A_i(t) := A_i(t; x(t-1))$. Under M1 and M3, the matrix

$$A_i(t) := [I_m - K_i(t)\Phi_{*i}] \quad (48)$$

is a positive definite matrix for all $i \in \mathcal{I}$ and $t \in \mathbb{Z}_{>0}$. All eigenvalues of $A_i(t)$ in (48) are nonnegative and strictly less than 1, i.e.,

$$\lambda_{\min}(A_i(t)) > 0, \lambda_{\max}(A_i(t)) < 1.$$

Hence, the induced matrix two norm of $A_i(t)$ is strictly less than 1:

$$\|A_i(t)\| < 1, \forall i \in \mathcal{I}, \forall t \in \mathbb{Z}_{>0}. \quad (49)$$

Proof: By the definition of A_i , it is a symmetrical matrix.

$$A_i(t) = I_m - P_i(t)\Phi_{*i}^T(I_s + \Phi_{*i}P_i(t)\Phi_{*i}^T)^{-1}\Phi_{*i}.$$

where $P_i(t) \succ 0$ is a positive definite matrix. From (30), notice that

$$P_i(t-1) - P_i(t) \succ 0, P_i(t) = A_i(t)P_i(t-1) \succ 0,$$

implies $P_i(t-1)(I_m - A_i(t)) \succ 0$.

Hence, we conclude that $0 \prec A_i(t) = A_i^T(t) \prec I_m$. Moreover, since $A_i(t) \succ 0$, there exists a square root matrix F so that $A_i(t) = F^T F$ and $F = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_m})R$ where R is the orthonormal matrix and $\lambda_1 = \lambda_{\max}(A_i(t)) > \lambda_2 > \dots > \lambda_m = \lambda_{\min}(A_i(t)) > 0$. Since $A_i(t) = F^T F \prec I_m$ implies that $\sqrt{\lambda_{\max}(F^T F)} < 1$, we have $\lambda_{\max}(A_i(t)) = \lambda_{\max}(F^T F) = \|F\|^2 < 1$ and $\|A_i(t)\| = \sqrt{\lambda_{\max}(A_i^T(t)A_i(t))} < 1$. \square

Lemma 8 Consider the transformed recursive algorithm after applying Lemma 5 under assumptions M1-M3. Then the algorithm is subject to the regularity conditions C1-C11, and $(\mathcal{M}^{N_s} \setminus Z) \times \mathcal{M}_p \subset D \subset D_R$, where $\mathcal{M}_p = [p_{\min}, p_{\max}]^{2N_s}$ and Z is the set defined by

$$Z := \left\{ q \in \mathcal{M}^{N_s} \mid \sum_{j \in \{i\} \cup \mathcal{N}(i, q)} \phi(q_j)\phi^T(q_j) \neq 0, \forall i \in \mathcal{I} \right\}. \quad (50)$$

Moreover, $f(x)$ in (42) of Corollary 4 is given by

$$f(x) = \begin{bmatrix} p \\ -\nabla U(q) - (\hat{L}(q) + K_d)p - \nabla C(q) \end{bmatrix}, \quad (51)$$

where $C(q) \in \mathbb{R}_{\geq 0}$ is the collective performance cost function defined by

$$C(q) := k_4 \sum_{i \in \mathcal{I}} [\mu_{\max} - \mu(q_i)], \quad (52)$$

here $k_4 \in \mathbb{R}_{>0}$ is a gain factor and $\mu_{\max} \in \mathbb{R}_{>0}$ is the maximum of the field μ .

Proof: Verifications of C1-C11 are as follows:

- **C1:** This is satisfied by the measurement noise assumption in (17) and (18) with Lemma 7 under M1 and M3.
- **C2:** This is satisfied due to the assumption M2 and smooth and bounded derivatives of radial basis functions in $\nabla \hat{C}(\varphi, q)$ with respect to q .
- **C3:** $A(t; \cdot)$ and $B(t; \cdot)$ in (46) are functions of smooth radial basis functions, therefore, they are smooth in D_R .
- **C4:** We take the similar argument used in [31].

Notice that:

$$\begin{aligned} \varphi(t) - \bar{\varphi}(t) &= A(t; x)|_{\bar{x}(t-1)}(\varphi(t-1) - \bar{\varphi}(t-1)) \\ &\quad + \frac{\partial g}{\partial x} \Big|_{\substack{\bar{x}(t-1) \\ \bar{\varphi}(t-1)}}(x(t-1) - \bar{x}) \\ &= \frac{\partial g}{\partial x} \Big|_{\substack{\bar{x}(t-1) \\ \bar{\varphi}(t-1)}}(x(t-1) - \bar{x}) \\ &\quad + \sum_{i=2}^{t-n} \frac{\partial g}{\partial x} \Big|_{\substack{\bar{x}(t-i) \\ \bar{\varphi}(t-i)}}(x(t-i) - \bar{x}) \left(\prod_{j=1}^{i-1} A(t; x)|_{\bar{x}(t-j)} \right), \end{aligned}$$

where the mean value theorem was used for a smooth g with respect to x and φ . $[\tilde{x}^T(s), \tilde{\varphi}^T(s)]^T$ is a point between $[x^T(s), \varphi^T(s)]^T$ and $[\bar{x}^T(s), \bar{\varphi}^T(s)]^T$. From Lemma 7 under M1 and M3, we have:

$$\|A(t; x)|_{\bar{x}}\| \leq 1 - \delta < 1, \forall t. \quad (45)$$

Therefore we obtain:

$$\begin{aligned} \|\varphi(t) - \bar{\varphi}(t)\| &\leq \left\| \frac{\partial g}{\partial x} \Big|_{\substack{\bar{x}(t-1) \\ \bar{\varphi}(t-1)}} \right\| \|x(t-1) - \bar{x}\| \\ &\quad + \sum_{i=2}^{t-n} \left(\prod_{j=1}^{i-1} \|A(t; x)|_{\bar{x}(t-j)}\| \right) \left\| \frac{\partial g}{\partial x} \Big|_{\substack{\bar{x}(t-i) \\ \bar{\varphi}(t-i)}} \right\| \|x(t-i) - \bar{x}\| \\ &\leq \sum_{i=1}^{t-n} (1 - \delta)^{i-1} \max_{n \leq s \leq t} \left\| \frac{\partial g}{\partial x}(s) \right\| \max_{n \leq s \leq t} \|x(s) - \bar{x}\| \\ &< \frac{1}{\delta} C \max_{n \leq s \leq t} \|x(s) - \bar{x}\| < C \max_{n \leq s \leq t} \|x(n) - \bar{x}\|. \end{aligned}$$

- C5: For a fixed \bar{x} , notice that:

$$\begin{aligned} \bar{\varphi}_i(t, \bar{x}) &= \prod_{k=s+1}^t A(k; \bar{x}) \bar{\varphi}_i(s, \bar{x}) \\ &+ \sum_{j=s+1}^t \left[\prod_{k=j+1}^t A(k; \bar{x}) \right] B(j; \bar{x}) e(j), \quad i \in \{1, 2\}. \end{aligned}$$

Under *M1* and *M3*, $\|A(k; \bar{x})\| < \lambda(\bar{x})$ for all $k \in \{s+1, \dots, t\}$, where $\lambda(\bar{x}) < 1$. Hence we have:

$$\|\bar{\varphi}_1(t, \bar{x}) - \bar{\varphi}_2(t, \bar{x})\| < \lambda^{t-s}(\bar{x}) \|\bar{\varphi}_1(s, \bar{x}) - \bar{\varphi}_2(s, \bar{x})\|,$$

for all $\bar{x} \in (\mathcal{M}^{N_s} \setminus Z) \times \mathcal{M}_p \subset D_S$, where Z is the set defined in (50).

- C6: Elements of Q in (45) are deterministic functions of $x \in D_R$ except for $\nabla \hat{C}(\varphi(t), q)$. Thanks to *M1* and (28), for a fixed q , we have

$$\lim_{t \rightarrow \infty} \mathbb{E}(\nabla \hat{C}(\varphi(t), q)) = \nabla C(q),$$

which proves C6 and (51) simultaneously.

- C7: This is satisfied due to the measurement noise assumption in (17).
- C8, C9, C10, C11: These are satisfied by the time varying gain sequence defined in (34). \square

Finally, the global performance cost that sensing agents to minimize, is defined as

$$V(q(\tau), p(\tau)) := U(q(\tau)) + \frac{p^T(\tau)p(\tau)}{2} + C(q(\tau)). \quad (53)$$

We have the following theorem regarding the convergence properties of the proposed multi-agent system.

Theorem 9 For any initial state $x_0 = \text{col}(q_0, p_0) \in D$, where D is a compact set as in (41), we consider the recursive coordination algorithm obtained by Lemma 5 under conditions from Lemma 8. Let $D_A := \{x \in D \mid V(x) \leq a\}$ be a level-set of the cost function in (53). Let D_c be the set of all points in D_A , where $\frac{d}{dt}V(x) = 0$. Then every solution starting from D_A approaches the largest invariant set D_M contained in D_c with probability one as $t \rightarrow \infty$, or $\{x(t)\}$ has a cluster point on the boundary ∂D of D . Moreover, if $\{x(t)\}$ does not have a cluster point on ∂D and $(\hat{L}(q) + K_d) \succ 0$, $\forall x \in D$, then any point $x^* = \text{col}(q^*, 0)$ in D_M is a critical point of the cost function $V(x)$, which yields either a (local) minimum of $V(x)$ or an inflection point, i.e.,

$$\left. \frac{\partial V(x)}{\partial x} \right|_{x=x^*} = 0.$$

Proof: From Lemma 5, Lemma 8 and Corollary 4, the asymptotic trajectory $x(\tau) := \text{col}(q(\tau), p(\tau)) \in D_R$ is given by the associated ODE

$$\frac{dx(\tau)}{d\tau} = f(x(\tau)). \quad (54)$$

Taking the derivative of $V(x(\tau))$ in (53) with respect to τ and using (54), we obtain

$$\begin{aligned} \frac{dV(x(\tau))}{d\tau} &= \left(\frac{\partial V(x)}{\partial x} \right)^T f(x(\tau)) \\ &= \begin{bmatrix} \nabla U(q(\tau)) + \nabla C(q(\tau)) \\ p(\tau) \end{bmatrix}^T \\ &\quad \begin{bmatrix} p(\tau) \\ -\nabla U(q(\tau)) - \nabla C(q(\tau)) - (\hat{L}(q(\tau)) + K_d)p(\tau) \end{bmatrix} \\ &= -p^T(\tau)(\hat{L}(q(\tau)) + K_d)p(\tau) \leq 0. \end{aligned} \quad (55)$$

From (9) and (53), we conclude that $V(x)$ is radially-unbounded, i.e., $V(x) \rightarrow \infty$ as $\|x\| \rightarrow \infty$. Then

$$D_A := \{x \mid V(x) \leq a\}$$

is a bounded set with $\frac{d}{dt}V(x) \leq 0$ for all $x \in D_A$ as in (55), which is a positively invariant set. By LaSalle's invariant principle and Corollary 4, $x(t)$ approaches the largest invariant set D_M contained in D_c given by

$$\left\{ x(\tau) \mid \dot{V}(x(\tau)) = -p^T(\tau)(\hat{L}(q(\tau)) + K_d)p(\tau) = 0 \right\}, \quad (56)$$

with probability one as $t \rightarrow \infty$.

If $(\hat{L}(q) + K_d) \succ 0 \forall x \in D$, from (56), any point x^* in D_M is the form of $x^*(t) = \text{col}(q^*(t), 0)$. Moreover, from (51), we have $\dot{q}^*(t) \equiv 0$ and $0 \equiv -\nabla U(q^*) - \nabla C(q^*)$, which verifies that x^* is a critical point of the cost function $V(x)$. Hence this completes the proof. Q.E.D.

5 Simulation Results

We applied the proposed multi-agent system to static fields, which are represented by twenty five radial basis functions as depicted in the left side of Fig. 7 (uni-modal) and Fig. 9 (bi-modal). The estimated field was updated once per iteration used for the coordination of agents. Twenty five agents were launched at random positions away from the maximum of the field in the simulation study. Parameters used for the numerical evaluation are given in Table 1. Simulation results are evaluated for different parameters and conditions.

Table 1

Parameters in the simulation.

Parameters	Values
Number of agents N_s	25
Number of basis functions m	25
Surveillance region \mathcal{M}	$[-5, 5]^2$
(d, d_0, d_1)	$(0.6, 1.5d, 3.5d)$ $(0.4, 1.62d, 3.5d)$
Transmission range r	$4d$
Noise level W	1
(k_1, k_2, k_3, k_4)	$(0.1, 10, 0.1, 0.1)$
K_d	$I_{2N_s}; 0.1I_{2N_s}; 0I_{2N_s}$
Saturation limit D	$[-5, 5]^{2N_s} \times [-1, 1]^{2N_s}$
$\gamma(0)$	0.2
$\Theta(0)$	$0_{m \times 1}$
$P(0)$	$3I_m$

5.1 Standard Conditions

We consider the proposed multi-agent system under the standard operating conditions (used in Theorem 9), which include the projection algorithm defined in (41), velocity feedback ($K_d \succ 0$ as defined in (45)), and an artificial potential wall. Fig. 6-(a) shows that the recursively estimated field by agent 1 at the iteration time $t = 20$ under a noise level $W = 1$. The swarming agents have the equilibrium distance of $d = 0.6$ as defined in (7). The estimation error field is also shown with colored contour lines as in Fig. 6-(a). Fig. 6-(b) illustrates the estimated field by agent 1 at iteration time $t = 200$. The true field is illustrated in the left-side of Fig. 7. As shown in Fig. 6-(b), twenty five swarming agents have located the maximum point of the field successfully. The right-side of Fig. 7 shows the root mean square (RMS) values of the spatially averaged error field achieved by all agents with respect to the iteration time. All agents managed to bring the RMS values of the estimation error down around 2 after 150 iterations. With a bit higher damping coefficients contained in $K_d = I_{2N_s}$, the rate of convergence to the maximum point was slow as shown in the right-hand side of Fig. 7. Hence, the group of agents does not show much overshoot and oscillatory behavior around the maximum point. Agents converge to a configuration near the maximum point as $t \rightarrow \infty$.

The proposed multi-agent system with a smaller communication range and $K_d = 0I_{2N_s}$ is applied to a bi-modal static field, which is shown in the left-side of Fig. 9. Fig. 8 reminds of the fact that the proposed agents can split into different groups according to the configuration of the global network cost function V defined in (53). It is straightforward to understand that agent 1 does not have information on the other mode located at the upper-right side of the surveillance region as shown in Fig. 8, which results in higher RMS estimation error values plotted in the right-hand side of Fig. 9

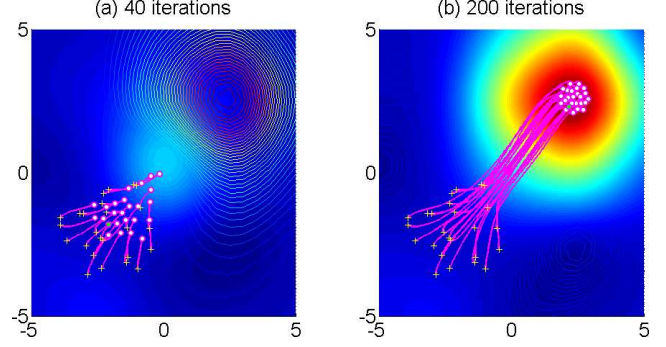


Fig. 6. Trajectories of twenty five learning agents for $d = 0.6$, $W = 1$ and $K_d = I_{2N_s}$, at iteration times $t = 40$ (a) and $t = 200$ (b) under the projection algorithm. The estimated field by agent 1 is shown as a background in colors. Agent 1 is plotted as a green dot. Thin contour lines represent the error field between the true field and the estimated field. (+) and (o) represent, respectively, initial and final locations. Solid lines represent trajectories of agents.

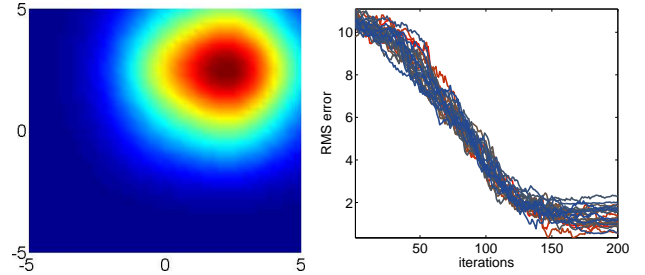


Fig. 7. A uni-modal field of interest (left). The root mean square (RMS) values of the spatially averaged error field achieved by all agents with respect to the iteration number (right). Parameters are $d = 0.6$, $W = 1$, and $K_d = I_{2N_s}$, and the projection was used.

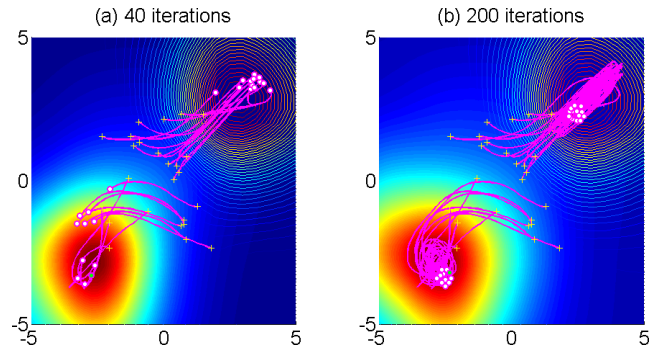


Fig. 8. (a), (b): The proposed agents are splitting into two groups for multi-modes under standard conditions. The estimated field by agent 1 is shown as a background in colors. Thin contour lines represent the error field between the true field and the estimated field.

as compared to those for the previous case (Fig. 7).

Fig. 10 illustrates a case without communication and the swarming capabilities of agents. Only a couple of agents manage to approach the maximum point with slow convergence rates as compared to the previous case in Fig. 6. The lowest RMS value of the estimation error achieved by agents

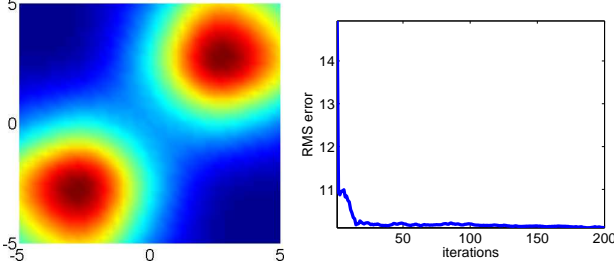


Fig. 9. A bi-modal field of interest (left). The root mean square (RMS) values of the spatially averaged error field achieved by agent 1 with respect to the iteration number for the bi-modal field of interest (right).

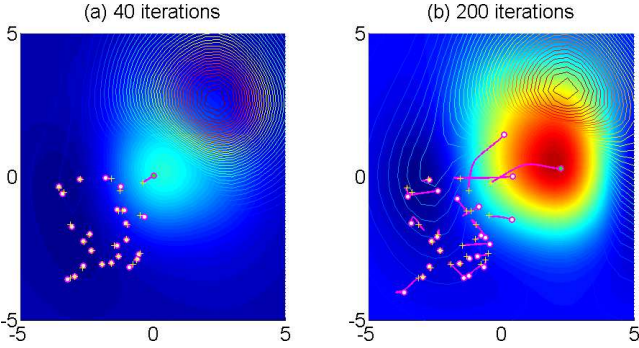


Fig. 10. (a), (b): Trajectories of agents without communication and the swarming algorithm for $d = 0.6$, $W = 1$, and $K_d = I_{2N_s}$, and the projection was not used.

was about 6. This simulation clearly justifies the usage of the communication network and swarming algorithms in this our proposed multi-agent system.

5.2 Without the Velocity Feedback

We consider a case without the velocity feedback (i.e., $K_d = 0I_{2N_s}$) for the uni-modal field of interest. Without the velocity feedback, there will be no dissipative terms once the consensus of velocities of agents is achieved, which explains the oscillatory behavior of agents in Fig. 11. The group disagreement function $\Psi_G(p(t)) = \frac{1}{2}p^T(t)\hat{L}(q(t))p(t)$ with respect to the iteration number is shown in Fig. 12.

We also consider a case without both the velocity feedback and the projection algorithm (i.e., no saturations on both positions and velocities) for the bi-modal field of interest. In this simulation, agents happened to locate two maximum points of the bi-modal field as depicted in Figs. 13 and 14. The group disagreement function and convergence rate of the agents are illustrated in Fig. 14. In this simulation, the artificial potential wall prevents agents from going outside of the compact surveillance region \mathcal{M} .

5.3 Without the Artificial Potential Wall

Finally, we consider a case without the potential wall and with the projection algorithm for the uni-modal field of in-

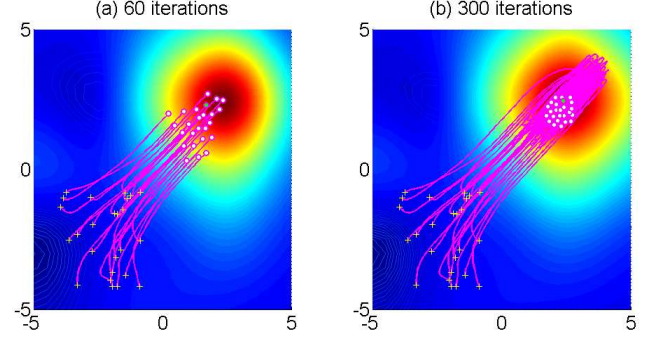


Fig. 11. (a), (b): Trajectories of agents for $d = 0.6$, $W = 1$, and $K_d = 0I_{2N_s}$, and the projection was used.

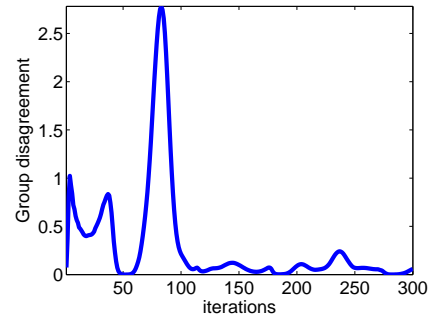


Fig. 12. The group disagreement function $\Psi_G(p(t))$ with respect to the iteration number. Parameters are $d = 0.6$, $W = 1$, and $K_d = 0I_{2N_s}$, and the projection was used.

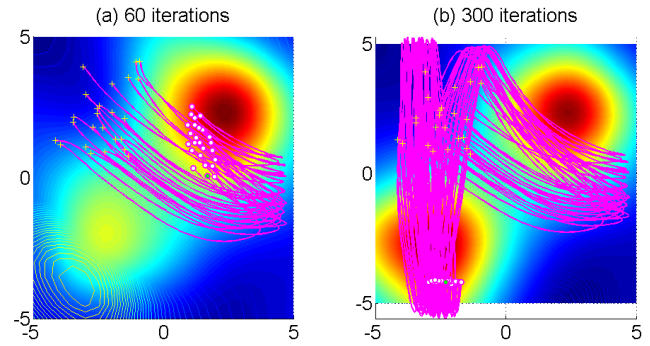


Fig. 13. (a), (b): Trajectories of agents for $d = 0.6$, $W = 1$, and $K_d = 0I_{2N_s}$, and the projection was not used.

terest. In addition, we relocate the maximum of the field at the boundary of the surveillance region. As can be seen in Fig. 15, agents with $K_d > 0$ have located the maximum point of the field and converge to a configuration around the boundary of the surveillance region. The projection algorithm ensures that agents stay inside of the compact set \mathcal{M} .

6 Conclusions

This paper presented a novel class of self-organizing sensing agents that form a swarm and learn through noisy measurements cooperatively with neighboring agents to estimate an

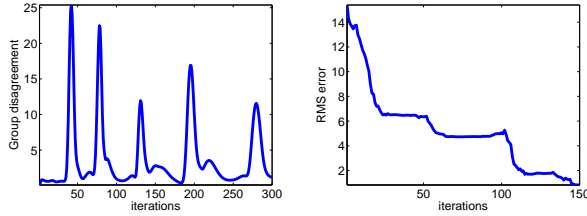


Fig. 14. The group disagreement function $\Psi_C(p(t))$ with respect to the iteration number (left). The root mean square (RMS) values of the spatially averaged error field achieved by agent 1 with respect to the iteration number (right). Parameters are $d = 0.6$, $W = 1$, and $K_d = 0I_{2N_s}$, and the projection was not used.

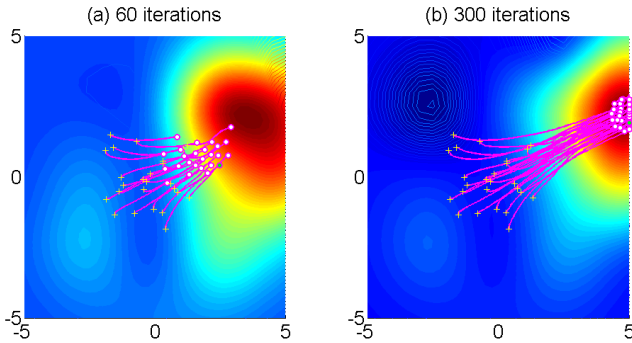


Fig. 15. (a), (b): The projection algorithm guarantees that agents are inside of the compact surveillance region $\mathcal{M} := [-5, 5]^2$ even without the artificial potential wall U_2 which pushes agents back into \mathcal{M} when they approach the boundary of \mathcal{R} .

unknown field of interest for gradient climbing. The proposed cooperatively learning control consists of motion coordination based on the recursive estimation of an unknown field of interest with measurement noise. Our strategy of the cooperative learning control can be applied to a large class of coordination algorithms for mobile agents in a situation where the field of interest is not known a priori and to be estimated. We have shown that the closed-loop dynamics of the proposed multi-agent system can be transformed into a form of a stochastic approximation algorithm. Hence, the convergence properties of the proposed multi-agent system were analyzed using the ODE approach and verified by a simulation study with respect to different parameters and conditions. Simulation study on the proposed multi-agent system and learning agents without communication and the swarming effort clearly demonstrated the advantage of the communication network and the swarming effort. A possible future work is to deploy heterogeneous mobile sensing agents (each with different parameters) to study how mixture of different types of agents can be coordinated for sub-optimality by a consensus type algorithm that enforces all agents converge to the same set of parameters.

Acknowledgements

The authors would like to thank professor Lennart Ljung, with the department of Automatic Control, Linköping Insti-

tute of Technology, Linköping, Sweden for kindly providing his technical report to the first author. The results from the report were frequently used in the paper. This work was partially supported by Intramural Research Grants Program (IRGP) from Michigan State University.

References

- [1] 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.
- [2] A. Jadbabie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, pp. 988–1001, June 2003.
- [3] H. G. Tanner, A. Jadbabaie, and G. J. Pappas, "Stability of flocking motion," technical report, University of Pennsylvania, 2003.
- [4] Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithm and theory," *IEEE Transactions on Automatic Control*, vol. 51, pp. 401–420, March 2006.
- [5] W. Ren and R. W. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Transactions on Automatic Control*, vol. 50, pp. 655–661, May 2005.
- [6] P. Ögren, E. Fiorelli, and N. E. Leonard, "Cooperative control of mobile sensor networks: Adaptive gradient climbing in a distributed environment," *IEEE Transaction on Automatic Control*, vol. 49, p. 1292, August 2004.
- [7] "DOD/ONR MURI: adaptive sampling and prediction project." Also available at <http://www.princeton.edu/~dcs1/asap/>.
- [8] "Center of Excellence for Great Lakes, Harmful Algal Bloom Event Response." Also available at <http://www.glerl.noaa.gov/res/Centers/HABS/habs.html>.
- [9] "Harmful Algal BloomS Observing System (HABSOS) by National Oceanic and Atmospheric Administration (NOAA)."
- [10] J. Adler, "Chemotaxis in bacteria," *Journal of Supramolecular Structure*, vol. 4, no. 3, pp. 305–317, 1966.
- [11] A. Dhariwal, G. S. Sukhatme, and A. A. G. Requicha, "Bacterium-inspired robots for environmental monitoring," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2004.
- [12] D. Grünbaum, "Schooling as a strategy for taxis in a noisy environment," *Evolutionary Ecology*, vol. 12, pp. 503–522, 1998.
- [13] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," *Computer Graphics*, vol. 21, no. 4, pp. 25–34, 1987.
- [14] R. C. Eberhart, Y. Shi, and J. Kennedy, *Swarm Intelligence*. The Morgan Kaufmann Series in Artificial Intelligence, Academic Press, 2001.
- [15] E. Bonabeau, M. Dorigo, and G. Theraulaz, *Swarm Intelligence: From Natural to Artificial Systems*. Santa Fe Institute Studies on the Sciences of Complexity, Oxford University Press, 1999.
- [16] R. Graham and J. Cortes, "Asymptotic optimality of multicenter voronoi configurations for random field estimation," *submitted to IEEE Transactions on Automatic Control*, August 2007.
- [17] L. Ljung, "Analysis of recursive stochastic algorithms," *IEEE Transactions on Automatic Control*, vol. 22, no. 4, pp. 551–575, 1977.
- [18] H. J. Kushner and G. G. Yin, *Stochastic Approximation Algorithms and Applications*. Springer, 1997.
- [19] C. Godsil and G. Royle, *Algebraic Graph Theory*, vol. 207 of *Graduate Text in Mathematics*. Springer-Verlag, 2001.

- [20] J. Choi, S. Oh, and R. Horowitz, "Cooperatively learning mobile agents for gradient climbing," in *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007.
- [21] J. Choi, J. Lee, and S. Oh, "Swarm Intelligence for Achieving the Global Maximum using Spatio-Temporal Gaussian Processes," to appear in the *2008 American Control Conference (ACC)*, 2008.
- [22] J. Choi, J. Lee, and S. Oh, "Biologically-inspired Navigation Strategies for Swarm Intelligence using Spatial Gaussian Processes," to appear in the *International Federation of Automatic Control (IFAC) World Congress*, 2008.
- [23] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, Cambridge, Massachusetts, London, England, 2006.
- [24] D. J. C. MacKay, "Introduction to Gaussian processes," in *Neural Networks and Machine Learning* (C. M. Bishop, ed.), NATO ASI Series, pp. 133–166, Kluwer, 1998.
- [25] N. A. C. Cressie, *Statistics for Spatial Data*. A Wiley-Interscience Publication, John Wiley and Sons, Inc., 1991.
- [26] N. Cressie and C. K. Wikle, *Encyclopedia of Environments*, vol. 4, ch. Space-time Kalman filter, pp. 2045–2049. John Wiley and Sons, Ltd, Chichester, 2002.
- [27] K. J. Åström and B. Wittenmark, *Adaptive Control*. Addison Wesley, 2 ed., 1995.
- [28] L. Ljung, "Theorems for the asymptotic analysis of recursive, stochastic algorithms," technical report 7522, Department of Automatic Control, Lund Institute of Technology, 1975.
- [29] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*. Cambridge, Massachusetts, London, England: The MIT Press, 1983.
- [30] T. Wigren, "Convergence analysis of recursive identification algorithms based on the nonlinear wiener model," *IEEE Transactions on Automatic Control*, vol. 39, pp. 2191–2206, November 1994.
- [31] L. Brus, "Recursive black-box identification of nonlinear state-space ode models.," it licentiate theses; 2006-001, Uppsala Universitet, 2006.