Automatica 45 (2009) 2802-2814

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica



Jongeun Choi^{a,b,*}, Songhwai Oh^c, Roberto Horowitz^d

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

^b Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824-1226, USA

^c School of Electrical Engineering and Computer Science, Seoul National University, Seoul, Republic of Korea

^d Department of Mechanical Engineering, University of California, Berkeley, CA 94720-1740, USA

ARTICLE INFO

Article history: Received 29 March 2008 Received in revised form 4 May 2009 Accepted 15 September 2009 Available online 7 October 2009

Keywords: Cooperative control Multi-agent systems Recursive parameter estimation

ABSTRACT

This paper presents an algorithm and analysis of distributed learning and cooperative control for a multiagent system so that a global goal of the overall system can be achieved by locally acting agents. We consider a resource-constrained multi-agent system, in which each agent has limited capabilities in terms of sensing, computation, and communication. 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 using collective measurements from itself and nearby agents. Each agent then moves towards peaks of the field using the gradient of its estimated field while avoiding collision and maintaining communication connectivity. The proposed algorithm is based on a recursive spatial estimation of an unknown field. We show that the closedloop dynamics of the proposed multi-agent system can be transformed into a form of a stochastic approximation algorithm and prove its convergence using Ljung's ordinary differential equation (ODE) approach. We also present extensive simulation results supporting our theoretical results.

© 2009 Elsevier Ltd. All rights reserved.

automatica

1. Introduction

Recently, there has been a growing interest in wireless sensor networks (Culler, Estrin, & Srivastava, 2004; Estrin, Culler, Pister, & Sukhatme, 2002). A wireless sensor network consists of a large number of sensor nodes. Each sensor node can perform sensing and computation, and sensor nodes form an ad hoc wireless network for communication. Applications of wireless sensor networks include, but not limited to, environment monitoring, building comfort control, traffic control, manufacturing and plant automation, and surveillance systems (Oh, Schenato, Chen, & Sastry, 2007, and the references therein). However, faced with the dynamic nature of environment, stationary sensor networks are sometimes inadequate and a mobile sensing technology shows superior performance in terms of its adaptability and high-resolution sampling capability (Singh et al., 2007).

Mobility in a sensor network can increase its sensing coverage both in space and time and robustness against dynamic changes in the environment. As in wireless sensor networks, each mobile agent is resource constrained; it operates under a short communication range, limited memory, and limited computational power. These mobile agents form an ad hoc wireless network for communication. Although each agent has limited capabilities, as a group, they can perform various tasks at a level which is compatible to a small number of high-end mobile agent. To perform various tasks such as exploration, surveillance, and environmental monitoring, distributed coordination of mobile sensing agents is required to achieve a global goal; and it has received significant attention recently (Cortes, Martinez, Karatas, & Bullo, 2004; Jadbabie, Lin, & Morse, 2003; Olfati-Saber, 2006; Ren & Beard, 2005; Tanner, Jadbabaie, & Pappas, 2003).

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 (Őgren, Fiorelli, & Leonard, 2004, DOD/ONR MURI). This is due to numerous applications of tracking toxins in a distributed environment. An interesting practical application is to trace harmful algal blooms in a lake. For certain environmental conditions, rapidly reproducing harmful algal blooms in lakes



[☆] The material in this paper was partially presented at the 46th IEEE Conference on Decision and Control, December 12–14, 2007 at the Hilton New Orleans Riverside in New Orleans, Louisiana USA. This paper was recommended for publication in revised form by Associate Editor Brett Ninness under the direction of Editor Torsten Söderström.

^{*} Corresponding author at: Department of Mechanical Engineering, Michigan State University, East Lansing, MI 48824-1226, USA. Tel.: +1 517 432 3164; fax: +1 517 353 1750.

E-mail addresses: jchoi@egr.msu.edu (J. Choi), songhwai@snu.ac.kr (S. Oh), horowitz@me.berkeley.edu (R. Horowitz).

^{0005-1098/\$ -} see front matter © 2009 Elsevier Ltd. All rights reserved. doi:10.1016/j.automatica.2009.09.025



Fig. 1. The estimated field of chlorophyll generated by the harmful algal blooms observation system (Harmful Algal BloomS Observing System) by NOAA. (Photo courtesy of NOAA).

and in oceans can produce cyanotoxins (Center of Excellence for Great Lakes). Exposure to water contaminated with algal cyanotoxins causes serious acute and chronic health effects to humans and adverse effects to aquatic life (Center of Excellence for Great Lakes; Harmful Algal BloomS Observing System). The level of chlorophyll is a measure closely related to algal and cyanobacterial biomass. 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, theses 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 biologically inspired *chemotaxis* (Adler, 1966; Dhariwal, Sukhatme, & Requicha, 2004), 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 DOD/ONR MURI and Őgren et al. (2004). 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 (1998) 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.

Olfati-Saber (2006) and Tanner et al. (2003) presented comprehensive analyses of the flocking algorithm by Reynolds (1987). 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 (Reynolds, 1987). 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 (Bonabeau, Dorigo, & Theraulaz, 1999; Eberhart, Shi, & Kennedy, 2001).

The state-of-the-art technology in spatial estimation by mobile sensor networks is as follows. A distributed interpolation scheme described in Martinez (in press) for field estimation by mobile sensor networks uses a distributed nonparametric inference method and made to be compatible with



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. In the color map, red color denotes the highest scalar value while blue color represents the lowest value. 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.

coverage control (e.g., Cortes et al., 2004). A distributed Kriged Kalman filter for robotic sensor networks to estimate a field of interest is described in Cortés (in press). In each iteration of this Kriged Kalman filter, agents execute a consensus algorithm based on new measurements, e.g., Olfati-Saber and Shamma (2005), for computing average values of interests. Then an iterative weighted least-squares algorithm is performed to compute state estimates of the field.

In this paper, we develop novel distributed learning and cooperative control algorithms for multi-agent systems by extending the recent development in the flocking algorithm (Olfati-Saber, 2006; Tanner et al., 2003). 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 of a surveillance region. Hence, agents cannot perform the coverage control as in Cortes et al. (2004), Graham and Cortés (2009) and Martinez (in press). Instead, each agent takes a measurement and also receives measurements from its neighboring agents within its communication range. Upon receiving collective measurements, each agent recursively updates its own estimate of an unknown static field of interest. The recursive estimation is based on a radial basis function network in order to represent a wide range of physical phenomena. To locate the maximum of the field, the sensing agent will climb the gradient of its own estimated field. In our proposed approach, each agent has its own recursive estimation of the field based on collective measurements from itself and its neighbors without requiring the consensus step of the Kriged Kalman filter (Cortés, in press). 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 the multi-agent system into a form of a stochastic approximation algorithm. Our theoretical results are based on the ordinary differential equation (ODE) approach (Kushner & Yin, 1997; Ljung, 1977). 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, $A \succeq 0$). $I_n \in \mathbb{R}^{n \times n}$ and $0_n \in \mathbb{R}^{n \times n}$ denote the identity and zero matrices, respectively. $0_{m \times n} \in \mathbb{R}^{m \times n}$ denotes the m by n zero matrix. However, the subscript can be omitted for the sake of brevity when it is obvious from the context. The gradient of a differentiable real function $\phi(q) : \mathbb{R}^{2n} \to \mathbb{R}$ with respect to its vector domain q is denoted by

$$\nabla \phi(q) := \begin{bmatrix} \frac{\partial \phi(q)}{\partial q_1}^{\mathrm{T}} & \frac{\partial \phi(q)}{\partial q_2}^{\mathrm{T}} & \cdots & \frac{\partial \phi(q)}{\partial q_n}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{2n},$$

where $q = col(q_1, ..., q_n) \in \mathbb{R}^{2n}$ and $q_i \in \mathbb{R}^2$. To be concise, the gradient of $\phi(q)$ with respect to q_i is denoted by

$$abla \phi(q_i) \coloneqq rac{\partial \phi(q)}{\partial q_i} \in \mathbb{R}^2.$$

The norm $\|\cdot\|$ will denote the standard Euclidean norm (or 2-norm) on vectors. The induced matrix 2-norm is defined as

$$||A|| := \sup_{x \neq 0} \frac{||Ax||}{||x||},$$

where $A \in \mathbb{R}^{m \times n}$. Other notation will be explained in due course.

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 $\mathfrak{l} := \{1, 2, \ldots, N_s\}$. 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 $q := \operatorname{col}(q_1, q_2, \ldots, q_{N_s}) \in \mathbb{R}^{2N_s}$ be the configuration of the multi-agent system. The discrete time, high-level dynamics of agent *i* is modeled by

$$q_i(t + \Delta_t) = q_i(t) + \Delta_t v_i(t), \tag{1}$$

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



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

of the field $\mu(q_i(t))$ and sensor noise w(t), at its position $q_i(t)$ and a sampled time t,

$$y(q_i(t)) \coloneqq \mu(q_i(t)) + w(t), \tag{2}$$

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

The proposed algorithm will be 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 will maintain its own local estimate of the field and will update the estimate using collective measurements from itself and nearby agents. Each agent will then be programmed to move towards peaks of the field using the gradient of its estimated field.

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 $G(q) := (\mathfrak{l}, \mathfrak{E}(q))$ 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 $||q_i(t) - q_i(t)|| \le r$. For example, as shown in Fig. 3, agent 2 communicates with and collects measurements from agents 1 and 3 within its 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 $a_{ii} = 0$ otherwise. Notice that an adjacency matrix A can be also defined in a smooth fashion in terms of q (Olfati-Saber, 2006). The scalar graph Lapla-cian $L = [l_{ij}] \in \mathbb{R}^{N_s \times N_s}$ 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 two-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 the example shown in Fig. 3 are:

$$A = k_3 \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad L = k_3 \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & -1 & 0 \\ -1 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$
$$\hat{L} = L \otimes I_2 = k_3 \begin{bmatrix} 2I_2 & -I_2 & -I_2 & 0_2 \\ -I_2 & 2I_2 & -I_2 & 0_2 \\ -I_2 & -I_2 & 2I_2 & 0_2 \\ 0_2 & 0_2 & 0_2 & 0_2 \end{bmatrix}.$$

Let $p_i \in \mathbb{R}^2$ be the state of agent *i* for $i \in \mathcal{I}$ under the 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 $\Psi_G: \mathbb{R}^{2N_s} \to \mathbb{R}_{\geq 0}$ evaluates the group disagreement in the network of agents:

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

where $p := \operatorname{col}(p_1, p_2, \ldots, p_{N_s}) \in \mathbb{R}^{2N_s}$. A disagreement function (Godsil & Royle, 2001; Olfati-Saber, 2006) can be obtained via the Laplacian \hat{L} :

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

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

$$\nabla \Psi_G(p) = Lp. \tag{5}$$

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 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 Choi, Oh, and Horowitz (2007), Olfati-Saber (2006) and Tanner et al. (2003) 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

$$U_{1}(q) := \sum_{i} \sum_{j \in \mathcal{N}(i,q), j \neq i} U_{ij}(\|q_{i} - q_{j}\|^{2})$$

=
$$\sum_{i} \sum_{j \in \mathcal{N}(i,q), j \neq i} U_{ij}(r_{ij}), \qquad (6)$$

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

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

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

$$\nabla U_{1}(q_{i}) = \frac{\partial U_{1}(q)}{\partial q_{i}} = \sum_{j \neq i} \frac{\partial U_{ij}(r)}{\partial r} \Big|_{r=r_{ij}} 2(q_{i} - q_{j})$$

$$= \begin{cases} \sum_{j \neq i} \frac{(r_{ij} - d^{2})(q_{i} - q_{j})}{(\alpha + r_{ij})^{2}} & \text{if } r_{ij} < d_{0}^{2} \\ \sum_{j \neq i} \rho \left(\frac{\sqrt{r_{ij}} - d_{0}}{|d_{1} - d_{0}|}\right) \frac{\|d_{0}^{2} - d^{2}\|}{(\alpha + d_{0}^{2})^{2}} (q_{i} - q_{j}) & \text{if } r_{ij} \ge d_{0}^{2}, \end{cases}$$
(8)

where $\rho : \mathbb{R}_{\geq 0} \rightarrow [0, 1]$ is the bump function (Olfati-Saber, 2006)

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



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

Notice that ρ varies smoothly from 1 to 0 as the scalar input increases. (6)-(8) will be used to 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 Eqs. (6)-(8), a non-zero gain factor α is introduced to prevent the reaction force from diverging at $r_{ii} = ||q_i - q_i||^2 = 0$. As illustrated in Fig. 4, this potential yields a reaction force that is attracting when the agents are apart 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 $\sqrt{r_{ii}}$ increases from zero. For $\sqrt{r_{ii}} > d_0$, the bump function will shape 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 i 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 continuously differentiable 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. U₂ enforces each agent to stay inside the closed and connected surveillance region in \mathcal{M} and prevents collisions with obstacles in \mathcal{M} . Define the total artificial potential by

$$U(q) \coloneqq k_1 U_1(q) + k_2 U_2(q),$$
 (9)

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

network of radial basis functions¹:

$$\mu(\nu) := \sum_{j=1}^{m} \phi_j(\nu) \theta_j = \phi^{\mathrm{T}}(\nu) \Theta, \qquad (10)$$

where $\phi^{T}(v)$ and Θ are defined respectively by

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

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

$$\phi_j(\nu) \coloneqq \frac{1}{\Gamma_j} \exp\left(\frac{-\|\nu - \kappa_j\|^2}{2\sigma_j^2}\right),\tag{11}$$

where σ_j is the width of the Gaussian basis and Γ_j is a normalizing constant. Centers of basis functions $\{\kappa_j \mid j \in \{1, ..., m\}\}$ are uniformly distributed in the surveillance region $\mathcal{M}. \Theta \in \mathbf{\Theta} \subset \mathbb{R}^m$ is the true parameter of the regression model in (10). From (2), we have observations through sensors at the location v_k , $y(v_k) = \phi^T(v_k)\Theta + w(k)$, where k is a measurement sampling index. Based on the observations and regressors $\{(y(v_k), \phi(v_k))\}_{k=1}^n$, our objective is to find $\hat{\Theta}$ which minimizes the least-squares error:

$$\sum_{k=1}^{n} |y(v_k) - \phi^{\mathsf{T}}(v_k)\hat{\Theta}|^2.$$
 (12)

Remark 1. Our environmental model in (10) can be viewed as a radial basis function network to model nonlinear spatial phenomena. Other popular approaches to model nonlinear spatial phenomena include Gaussian processes (MacKay, 1998; Rasmussen & Williams, 2006) and kriging models (Cressie, 1991). A dynamical version of (10) was used to represent a time-varying trend in the space-time Kalman filter model (Cressie & Wikle, 2002) for modeling spatiotemporal random fields. Recently, these approaches have been frequently adopted for mobile sensor networks (Cortés, in press ; Graham & Cortés, 2009; Martinez, in press).

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 Choi et al. (2008a,b). For a set $\{(y(v_k), \phi(v_k))\}_{k=1}^n$, the optimal least-squares estimation solution that minimizes the error function in (12) is well known (Åström & Wittenmark, 1995):

$$\hat{\Theta}(n) = P(n, 1)\Phi^{\mathrm{T}}(n, 1)Y(n, 1),$$
(13)

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

$$Y(n,s) := \begin{bmatrix} y(s) & y(s+1) & \cdots & y(n) \end{bmatrix}^{T} \in \mathbb{R}^{n-s+1},$$

$$\Phi(n,s) := \begin{bmatrix} \phi(s) & \cdots & \phi(n) \end{bmatrix}^{T} \in \mathbb{R}^{n-s+1 \times m},$$

$$P(n,s) := \begin{bmatrix} \Phi^{T}(n,s)\Phi(n,s) \end{bmatrix}^{-1}$$

$$= \left[\sum_{k=s}^{n} \phi(k)\phi^{T}(k) \right]^{-1} \in \mathbb{R}^{m \times m}.$$

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 a number *s* 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 collectively measured samples at a number *s* of points. We then have the following algorithm. Assume that $\Phi^{T}(t)\Phi(t)$ is nonsingular for all *t*. For the collected number *s* of observations and regressors $\{(y(k), \phi(k))\}_{k=n-s+1}^{n}$, consider the recursive algorithm given as

$$K(n) = P(n-s)\Phi_*^{\mathsf{T}} \left[I_s + \Phi_* P(n-s)\Phi_*^{\mathsf{T}} \right]^{-1},$$

$$P(n) = \left[I_m - K(n)\Phi_* \right] P(n-s),$$

$$\hat{\Theta}(n) = \hat{\Theta}(n-s) + K(n) \left[Y_* - \Phi_* \hat{\Theta}(n-s) \right],$$

$$\hat{\mu}(\nu) := \phi^{\mathsf{T}}(\nu) \hat{\Theta}(n),$$

(14)

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 (14) is the least-squares estimation that minimizes the error function in (12).

Remark 2. $\Phi^{T}(n)\Phi(n)$ is always singular for n < m. $\Phi^{T}(n)\Phi(n)$ is nonsingular for $n \ge m$ 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 $\phi_i(v) = \phi_j(v)$. In practice, each agent starts the recursive LSE algorithm in (14) 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 the agent keeps running the recursive algorithm with new measurements. With these initial values, we have

$$P^{-1}(n) := P^{-1}(0) + \Phi^{\mathrm{T}}(n)\Phi(n) > 0.$$
(15)

In the next subsection, we elaborate on the case of noisy observations 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:

$$\mathbb{E}(w(k)) = 0, \qquad \mathbb{E}(w(k)w(\ell)) = \begin{cases} W > 0 & \text{if } k = \ell \\ 0 & \text{if } k \neq \ell, \end{cases}$$
(16)

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

$$w(k)| < L$$
 with probability one (w.p.1) $\forall k$. (17)

Given the measurement data set

$$\{y(\mu) \mid \mu \in S\}$$
, where $S = \{v_k \mid 1 \le k \le n\}$

and the sensor noise $\{w(k) \mid k \in \{1, ..., n\}\}$ defined in (16) and (17), an agent will estimate $\hat{\Theta}(n)$ using the recursive LSE algorithm in (14). 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 $v \in \mathcal{M}$ by

$$\tilde{\mu}(S,\nu) \coloneqq \hat{\mu}(S,\nu) - \mu(\nu) = \phi^{\mathrm{T}}(\nu)\tilde{\Theta}(|S|), \qquad (18)$$

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

$$\tilde{\mu}(S,\nu) = \mathbb{E}(\tilde{\mu}(S,\nu)) + \epsilon(S,\nu), \tag{19}$$

¹ We have considered a 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 Cortés (in press), Choi, Lee, and Oh (2008a,b).

where

$$\mathbb{E}(\tilde{\mu}(S, \nu)) \coloneqq \phi^{\mathrm{T}}(\nu) \left[P(|S|) \sum_{\nu_t \in S} \phi(\nu_t) \phi^{\mathrm{T}}(\nu_t) - I_m \right] \Theta,$$

$$\epsilon(S, \nu) \coloneqq \phi^{\mathrm{T}}(\nu) \left[P(|S|) \sum_{t=1}^{|S|} \phi(\nu_t) w(t) \right],$$

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

$$\lim_{|S| \to \infty} \mathbb{E}(\tilde{\mu}(S, \nu)) = 0, \quad \forall \nu \in \mathcal{M}.$$
 (20)

The variance of the estimation error is given by

$$\mathbb{E}(\epsilon(S, \nu)\epsilon^{\mathrm{T}}(S, \nu)) = \phi^{\mathrm{T}}(\nu)WP(|S|)\phi(\nu),$$

$$= \phi^{\mathrm{T}}(\nu)\frac{W}{|S|}R^{-1}(S)\phi(\nu),$$
 (21)

where *R*(*S*) is defined by

$$R(S) := \left[\frac{P^{-1}(0)}{|S|} + \frac{1}{|S|} \sum_{\nu_k \in S} \phi(\nu_k) \phi^{\mathrm{T}}(\nu_k)\right].$$
(22)

Remark 3. From (21), 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

$$\nabla \mu(\nu) := \left. \frac{\partial \mu(x)}{\partial x} \right|_{x=\nu}.$$
(23)

From (10), we have

$$\nabla \mu(\nu) = \left. \frac{\partial \phi^{\mathrm{T}}(x)}{\partial x} \right|_{x=\nu} \Theta \eqqcolon \phi^{/\mathrm{T}}(\nu) \Theta \in \mathbb{R}^{2 \times 1},$$
(24)

where $\phi^{T}(v) \in \mathbb{R}^{2 \times m}$. Thus, the gradient of the estimated field based on observations $S := \{v_k\}_{k=1}^n$ and $\{y(\mu)\}_{\mu \in S}$ is given by

$$\nabla \hat{\mu}(S, \nu) := \phi^{T}(\nu) \hat{\Theta}(|S|) \in \mathbb{R}^{2 \times 1}.$$
(25)

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

$$\nabla \tilde{\mu}(S, \nu) := \phi^{T}(\nu) \hat{\Theta}(|S|) - \nabla \mu(\nu) = \phi^{T}(\nu) \tilde{\Theta}(|S|)$$
$$= \mathbb{E}(\nabla \tilde{\mu}(S, \nu)) + \nabla \epsilon(S, \nu), \qquad (26)$$

where

$$\mathbb{E}(\nabla \tilde{\mu}(S, \nu)) = \phi^{T}(\nu) \left[P(|S|) \sum_{\nu_{k} \in S} \phi(\nu_{k}) \phi^{T}(\nu_{k}) - I_{m} \right] \Theta,$$

$$\nabla \epsilon(S, \nu) \coloneqq \phi^{T}(\nu) \left[P(|S|) \sum_{k=1}^{|S|} \phi(\nu_{k}) w(k) \right].$$

Analogous to (20) and (21), for $\Phi_*^T \Phi_* \succ 0$, the gradient estimator is asymptotically unbiased

$$\lim_{|S| \to \infty} \mathbb{E}(\nabla \tilde{\mu}(S, \nu)) = 0, \quad \forall \nu \in \mathcal{M},$$
(27)

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

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

where R(S) is defined in (22). Now we present our cooperatively learning control protocol.

3.2. Cooperative control

Each mobile agent receives measurements from neighbors. Then it updates its gradient of the estimated field using $\hat{\Theta}$ from the recursive algorithm presented in (14). 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 (14). 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 (14) such that the resulting recursive algorithm with the new time index for agent *i* at its position $q_i(t)$ is given by

$$K_{i}(t+1) = P_{i}(t)\Phi_{*i}^{T} \left(I_{s} + \Phi_{*i}P_{i}(t)\Phi_{*i}^{T} \right)^{-1},$$

$$P_{i}(t+1) = (I_{m} - K_{i}(t+1)\Phi_{*i})P_{i}(t),$$

$$\hat{\Theta}_{i}(t+1) = \hat{\Theta}_{i}(t) + K_{i}(t+1) \left[Y_{*i} - \Phi_{*i}\hat{\Theta}_{i}(t) \right],$$

$$\nabla \hat{\mu}_{i}(t, q_{i}(t)) = \phi^{T}(q_{i}(t))\hat{\Theta}_{i}(t+1),$$
(28)

where $\nabla \hat{\mu}_i(t, v) : \mathbb{Z}_{\geq 0} \times \mathcal{M} \to \mathbb{R}^2$ denotes the gradient of the estimated field at v 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 (14). Y_{*i} is the collection of cooperatively measured data. From (2), for all $j \in \mathcal{N}(i, q(t)) \cup \{i\}$, we have

$$Y_{*i} = \Phi_{*i}\Theta + \begin{bmatrix} \vdots \\ w_j(k) \\ \vdots \end{bmatrix} =: \Phi_{*i}\Theta + w_{*i}(t),$$
(29)

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 $j \in I$. We also define a new variable $w_{*i}(t)$ as in (29) 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

$$v_i(t+1) := \frac{\gamma(t+1)}{\Delta_t} \left[\frac{\Delta_t}{\gamma(t)} v_i(t) + \gamma(t) u_i(t) \right],$$
(30)
with

$$u_{i}(t) \coloneqq -\nabla U(q_{i}(t)) - k_{di} \frac{\Delta_{t}}{\gamma(t)} v_{i}(t) + \sum_{j \in \mathcal{N}(i,q(t))} a_{ij}(q(t)) \left(\frac{\Delta_{t}(v_{j}(t) - v_{i}(t))}{\gamma(t)} \right) + k_{4} \nabla \hat{\mu}_{i}(t, q_{i}(t)),$$
(31)

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 righthand side of (31) is the gradient of the artificial potential defined in (9) which attracts agents while avoiding collisions among them.

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

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 (31) provides damping. The third term in (31) 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 (30) with a standard adaptive gain sequence $\gamma(t)$ that satisfies the following properties

$$\gamma(t) > 0, \quad \sum_{t=1}^{\infty} \gamma(t) = \infty, \qquad \sum_{t=1}^{\infty} \gamma^2(t) < \infty,$$
(32)

 $\lim \sup \left[1/\gamma(t) - 1/\gamma(t-1) \right] < \infty.$

This gain sequence is often used for stochastic approximation algorithms (Kushner & Yin, 1997; Ljung, 1977) and enables us to apply the ODE approach (Ljung, 1977, 1975; Ljung & Söderström, 1983) 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 $v_i(t)$:

$$p_i(t) := \frac{\Delta_t}{\gamma(t)} v_i(t), \tag{33}$$

where $v_i(t)$ is the control input to agent *i* as defined in (30). After the change of variables in (33), the resulting dynamics of agent *i* is given by

$$\begin{cases} q_i(t+1) = q_i(t) + \gamma(t)p_i(t), \\ p_i(t+1) = p_i(t) + \gamma(t)u_i(t), \end{cases}$$
(34)

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

Incorporating the discrete time model in (34) along with the proposed control in (30) and (31) gives

$$q_{i}(t+1) = q_{i}(t) + \gamma(t)p_{i}(t),$$

$$p_{i}(t+1) = p_{i}(t) + \gamma(t) \{-\nabla U(q_{i}(t)) - k_{di}p_{i}(t) - \nabla \Psi_{G}(p_{i}(t)) + k_{4}\phi^{/T}(q_{i}(t))\hat{\Theta}_{i}(t+1)\},$$
(35)

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 Section 4, we will transform our multi-agent system into a recursive stochastic algorithm with states

$$\mathbf{x}(t) := \operatorname{col}(q_1, \ldots, q_{N_s}(t), p_1(t), \ldots, p_{N_s}(t)),$$

and

$$\varphi(t) \coloneqq \operatorname{col}(\tilde{\Theta}_1(t), \ldots, \tilde{\Theta}_{N_s}(t)).$$

4. Convergence analysis

In order to analyze the convergence properties of (28), (35) and (32), we utilize Ljung's ordinary differential equation (ODE) approach developed in Ljung (1977, 1975) and Ljung and Söderström (1983). In particular, Ljung (1977, 1975) presented an analysis technique of general recursive stochastic algorithms in the canonical form of

$$x(t) = x(t-1) + \gamma(t)Q(t; x(t-1), \varphi(t)),$$
(36)

along with the observation process

$$\varphi(t) = g(t; x(t-1), \varphi(t-1), e(t)).$$
(37)

In order to use the ODE approach, for this nonlinear observation process in (37), the following regularity conditions in Ljung (1975) need to be satisfied. Let D_R be a subset of the x space in (36), 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

$$\bar{\varphi}(t,\bar{x}) = g(t;\bar{x},\varphi(t-1,\bar{x}),e(t)), \quad \bar{\varphi}(0,\bar{x}) = 0,$$
 (38)

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

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

if $\bar{\varphi}(n, \bar{x}) = \varphi(n)$. This means that small variations in x in (37) 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 (38) with $\bar{\varphi}_1(s, \bar{x}) :=$ φ_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 (37).

- C6: $\lim_{t\to\infty} \mathbb{E}Q(t, \bar{x}, \bar{\varphi}(t, \bar{x}))$ 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: $\lim_{t\to\infty} \sup[1/\gamma(t) 1/\gamma(t-1)] < \infty$.

For practical algorithm implementation, the projection or saturation is often introduced (Kushner & Yin, 1997; Ljung, 1977) to meet the boundedness condition required in the ODE approach (Ljung, 1977). Since dynamics of agents are given by a single integrator, 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 then apply the usual saturation given by $[\cdot]_D$

$$\mathbf{x}(t) = [\Omega(t)]_D = \begin{cases} \Omega(t), & \Omega(t) \in D\\ \mathbf{x}(t-1), & \Omega(t) \notin D, \end{cases}$$
(39)

where *D* is a compact subject of D_R in which the regularity conditions hold. x(t) and $\Omega(t)$ denote the left- and right-hand sides of (36) 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 (35) will be converted to the canonical form in (36). Throughout the paper, we assume that the projection is applied to the resulted algorithm in the form of (36). 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 (36). For more details, see Ljung and Söderström (1983) and Wigren (1994) and the references therein.

We will then utilize the following corollary reported in Ljung and Söderström (1983).

Corollary 4 (Ljung & Söderström, 1983). Consider the algorithms (36), (37) and (39) subject to the regularity conditions C1-C11. Let D_R be an open connected subset of D_S . Let D in (39) be the compact subset of D_R such that the trajectories of the associated ODE

$$\frac{\mathrm{d}}{\mathrm{d}\tau}x(\tau) = f(x(\tau)) \tag{40}$$

where

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

that start in D remain in a closed subset \bar{D}_R of D_R for $\tau > 0$. Assume that the differential equation (40) has an invariant set D_c with a domain of attraction $D_A \supset D$.

Then either

$$x(t) \to D_c$$
, with probability one as $t \to \infty$, (41)

 $x(t) \rightarrow \partial D$, with probability one as $t \rightarrow \infty$, (42)

where ∂D is the boundary of D.

The conclusion (42) is possible only if there is a trajectory of the differential equation in (40) that leaves D in (39).

Now we present our main results. The following lemma shows how to transform our coordination and estimation algorithms to the canonical forms in (36) and (37).

Lemma 5. The algorithms (35) and (28) can be transformed into the forms of (36) and (37) respectively, using the following definitions;

$$\begin{aligned} q(t) &:= \operatorname{col}(q_{1}(t), \dots, q_{N_{s}}(t)) \in \mathbb{R}^{2N_{s}}, \\ p(t) &:= \operatorname{col}(p_{1}(t), \dots, p_{N_{s}}(t)) \in \mathbb{R}^{2N_{s}}, \\ x(t) &:= [q^{\mathrm{T}}(t), p^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^{4N_{s}}, \\ Q(t; x(t-1), \varphi(t)) \\ &:= \begin{bmatrix} p \\ -\nabla U(q) - (\hat{L}(q) + K_{d})p - \nabla \hat{C}(\varphi, q) \end{bmatrix}, \end{aligned}$$
(43)

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

$$\nabla C(\varphi(t), q(t-1)) = -k_4 \text{col}(\nabla \hat{\mu}_1(t-1, q_1(t-1)), \dots, \nabla \hat{\mu}_{N_s}(t-1, q_{N_s}(t-1)))$$

= $-k_4 \text{col}(\phi^T(q_1(t-1))\hat{\Theta}_1(t), \dots, \phi^T(q_{N_s}(t-1))\hat{\Theta}_{N_s}(t)).$

For the observation process in (37), we have:

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

= $A(t; x(t-1))\varphi(t-1) + B(t; x(t-1))e(t),$ (44)

where

$$\begin{split} \varphi(t) &:= \operatorname{col}(\tilde{\Theta}_1(t), \dots, \tilde{\Theta}_{N_s}(t)) \in \mathbb{R}^{mN_s}, \\ A(t; x(t-1)) &:= \operatorname{diag}\left(I_m - K_1(t)\Phi_{*1}, \dots, I_m - K_{N_s}(t)\Phi_{*N_s}\right) \in \mathbb{R}^{mN_s \times mN_s}, \\ B(t; x(t-1)) &:= \operatorname{diag}\left(K_1(t), \dots, K_{N_s}(t)\right) \in \mathbb{R}^{mN_s \times 0}, \\ e(t) &:= \operatorname{col}(w_{*1}(t-1), \dots, w_{*N_s}(t-1)) \in \mathbb{R}^0, \end{split}$$

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

Proof. From (28), notice that:

$$\Theta_i(t) = [I_m - K_i(t)\Phi_{*i}]\Theta_i(t-1) + K_i(t)w_{*i}(t-1).$$
(45)

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

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

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

$$\sum_{k=1}^{3} \phi(\nu_k) \phi^{\mathrm{T}}(\nu_k) \succ 0,$$

where m is in (10).

- M2: The artificial potential force and the adjacency matrix are continuously differentiable with respect to q and derivatives are bounded.
- M3: The projection algorithm (39) is applied to the coordination algorithm (36). Let D in (39) be a convex and compact set defined by $D := \mathcal{M}^{Ns} \times \mathcal{M}_p$, where $\mathcal{M}_p := [p_{\min}, p_{\max}]^{2N_s}$.

Remark 6. M1 can be viewed as a persistent excitation condition in adaptive control (Åström & Wittenmark, 1995). M2 can be satisfied, for instance, see (8) and Olfati-Saber (2006). 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 $A_i(t) := [I_m - K_i(t)\Phi_{*i}]$ (46)

is a positive definite matrix for all $i \in I$ and $t \in \mathbb{Z}_{>0}$. All eigenvalues of $A_i(t)$ in (46) are non-negative and strictly less than 1, i.e.,

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

Hence, the induced matrix 2-norm of $A_i(t)$ is strictly less than 1:

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

$$\tag{47}$$

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

$$A_{i}(t) = I_{m} - P_{i}(t)\Phi_{*i}^{\mathrm{T}}(I_{s} + \Phi_{*i}P_{i}(t)\Phi_{*i}^{\mathrm{T}})^{-1}\Phi_{*i}.$$

where $P_i(t) > 0$ is a positive definite matrix. From (28), notice that

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

which implies that $P_i(t - 1)(I_m - A_i(t)) > 0$. Hence, we conclude that $0 \prec A_i(t) = A_i^{T}(t) \prec I_m$. Moreover, since $A_i(t) > 0$, there exists a square root matrix F so that $A_i(t) =$ $F^{\mathrm{T}}F$ and $F = \operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_m})R$ where R is the orthonormal matrix and $\lambda_1 = \lambda_{\max}(A_i(t)) > \lambda_2 > \cdots > \lambda_m = \lambda_{\min}(A_i(t)) > 0$. 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 $||A_i(t)|| =$ $\sqrt{\lambda_{\max}(A_i^{\mathrm{T}}(t)A_i(t))} < 1.$

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

$$Z := \left\{ q \in \mathcal{M}^{N_{s}} \middle| \sum_{j \in \{i\} \cup \mathcal{N}(i,q)} \phi(q_{j}) \phi^{\mathsf{T}}(q_{j}) \neq 0, \ \forall i \in \mathcal{I} \right\}.$$
(48)

Moreover, f(x) in (40) of Corollary 4 is given by

$$f(x) = \begin{bmatrix} p \\ -\nabla U(q) - (\hat{L}(q) + K_d)p - \nabla C(q) \end{bmatrix},$$
(49)

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

$$C(q) \coloneqq k_4 \sum_{i \in \mathcal{I}} [\mu_{\max} - \mu(q_i)], \tag{50}$$

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 (16) and (17) 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 (44) are functions of smooth radial basis functions, therefore, they are smooth in D_{R} .
- C4: We use the similar argument used in Brus (2006). Notice that:

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

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:

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

Therefore we obtain:

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

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

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

Under M1 and M3, $||A(k; \bar{x})|| < \lambda(\bar{x})$ for all $k \in \{s + 1, ..., t\}$, 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 (48).

• C6: Elements of *Q* in (43) are deterministic functions of $x \in D_R$ except for $\nabla \hat{C}(\varphi(t), q)$. Thanks to M1 and (27), for a fixed *q*, we have

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

which proves C6 and (49) simultaneously.

- C7: This is satisfied due to the measurement noise assumption in (16).
- C8, C9, C10, C11: These are satisfied by the time-varying gain sequence defined in (32).

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

$$V(q(\tau), p(\tau)) \coloneqq C(q(\tau)) + U(q(\tau)) + \frac{p^{\mathrm{T}}(\tau)p(\tau)}{2},$$
(51)

where C(q) is a cost function of q to make agents to trace peaks of the field. U(q) is obtained from (9) and is a cost function of q to enforce swarming behavior, obstacle avoidance, and artificial walls etc. to the multi-agent system. The last term on the right-hand side of (51) is the kinetic energy of the multi-agent system and is a cost function of p alone.

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

Theorem 9. For any initial state $x_0 = \operatorname{col}(q_0, p_0) \in D$, where *D* is a compact set as in (39), 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 (51). Let D_c be the set of all points in D_A , where $\frac{d}{d\tau}V(x) = 0$. Then every solution starting from D_A approaches the largest invariant set D_M contained in D_c with probability one as $t \to \infty$, 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 $x^* = \operatorname{col}(q^*, 0)$ 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.,

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

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

$$\frac{\mathrm{d}x(\tau)}{\mathrm{d}\tau} = f(x(\tau)). \tag{52}$$

Taking the derivative of $V(x(\tau))$ in (51) with respect to τ and using (52), we obtain

$$\frac{dV(x(\tau))}{d\tau} = \left(\frac{\partial V(x)}{\partial x}\right)^{\mathrm{T}} f(x(\tau))$$

$$= \begin{bmatrix} \nabla U(q(\tau)) + \nabla C(q(\tau)) \\ p(\tau) \end{bmatrix}^{\mathrm{T}}$$

$$\times \begin{bmatrix} p(\tau) \\ -\nabla U(q(\tau)) - \nabla C(q(\tau)) - (\hat{L}(q(\tau)) + K_d)p(\tau) \end{bmatrix}$$

$$= -p^{\mathrm{T}}(\tau)(\hat{L}(q(\tau)) + K_d)p(\tau) \leq 0.$$
(53)

 $D_A := \{x \in D \mid V(x) \le a\}$ is a bounded set with $\frac{d}{d\tau}V(x) \le 0$ for all $x \in D_A$ as in (53), 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

$$\left\{ x(\tau) \mid \dot{V}(x(\tau)) = -p^{\mathrm{T}}(\tau)(\hat{L}(q(\tau)) + K_d)p(\tau) = 0 \right\},$$
(54)

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

If $(\hat{L}(q) + K_d) > 0 \ \forall x \in D$, from (54), any point x^* in D_M is the form of $x^*(t) = \operatorname{col}(q^*(t), 0)$. Moreover, from (49), 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. \Box

Table 1

Parameters in the simulation.

Parameters	Values
Number of agents N_s	25
Number of basis functions m	25
Surveillance region 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. $1I_{2N_s}$; 0 _{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)	$3I_m$



Fig. 5. 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. Red color denotes the highest scalar value while blue color represents the lowest value. 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.

5. Simulation results

We applied the proposed multi-agent system to static fields $\mu(s)$, which are represented by twenty five radial basis functions, as depicted in the left side of Fig. 6 (uni-modal) and Fig. 8 (bi-modal). Each agent updates the estimated field for the coordination once per iteration. 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.

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 (39), velocity feedback ($K_d > 0$ as defined in (43)), and an artificial potential wall. Fig. 5(a) shows that the recursively estimated field by agent 1 at the iteration time t = 40 under the noise level of 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. 5(a). Fig. 5(b) illustrates the estimated field by agent 1 at iteration time t = 200. As shown in Fig. 5(b), twenty five swarming agents have located the maximum point of the field successfully. The right side of Fig. 6 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.



Fig. 6. A uni-modal field is considered as a true field (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. 7. 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.



Fig. 8. A bi-modal field is considered as a true field (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).

With a bit higher damping coefficients contained in $K_d = I_{2N_s}$, the rate of convergence to the maximum point was slow. 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 $K_d = 0_{2N_s}$ is applied to a bi-modal static field, which is shown in the left side of Fig. 8. Fig. 7 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 (51). 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, which results in higher RMS estimation error values plotted in the right-hand side of Fig. 8 as compared to those for the previous case (Fig. 6).

Fig. 9 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. 5. The lowest RMS value of the



Fig. 9. 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.



Fig. 10. Trajectories of agent 1 (green dot) with a faulty sensor and all other 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.



Fig. 11. The root mean square (RMS) values of the spatially averaged error field achieved by agent 1 with a faulty sensor (red-solid line) and all other agents with respect to the iteration number.

estimation error achieved by agents was about 6. This simulation clearly justifies the usage of the communication network and swarming algorithms in the proposed multi-agent system.

We now consider a case in which agent 1 has a faulty sensor, e.g., the measurement value of agent 1 is a constant value. Fig. 10 depicts a typical collective behavior of agent 1 with a faulty sensor and all other agents for the same uni-modal true field shown in the left side of Fig. 7. In this simulation, we consider that the faulty sensor of agent 1 produces a higher value, e.g., $y(q_1(t)) = 10$ for all $t \in \mathbb{Z}_{\geq 0}$, which gave us most detrimental effect to the behavior of the multi-agent system. This consistently wrong measurement for agent 1 made its estimated field far away from the true field (Fig. 11) and misguided agent 1 to a wrongfully estimated maximum point by agent 1 (Fig. 10). In all simulations, all other agents using our algorithm except for agent 1 seem to be able to locate the



Fig. 12. Trajectories of agents for d = 0.6, W = 1, and $K_d = 0_{2N_s}$, and the projection was used.



Fig. 13. The group disagreement function $\Psi_G(p(t))$ with respect to the iteration number. Parameters are d = 0.6, W = 1, and $K_d = 0_{2N_s}$, and the projection was used.

peak of the true field as typically shown in Fig. 10. In general, the robustness of the multi-agent system with respect to faulty sensors will decreases as the number of faulty sensors increases.

5.2. Without the velocity feedback

We consider a case without the velocity feedback (i.e., $K_d = 0_{2N_s}$) 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. 12. The group disagreement function $\Psi_G(p(t)) = \frac{1}{2}p^{\mathrm{T}}(t)\hat{L}(q(t))p(t)$ with respect to the iteration number is shown in Fig. 13.

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 discover two maximum points of the bimodal field as depicted in Figs. 14 and 15. The group disagreement function and convergence rate of the agents are illustrated in Fig. 15. 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 interest. In addition, we relocate the maximum of the field at the boundary of the surveillance region. As can be seen in Fig. 16, 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} .

2812



Fig. 14. Trajectories of agents for d = 0.6, W = 1, and $K_d = 0_{2N_s}$, and the projection was not used.



Fig. 15. The group disagreement function $\Psi_G(p(t))$ 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 $K_d = 0_{2N_s}$, and the projection was not used.



Fig. 16. 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} .

6. Conclusions

This paper presented a novel class of self-organizing sensing agents that form a swarm and learn through noisy measurements collectively from neighboring agents to estimate an 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 results for 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. Motivated by techniques developed in this paper, we are currently developing a systematic way that allows us to efficiently design and analyze a class of practical algorithms for distributed learning and control of multi-agent systems. A possible future work is to investigate how mixture of heterogeneous learning agents can be optimally coordinated by consensus type algorithms for learning unknown fields.

Acknowledgements

The authors would like to thank Professor Lennart Ljung, with the department of Automatic Control, Linköping Institute of Technology, Linköping, Sweden for kindly providing the first author with the technical report (Ljung, 1975).

Jongeun Choi has been supported, in part, by Intramural Research Grants Program (IRGP) from Michigan State University and the National Science Foundation through CAREER Award CMMI-0846547.

Songhwai Oh was in part supported by ARL Micro Autonomous Systems and Technology Collaborative Technology Alliance (W91INF-08-2-0004).

These supports are gratefully acknowledged.

References

- Adler, J. (1966). Chemotaxis in bacteria. Journal of Supramolecular Structure, 4(3), 305–317.
- Åström, K. J., & Wittenmark, B. (1995). Adaptive control (2nd ed.). Addison Wesley.
- Bonabeau, E., Dorigo, M., & Theraulaz, G. (1999). Swarm intelligence: From natural to artificial systems. Santa Fe Institute Studies on the Sciences of Complexity, Oxford University Press.
- Brus, L. (2006). Recursive black-box identification of nonlinear state-space ODE models, it licentiate theses; 2006-001, Uppsala Universitet.
- Center of Excellence for Great Lakes. Harmful algal bloom event response. Also available at: http://www.glerl.noaa.gov/res/Centers/HABS/habs.html.
- Choi, J., Lee, J., & Oh, S. (2008a). Swarm intelligence for achieving the global maximum using spatio-temporal Gaussian processes. In Proceedings of the 2008 American control conference ACC.
- Choi, J., Lee, J., & Oh, S. (2008b). Biologically-inspired navigation strategies for swarm intelligence using spatial Gaussian processes. In Proceedings of the international federation of automatic control IFAC world congress.
- Choi, J., Oh, S., & Horowitz, R. (2007). Cooperatively learning mobile agents for gradient climbing. In Proceedings of the 46th IEEE conference on decision and control.
- Cortés, J. (2010). Distributed kriged Kalman filter for spatial estimation, IEEE Transactions on Automatic Control, 55(4) (in press).
- Cortes, J., Martinez, S., Karatas, T., & Bullo, F. (2004). Coverage control for mobile sensing networks. IEEE Transactions on Robotics and Automation, 20(2), 243–255.
- Cressie, N. A. C. (1991). Statistics for spatial data. A Wiley-Interscience Publication, John Wiley and Sons, Inc..
- Cressie, N., & Wikle, C. K. (2002). ch. Space-time Kalman filter. In Encyclopedia of environments: Vol. 4 (pp. 2045–2049). Chichester: John Wiley and Sons, Ltd..
- Culler, D., Estrin, D., & Srivastava, M. (2004). Overview of sensor networks. IEEE Computer, Special Issue in Sensor Networks,.
- Dhariwal, A., Sukhatme, G. S., & Requicha, A. A. G. (2004). Bacterium-inspired robots for environmental monitoring. In Proceedings of the IEEE international conference on robotics and automation.
- DOD/ONR MURI: Adaptive sampling and prediction project. Also available at: http://www.princeton.edu/~dcsl/asap/.
- Eberhart, R. C., Shi, Y., & Kennedy, J. (2001). The Morgan Kaufmann series in artificial intelligence, Swarm intelligence. Academic Press.
- Estrin, D., Culler, D., Pister, K., & Sukhatme, G. (2002). Connecting the physical world with pervasive networks. *IEEE Pervasive Computing*, 1, 59–69.
- Godsil, C., & Royle, G. (2001). Graduate text in mathematics: Vol. 207. Algebraic graph theory. Springer-Verlag.
- Graham, R., & Cortés, J. (2009). Asymptotic optimality of multicenter voronoi configurations for random field estimation. *IEEE Transactions on Automatic Control*, 54(1), 153–158.
- Grünbaum, D. (1998). Schooling as a strategy for taxis in a noisy environment. *Evolutionary Ecology*, *12*, 503–522.
- Harmful Algal BloomS Observing System (HABSOS) by National Oceanic and Atmospheric Administration NOAA, available at:http://habsos.noaa.gov/.
- Jadbabie, A., Lin, J., & Morse, A. S. (2003). Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48, 988–1001.

- Kushner, H. J., & Yin, G. G. (1997). Stochastic approximation algorithms and applications. Springer.
- Ljung, L. (1977). Analysis of recursive stochastic algorithms. *IEEE Transactions on Automatic Control*, 22(4), 551–575.
- Ljung, L. (1975). Theorems for the asymptotic analysis of recursive, stochastic algorithms, technical report 7522, Department of Automatic Control, Lund Institute of Technology.
- Ljung, L., & Söderström, T. (1983). *Theory and practice of recursive identification*. Cambridge, Massachusetts, London, England: The MIT Press.
- MacKay, D. J. C. (1998). Introduction to Gaussian processes. In C. M. Bishop (Ed.)., NATO ASI series, Neural networks and machine learning (pp. 133–166). Kluwer.
- Martinez, S. (2009). Distributed interpolation schemes for field estimation by mobile sensor networks, *IEEE Transactions on Control Systems Technology* (in press).
- Ögren, P., Fiorelli, E., & Leonard, N. E. (2004). Cooperative control of mobile sensor networks: Adaptive gradient climbing in a distributed environment. *IEEE Transaction on Automatic Control*, 49, 1292.
- Oh, S., Schenato, L., Chen, P., & Sastry, S. (2007). Tracking and coordination of multiple agents using sensor networks: System design, algorithms and experiments. *Proceedings of the IEEE*, 95, 234–254.
- Olfati-Saber, (2006). Flocking for multi-agent dynamic systems: Algorithm and theory. *IEEE Transactions on Automatic Control*, 51, 401–420.
- Olfati-Saber, R., & Shamma, J. (2005). Consensus filters for sensor networks and distributed sensor fusion. In Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on (pp. 6698–6703).
- Rasmussen, C. E., & Williams, C. K. I. (2006). Gaussian processes for machine learning. Cambridge, Massachusetts, London, England: The MIT Press.
- Ren, W., & Beard, R. W. (2005). Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transactions on Automatic Control*, 50, 655–661.
- Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. Computer Graphics, 21(4), 25–34.
- Singh, A., Batalin, M., Stealey, M., Chen, V., Lam, Y., Hansen, M., Harmon, T., Sukhatme, G. S., & Kaiser, W. (2007). Mobile robot sensing for environmental applications, In Proceedings of the International Conference on Field and Service Robotics.
- Tanner, H. G., Jadbabaie, A., & Pappas, G. J. (2003). Stability of flocking motion, technical report, University of Pennsylvania.
- Wigren, T. (1994). Convergence analysis of recursive identification algorithms based on the nonlinear Wiener model. *IEEE Transactions on Automatic Control*, 39, 2191–2206.



Jongeun Choi received his Ph.D. and M.S. degrees in Mechanical Engineering from the University of California at Berkeley in 2006 and 2002 respectively. He also received a B.S. degree in Mechanical Design and Production Engineering from Yonsei University at Seoul, Republic of Korea in 1998. He is currently an Assistant Professor with the Departments of Mechanical Engineering and Electrical and Computer Engineering at the Michigan State University. His research interests include adaptive, distributed and robust control and statistical learning algorithms, with applications to self-organizing systems, mobile robotic sensors,

environmental adaptive sampling, and biomedical problems. He was a recipient of an NSF CAREER Award in 2009. Dr. Choi is a member of ASME and IEEE.



Songhwai Oh is an assistant professor in the School of Electrical Engineering and Computer Science at the Seoul National University. He received all his degrees in Electrical Engineering and Computer Sciences (EECS) at the University of California, Berkeley (B.S. with highest honors in 1995, M.S. in 2003, and Ph.D. in 2006). Before his Ph.D. studies, he worked as a senior software engineer at Synopsys, Inc. and a microprocessor design engineer at Intel Corporation. In 2007, he was a postdoctoral researcher in EECS at UC Berkeley. From 2007 to 2009, he was an assistant professor of Electrical Engineering and

Computer Science in the School of Engineering at the University of California, Merced. His research interests include cyberphysical systems, wireless sensor networks, robotics, estimation and control of stochastic systems, multi-modal sensor fusion, and machine learning.



Roberto Horowitz received a B.S. degree with highest honors in 1978 and a Ph.D. degree in 1983 in mechanical engineering from the University of California at Berkeley. In 1982 he joined the Department of Mechanical Engineering at the University of California at Berkeley, where he is currently a Professor and the James Fife Endowed Chair. Dr. Horowitz teaches and conducts research in the areas of adaptive, learning, nonlinear and optimal control, with applications to MicroElectromechanical Systems (MEMS), computer disk file systems, robotics,

mechatronics and Intelligent Vehicle and Highway Systems (IVHS). Dr. Horowitz is a member of IEEE and ASME.