Cooperatively Learning Mobile Agents for Gradient Climbing

Jongeun Choi, Songhwai Oh and Roberto Horowitz

***Abstract*— This paper presents a novel class of self-organizing autonomous sensing agents that form a swarm and learn the static field of interest through noisy measurements from neighbors for gradient climbing. In particular, each sensing agent maintains its own smooth map which estimates the field. It updates its map using measurements from itself and its neighbors and simultaneously moves toward a maximum of the field using the gradient of its map. The proposed cooperatively learning control consists of motion coordination based on the recursive spatial estimation of an unknown field of interest with measurement noise. The convergence properties of the proposed coordination algorithm are analyzed using the ODE approach and verified by a simulation study.**

I. 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], [2], [3], [4], [5]. Mobile sensing agents form an ad-hoc wireless communication network in which each agent operates usually under a short communication range, limited memory and computational power. To perform various tasks such as exploration, surveillance, and environmental monitoring, mobile sensing agents require distributed coordination to adapt to environments to achieve a global goal. Among challenging problems of distributed coordination of mobile sensing agents, gradient climbing over an unknown field of interest has attracted much attention of control engineers. This has numerous applications including homeland security, toxic-chemical plume tracing and environmental monitoring. For instance, the most common approach to toxic-chemical plume tracing has been biologically inspired *chemotaxis* [6], [7], in which a mobile sensing agent is driven according to a local gradient of the chemical plume concentration. However, with this approach, the convergence rate can be slow and the mobile robot may get stuck in the local maxima of chemical plume concentration. The cooperative network of agents that performs adaptive gradient climbing in a distributed environment was presented in [8], [9]. 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 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 orient with respect to local gradients in environmental properties. Grünbaum [10] showed that schooling behavior can improve the ability of performing taxis to climb gradients, since the swarming alignment tendency averages out the stochastic effects of individual sampling errors. The collective swarm behaviors of birds/fish/ants/bees are known to be the outcomes of natural optimization.

Tanner [4] and Olfati-Saber [5] presented comprehensive analyses of the flocking algorithm by Reynolds [11]. 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 [11]. 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.

The contribution of this paper is to combine the recent swarming system theoretical results [4] [5] with a cooperatively learning mechanism to form an "artificial intelligence" of each individual, which results in "Cooperatively Learning Mobile Agents" (CoLMAs). In particular, the sensing agent will receive collective measurements from its neighboring agents within a limited transmission range. Upon receiving cooperative measurements, each mobile sensing agent will recursively update the image of an unknown static field. In this paper, the recursive estimation is based on radial basis function learning. To locate the maximum (or source) of the field, the sensing agent will climb the gradient of its own updated image of the unknown field. Equipped with a swarming behavior, access to cooperatively sensed measurements and the individual *learning* capability, CoLMAs are expected to be resilient not only to measurement noise but also to local maxima of a field. The proposed CoLMAs exactly mimic 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 drug sniffing dogs, have provided strong incentives to invent cooperatively learning mobile sensing agents.

In this way, the distributed and scalable control law can be derived without the knowledge of the density of interest in the environment, which is the main difference from other coordination algorithms. The proposed cooperatively learning control consists of motion coordination based on the recursive estimation of an unknown field of interest with

Jongeun Choi is with the Department of Mechanical Engineering, Michigan State University jchoi@egr.msu.edu

Songhwai Oh is with the School of Engineering, University of California, Merced songhwai.oh@ucmerced.edu

Roberto Horowitz is with the Department of Mechanical Engineering, University of California at Berkeley horowitz@me.berkeley.edu

measurement noise. Our strategy of the cooperative learning control can be applied to a large class of coordination algorithms for mobile agents to deal with the field of interest that requires to be recursively estimated.

This paper is organized as follows. In Section II, 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 III. In Section IV, cooperatively learning control is described with a scheduled sampling rate. Section V analyzes the convergence properties of the proposed coordination algorithm based on the ODE approach. As an illustrative example, we apply CoLMAs to a static field with measurement noise.

II. Preliminaries

In this section, we explain notations and concepts that will arise throughout the paper.

A. Mobile Sensing Agent Network

First, we explain the mobile sensing network and sensor models used in this paper. Let N_s be the number of sensing agents distributed over the surveillance region $\mathcal{R} \in \mathbb{R}^2$, which is a compact set. The identity of each agent is indexed by $\mathcal{I} := \{1, 2, \cdots, N_s\}$. Let $q_i(t) \in \mathcal{R}$ be the location of the i-th sensing agent at time $t \in \mathbb{R}_+$ and let $q := \text{col}(q_1, q_2, \cdots, q_{N_s}) \in \mathbb{R}^{2N_s}$ be the configuration of the swarm system. The discrete time, high-level dynamics of agent i is modeled by

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

where $q_i, p_i, u_i \in \mathbb{R}^2$ are, respectively, the position, the velocity, and the input of the mobile agent and δ is the iteration step size (or sampling time) which will be described in detail in Section IV. We assume that the measurement $y(q_i(k))$ of the i-th sensor includes the scalar value of the field $c(k)$ and sensor noise $w(k)$, at its position $q_i(k)$ and some measurement sampling time index k,

$$y(q_i(k)) := c(q_i(k), k) + w(k), \tag{2}$$

where $c : \mathcal{R} \times \mathbb{R}_+ \to [0, c_{max}]$ is the field of interest.

B. A Graph

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 a 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. 1. 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. 1 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 $\mathcal{N}(i, q) := \{j : (i, j) \in \mathcal{E}(q), i \in \mathcal{I}\}$. The adjacency matrix $A := [a_{ij}]$ of an undirected graph G is a symmetric matrix such that $a_{ij} = k_3 > 0$ if vertex i and vertex j are neighbors and $a_{ij} = 0$ otherwise. The scalar graph Laplacian $L = [l_{ij}] \in \mathbb{R}^{N_s \times N_s}$ is a matrix defined as $L := \Delta(A) - A$, where $\Delta(A)$ is a diagonal matrix whose diagonal entries are row sums of A, i.e., $\Delta(A) := \text{diag}(\sum_{j=1}^{N_s} a_{ij})$. The 2-dimensional graph Laplacian is defined as $\hat{L} := L \otimes I_2$, where $\otimes$ is the Kronecker product. A quadratic disagreement function [5] can be obtained via the Laplacian $\hat{L}$:

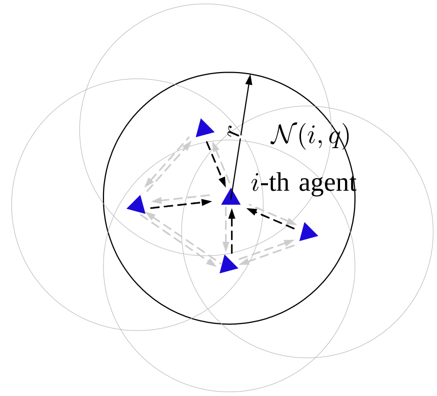

Fig. 1. The model of the mobile sensing agent network. The i-th agent gathers measurements from all four neighboring sensing agents in a r interactive range.

$$p^T \hat{L} p = \frac{1}{2} \sum_{(i,j) \in \mathcal{E}(q)} a_{ij} ||p_j - p_i||^2, \tag{3}$$

where $p := \text{col}(p_1, p_2, \cdots, p_{N_s}) \in \mathbb{R}^{2N_s}$.

C. A Swarming Behavior

We use attractive and repulsive smooth potentials similar to those used in [4], [5] 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 collective potential

$$U_1(q) := \sum_i \sum_{j \neq i} U_{ij}(r_{ij}), \tag{4}$$

where $r_{ij} := \|q_i - q_j\|^2$ and V_{ij} in (4) is defined by

$$U_{ij} := \begin{cases} \frac{1}{2}\left(\log(\alpha + r_{ij}) + \frac{\alpha + d^2}{\alpha + r_{ij}}\right) & \text{if } r_{ij} < d_0^2 \\ \frac{1}{2}\left(\log(\alpha + d_0^2) + \frac{\alpha + d^2}{\alpha + d_0^2}\right) & \text{otherwise,} \end{cases} \tag{5}$$

here $d < d_0$. The gradient of the potential for agent i is

$$\nabla U_1(q_i) = \begin{cases} \sum_{j \neq i} \frac{(r_{ij} - d^2)(q_i - q_j)}{(\alpha + r_{ij})^2} & \text{if } r_{ij} < d_0^2 \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

In equations (4), (5), (6), α was introduced to prevent the reaction force from diverging at $r_{ij} = \|q_i - q_j\|^2 = 0$. As illustrated in Fig. 2-(a), this potential yields a reaction force that is attracting when the agents are too far and repelling when a pair of two agents are too close. It has an equilibrium point at a distance of d. We also introduce a potential U_2 to

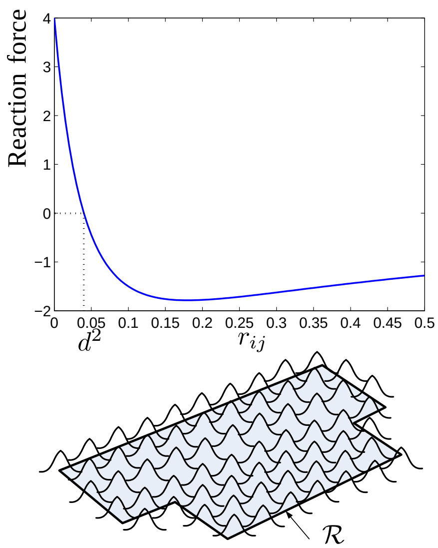

Fig. 2. (a)-up: Reaction force generated by potential with respect to r_{ij}. (b)-down: Uniformly and densely distributed Gaussian bases over a surveillance region $\mathcal{R}$.

model the environment. U_2 enforces each agent to stay inside the closed and connected surveillance region $\mathcal{R}$ and prevents collisions with obstacles in $\mathcal{R}$. We construct U_2 such that it is radially unbounded in q, i.e.,

$$U_2(q) \to \infty \text{ as } \|q\| \to \infty. \tag{7}$$

Define the total artificial potential by

$$U(q) := k_1 U_1(q) + k_2 U_2(q), \tag{8}$$

where $k_1, k_2 > 0$ are weighting factors.

III. Learning Mobile Sensing Agents

In this section, we introduce a learning algorithm for each mobile sensing agent to estimate the spatial function c. Suppose that the scalar field $c(\nu)$ is generated by a network of radial basis functions:

$$c(\nu) := \sum_{j=1}^{m} \phi_j(\nu)\theta_j = \phi^T(\nu)\Theta, \tag{9}$$

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

$$\phi^T(\nu) := \begin{pmatrix} \phi_1(\nu) & \phi_2(\nu) & \cdots & \phi_m(\nu) \end{pmatrix}$$
$$\Theta := \begin{pmatrix} \theta_1 & \theta_2 & \cdots & \theta_m \end{pmatrix}^T.$$

$\phi_j(\nu)$ are Gaussian basis functions given by

$$\phi_j(\nu) := \frac{1}{Z} \exp \frac{-(\nu - \kappa_j)^T(\nu - \kappa_j)}{\sigma^2}, \tag{10}$$

where σ is the width of the Gaussian basis and Z is a normalizing constant. κ_j for $j \in \{1, \cdots, m\}$ are uniformly distributed in the surveillance region $\mathcal{R}$ as shown in Fig. 2-(b). $\Theta \in \mathbb{R}^m$ is the parameter of the regression model in (9). 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 $\{(y(\nu_k), \phi(\nu_k))\}_{k=1}^n$, the parameter Θ can be estimated to minimize the least-squares error

$$\sum_{k=1}^{n} (y(\nu_k) - \phi^T(\nu_k)\Theta)^2. \tag{11}$$

For now, let us consider (2) without the sensor noise $w(k)$. For a set $\{(y(\nu_k), \phi(\nu_k))\}_{k=1}^n$, the optimal least-squares estimation solution is well-known [12] to be

$$\hat{\Theta}(n) = P(n,1)\Phi^T(n,1)Y(n,1), \tag{12}$$

where abusing notations slightly by $y(k) := y(\nu_k)$ and $\phi(k) := \phi(\nu_k)$ for simplicity, we define

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

During a time interval between the coordination iteration indices t and $t+\delta$ as in (1), we suppose that a sensing agent has collected s samples from itself and $s-1$ neighbors within the transmission range. Suppose at previous iteration, the agent has already updated the field $\hat{c}$ based on the previous data set $\{(y(k), \phi(k))\}_{k=1}^{n-s}$, where $n-s$ is the total number of past measurement points. Now the sensing agent needs to update the field $\hat{c}$ upon receiving cooperatively measured s number of points $\{(y(k), \phi(k))\}_{k=n-s+1}^{n}$, where $1 \leq s \leq N_s$. Then we have the following lemma.

Lemma 1: 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

$$\begin{aligned}
K(n) &= P(n-s)\Phi_*^T \left(I + \Phi_* P(n-s)\Phi_*^T\right)^{-1}, \\
P(n) &= (I - K(n)\Phi_*)P(n-s), \\
\hat{\Theta}(n) &= \hat{\Theta}(n-s) + K(n)\left[Y_* - \Phi_*\hat{\Theta}(n-s)\right], \\
\hat{c}(n,\nu) &:= \phi^T(\nu)\hat{\Theta}(n),
\end{aligned} \tag{13}$$

where some abbreviations are defined: $Y_* := Y(n, n-s+1)$, $\Phi_* = \Phi(n, n-s+1)$, $\Phi^T(n) := \Phi^T(n,1)$, $Y(n) := Y(n,1)$ and $P(n) := P(n,1)$. Then the recursive estimation presented in (13) is the least-squares estimation that minimizes the error function in (11).

Proof: It is straightforward and so omitted for brevity.

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 symmetrically such that $\phi_i(\nu) = \phi_j(\nu)$. In practice, we start the recursive LSE algorithm in (13) with initial $\hat{\Theta}(0)$ and $P(0) \succ 0$ which corresponds to the situation in which the parameters have an initial distribution and keep running the recursive algorithm with new measurements. Along this line, we define $P_+(\cdot)$ by

$$P_+^{-1}(n) := P^{-1}(0) + \Phi^T(n)\Phi(n) \succ 0. \quad (14)$$

Now we consider the measurement model (2) with the sensor noise $w(k)$. $w(k)$ is assumed to be a white noise sequence with variance W given by

$$\mathbb{E}(w(k)) = 0, \quad \mathbb{E}(w(k)w(z)) = \begin{cases} W & \text{if } k = z \\ 0 & \text{if } k \neq z \end{cases}, \quad (15)$$

where $\mathbb{E}$ denotes expectation. Moreover, we assume that

$$|w(k)| < L \text{ with probability one (w.p.1) } \forall k. \quad (16)$$

Let the estimation error vector of the parameter Θ be $\tilde{\Theta}(n) := \hat{\Theta}(n) - \Theta$. We define the gradient of a field of interest by $\nabla c(\nu) := \frac{\partial c(x)}{\partial x}\Big|_{x=\nu}$. From (9), we have

$$\nabla c(\nu) = \frac{\partial \phi^T(x)}{\partial x}\Big|_{x=\nu} \Theta =: \phi'^T(\nu)\Theta \in \mathbb{R}^{2\times 1}, \quad (17)$$

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

$$\nabla \hat{c}(n, S, \nu) := \phi'^T(\nu)\hat{\Theta}(n, S) \in \mathbb{R}^{2\times 1}. \quad (18)$$

The estimation error of the gradient can be obtained by

$$\begin{aligned} \varepsilon(n, S, \nu) &:= \phi'^T(\nu)\hat{\Theta}(n, S) - \nabla c(\nu) = \phi'^T(\nu)\tilde{\Theta}(n, S) \\ &= \mathbb{E}(\varepsilon(n, S, \nu)) + \xi(n, S, \nu), \end{aligned} \quad (19)$$

where

$$\mathbb{E}(\varepsilon(n, S, \nu)) = \phi'^T(\nu)\left[P_+(n)\sum_{k=1}^n \phi(\nu_k)\phi^T(\nu_k) - I\right]\Theta,$$

$$\xi(n, S, \nu) := \phi'^T(\nu)\left[P_+(n)\sum_{k=1}^n \phi(\nu_k)w(k)\right].$$

For $\Phi^T(n)\Phi(n) \succ 0$, the gradient estimator is unbiased asymptotically

$$\lim_{n\to\infty} \mathbb{E}(\varepsilon(n, S, \nu)) = 0. \quad (20)$$

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

$$\phi'^T(\nu)\frac{W}{n}R^{-1}(n, S)\phi'(\nu) \succeq 0, \quad (21)$$

where $R(n, S)$ is defined by

$$R(n, S) := P(0)^{-1}/n + \frac{1}{n}\sum_{k=1}^n \phi(\nu_k)\phi^T(\nu_k)$$

$R(n, S)$ asymptotically serves as a time average of outer products of the collection of basis functions evaluated at the measurement points S. From (21), it is straightforward to see that the error covariance matrix is a function of the evaluated position ν in $\mathcal{R}$, is proportional to noise to signal ratio W, and decreases at the rate of $1/n$ and $R^{-1}(n, S)$. Now we present our collaboratively learning control protocol.

IV. Cooperatively Learning Control

Each of mobile vehicles receives measurements from neighbors, then updates its estimation of the gradient via the recursive algorithm presented in (13). Subsequently, based on this new gradient, the control for its coordination will be decided. We apply a new time notation to the recursive algorithm in (13) according to the coordination time notation. In particular, we replace $n - s \in \mathbb{Z}_+$ by $t \in \mathbb{Z}_+$ and $n \in \mathbb{Z}_+$ by $t + 1 \in \mathbb{Z}_+$ in (13) such that the resulting recursive algorithm with the new time index for agent i at its position $q_i(t)$ is given by

$$\begin{aligned} K_i(t+1) &= P_i(t)\Phi_{*i}^T\left(I + \Phi_{*i}P_i(t)\Phi_{*i}^T\right)^{-1}, \\ P_i(t+1) &= (I - K_i(t+1)\Phi_{*i})P_i(t), \\ \hat{\Theta}_i(t+1) &= \hat{\Theta}_i(t) + K_i(t+1)\left[Y_{*i} - \Phi_{*i}\hat{\Theta}_i(t)\right], \\ \nabla \hat{c}_i(t, q_i(t)) &= \phi'^T(q_i(t))\hat{\Theta}_i(t+1), \end{aligned} \quad (22)$$

where Y_{*i}, Φ_{*i} of agent i are defined in the same way as Y_*, Φ_* are defined in (13). Y_{*i} is the collection of collaboratively measured data. From (2), for all $j \in \mathcal{N}(i, q(t)) \cup \{i\}$, we have

$$Y_{*i} = \begin{bmatrix} \vdots \\ c(q_j(t)) \\ \vdots \end{bmatrix} + \begin{bmatrix} \vdots \\ w_j(t) \\ \vdots \end{bmatrix} =: c_{*i}(t) + w_{*i}(t), \quad (23)$$

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 $t + 1$. $w_j(t)$ is the measurement noise of sensor j, and is independently and identically distributed over j. We also define new variables $c_{*i}(t)$ and $w_{*i}(t)$ as in (23) for later use.

Based on the latest update of the gradient estimate $\nabla \hat{c}_i(t, q_i(t))$, a distributed control for agent i is decided by

$$\begin{aligned} u_i(t) &:= -\nabla U(q_i(t)) \\ &+ \sum_{j\in\mathcal{N}(i,q(t))} a_{ij}(q(t))(p_j(t) - p_i(t)) + k_4\nabla \hat{c}_i(t, q_i(t)), \end{aligned} \quad (24)$$

where $k_4 \in \mathbb{R}_+$ is a gain factor. The first term in (24) is the gradient of (8) which attracts agents while avoiding collisions among them. Also it restricts the movements of agents inside $\mathcal{R}$. Appropriate artificial potentials can be added to $U(q_i)$ for agents to avoid obstacles in $\mathcal{R}$. The second term in (24) is an effort for agent i to match its velocity with those of neighbors. This term is also called a "velocity consensus" and serves as a damping force.

Incorporating the closed-loop discrete time model in (1) along with the proposed control in (24) gives

$$\begin{aligned} q_i(t+1) =& q_i(t) + \gamma(t)p_i(t) \\ p_i(t+1) =& p_i(t) + \gamma(t)\Big\{ -\nabla U(q_i(t)) \\ &+ \sum_{j\in\mathcal{N}(i,q(t))} a_{ij}(q(t))(p_j(t)-p_i(t)) \\ &+ k_4\phi'^T(q_i(t))\hat{\Theta}_i(t+1)\Big\}, \end{aligned} \tag{25}$$

where we applied notations to (1) by replacing δ by $\gamma(t)$, $t+\delta \in \mathbb{R}_+$ by $t+1 \in \mathbb{Z}_+$ and $t \in \mathbb{R}_+$ by $t \in \mathbb{Z}_+$. The sampling rate of the coordination of CoLMAs will be gradually increased for perfect tracking of the maximum of an unknown field. In particular, we propose the control protocol in Eq. (24) with the scheduling of the sampling time

$$\begin{gathered} \gamma(t) > 0, \quad \sum_{t=1}^{\infty}\gamma(t) = \infty, \quad \sum_{t=1}^{\infty}\gamma^2(t) < \infty, \\ \lim_{t\to\infty}\sup[1/\gamma(t) - 1/\gamma(t-1)] < \infty. \end{gathered} \tag{26}$$

In this protocol, the sampling time size deceases gradually $\gamma(t) \to 0$, as $t \to \infty$ with properties in (26), which let us apply the ODE approach [13], [14], [15] for convergence analysis.

V. Convergence Analysis

In this section, we study the convergence properties of CoLMAs. In order to analyze the convergence properties of (22), (25) and (26), we utilize Ljung's ordinary differential equation (ODE) approach developed in [13], [14], [15]. In particular, Ljung [13], [14] developed an analysis technique of general recursive stochastic algorithms in the form of

$$x(t) = x(t-1) + \gamma(t)Q(t; x(t-1), \varphi(t)), \tag{27}$$

along with the observation process

$$\varphi(t) = g(t; x(t-1), \varphi(t-1), e(t)). \tag{28}$$

By using $x(t) := [q(t)^T, p(t)^T]^T$, where $q(t) := \mathrm{col}(q_1(t), \cdots, q_{N_s}(t)), p(t) := \mathrm{col}(p_1(t), \cdots, p_{N_s}(t))$, we can transform (25) and (22) into (27) and (28) respectively. It can be shown that the regularity conditions in [14] are satisfied 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 (9).
- *M2* Artificial potentials and the adjacency matrix are continuously differentiable w.r.t q and derivatives are bounded[1].

Let $D_R \ni x$ be an open connected set where the regularity conditions [14] are valid.

We will utilize the following theorem introduced in [13], [14].

Theorem 3: (Ljung [13], [14]) Consider the algorithm (27) and (28) subject to the regularity conditions [14]. Let $\bar{D}$ be a compact subset of D_R such that the trajectories of the associated ODE

$$\frac{d}{d\tau}x(\tau) = f(x(\tau)) \tag{29}$$

where $f(x) := \lim_{t\to\infty}\mathbb{E}Q(t; x, \bar{\varphi}(t,x))$, that start in $\bar{D}$ remain in a closed subset $\bar{D}_R$ of D_R for $\tau > 0$. Assume that

1) there is a random variable L such that

$$x(t) \in \bar{D} \text{ and } |\varphi| < L \text{ infinitely often w.p.1} \tag{30}$$

2) the differential equation (29) has an invariant set D_c with domain of attraction $D_A \supset \bar{D}$.

Then $x(t) \to D_c$ with probability one as $t \to \infty$.

Proof: See [14].

Remark 4: Due to the assumption *M1* and the unbiased estimates (20), $f(x)$ in (29) of Theorem 3 is obtained by

$$f(x) = \begin{bmatrix} p \\ -\nabla U(q) - \hat{L}(q)p - \nabla C(q) \end{bmatrix}, \tag{31}$$

where $C(q) \in \bar{\mathbb{R}}_+$ is defined by

$$C(q) := k_4\sum_{i\in\mathcal{I}}[c_{max} - c(q_i)], \quad k_3 > 0, \tag{32}$$

here c_{max} is the maximum of the entity of interest and is assumed to be bounded.

The global performance cost that serves as the global goal of CoLMAs, is defined as

$$V(q(\tau), p(\tau)) := U(q(\tau)) + \frac{p^T(\tau)p(\tau)}{2} + C(q(\tau)). \tag{33}$$

We have the following theorem.

Theorem 5: For any initial $x_0 = \mathrm{col}(q_0, p_0) \in D_R$, we consider the recursive coordination algorithm transformed in terms of (27) and (28) under regularity conditions [14]. Let $D_A := \{\, x \in D_R \mid V(x) \leq a \,\}$ be a level-set of the cost function in (33). 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 in D_c with probability one as $t \to \infty$.

Proof: From Theorem 3, the asymptotic trajectory $x(\tau) := \mathrm{col}(q(\tau), p(\tau)) \in D_R$ is given by the associated ODE

$$\frac{dx(\tau)}{d\tau} = f(x(\tau)). \tag{34}$$

Taking the derivative of $V(x(\tau))$ in (33) with respect to τ and using (34), we obtain

$$\begin{aligned} \frac{dV(x(\tau))}{d\tau} &= \left(\frac{\partial V(x)}{\partial x}\right)^T f(x(\tau)) \\ &= -p^T(\tau)\hat{L}(q(\tau))p(\tau) \leq 0. \end{aligned} \tag{35}$$

From (7) and (33), we conclude that $V(x)$ is radially-unbounded, i.e., $V(x) \to \infty$ as $||x|| \to \infty$. Then

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

is a bounded set with $\frac{d}{d\tau}V(x) \leq 0$ for all $x \in D_A$, which is a positively invariant set. By LaSalle's invariant principle and Theorem 3, Theorem 5 follows. □

[1] This can be done. See [5]

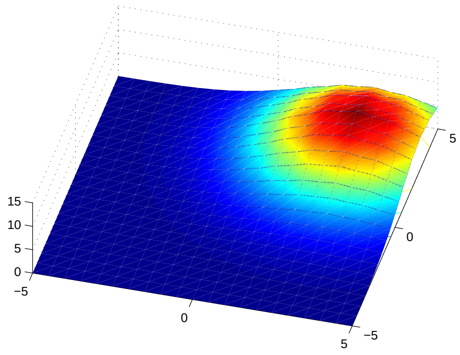

Fig. 3. True field of interest to be learned.

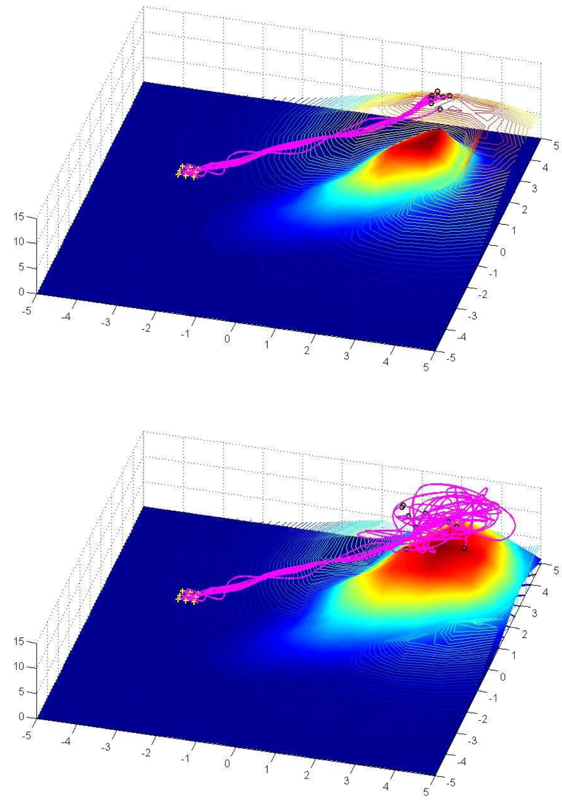

Fig. 4. (a)-up: Trajectories of nine agents with $d = 0.2$ printed at the height of 15 at iteration time $t = 80$ (b)-down: Trajectories of agents at iteration time $t = 160$. The emerging background in different colors represents the learned field by agent 1. Contour lines represent the error field between the true field and the estimated field by agent 1.

VI. Simulation Results

We applied CoLMAs to the static field depicted in Fig. 3. The estimate of the unknown field was updated once per iteration for coordination. Nine agents were launched with the equilibrium distance $d = 0.2$ in Fig. 4. Fig. 4-(a) shows that the recursively estimated image of the field by agent 1 at iteration time $t = 80$. This figure also shows that the error field with large values (depicted by the colored lines) at regions that were not sampled by agent 1 and its neighbors. Fig. 4-(b) illustrates the updated image of the field by agent 1 at iteration time $t = 160$. Nine agents have located the maximum of the field successfully.

VII. Conclusions

This paper presented a novel class of self-organizing autonomous sensing agents that form a swarm and learn through noisy cooperative measurements 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 for their mobility.

VIII. Acknowledgment

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] J. Cortes, S. Martinez, and F. Bullo, "Spatially-distributed coverage optimization and control with limited-range interactions," *ESAIM. Control, Optimisation and Calculus of Variations*, vol. 11, pp. 691–719, 2005.

[3] A. Jadbabie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomuous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 988–1001, June 2003.

[4] H. G. Tanner, A. Jadbabaie, and G. J. Pappas, "Stability of flocking motion," University of Pennsylvania," Technical Report, 2003.

[5] Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithm and theory," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 401–420, March 2006.

[6] J. Adler, "Chemotaxis in bacteria," *Journal of Supramolecular Structure*, vol. 4, no. 3, pp. 305–317, 1966.

[7] 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.

[8] P. Őgren, E. Fiorelli, and N. E. Leonard, "Cooperative control of mobile sensor networks: Adaptive gradient climing in a distributed envrironment," *IEEE Transaction on Automatic Control*, vol. 49, no. 8, p. 1292, August 2004.

[9] N. Leonardo and A. Robinson, "Adaptive sampling and forecasting plan AOSN-II MB'03 project," Available: http://www.princeton.edu/ dcsl/aosn/documents/ASFP.pdf, Tech. Rep., 2003.

[10] D. Grünbaum, "Schooling as a strategy for taxis in a noisy environment," *Evolutionary Ecology*, vol. 12, pp. 503–522, 1998.

[11] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model." *Computer Graphics*, vol. 21, no. 4, pp. 25–34, 1987.

[12] K. J. Åström and B. Wittenmark, *Adaptive Control*, 2nd ed. Addison Wesley, 1995.

[13] L. Ljung, "Analysis of recursive stochastic algorithms," *IEEE Transactions on Automatic Control*, vol. 22, no. 4, pp. 551–575, 1977.

[14] ——, "Theorems for the asymptotic analysis of recursive, stochastic algorithms," Department of Automatic Control, Lund Institute of Technology," Technical Report 7522, 1975.

[15] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*. Cambridge, Massachusetts, London, England: The MIT Press, 1983.