

Learning Coverage Control of Mobile Sensing Agents in One-Dimensional Stochastic Environments

Jongun Choi, *Member, IEEE*, and Roberto Horowitz, *Member, IEEE*

Abstract—This technical note presents learning coverage control of mobile sensing agents without *a priori* statistical information regarding random signal locations in a one-dimensional space. In particular, the proposed algorithm controls the usage probability of each agent in a network while simultaneously satisfying an overall network formation topology. The proposed control algorithm is rather direct, not involving any identification of an unknown probability density function associated to random signal locations. Our approach builds on diffeomorphic function learning with kernels. The almost sure convergence properties of the proposed control algorithm are analyzed using the ODE approach. Numerical simulations for different scenarios demonstrate the effectiveness of the proposed approach.

Index Terms—Coverage control, learning with kernels, mobile sensing agents.

I. INTRODUCTION

Coordination of autonomous agents and distributed mobile sensor networks have increasingly drawn the attention of engineers and scientists [1]–[6]. Mobile sensing agents form an ad-hoc wireless communication network in which each agent operates usually under a short communication range, a limited memory storage, and limited computational power. Sensing agents are often spatially distributed in an uncertain surveillance environment; can sense, communicate, and take control actions locally in order to achieve a global goal. One of the challenging problems in the coordination of sensing agents is to allocate coverage regions to agents optimally, which will be referred to as the coverage control problem. In [2], a distributed coordination algorithm for sensing agents was derived and analyzed based on the classic Lloyd algorithm [7], which requires the knowledge of the probability density function associated to random signal locations. However, in practice, the exact knowledge of a statistical distribution of signal locations including its support may not be available *a priori*. This coverage control strategy is extended by [8] using a deterministic adaptive control approach assuming that the true density related to the cost function can be measured by sensing agents. Dynamic vehicle routing problems were studied in [4], [5], in which mobile agents in a fixed known convex region must visit event points generated by an (unknown) spatio-temporal Poisson point process. Arsie and Frazzoli [5] introduced strategies to minimize the expected time between the appearance of a target point and the time it is visited by one of the agents. The policy was similar to the MacQueen's [9] learning vector quantization algorithm in that it does not rely on the knowledge of the underlying stochastic process. Due to recent advances in micro-electro-mechanical systems (MEMS)

Manuscript received December 10, 2008; revised August 05, 2009. First published February 05, 2010; current version published March 10, 2010. Recommended by Associate Editor M. Prandini.

J. Choi is with the Departments of Mechanical Engineering, and Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824-1226 USA (e-mail: jchoi@egr.msu.edu).

R. Horowitz is with the Department of Mechanical Engineering, University of California, Berkeley, CA 94720-1740 USA (e-mail: horowitz@me.berkeley.edu).

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

Digital Object Identifier 10.1109/TAC.2010.2040510

technology [10], each agent can afford a particular set of sensors among different types. Measurements from heterogeneous sensors in different locations will provide statistically rich information in the sense of redundancy and complementarity [11], [12]. Such collective measurements along with multisensor fusion algorithms [13] will improve the performance of the sensor network significantly regarding estimation, prediction and tracking of a process of interest. In [14], heterogeneous robots with different configurations use their special capabilities collaboratively to accomplish localization and mapping tasks. Motivated by current research trends and needs, we propose a class of self-organizing sensing agents with the following properties in a one-dimensional space. First, a network of sensing agents should perform the coverage control without the statistical knowledge of random signal locations. Second, frequencies of random events or signal occurrences covered by agents are to be controlled according to each agent's limited capability and resources. To this end, we introduce a concept of the usage frequency of an agent, which will be referred to as *the usage probability* of the agent. Finally, the formation topology of the sensor network should be controlled so that each sensing agent can select specific neighbors equipped with functionally complementary sensor configurations with respect to its own configuration.

Standard notation is used throughout the note. Let \mathbb{R} , $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{> 0}$, $\mathbb{Z}_{\geq 0}$, and $\mathbb{Z}_{> 0}$ denote, respectively, the set of real, non-negative real, positive real, non-negative integer, and positive integer numbers. The positive definiteness (respectively, semi-definiteness) of a matrix A is denoted by $A \succ 0$ (respectively, $A \succeq 0$). The relative complement of a set A in a set B is denoted by $B \setminus A := B \cap A^c$, where A^c is the complement of A . The derivative of a column vector $y \in \mathbb{R}^m$ with respect to a column vector $x \in \mathbb{R}^n$ is defined by the matrix $\partial y / \partial x \in \mathbb{R}^{n \times m}$ whose (i, j) th entry is given by $\partial y_i / \partial x_j$. Other notation will be explained in due course.

II. PROBLEM STATEMENT

Consider N number of agents with a set of identities denoted by $\mathcal{I} = \{1, \dots, N\}$. Let $q_\gamma(t)$ be the location of agent γ at time t . The collection of agent's locations is denoted by $q(t) = [q_1(t) \cdots q_N(t)]^T \in \mathbb{R}^N$. In each discrete time iteration $t \in \mathbb{Z}_{> 0}$, an event occurs at a stochastic location, generated by the stationary random process $u : \mathbb{Z}_{> 0} \rightarrow \mathcal{U}$, where $\mathcal{U} = (u_{\min}, u_{\max}) \subset \mathbb{R}$ is the signal locational space in which events or signals of interest occur. Each sensing agent will detect an event or a signal and its corresponding location over the surveillance region in charge. We assume that agent γ takes charge of measuring signals and getting necessary tasks done in its *coverage region* R_γ determined by the nearest neighbor rule [2]. The coverage region R_γ is given by the Voronoi cell [15] of agent γ

$$R_\gamma := \{u \in \mathcal{U} \mid |u - q_\gamma| \leq |u - q_i|, \forall i \neq \gamma\} \quad (1)$$

where $|\cdot|$ is the Euclidean norm, and u is the location of the signal. For the given configuration $q(t)$ and the signal location $u(t)$, the *winning index* $w(\cdot, \cdot) : \mathbb{R}^N \times \mathcal{U} \rightarrow \mathcal{I}$ is defined by

$$w(q(t), u(t)) := \arg \left\{ \min_{i \in \mathcal{I}} |u(t) - q_i(t)| \right\}. \quad (2)$$

When there are multiple minimizers in (2), the function will select the smallest index. Throughout the note, the winning index in (2) will be often written as $w(u(t))$, or $w(t)$ for notational simplicity in different contexts. The sequence $w(t)$ in (2) is then a random sequence with a discrete probability distribution that is induced by the pdf f_U . A vector

$p = [p_1 \cdots p_N]^T \in \mathbb{R}_{\geq 0}^N$ is referred to as the *usage probability distribution* whose elements are defined by the *usage probabilities* of agents

$$p_\gamma := \int_{u \in R_\gamma} f_U(u) du \\ = \Pr[w(q, u(t)) = \gamma] \geq 0, \forall \gamma \in \mathcal{I} \quad (3)$$

where $\sum_{\gamma=1}^N p_\gamma = 1$. Hence, p_γ provides a direct measure of how frequently agent γ is being used.

Assume that there exists a cost function J of the sensor network, which may be related to the functional lifetime, limited resources, and/or capabilities of sensing agents. Suppose that $J : \mathbb{R}_{\geq 0}^N \rightarrow \mathbb{R}_{\geq 0}$ is a function of the usage probability distribution p . Then, the optimal (or target) usage probability distribution that minimizes J , will be denoted by

$$p^\circ = [p_1^\circ \cdots p_N^\circ]^T \in \mathbb{R}_{> 0}^N, \quad \sum_{\gamma=1}^N p_\gamma^\circ = 1. \quad (4)$$

p° will be generated by (3) with the pdf f_U and optimal locations¹ of agents, denoted by $\bar{q}^\circ = [q_1^\circ \cdots q_N^\circ]^T \in \mathcal{U}^N$. The problem is then formulated as follows.

Problem: Consider the random variable of signal locations generated by a pdf f_U . We assume that f_U is continuous and its support is connected and finite. However, f_U and its support are not known a priori. Assume that agents sense signals and their locations based on the nearest neighbor rule in (1). For a given p° in (4), design a learning coverage algorithm that coordinates sensing agents to achieve the following asymptotic properties:

$$\lim_{t \rightarrow \infty} p_\gamma(q(t)) = p_\gamma^\circ, \forall \gamma \in \mathcal{I}, \\ \text{subject to } \lim_{t \rightarrow \infty} q_1(t) < \lim_{t \rightarrow \infty} q_2(t) < \cdots < \lim_{t \rightarrow \infty} q_N(t). \quad (5)$$

Remark 1: The constraint on the *formation order* of agents in (5) will predecide the neighbors of each agent, since some agents prefer particular agents to be its neighbors among agents equipped with heterogeneous sensing devices.

III. DIFFEOMORPHISM LEARNING WITH KERNELS

In this section, we explain a diffeomorphism that maps a domain containing the indices of agents to the signal locational space. This map plays a central role in providing a structure in our learning coverage algorithm. We introduce a fictitious random sequence $x : \mathbb{Z}_{> 0} \rightarrow \mathcal{X} \subset \mathbb{R}$, where $\mathcal{X} = (x_{\min}, x_{\max})$ is a *specified* finite open interval. We set $\mathcal{X} = (1/2, N + 1/2)$, so that $\mathcal{I} \subset \mathcal{X}$. Let $u : \mathcal{X} \rightarrow \mathcal{U}$ be a mapping from \mathcal{X} to the locational space \mathcal{U} . We now assume that u is actually a diffeomorphic function of x , i.e., $u : \mathcal{X} \rightarrow \mathcal{U}$ is a differentiable bijection and has a differentiable inverse such that the time samples of the locational random variable, $u(t)$, are generated by $u(t) = u(x(t))$. Thus, the pdf of x , $f_X : \mathcal{X}_{\mathbb{R}} \rightarrow \mathbb{R}_{> 0}$, is given by

$$f_X(x) = \left| \frac{du}{dx} \right| f_U(u(x)) \quad \text{for all } x \in \mathcal{X}. \quad (6)$$

The diffeomorphism $u : \mathcal{X} \rightarrow \mathcal{U}$, induces the pdf f_X from the unknown pdf f_U via (6). This map will be subsequently referred to as the

¹The map from a signal location u to the location of the responsible sensing agent q_u can be viewed as a *quantizer* in quantization theory [7]. p° must be produced by a quantizer with an optimal location codebook vector \bar{q}° .

reference or the target map. Since $u : \mathcal{X} \rightarrow \mathcal{U}$ is a diffeomorphism, the collection of optimal sensor locations in (5) becomes

$$\bar{q}^\circ = [q_1^\circ \cdots q_N^\circ]^T = [u(1) \cdots u(N)]^T \in \mathbb{R}^N. \quad (7)$$

Suppose that the target map $u(x)$ can be obtained by solving an integral equation of the first kind with a known smooth scalar symmetric kernel $\mathcal{K}(\cdot, \cdot) : \mathcal{X}^e \times \mathcal{X}^e \rightarrow \mathbb{R}_{\geq 0}$ and unknown influence coefficients c_ν° 's $\in \mathbb{R}$ that satisfy

$$u(x) = \sum_{\nu \in \mathcal{I}^e} \mathcal{K}(x, \nu) c_\nu^\circ \quad (8)$$

where

$$\mathcal{I}^e = \{-(m-1), \dots, N+m\}, \\ \mathcal{X}^e = \left(\frac{-1}{2} - (m-1), N+m + \frac{1}{2} \right) \quad (9)$$

for some integer $0 < m \ll N$. To obtain a distributed coordination algorithm, the support of the kernel has to be finite. We assume that the kernel in (8) is a radial basis function that is (at least) \mathcal{C}^1 differentiable with a compact support. Then the resulting $u(x)$ is a \mathcal{C}^1 diffeomorphism with an appropriate set of c_ν° . The elements of the vector $c^\circ = [c_{-(m-1)}^\circ \cdots c_{N+m}^\circ]^T \in \mathbb{R}^{N+2m}$ are the *unknown* optimal influence coefficients that satisfy $q_\gamma^\circ = u(\gamma)$ for all $\gamma \in \mathcal{I}$. Hence, a logical approach to deal with our problem is to coordinate sensing agents according to the diffeomorphism function learning with kernels. The time varying outputs of the learning algorithm will directly update the locations of agents given as

$$q_\gamma(t) = q(\gamma, t), \forall \gamma \in \mathcal{I}^e \quad (10)$$

where $q(x, t)$ is produced by the estimates of influence coefficients $\hat{c}_\nu(t)$

$$q(x, t) = \sum_{\nu \in \mathcal{I}^e} \mathcal{K}(x, \nu) \hat{c}_\nu(t), \text{ for all } x \in \mathcal{X}^e. \quad (11)$$

Here $\{q(x, t)\}$ in (11) is a parameterized family of smooth functions that contains the diffeomorphism of interest in (8). For given time t , we define the extended locational space \mathcal{U}^e by the union of \mathcal{U} and the range space of $q(x, t)$ (denoted by $\mathbb{R}(q(\mathcal{X}, t))$), i.e., $\mathcal{U}^e := \mathcal{U} \cup \mathbb{R}(q(\mathcal{X}, t))$. We define the influence coefficient estimate vector by

$$\hat{c}(t) := [\hat{c}_{-(m-1)}(t) \cdots \hat{c}_{N+m}(t)]^T \in \mathbb{R}^{N+2m}. \quad (12)$$

Equation (10) can then be rewritten as

$$q^e(t) := [q_{-(m-1)}(t) \cdots q_{N+m}(t)]^T = K \hat{c}(t) \in \mathbb{R}^{N+2m}$$

where $K \in \mathbb{R}^{(N+2m) \times (N+2m)}$ is the kernel matrix with (i, j) element $K_{i,j} := \mathcal{K}(i-m, j-m)$, which must be rank $N+2m$. For the function $q(x, t)$ in (11) to converge to an orientation preserving diffeomorphism, it is necessary to have

$$\lim_{t \rightarrow \infty} q'(x, t) = \lim_{t \rightarrow \infty} \frac{\partial}{\partial x} q(x, t) > 0.$$

Define the vector of partial derivatives of $q(x, t)$ with respect to x evaluated at $\gamma \in \mathcal{I}^e$ by

$$q^e(t) = [q'(-(1-m), t) \cdots q'(N+m, t)]^T \\ = K' \hat{c}(t) \quad (13)$$

where $K' \in \mathbb{R}^{(N+2m) \times (N+2m)}$ is the matrix whose (i, j) element is given by $K'_{i,j} := \partial \mathcal{K}(x, \lambda) / \partial x|_{x=i-m, \lambda=j-m}$. K (respectively K') is the collection of kernel vectors k_γ (respectively k'_γ) as defined by

$$\begin{aligned} K^T &= [k_{-(1-m)} \quad \cdots \quad k_{N+m}], \\ K'^T &= [k'_{-(1-m)} \quad \cdots \quad k'_{N+m}]. \end{aligned}$$

Therefore, for $\gamma \in \mathcal{I}^e$, we have

$$\begin{aligned} q_\gamma(t) &= k_\gamma^T \hat{c}(t), \\ q'_\gamma(t) &= \left. \frac{\partial q(x, t)}{\partial x} \right|_{x=\gamma} = k'_\gamma{}^T \hat{c}(t). \end{aligned}$$

IV. LEARNING COVERAGE CONTROL

The following discrete-time control system describes the motion of mobile agents:

$$q_i(t+1) = q_i(t) + \zeta_i(t), \quad \forall i \in \mathcal{I}^e \quad (14)$$

where $q_i(t)$ and $\zeta_i(t)$ are, respectively, the position and the control of agent i at time index $t \in \mathbb{Z}_{\geq 0}$. For a sampled input location $u(t)$ at time t , the control of each sensing agent will have the form of

$$\zeta_i(t) = \alpha(t) k_i^T [-\delta \hat{c}_1(t) - \delta \hat{c}_2(t)], \quad \forall i \in \mathcal{I}^e \quad (15)$$

where $\alpha(t)$ is the monotonically decreasing gain sequence often used in stochastic approximation theory [16], [17] and it satisfies the following properties:

$$\alpha(t) > 0, \quad \sum_{t=1}^{\infty} \alpha(t) = \infty, \quad \sum_{t=1}^{\infty} \alpha^2(t) < \infty. \quad (16)$$

This gain sequence $\alpha(t)$ is introduced to guarantee that sensing agents converge to an optimal configuration in spite of stochastic locational signals. This sufficiently slow vanishing rate of the sequence is a key mechanism to ensure the almost sure convergence of states by slowly attenuating the effects of randomness [16], [17]. $\delta \hat{c}_1(t)$ and $\delta \hat{c}_2(t)$ in (15) will be provided shortly. To parameterize a family of slopes of $q(x, t)$ properly at the boundary of \mathcal{X} , agent 1 (respectively, agent N) needs to memorize and update the positions of fictitious agents $-(m-1) \cdots 0$ (respectively, agents $(N+1) \cdots (N+m)$) according to (15). These fictitious agents do not have sensors and are only passively updated by either agents 1 or N .

We first define some notation. Let $\bar{\varrho}x$ and $\underline{\varrho}x$ be the indices associated to the extremum values of $\{q_\gamma \mid \gamma \in \mathcal{I}\}$ defined, respectively, by

$$\bar{\varrho}x := \arg \left\{ \max_{\gamma \in \mathcal{I}} q_\gamma \right\}, \quad \underline{\varrho}x := \arg \left\{ \min_{\gamma \in \mathcal{I}} q_\gamma \right\}. \quad (17)$$

The indices of local neighbors in \mathcal{U}^e , $\bar{\varphi} : \mathcal{I} \rightarrow \mathcal{I}$ and $\underline{\varphi} : \mathcal{I} \rightarrow \mathcal{I}$ are defined, respectively, by

$$\bar{\varphi}(w) := \arg \min_{\gamma \in \mathcal{I} \setminus \{w\}} \{|q_\gamma - q_w|\}, \quad \text{subject to } q_\gamma \geq q_w, \quad (18)$$

$$\underline{\varphi}(w) := \arg \min_{\gamma \in \mathcal{I} \setminus \{w\}} \{|q_\gamma - q_w|\}, \quad \text{subject to } q_w \geq q_\gamma. \quad (19)$$

The first term $\delta \hat{c}_1$ in (15) is designed for the usage probability $\{p_\gamma(t) \mid \gamma \in \mathcal{I}\}$ to track the target usage probability $\{p'_\gamma \mid \gamma \in \mathcal{I}\}$. $\delta \hat{c}_1$ is given by

$$\delta \hat{c}_1(t) = \beta_1 \frac{\delta \hat{c}_1^X(w(t))}{p_{w(t)}^o} \quad (20)$$

where $\beta_1 > 0$ is a gain and $p_{w(t)}^o$ is the target usage probability of the winning index w at time t given by (2) and (4). The function $\delta \hat{c}_1^X(\cdot) : \mathcal{I} \rightarrow \mathbb{R}^{N+2m}$ is defined by

$$\delta \hat{c}_1^X(w) := \begin{cases} \frac{(k_{\bar{\varphi}(w)} - k_{\underline{\varphi}(w)})}{2}, & \text{if } w \in \mathcal{I} \setminus \{\bar{\varrho}x, \underline{\varrho}x\}, \\ \frac{(k_w + k_{\bar{\varphi}(w)})}{2}, & \text{if } w = \underline{\varrho}x, \\ -\frac{(k_w + k_{\underline{\varphi}(w)})}{2}, & \text{if } w = \bar{\varrho}x. \end{cases} \quad (21)$$

The second term $\delta \hat{c}_2$ in (15) controls the orientation of the map $q(\cdot, t)$ in (11), and it is given by

$$\delta \hat{c}_2(t) := \beta_2 k'_{w(t)} (|q'_{w(t)}| - q'_{w(t)}) \frac{[\text{sign}(q'_{w(t)}) - 1]}{p_{w(t)}^o} \quad (22)$$

where $\beta_2 > 0$ is a gain and $q'_{w(t)} = q'(w(t), t)$ as defined in (13). $\text{sign}(\cdot)$ is defined by

$$\text{sign}(y) := \begin{cases} 1, & \text{if } y > 0, \\ 0, & \text{if } y = 0, \\ -1, & \text{if } y < 0. \end{cases}$$

To calculate $\delta \hat{c}_2(t)$ in (22), agent γ should update the slope of the map $q(x, t)$ at $x = \gamma$ and keep the updated slope for the next iteration. Hence, for agent γ , the proposed learning coordination algorithm is summarized as follows:

$$\begin{aligned} q_\gamma(t+1) &= q_\gamma(t) + \alpha(t) \\ &\quad \times k_\gamma^T [-\delta \hat{c}_1(t) - \delta \hat{c}_2(t)], \quad \forall \gamma \in \mathcal{I}^e, \\ q'_\gamma(t+1) &= q'_\gamma(t) + \alpha(t) \\ &\quad \times k'_\gamma{}^T [-\delta \hat{c}_1(t) - \delta \hat{c}_2(t)], \quad \forall \gamma \in \mathcal{I}. \end{aligned} \quad (23)$$

Since $q^e(t) = K \hat{c}(t)$ and K is bijective, it is easy to see that the overall dynamics of agents in (23) can be rewritten as

$$\hat{c}(t+1) = \hat{c}(t) + \alpha(t) [-\delta \hat{c}_1(t) - \delta \hat{c}_2(t)] \quad (24)$$

where $\hat{c}(t)$ is the influence coefficient estimate defined in (12). For convergence analysis, we will consider the learning coordination algorithm in the form of the centralized adaptation in (24).

V. THE MAIN RESULT

We use Ljung's ordinary differential equation (ODE) approach developed in [16], [18] to analyze the convergence properties of our new learning algorithm. Equation (24) can be expressed as

$$\dot{\hat{c}}(t+1) = \hat{c}(t) + \alpha(t) F(t, \hat{c}(t), u(t)) \quad (25)$$

where $F(t, \hat{c}(t), u(t)) := -\delta \hat{c}_1(\hat{c}(t), u(t)) - \delta \hat{c}_2(\hat{c}(t), u(t))$. The ODE associated to (25) is

$$\begin{aligned} \dot{\hat{c}}(\tau) &= f(\tau, \hat{c}(\tau)) = \mathbb{E}^{\hat{c}(\tau)} \{F(\hat{c}(\tau), u(t))\} \\ &= \int_{\mathcal{X}} F(\hat{c}(\tau), u(x)) f_X(x) dx \end{aligned} \quad (26)$$

where $\hat{c}(\tau)$ is kept constant at the frozen time τ in the calculation of $\mathbb{E}^{\hat{c}(\tau)} \{F(\hat{c}(\tau), u(t))\}$. Two of the nontrivial sufficient conditions for the ODE [16] approach to be applicable are that $F(t, \hat{c}, u)$ must be Lipschitz continuous in \hat{c} and u (B.3 in [16]), and the Lipschitz constants must be Lipschitz continuous in \hat{c} and u (B.4 in [16]). These conditions are verified by the following Lemma.

Lemma 2: For the input signal u , let $w(\hat{c}, u)$ be the value determined by (2) and $q(x, t)$ that builds on \hat{c} as in (11). Given the function $w(\hat{c}, u)$, except for a set $\{\hat{c}, u\}$ of measure zero, there exists a sufficiently small

$\delta > 0$ such that for any bounded \hat{c}_1, \hat{c}_2 and u_1, u_2 , if $\|\hat{c}_1 - \hat{c}_2\|_\infty < \delta$ and $|u_1 - u_2| < \delta$ then $|w(\hat{c}_1, u_1) - w(\hat{c}_2, u_2)| = 0$.

Proof: See [19].

Following the ODE approach, we freeze the parameter $\hat{c}(\tau)$ at time τ and take the average of $F(\hat{c}(\tau), u(t))$ over the random variable $u(t)$ as in (26). For this purpose, the winning index $w(\cdot)$ in (2) will be represented as $w(x, \tau)$ which is a function of the random variable x and the frozen time τ (and $\hat{c}(\tau)$). However, we will often omit τ for convenience. Even though we sample the random variable $u(t)$, we take average with respect to the random variable x , using the reference diffeomorphism $u(x)$ as in (6). For instance, we define $\Delta\hat{c}_1(\tau)$ and $\Delta\hat{c}_2(\tau)$, respectively, by

$$\begin{aligned}\Delta\hat{c}_1(\tau) &:= \int_{\mathcal{X}} \delta\hat{c}_1(x, \tau) f_X(x) dx, \\ \Delta\hat{c}_2(\tau) &:= \int_{\mathcal{X}} \delta\hat{c}_2(x, \tau) f_X(x) dx.\end{aligned}\quad (27)$$

We summarize sufficient conditions for the correct convergence of the learning coverage control algorithm.

A.1 p° in (4), random signal locations $u(t) \in \mathcal{U}$ with an associated unknown pdf f_U , and the kernel function \mathcal{K} in (11) are compatible in the sense that the family of smooth functions in (11) contains an optimal configuration in (7). Moreover, if $q'_\gamma > 0$, $\forall \gamma \in \mathcal{I}$, $q(\cdot, t)$ is an orientation preserving map.

A.2 p° , f_U , the kernel function \mathcal{K} , β_1 in (20) and β_2 in (22) satisfy that $\Delta\hat{c}_1 \neq -\Delta\hat{c}_2$, for any $\Delta\hat{c}_2 \neq 0$ where $\Delta\hat{c}_1$ and $\Delta\hat{c}_2$ are defined in (27).

Remark 3: In general, $q_\gamma(0)$ for all $\gamma \in \mathcal{I}$ may be initially far away from the support of the pdf f_U . However, it is straightforward to see that the algorithm attracts the agents into the support of the pdf f_U . Thus, in the following arguments, we assume that positions of winning agents whose indices are not extremum values ($w \in \mathcal{I} \setminus \{\underline{x}, \bar{x}\}$) are contained in the support of the pdf f_U , i.e., $q_w \in \text{Supp}(f_U)$, where $w \in \mathcal{I} \setminus \{\underline{x}, \bar{x}\}$.

For convergence analysis, we need to calculate changes in the usage probability distribution $\{p_\gamma \mid \gamma \in \mathcal{I}\}$ caused by changes in the influence coefficient vector $\hat{c}(\tau)$. The relationship is elaborated in the following lemma.

Lemma 4: The time derivative of the usage probability distribution $\dot{p}_{w(\tau)}$ is related to the time derivative of the influence coefficient vector $\dot{\hat{c}}(\tau)$ by

$$\dot{p}_{w(\tau)} \simeq f_U^*(q_{w(\tau)}, w(\tau)) \left[\delta\hat{c}_1^{X(w(\tau))} \right]^T \dot{\hat{c}}(\tau) \quad (28)$$

where $\delta\hat{c}_1^{X(w)}$ is defined in (21) and $f_U^* : \mathcal{U} \times \mathcal{I} \rightarrow \mathbb{R}_{\geq 0}$ is defined by

$$f_U^*(q_w, w) := \begin{cases} f_U(q_w), & \text{if } w \in \mathcal{I} \setminus \{\underline{x}, \bar{x}\} \\ f_U\left(\frac{q_w + q_{\bar{x}}(w)}{2}\right), & \text{if } w = \underline{x}, \\ f_U\left(\frac{q_w + q_{\underline{x}}(w)}{2}\right), & \text{if } w = \bar{x}. \end{cases} \quad (29)$$

Moreover, the approximation symbol used in (28) is replaced with an equal sign for the case of uniform pdfs.

Proof: See [19].

We introduce our main result for the uniform pdf f_U case.

Theorem 5: Consider the proposed learning coordination algorithm in (23) under conditions A.1 and A.2 with a uniform pdf f_U . Then the locational vector $q(t)$ of the sensing agents converges to an optimal codebook vector \bar{q}° almost surely, satisfying (7).

Now we present the proof of our main result.

Proof: Define the functions $q^\circ : \mathcal{U} \rightarrow \mathcal{U}$ by $q^\circ(u) := \arg \left\{ \min_{q_\gamma^\circ, \forall \gamma \in \mathcal{I}} |u - q_\gamma^\circ| \right\}$, where q_γ° is defined in (7) and

$w^\circ : \mathcal{U} \rightarrow \mathcal{I}$ by $w^\circ(u) := \arg \left\{ \min_{\gamma \in \mathcal{I}} |u - q_\gamma^\circ| \right\}$. Let us define the lower bounded functionals

$$\begin{aligned}V(\hat{c}(\tau)) &:= V_1(\hat{c}(\tau)) + V_2(\hat{c}(\tau)), \\ V_1(\hat{c}(\tau)) &:= \beta_1 \int_{\mathcal{X}} \frac{P_{w(x)}}{p_{w(x)}^\circ f_U^*(q^\circ(u(x)), w^\circ(u(x)))} f_X(x) dx, \\ V_2(\hat{c}(\tau)) &:= \frac{\beta_2}{2} \int_{\mathcal{X}} \frac{(|q'_{w(x)}| - q'_{w(x)})^2}{p_{w(x)}^\circ} f_X(x) dx\end{aligned}\quad (30)$$

where f_U^* is defined in (29) and $w(x) = w(u(x))$ is based on $q_\gamma(\tau)$ in (10) and (11) given by (2). p_γ° is the predefined optimal usage target probability distribution, as defined in (4). $\beta_1 > 0$ and $\beta_2 > 0$ are the weighting gains.

Applying the ODE approach to (2), (20), (22), and (24), we obtain

$$\dot{\hat{c}}(\tau) = -\Delta\hat{c}_1(\tau) - \Delta\hat{c}_2(\tau) \quad (31)$$

where $\Delta\hat{c}_1(\tau)$ and $\Delta\hat{c}_2(\tau)$ are defined by (27).

Differentiating $V_1(\hat{c}(\tau))$ in (30) with respect to time τ , and utilizing (28) in Lemma 4, we obtain

$$\begin{aligned}\dot{V}_1(\hat{c}(\tau)) &= \int_{\mathcal{X}} \frac{\beta_1 \left(\frac{\partial p_{w(x)}}{\partial \hat{c}} \right)^T f_X(x)}{p_{w(x)}^\circ f_U^*(q^\circ(u(x)), w^\circ(u(x)))} dx \dot{\hat{c}}(\tau) \\ &= [\Delta\hat{c}_1(\tau)]^T \dot{\hat{c}}(\tau).\end{aligned}\quad (32)$$

Taking the second partial derivative of $V_1(\hat{c})$ with respect to \hat{c} , we obtain

$$\frac{\partial^2 V_1(\hat{c})}{\partial \hat{c}^2} = 0 \in \mathbb{R}^{N+2m} \quad (33)$$

where \hat{c}_m is the m th element of \hat{c} . As can be seen in (33), the second derivative of $V_1(\hat{c})$ with respect to \hat{c} is a zero matrix, $\partial^2 V_1(\hat{c}) / \partial \hat{c}^2 = 0$. Taking the time derivative of $V_2(\hat{c}(\tau))$ with respect to time τ , $\dot{V}_2(\hat{c}(\tau))$ is obtained by

$$\begin{aligned}\dot{V}_2(\hat{c}(\tau)) &= \beta_2 \int_{\mathcal{X}} \frac{f_X(x)}{p_{w(x)}^\circ} (|q'_{w(x)}| - q'_{w(x)}) \\ &\quad [\text{sign}(q'_{w(x)}) - 1] k'_{w(x)} dx \dot{\hat{c}}(\tau) \\ &= \left(\frac{\partial V_2(\hat{c}(\tau))}{\partial \hat{c}(\tau)} \right)^T \dot{\hat{c}}(\tau) = \left(\int_{\mathcal{X}} \delta\hat{c}_2(x) f_X(x) dx \right)^T \dot{\hat{c}}(\tau).\end{aligned}\quad (34)$$

The matrix of the second derivative of $V_2(\hat{c})$ with respect to \hat{c} is positive semi-definite

$$\begin{aligned}\frac{\partial^2 V_2(\hat{c})}{\partial \hat{c}^2} &= \beta_2 \int_{\mathcal{X}} k'_{w(x)} k'^T_{w(x)} [\text{sign}(q'_{w(x)}) - 1]^2 \\ &\quad \times \frac{f_X(x)}{p_{w(x)}^\circ} dx \succeq 0.\end{aligned}\quad (35)$$

From (32) and (34), we have

$$\begin{aligned}\dot{V}_1(\hat{c}(\tau)) &= \left(\frac{\partial V_1(\hat{c})}{\partial \hat{c}} \right)^T \dot{\hat{c}}(\tau) = \Delta\hat{c}_1(\tau)^T \dot{\hat{c}}(\tau), \\ \dot{V}_2(\hat{c}(\tau)) &= \left(\frac{\partial V_2(\hat{c})}{\partial \hat{c}} \right)^T \dot{\hat{c}}(\tau) = \Delta\hat{c}_2^T(\tau) \dot{\hat{c}}(\tau).\end{aligned}\quad (36)$$

From (30), (31) and (36), $\dot{V}(\hat{c}(\tau))$ can be represented as

$$\begin{aligned}\dot{V}(\hat{c}(\tau)) &= \dot{V}_1(\hat{c}(\tau)) + \dot{V}_2(\hat{c}(\tau)) \\ &= -\|\Delta\hat{c}_1(\tau) + \Delta\hat{c}_2(\tau)\|^2.\end{aligned}\quad (37)$$

From (37), $\dot{V}(\hat{c}(\tau))$ is negative semi-definite. Integrating (37) with respect to time, for all $T \geq 0$, $V(\hat{c}(T)) - V(\hat{c}(0)) = \int_0^T \dot{V}(\hat{c}(\tau)) d\tau$. This implies that $V(\hat{c}(T)) \leq V(\hat{c}(0))$, $V(\hat{c}) \in \mathcal{L}_\infty$, and $V_1(\hat{c})$ and $V_2(\hat{c})$ are bounded. Notice that $\Delta \hat{c}_1 \in \mathcal{L}_\infty$. From (30) and (34), utilizing Cauchy–Schwartz inequality we obtain

$$\|\Delta \hat{c}_2(\tau)\| \leq 2\sqrt{2\beta_2} \left\{ \int_{\mathcal{X}} k_{w(x)}'^T k_{w(x)}' dx \right\}^{1/2} V_2^{1/2}(\hat{c}(\tau)). \quad (38)$$

Thus, $\Delta \hat{c}_2(\tau) \in \mathcal{L}_\infty$ since $V_2(\hat{c}) \in \mathcal{L}_\infty$. Now we obtain

$$\dot{\hat{c}}(\tau) = -(\Delta \hat{c}_1(\tau) + \Delta \hat{c}_2(\tau)) \in \mathcal{L}_\infty. \quad (39)$$

From (36) we obtain

$$\begin{aligned} \frac{d\Delta \hat{c}_1(\tau)}{d\tau} &= \left(\frac{\partial^2 V_1(\hat{c})}{\partial \hat{c}^2} \right)^T \dot{\hat{c}}(\tau), \\ \frac{d\Delta \hat{c}_2(\tau)}{d\tau} &= \left(\frac{\partial^2 V_2(\hat{c})}{\partial \hat{c}^2} \right)^T \dot{\hat{c}}(\tau). \end{aligned}$$

$\dot{\hat{c}}(\tau)$, $(\partial^2 V_1(\hat{c})/\partial \hat{c}^2)$ and $(\partial^2 V_2(\hat{c})/\partial \hat{c}^2)$ are bounded from (39), (33) and (35). By differentiating $\dot{\hat{c}}(\tau)$ and $\dot{V}(\hat{c})$ in (36) and (39), respectively, with respect to time τ , we can conclude that $\dot{\hat{c}}(\tau) \in \mathcal{L}_\infty$ and $\dot{V}(\tau) \in \mathcal{L}_\infty$. Thus, $\dot{V}(\tau)$ is uniformly continuous in time τ . By Barbalat's lemma [20], we conclude that

$$\lim_{\tau \rightarrow \infty} \dot{V}(\tau) = - \lim_{\tau \rightarrow \infty} \|\Delta \hat{c}_1(\tau) + \Delta \hat{c}_2(\tau)\|^2 = 0 \text{ a.s.} \quad (40)$$

Due to the condition A.2, (40) implies that $\lim_{\tau \rightarrow \infty} \Delta \hat{c}_1(\tau) = 0$, $\lim_{\tau \rightarrow \infty} \Delta \hat{c}_2(\tau) = 0$ a.s. $\Delta \hat{c}_1(\tau)$ can then be rewritten as

$$\begin{aligned} \int_{\mathcal{X}} \delta \hat{c}_1(x) f_X(x) dx &= \int_{\mathcal{X}} \frac{\beta_1 \delta \hat{c}_1^X(w(x)) f_X(x)}{p_{w(x)}^o} dx \\ &= \frac{\beta_1 p_1}{p_1^o} \left[\frac{k_1 + k_2}{2} \right] + \frac{\beta_1 p_2}{p_2^o} \left[\frac{k_3 - k_1}{2} \right] + \dots \\ &+ \frac{\beta_1 p_{N-1}}{p_{N-1}^o} \left[\frac{k_N - k_{N-2}}{2} \right] + \frac{\beta_1 p_N}{p_N^o} \left[\frac{-k_N - k_{N-1}}{2} \right]. \end{aligned} \quad (41)$$

Since k_γ 's in (41) are discretized radial basis kernels centered at $\gamma \in \mathcal{I}$, from (41), we conclude that $\Delta \hat{c}_1(\tau) = \Delta \hat{c}_2(\tau) = 0$ and $\sum_{i \in \mathcal{I}} p_i = \sum_{i \in \mathcal{I}} p_i^o = 1$ imply that the usage probability p_γ in (3) is equal to the target probability p_γ^o in (4) for all $\gamma \in \mathcal{I}$, i.e., $p = p^o$. $\Delta \hat{c}_2(\tau) = 0$ along with $\Delta \hat{c}_1(\tau) = 0$ implies that q_γ is monotonically non-decreasing and $q(x, t)$ is orientation preserving (A.1). By A.1, the locational vector $q(t)$ converges to an optimal codebook vector \bar{q}^o almost surely. This completes the proof of Theorem 5. ■

VI. SIMULATION RESULTS AND DISCUSSION

Numerical simulations were carried out for different pdfs (f_U) and target probability distributions (p^o). A total of 34 agents and 4 fictitious agents ($m = 2$), as in (9) were used in all simulations. The local kernel used in this simulation study is given by

$$\mathcal{K}(x, \nu) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\sigma^2(x-\nu)^2}, & \text{if } |x - \nu| < \underline{\sigma}, \\ b(x, \nu), & \text{if } \underline{\sigma} \leq |x - \nu| \leq \bar{\sigma}, \\ 0, & \text{if } |x - \nu| > \bar{\sigma} \end{cases} \quad (42)$$

where $b(x, \nu)$ is a (C^1) bump function that makes the kernel smoothly diminish to 0, and $\bar{\sigma} > \underline{\sigma} > 3\sigma > 0$ and $\bar{\sigma} \simeq \underline{\sigma} \simeq 3\sigma$. A value of $\sigma = 1.5$ was used for the kernel given by (42). The initial positions of agents $q_\gamma(0)$'s were randomly distributed. The performance of the proposed scheme for different situations will be analyzed in the following systematic manner. We will compare the partitioning border

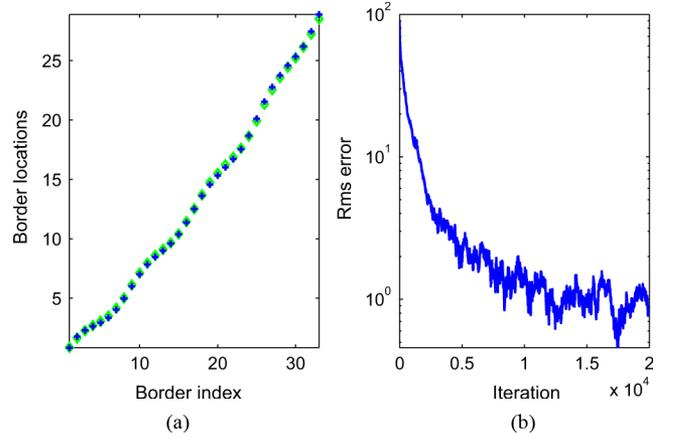


Fig. 1. Results for a uniform pdf f_U and a target usage probability p^o in (43). (a) The analytically calculated border locations b^o (diamonds \diamond) and border locations of agents $b(t)$ calculated from the new learning algorithm (pluses $+$) after 20000 iterations. (b) The RMS error value between $b(t)$ and b^o v.s. iteration time. The final RMS error value is 0.8185. (a) 20000 iterations (b) 20000 iterations.

vector $b(t) := [b_1(t) \cdots b_{N-1}(t)]^T$ calculated from the locations of agents at iteration time t , $b_i(t) := q_i(t) + q_{i+1}(t)/2$, $\forall i \in \mathcal{I} \setminus \{N\}$, with respect to the analytically calculated optimal border vector $2b^o$ for the corresponding pdf f_U and the target probability distribution p^o . The root mean square (RMS) error value between $b(t)$ and b^o will be computed. We will compare the usage probability distribution p with respect to the target usage probability distribution p^o in terms of the Kullback-Leibler (KL) measure of cross-entropy between p and p^o . Since p_γ is not available, it will be estimated by the following algorithm: $\hat{p}_\gamma(t+1) = \hat{p}_\gamma(t) + \alpha(t)(-\hat{p}_\gamma(t) + \delta_{\gamma, w(t)})$, $\forall \gamma \in \mathcal{I}$, $\hat{p}_\gamma(0) \geq 0 \forall \gamma \in \mathcal{I}$, $\sum_{\gamma=1}^N \hat{p}_\gamma(0) = 1$, where $w(t)$ is the winning index given by (2), $\delta_{i,j}$ is the Kronecker delta function, and $\alpha(t)$ is the stochastic approximation gain introduced in (16). The KL measure between $\hat{p} := [\hat{p}_1 \cdots \hat{p}_N]^T$ and p^o is given by $D(\hat{p}, p^o) = \sum_{\gamma=1}^N \hat{p}_\gamma \ln \hat{p}_\gamma / p_\gamma^o$, where $D(\hat{p}, p^o) \geq 0$, and $D(\hat{p}, p^o)$ vanishes if and only if $\hat{p} = p^o$.

Consider a uniform pdf f_U with $u \sim \mathcal{U}[0, 30]$ and the target probability distribution $\{p_\gamma^o \mid \gamma \in \mathcal{I}\}$ (see green circles in Fig. 2(a))

$$p_\gamma^o := \frac{\cos\left(\frac{8\pi\gamma}{N-1}\right) + 2\left(1 + \frac{\gamma}{N-1}\right)}{\sum_{j=0}^{N-1} \left[\cos\left(\frac{8\pi j}{N-1}\right) + 2\left(1 + \frac{j}{N-1}\right)\right]}, \quad \forall \gamma \in \mathcal{I}. \quad (43)$$

In this case, the border vector $b(t)$ of agents from the proposed algorithm successfully converges to the analytically calculated optimal border vector b^o for f_U and p^o after 20000 iterations as depicted in Fig. 1(a). The convergence rate in terms of the RMS error value between $b(t)$ and b^o v.s. iteration time is plotted in a logarithmic scale to magnify the transient behavior of the algorithm as shown in Fig. 1(b). The final RMS error value between $b(t)$ and b^o at $t = 20000$ is 0.8185.

Fig. 2(a) depicts the estimated usage probability \hat{p}_γ from the new learning law. This plot shows that $p \approx \hat{p}$ converges to p^o defined in (43) after 20000 iterations. Fig. 2(b) illustrates the KL measure between \hat{p} and p^o v.s. iteration time, showing that $\hat{p} \rightarrow p^o$ a.s. as $t \rightarrow \infty$.

To test our algorithm to a nonuniform pdf f_U , we consider a bimodal mixture model with two normal components. The locational random sequence u is assigned to one of the two normal densities, each of

²Notice that for given f_U and p^o , only b^o can be determined uniquely.

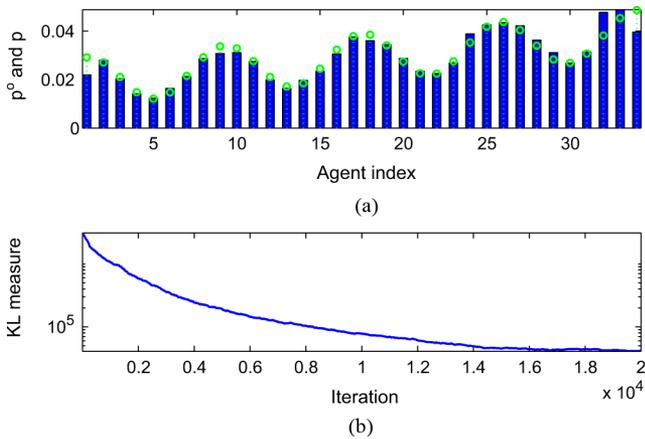


Fig. 2. Uniform pdf f_U : (a) The usage probability distribution estimate \hat{p} (solid bars), and the target usage probability distribution p^o (green circles) in (43). (b) The Kullback–Leibler (KL) measure of cross-entropy between \hat{p} and p^o v.s. iteration time. (a) 20000 iterations (b) 20000 iterations.

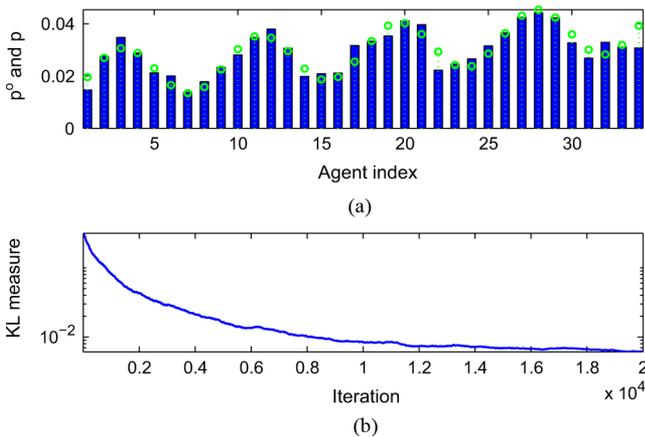


Fig. 3. Bimodal mixture model distribution: (a) The usage probability distribution estimate \hat{p} (solid bars), and the uniform target usage probability distribution p^o (green circles). (b) The KL measure of cross-entropy between \hat{p} and p^o v.s. iteration time. (a) 20000 iterations (b) 20000 iterations.

which has a prior discrete probability P_1 or P_2 . The joint probability density function of the random sequence u is given by

$$f_U(u|\{\theta_1, \theta_2\}) = f_{U_1}(u|\theta_1)P_1 + f_{U_2}(u|\theta_2)P_2 \quad (44)$$

where $\theta_i = [m_i, \sigma_i]^T$ is the sufficient statistics, $f_{U_i}(u|\theta_i) = 1/\sigma_i\sqrt{2\pi}e^{-1/2\sigma_i^2(u-m_i)^2}$ is the i th conditional probability density function, and $P_1 = 1/2$ and $P_2 = 1/2$ are the mixing probabilistic weights. We used that $m_1 = 8$, $\sigma_1 = 3$, $m_2 = -8$, and $\sigma_2 = 6$ for this case.

Consider the target probability distribution $\{p_\gamma^o | \gamma \in \mathcal{I}\}$ (see green circles in Fig. 3(a))

$$p_\gamma^o := \frac{\sin\left(\frac{8\pi\gamma}{N-1}\right) + 2\left(1 + \frac{\gamma}{N-1}\right)}{\sum_{j=0}^{N-1} \left[\sin\left(\frac{8\pi j}{N-1}\right) + 2\left(1 + \frac{j}{N-1}\right)\right]}, \quad \forall \gamma \in \mathcal{I}. \quad (45)$$

Fig. 3(a) depicts the estimated usage probability \hat{p}_γ of the new learning law, which shows that $p \approx \hat{p}$ converges to p^o in (45) after 20000 iterations. Fig. 3(b) also presents the KL measure between \hat{p}

and p^o v.s. iteration time, validating that $\hat{p} \rightarrow p^o$ a.s. as $t \rightarrow \infty$. In this case, the final RMS error value between $b(t)$ and b^o is 1.4105.

VII. CONCLUSION

A new formulation of learning coverage control for distributed mobile sensing agents was presented. This new one-dimensional coordination algorithm enables us to control the agent's usage probability and a formation order for given unknown statistics of random signal locations. The almost sure convergence properties of the proposed coordination algorithm were analyzed using the ODE approach for random signal locations with uniform pdfs. Successful simulation results for cases with a uniform pdf and a bimodal mixture model demonstrated the effectiveness of the proposed approach.

REFERENCES

- [1] J. Cortes and F. Bullo, "Coordination and geometric optimization via distributed dynamical systems," *SIAM J. Control Optim.*, vol. 44, no. 5, pp. 1543–1574, 2005.
- [2] J. Cortes, S. Martinez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Trans. Robot. Autom.*, vol. 20, no. 2, pp. 243–255, Feb. 2004.
- [3] Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithm and theory," *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 401–420, Mar. 2006.
- [4] E. Frazzoli and F. Bullo, "Decentralized algorithms for vehicle routing in a stochastic time-varying environment," in *Proc. Conf. Decision Control*, Paradise Island, Bahamas, Dec. 2004, pp. 3357–3363.
- [5] A. Arsie and E. Frazzoli, "Efficient routing of multiple vehicles with no communications," *Int. J. Robust Nonlin. Control*, vol. 18, no. 2, pp. 154–164, 2008.
- [6] J. Choi, S. Oh, and R. Horowitz, "Distributed learning and cooperative control for multi-agent systems," *Automatica*, vol. 45, no. 12, pp. 2802–2814, Dec. 2009.
- [7] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer Academic Publishers, 2001.
- [8] M. Schwager, J.-J. Slotine, and D. Rus, "Decentralized, adaptive control for coverage with networked robots," in *Proc. IEEE Int. Conf. Robot. Autom.*, Apr. 2007, pp. 3289–3294.
- [9] J. B. MacQueen, "Some methods of classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Stat. Prob.*, 1967, pp. 281–297.
- [10] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A Survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, Mar. 2002.
- [11] M. A. Abidi and R. C. Gonzales, Eds., *Data Fusion in Robotics and Machine Intelligence*. New York: Academic, 1992.
- [12] M. Chu, H. Haussecker, and F. Zhao, "Scalable information-driven sensor querying and routing for ad hoc heterogeneous sensor networks," *Int. J. High Performance Comput. Appl.*, vol. 16, no. 3, pp. 293–313, 2002.
- [13] S. Oh, L. Schenato, P. Chen, and S. Sastry, "Tracking and coordination of multiple agents using sensor networks: System design, algorithms and experiments," *Proc. IEEE*, vol. 95, no. 1, pp. 234–254, Jan. 2007.
- [14] R. Grabowski, L. Navarro-Serment, C. Paredis, and P. Khosla, "Heterogeneous teams of modular robots for mapping and exploration," *Autonomous Robots*, vol. 8, no. 3, pp. 293–308, 2000.
- [15] M. de Berg, O. Schwarzkopf, M. van Kreveld, and M. Overmars, *Computational Geometry: Algorithms and Applications*, 2nd ed. New York: Springer-Verlag, 2000.
- [16] L. Ljung, "Analysis of recursive stochastic algorithms," *IEEE Trans. Autom. Control*, vol. AC-22, no. 4, pp. 551–575, Aug. 1977.
- [17] H. J. Kushner and G. G. Yin, *Stochastic Approximation Algorithms and Applications*. New York: Springer, 1997.
- [18] L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification*. London, U.K.: The MIT Press, 1983.
- [19] J. Choi and R. Horowitz, Learning Coverage Control of Mobile Sensing Agents in One-Dimensional Stochastic Environments Dept. Mech. Eng., Michigan State Univ., Tech. Rep. CRL 2009-7-11, 2009.
- [20] H. K. Khalil, *Nonlinear Systems*, 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 1996.