

# Distributed coverage games for mobile visual sensors (II) : Reaching the set of global optima

Minghui Zhu and Sonia Martínez

**Abstract**—We formulate a coverage optimization problem for mobile visual sensor networks as a repeated multi-player game. Each visual sensor tries to optimize its own coverage while minimizing the processing cost. The rewards for the sensing are not prior information to the agents. We present an asynchronous distributed learning algorithm where each sensor only remembers the utility values obtained by its neighbors and itself, and the actions it played during the last two time steps when it was active. We show that this algorithm is convergent in probability to the set of global optima of certain coverage performance metric.

## I. INTRODUCTION

A substantial body of research on sensor networks has concentrated on simple sensors that can collect scalar data; e.g. temperature, humidity or pressure data. Thus, a main objective is the design of algorithms that can lead to optimal collective sensing through efficient motion control and communication schemes. However, scalar measurements can be insufficient in many situations; e.g. in automated surveillance or traffic monitoring applications. In contrast, cameras can collect visual data that are rich in information, thus having tremendous potential for monitoring applications, but at the cost of a higher processing overhead.

Precisely, this paper, part II, and its companion, part I, aim to solve a coverage optimization problem that takes into account part of the sensing/processing trade-off. Coverage optimization problems have mainly been formulated as cooperative problems where each sensor benefits from sensing the environment as a member of the group. However, sensing may also require expenditure; e.g. the energy consumed by image processing in visual networks. Because of this, we will endow each sensor with a utility function that quantifies this trade-off, formulating a coverage problem as a variation of congestion games in [21].

*Literature review.* In broad terms, the problem studied here is related to a bevy of sensor location and planning problems in the computational geometry, geometric optimization, and robotics literature. For example, different variations on the (combinatorial) Art Gallery problem include [20][23][26]. The objective here is how to find the optimum number of guards in a non-convex environment so that each point is visible from at least one guard. A related set of references for the deployment of mobile robots with omnidirectional cameras includes [9][8]. Unlike the Art Gallery classic algorithms, the latter papers only assume that robots have

local knowledge of the environment and no recollection of the past. Other related references on robot deployment in convex environments include [5][13] for anisotropic and circular footprints.

The paper [1] is an excellent survey on multimedia sensor networks where the state of the art in algorithms, protocols, and hardware is surveyed, and open research issues are discussed in detail. The investigation of coverage problems for static visual sensor networks is conducted in [4][10][24].

Another set of relevant references to this paper comprise those on the use of game-theoretic tools to (i) solve static target assignment problems, and (ii) devise efficient and secure algorithms for communication networks. In [15], the authors present a game-theoretic analysis of a coverage optimization for static sensor networks. This problem is equivalent to the weapon-target assignment problem in [19] which is nondeterministic polynomial-time-complete. In general, the solution to assignment problems is hard from a combinatorial optimization viewpoint.

Game Theory and Learning in Games are used to analyze a variety of fundamental problems in; e.g. wireless communication networks and the Internet. An incomplete list of references includes [2] on power control, [22] on routing, and [25] on flow control. However, there has been limited research on how to employ Learning in Games to develop distributed algorithms for mobile sensor networks. One exception is the paper [14] where the authors establish a link between cooperative control problems (in particular, consensus problems) and games (in particular, potential games and weakly acyclic games).

*Statement of contributions.* The contributions of this part II paper pertain to both coverage optimization problems and Learning in Games. Compared with [12][13], this paper employs a more accurate model of visual sensor and the results can be extended to deal with non-convex environments and collision avoidance between agents. Contrary to [12], we do not consider energy expenditure from sensor motions. Despite their distributed and adaptive properties, the coordination algorithms proposed in [12][13] and related ones are gradient-based. Thus, the corresponding emerging multi-agent behavior is easily trapped in a local maximum of certain coverage performance metric. In this paper, we develop an asynchronous distributed learning algorithm which enables sensors to asymptotically reach in probability to the set of global optima of certain coverage performance metric.

Regarding Learning in Games, we extend the use of the payoff-based learning dynamics in [16][17]. In our problem, each agent is unable to access the utility values induced by

This work was supported in part by NSF Career Award CMS-0643673 and NSF IIS-0712746. The authors are with Department of Mechanical and Aerospace Engineering, University of California, San Diego, 9500 Gilman Dr, La Jolla CA, 92093, {mizhu, soniamd}@ucsd.edu

alternative actions because motion and sensing capacities of each agent are limited and the rewards are not priori information to each agent. Furthermore, we aim to optimize the sum of all local utility function which captures the trade-off between the overall network benefits from sensing and the total energy the network consumes. To tackle these two challenges, we develop an asynchronous distributed learning algorithm concisely described as follows: At each time step, only one sensor is active and updates its state by either trying some new action or selecting an action according to a Gibbs-like distribution from those played in last two time steps when it was active. The algorithm is shown to be convergent in probability to the set of global maxima of our coverage performance metric. Compared with the payoff-based log-linear learning algorithm in [16], our algorithm optimizes a different global function, and has stronger convergence properties; see also Remark 4.1.

## II. PROBLEM FORMULATION AND LEARNING ALGORITHM

Here, we first review some basic game-theoretic concepts; see, for example [7]. This will allow us to formulate subsequently an optimal coverage problem for mobile visual sensor networks as a repeated multi-player game. We then present an algorithm to solve the coverage game, and introduce notation used throughout the paper.

### A. Background in Game Theory

A strategic game  $\Gamma := \langle V, A, U \rangle$  has three components:

1. A set  $V$  enumerating players  $i \in V := \{1, \dots, N\}$ .
2. An action set  $A := \prod_{i=1}^N A_i$  is the space of all actions vectors, where  $s_i \in A_i$  is the action of player  $i$  and an (multi-player) action  $s \in A$  has components  $s_1, \dots, s_N$ .
3. The collection of utility functions  $U$ , where the utility function  $u_i : A \rightarrow \mathbb{R}$  models player  $i$ 's preferences over action profiles.

Denote by  $s_{-i}$  the action profile of all players other than  $i$ , and by  $A_{-i} = \prod_{j \neq i} A_j$  the set of action profiles for all players except  $i$ . In conventional Non-cooperative Game Theory, all the actions in  $A_i$  always can be selected by player  $i$  in response to other players' actions. However, in the context of motion coordination, the actions available to player  $i$  will often be restricted to a state-dependent subset of  $A_i$ . In particular, we denote by  $F_i(s_i, s_{-i}) \subseteq A_i$  the set of feasible actions of player  $i$  when the action profile is  $s := (s_i, s_{-i})$ . We assume that  $F_i(s_i, s_{-i}) \neq \emptyset$ . Denote  $F(s) := \prod_{i \in V} F_i(s) \subseteq A$ ,  $\forall s \in A$  and  $F := \cup \{F(s) \mid s \in A\}$ . The introduction of  $F$  leads naturally to the notion of restricted strategic game  $\Gamma_{\text{res}} := \langle V, A, U, F \rangle$ .

### B. Coverage problem formulation

1) *Mission space*: We consider a convex 2-D mission space that is discretized into a (squared) lattice. We assume that each square of the lattice has unit dimensions. Each square will be labeled with the coordinate of its center  $q = (q_x, q_y)$ , where  $q_x \in [q_{x_{\min}}, q_{x_{\max}}]$  and  $q_y \in [q_{y_{\min}}, q_{y_{\max}}]$ , for some integers  $q_{x_{\min}}, q_{y_{\min}}, q_{x_{\max}}, q_{y_{\max}}$ . Denote by  $\mathcal{Q}$  the collection of all squares in the lattice.

We now define an associated location graph  $\mathcal{G}_{\text{loc}} := (\mathcal{Q}, E_{\text{loc}})$  where  $((q_x, q_y), (q_{x'}, q_{y'})) \in E_{\text{loc}}$  if and only if  $|q_x - q_{x'}| + |q_y - q_{y'}| = 1$  for  $(q_x, q_y), (q_{x'}, q_{y'}) \in \mathcal{Q}$ . Note that the graph  $\mathcal{G}_{\text{loc}}$  is undirected; i.e.,  $(q, q') \in E_{\text{loc}}$  if and only if  $(q', q) \in E_{\text{loc}}$ . The set of neighbors of  $q$  in  $E_{\text{loc}}$  is given by  $\mathcal{N}_q^{\text{loc}} := \{q' \in \mathcal{Q} \setminus \{q\} \mid (q, q') \in E_{\text{loc}}\}$ . We assume that the location graph  $\mathcal{G}_{\text{loc}}$  is fixed and connected, and denote its diameter by  $D$ .

Agents are deployed in  $\mathcal{Q}$  to detect certain events of interest. As agents move in  $\mathcal{Q}$  and process measurements, they will assign a numerical value  $W_q \geq 0$  to the events in each square (with center)  $q \in \mathcal{Q}$ . If  $W_q = 0$ , then there is no event of interest at the square  $q$ . The larger the value of  $W_q$  is, the more interest the set of events at the square  $q$  will have. Later, the amount  $W_q$  will be identified with a benefit of observing the point  $q$ . In this set-up, we assume the values  $W_q$  to be constant in time.

2) *Modeling of the visual sensor nodes*: Each mobile agent  $i$  is modeled as a point mass in  $\mathcal{Q}$ , with location  $a_i := (x_i, y_i) \in \mathcal{Q}$ . Each agent has mounted a pan-tilt-zoom camera, and can adjust its orientation and focal length.

The visual sensing range of a camera is directional, limited-range, and has a finite angle of view. Following a geometric simplification, we model the visual sensing region of agent  $i$  as an annulus sector in the 2-D plane; see Figure 1. The visual sensor footprint is completely characterized by

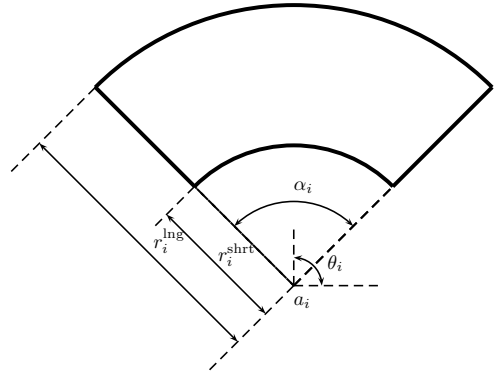


Fig. 1. Visual sensor footprint

the following parameters: the position of agent  $i$ ,  $a_i \in \mathcal{Q}$ , the camera orientation,  $\theta_i \in [0, 2\pi)$ , the camera angle of view,  $\alpha_i \in [\alpha_{\min}, \alpha_{\max}]$ , and the shortest range (resp. longest range) between agent  $i$  and the nearest (resp. farthest) object that can be recognized from the image,  $r_i^{\text{shrt}} \in [r_{\min}, r_{\max}]$  (resp.  $r_i^{\text{lng}} \in [r_{\min}, r_{\max}]$ ). The parameters  $r_i^{\text{shrt}}, r_i^{\text{lng}}, \alpha_i$  can be tuned by changing the focal length  $\text{FL}_i$  of agent  $i$ 's camera. In this way,  $c_i := (\text{FL}_i, \theta_i) \in [0, \text{FL}_{\max}] \times [0, 2\pi)$  is the camera control vector of agent  $i$ . In what follows, we will assume that  $c_i$  takes values in a finite subset  $\mathcal{C} \subset [0, \text{FL}_{\max}] \times [0, 2\pi)$ . An agent action is thus a vector  $s_i := (a_i, c_i) \in \mathcal{A}_i := \mathcal{Q} \times \mathcal{C}$ , and a multi-agent action is denoted by  $s = (s_1, \dots, s_N) \in \mathcal{A} := \prod_{i=1}^N \mathcal{A}_i$ .

Let  $\mathcal{D}(a_i, c_i)$  be the visual sensor footprint of agent  $i$ . Now we can define a proximity sensing graph [3]  $\mathcal{G}_{\text{sen}}(s) :=$

$(V, E_{\text{sen}}(s))$  as follows: the set of neighbors of agent  $i$ ,  $\mathcal{N}_i^{\text{sen}}(s)$ , is given as  $\mathcal{N}_i^{\text{sen}}(s) := \{j \in V \setminus \{i\} \mid \mathcal{D}(a_i, c_i) \cap \mathcal{D}(a_j, c_j) \cap \mathcal{Q} \neq \emptyset\}$ .

Each agent is able to communicate with others to exchange information. We assume that the communication range of agents is  $2r_{\text{max}}$ . This induces a  $2r_{\text{max}}$ -disk communication graph  $\mathcal{G}_{\text{comm}}(s) := (V, E_{\text{comm}}(s))$  as follows: the set of neighbors of agent  $i$  is given by  $\mathcal{N}_i^{\text{comm}}(s) := \{j \in V \setminus \{i\} \mid (x_i - x_j)^2 + (y_i - y_j)^2 \leq (2r_{\text{max}})^2\}$ . Note that  $\mathcal{G}_{\text{comm}}(s)$  is undirected and that  $\mathcal{G}_{\text{sen}}(s) \subseteq \mathcal{G}_{\text{comm}}(s)$ .

The motion of agents will be limited to a neighboring point in  $\mathcal{G}_{\text{loc}}$  at each time step. Thus, an agent feasible action set will be given by  $\mathcal{F}(a_i) := (\{a_i\} \cup \mathcal{N}_{a_i}^{\text{loc}}) \times \mathcal{C}$ .

3) *Coverage game*: We now proceed to formulate a coverage optimization problem as a restricted strategic game. For each  $q \in \mathcal{Q}$ , we denote  $n_q(s)$  as the cardinality of the set  $\{k \in V \mid q \in \mathcal{D}(a_k, c_k) \cap \mathcal{Q}\}$ ; i.e., the number of agents which can observe the point  $q$ . The ‘‘profit’’ given by  $W_q$  will be equally shared by agents that can observe the point  $q$ . The benefit that agent  $i$  obtains through sensing is thus defined by  $\sum_{q \in \mathcal{D}(a_i, c_i) \cap \mathcal{Q}} \frac{W_q}{n_q(s)}$ . In our set-up, we assume that  $W_q$  is unknown to each agent  $i$  unless agent  $i$  senses  $q$ .

On the other hand, and as argued in [18], the processing of visual data can incur a higher cost than that of communication. This is in contrast with scalar sensor networks, where the communication cost dominates. With this observation, we model the energy consumption of agent  $i$  by  $f_i(c_i) := \frac{1}{2}\alpha_i((r_i^{\text{lng}})^2 - (r_i^{\text{shrt}})^2)$ . This measure corresponds to the area of the visual sensor footprint and can serve to approximate the energy consumption or the cost incurred by image processing algorithms.

We will endow each agent with a utility function that aims to capture the above sensing/processing trade-off. In this way, we define a utility function for agent  $i$  by

$$u_i(s) = \sum_{q \in \mathcal{D}(a_i, c_i) \cap \mathcal{Q}} \frac{W_q}{n_q(s)} - f_i(c_i).$$

Note that the utility function  $u_i$  is distributed over the visual sensing graph  $\mathcal{G}_{\text{sen}}(s)$ ; i.e.,  $u_i$  is only dependent on the points  $q$  within its sensing range  $\mathcal{D}(a_i, c_i)$  and the actions of  $\{i\} \cup \mathcal{N}_i^{\text{sen}}(s)$ . With the set of utility functions  $U_{\text{cov}} = \{u_i\}_{i \in V}$ , and feasible action set  $\mathcal{F}_{\text{cov}} = \prod_{i=1}^N \bigcup_{a_i \in \mathcal{A}_i} \mathcal{F}_i(a_i)$ , we now have all the ingredients to introduce the coverage game  $\Gamma_{\text{cov}} := \langle V, \mathcal{A}, U_{\text{cov}}, \mathcal{F}_{\text{cov}} \rangle$ . This game is a variation of the congestion games introduced in [21]. In our companion paper [28], it is shown that the coverage game  $\Gamma_{\text{cov}}$  is a restricted potential game with potential function  $\phi(s) := \sum_{q \in \mathcal{Q}} \sum_{\ell=1}^{n_q(s)} \frac{W_q}{\ell} - \sum_{i=1}^N f_i(c_i)$ . However, the potential function  $\phi(s)$  is not a straightforward measure of the network performance. On the other hand, the objective function  $U_g(s) := \sum_{i \in V} u_i(s)$  captures the trade-off between the overall network benefit from sensing and the total energy the network consumes. In what follows,  $U_g(s)$  is perceived as the coverage performance metric. Finally, we let  $S^*$  denote  $S^* := \{s \mid \arg\max_{s \in \mathcal{A}} U_g(s)\}$ .

**Remark 2.1:** The assumptions of our problem formulation admit several extensions. For example, it is straightfor-

ward to extend our results to non-convex 3-D spaces. This is because the results that follow can also handle other shapes of the sensor footprint; e.g., a complete disk, a subset of the annulus sector. In addition, collision avoidance between robots can also be guaranteed. To do this, it is enough to remove from the feasible action set the neighboring locations where other agents are located. Furthermore, the coverage problem can be interpreted as a target assignment problem—here, the value  $W_q \geq 0$  would be associated with the value of a target located at the point  $q$ .

### C. Inhomogeneous asynchronous learning algorithm

The agents aim at maximizing the coverage performance metric  $U_g(s)$ . In our problem, motion and sensing capacities of each agent are limited and  $W_q$  is not priori information to each agent. This leads to the fact that each agent is unable to access to the utility values induced by alternative actions. To tackle this challenge, we present a distributed learning algorithm, say the Inhomogeneous Asynchronous Learning (IAL) Algorithm, which only requires each sensor to remember utility values obtained by its neighbors and itself, and actions it played during the last two time steps when it was active.

We next introduce some notations to present the IAL Algorithm. Denote by  $\mathcal{B}$  the space  $\mathcal{B} := \{(s, s') \in \mathcal{A} \times \mathcal{A} \mid s_{-i} = s'_{-i}, s'_i \in \mathcal{F}_i(a_i) \text{ for some } i \in V\}$ . For any  $s^0, s^1 \in \mathcal{A}$  with  $s_{-i}^0 = s_{-i}^1$  for some  $i \in V$ , we denote

$$\Delta_i(s^1, s^0) := \frac{1}{2} \sum_{q \in \Omega_1} \frac{W_q}{n_q(s^1)} - \frac{1}{2} \sum_{q \in \Omega_2} \frac{W_q}{n_q(s^0)},$$

where  $\Omega_1 := \mathcal{D}(a_i^1, c_i^1) \setminus \mathcal{D}(a_i^0, c_i^0) \cap \mathcal{Q}$  and  $\Omega_2 := \mathcal{D}(a_i^0, c_i^0) \setminus \mathcal{D}(a_i^1, c_i^1) \cap \mathcal{Q}$ , and

$$\begin{aligned} \rho_i(s^0, s^1) &:= (u_i(s^1) - \Delta_i(s^1, s^0)) - (u_i(s^0) - \Delta_i(s^0, s^1)) \\ \Psi_i(s^0, s^1) &:= \max\{u_i(s^0) - \Delta_i(s^0, s^1), u_i(s^1) - \Delta_i(s^1, s^0)\}, \\ m^* &:= \max_{(s^0, s^1) \in \mathcal{B}, s_i^0 \neq s_i^1} \{\Psi_i(s^0, s^1) - (u_i(s^0) - \Delta_i(s^0, s^1)), \frac{1}{2}\}. \end{aligned}$$

It is easy to check that  $\Delta_i(s^1, s^0) = -\Delta_i(s^0, s^1)$  and  $\Psi_i(s^0, s^1) = \Psi_i(s^1, s^0)$ . Assume that at each time instant, one of agents becomes active with equal probability. Denote by  $\gamma_i(t)$  the last time instant before  $t$  when agent  $i$  was active. We then denote  $\gamma_i^{(2)}(t) := \gamma_i \circ \gamma_i(t)$ . The main steps of the IAL Algorithm are described in the following.

**Initialization** At  $t = 0$ , all agents are uniformly placed in  $\mathcal{Q}$ .

Each agent  $i$  uniformly chooses the camera control vector  $c_i$  from the set  $\mathcal{C}$ , and then communicates with agents in  $\mathcal{N}_i^{\text{sen}}(s(0))$  and computes  $u_i(s(0))$ . Furthermore, each agent  $i$  chooses  $m_i \in (2m^*, Km^*]$  for some  $K \geq 2$ . At  $t = 1$ , all the sensors keep their actions.

**Update** Assume that agent  $i$  is active at time  $t \geq 2$ . Then agent  $i$  updates its state according to the following rules:

- Agent  $i$  chooses the exploration rate  $\epsilon(t) = t^{-\frac{1}{(D+1)(K+1)m^*}}$ .
- With probability  $\epsilon(t)^{m_i}$ , agent  $i$  experiments and uniformly chooses  $s_i^{\text{tp}} := (a_i^{\text{tp}}, c_i^{\text{tp}})$  from the action set  $\mathcal{F}(a_i(t)) \setminus \{s_i(t), s_i(\gamma_i^{(2)}(t) + 1)\}$ .

- With probability  $1 - \epsilon(t)^{m_i}$ , agent  $i$  does not experiment and chooses  $s_i^{\text{tp}}$  according to the following probability distribution:

$$\mathbb{P}(s_i^{\text{tp}} = s_i(t)) = \frac{1}{1 + \epsilon(t)^{\rho_i(s_i(\gamma_i^{(2)}(t)+1), s_i(t))}},$$

$$\mathbb{P}(s_i^{\text{tp}} = s_i(\gamma_i^{(2)}(t) + 1)) = \frac{\epsilon(t)^{\rho_i(s_i(\gamma_i^{(2)}(t)+1), s_i(t))}}{1 + \epsilon(t)^{\rho_i(s_i(\gamma_i^{(2)}(t)+1), s_i(t))}}.$$

- After  $s_i^{\text{tp}}$  is chosen, agent  $i$  moves to the position  $a_i^{\text{tp}}$  and sets its camera control vector to be  $c_i^{\text{tp}}$ .

**Communication and computation** At position  $a_i^{\text{tp}}$ , the active agent  $i$  communicates with agents in  $\mathcal{N}_i^{\text{sen}}(s_i^{\text{tp}}, s_{-i}(t))$ , and computes  $u_i(s_i^{\text{tp}}, s_{-i}(t))$ ,  $\Delta_i((s_i^{\text{tp}}, s_{-i}(t)), s(\gamma_i(t) + 1))$ ,  $\mathcal{F}(a_i^{\text{tp}})$ .

4: Repeat Step 2 and 3.

**Remark 2.2:** A variation of the previous algorithm corresponds to  $\epsilon(t) = \epsilon \in (0, \frac{1}{2}]$  constant for all  $t \geq 2$ . If this is the case, we will refer to the algorithm as the Homogeneous Asynchronous Learning (HAL, for short) Algorithm. Later, the convergence analysis of the IAL will be based on the analysis of the HAL.

#### D. Notations

The notation  $O(\epsilon^k)$  for some  $k \geq 0$  implies that  $0 < \lim_{\epsilon \rightarrow 0^+} \frac{O(\epsilon^k)}{\epsilon^k} < \infty$ . We denote by  $\text{diag}(\mathcal{A}) := \{(s, s) \in \mathcal{A}^2 \mid s \in \mathcal{A}\}$  and  $\text{diag}(S^*) := \{(s, s) \in \mathcal{A}^2 \mid s \in S^*\}$ .

Consider  $a, a' \in \mathcal{Q}$  where  $a_i \neq a'_i$  and  $a_{-i} = a'_{-i}$  for some  $i \in V$ . The transition  $a \rightarrow a'$  is feasible if and only if  $(a_i, a'_i) \in E_{\text{loc}}$ . If there is a feasible path, consisting of multiple feasible transitions, from  $a$  to  $a'$ , then we denote  $a \Rightarrow a'$ . We denote the reachable set from the state  $a$  by  $\diamond a := \{a' \in \mathcal{Q} \mid a \Rightarrow a'\}$ .

Consider  $s := (a, c), s' := (a', c') \in \mathcal{A}$  where  $a_i \neq a'_i$  and  $a_{-i} = a'_{-i}$  for some  $i \in V$ . The transition  $s \rightarrow s'$  is feasible if and only if  $s'_i \in \mathcal{F}(a)$ . If there is a feasible path, consisting of multiple feasible transitions, from  $s$  to  $s'$ , then we denote  $s \Rightarrow s'$ . We denote the reachable set from the state  $s$  by  $\diamond s := \{s' \in \mathcal{A} \mid s \Rightarrow s'\}$ .

### III. PRELIMINARIES TO CONVERGENCE ANALYSIS

For the sake of completeness, we include here some background in the Theory of Resistance Trees [27]. This section also includes a sufficient condition on the convergence of a class of time-inhomogeneous Markov chains that will be used in the general algorithm proof later.

#### A. Background in the Theory of Resistance Trees

Let  $P^0$  be the transition matrix of the time-homogeneous Markov chain  $\{\mathcal{P}_t^0\}$  on a finite state space  $X$ . And let  $P^\epsilon$  be the transition matrix of a *perturbed Markov chain*, say  $\{\mathcal{P}_t^\epsilon\}$ . With probability  $1 - \epsilon$ ,  $\{\mathcal{P}_t^\epsilon\}$  evolves according to  $P^0$ , while with probability  $\epsilon$ , the transitions do not follow  $P^0$ .

A family of stochastic processes  $\{\mathcal{P}_t^\epsilon\}$  is called a *regular perturbation* of  $\{\mathcal{P}_t^0\}$  if the following holds  $\forall x, y \in X$ :

(A1) For some  $\varsigma > 0$ , the Markov chain  $\{\mathcal{P}_t^\epsilon\}$  is irreducible and aperiodic for all  $\epsilon \in (0, \varsigma]$ .

(A2)  $\lim_{\epsilon \rightarrow 0^+} P_{xy}^\epsilon = P_{xy}^0$ .

(A3) If  $P_{xy}^\epsilon > 0$  for some  $\epsilon$ , then there exists a real number  $\chi(x \rightarrow y) \geq 0$  such that  $\lim_{\epsilon \rightarrow 0^+} P_{xy}^\epsilon / \epsilon^{\chi(x \rightarrow y)} \in (0, \infty)$ .

In (A3), the nonnegative real number  $\chi(x \rightarrow y)$  is called the *resistance* of the transition from  $x$  to  $y$ .

Let  $H_1, H_2, \dots, H_J$  be the recurrent communication classes of the Markov chain  $\{\mathcal{P}_t^0\}$ . Note that within each class  $H_\ell$ , there is a path of zero resistance from every state to every other. Given any two distinct recurrence classes  $H_\ell$  and  $H_k$ , consider all paths which start from  $H_\ell$  and end at  $H_k$ . Denote  $\chi_{\ell k}$  by the least resistance among all such paths.

Now define a complete directed graph  $\mathcal{G}$  where there is one vertex  $\ell$  for each recurrent class  $H_\ell$ , and the resistance on the edge  $(\ell, k)$  is  $\chi_{\ell k}$ . An  $\ell$ -*tree* on  $\mathcal{G}$  is a spanning tree such that from every vertex  $k \neq \ell$ , there is a unique path from  $k$  to  $\ell$ . Denote by  $G(\ell)$  the set of all  $\ell$ -trees on  $\mathcal{G}$ . The resistance of an  $\ell$ -tree is the sum of the resistances of its edges. The *stochastic potential* of the recurrent class  $H_\ell$  is the least resistance among all  $\ell$ -trees in  $G(\ell)$ .

**Theorem 3.1 ([27]):** Let  $\{\mathcal{P}_t^\epsilon\}$  be a regular perturbation of  $\{\mathcal{P}_t^0\}$ , and for each  $\epsilon > 0$ , let  $\mu(\epsilon)$  be the unique stationary distribution of  $\{\mathcal{P}_t^\epsilon\}$ . Then  $\lim_{\epsilon \rightarrow 0^+} \mu(\epsilon)$  exists and the limiting distribution  $\mu(0)$  is a stationary distribution of  $\{\mathcal{P}_t^0\}$ . The stochastically stable states (i.e., the support of  $\mu(0)$ ) are precisely those states contained in the recurrence classes with minimum stochastic potential.

#### B. A class of time-inhomogeneous Markov chains

Here, we derive sufficient conditions for a class of time-inhomogeneous Markov chains to converge. The main references include [6] and [11].

Consider a time-inhomogeneous Markov chain  $\{\mathcal{P}_t\}$  on a finite state space  $X$  with transition matrix  $P^{\epsilon(t)}$  where  $\epsilon(t) \in (0, \varsigma]$  for some  $\varsigma > 0$ . Let  $P^\epsilon$  be the transition matrix if  $\epsilon(t)$  is a constant  $\epsilon \in (0, \varsigma]$  for all  $t \geq 1$ . Denote by  $\{\mathcal{P}_t^\epsilon\}$  the time-homogeneous Markov chain which evolves according to  $P^\epsilon$ .

**Proposition 3.1:** Assume that,  $\{\mathcal{P}_t^\epsilon\}$  is a regular perturbation of  $\{\mathcal{P}_t^0\}$ . The time-inhomogeneous Markov chain  $\{\mathcal{P}_t\}$  is strongly ergodic if the following conditions hold:

(C1) The Markov chain  $\{\mathcal{P}_t\}$  is weakly ergodic.

(C2)  $\epsilon(t) > 0$  and is strictly decreasing.

(C3) If  $P_{xy}^{\epsilon(t)} > 0$ , then  $P_{xy}^{\epsilon(t)} = \alpha_{xy}(\epsilon(t)) / \beta_{xy}(\epsilon(t))$  for some polynomials  $\alpha_{xy}(\epsilon(t))$  and  $\beta_{xy}(\epsilon(t))$  in  $\epsilon(t)$ .

*Proof:* We omit the proof due to space limits. ■

**Remark 3.1:** In Proposition 3.1, (C3) can be replaced by the following. (C3') If  $P_{xy}^{\epsilon(t)} > 0$ , then  $P_{xy}^{\epsilon(t)} = \alpha_{xy}(\epsilon(t)) / \beta_{xy}(\epsilon(t))$  where  $\alpha_{xy}(\epsilon(t))$  and  $\beta_{xy}(\epsilon(t))$  are smooth at the origin. Following along the same lines as in Proposition 3.1, one can complete the proof by using the Taylor expansions of  $\alpha_{xy}(\epsilon(t))$  and  $\beta_{xy}(\epsilon(t))$  at the origin.

### IV. CONVERGENCE ANALYSIS OF THE IAL ALGORITHM

In this section, we show the convergence of the IAL Algorithm to  $S^*$  by appealing to the results in Section III. To do this, we first analyze the HAL Algorithm next.

### A. Convergence analysis of the HAL Algorithm

The convergence property of the HAL Algorithm will be studied by using Proposition 3.1. To simplify notations, we denote  $s_i(t-1) := s_i(\gamma_i^{(2)}(t) + 1)$  in the remainder of this section. Observe that  $z(t) := (s(t-1), s(t))$  in the HAL Algorithm constitutes a Markov chain  $\{\mathcal{P}_t^\epsilon\}$  on the space  $\mathcal{B} := \{(s, s') \in \mathcal{A} \times \mathcal{A} \mid s'_i \in \mathcal{F}(a_i), \forall i \in V\}$ .

**Lemma 4.1:**  $\{\mathcal{P}_t^\epsilon\}$  is a regular perturbation of  $\{\mathcal{P}_t^0\}$ .

*Proof:* We omit the proof due to space limits. ■

A direct result of Lemma 4.1 is that for each  $\epsilon > 0$ , there exists a unique stationary distribution of  $\{\mathcal{P}_t^\epsilon\}$ , say  $\mu(\epsilon)$ . From the proof of Lemma 4.1, we can see that the resistance of an experiment is  $m_i$  if sensor  $i$  is the unilateral deviator. We now utilize Theorem 3.1 to characterize  $\lim_{\epsilon \rightarrow 0^+} \mu(\epsilon)$ .

**Proposition 4.1:** Consider the regular perturbed Markov process  $\{\mathcal{P}_t^\epsilon\}$ . Then  $\lim_{\epsilon \rightarrow 0^+} \mu(\epsilon)$  exists and the limiting distribution  $\mu(0)$  is a stationary distribution of  $\{\mathcal{P}_t^0\}$ . Furthermore, the stochastically stable states (i.e., the support of  $\mu(0)$ ) are contained in the set  $\text{diag}(S^*)$ .

*Proof:* The unperturbed Markov chain corresponds to the HAL Algorithm with  $\epsilon = 0$ . Hence, the recurrent communication classes of the unperturbed Markov chain are contained in the set  $\text{diag}(\mathcal{A})$ . We will construct resistance trees over vertices in the set  $\text{diag}(\mathcal{A})$ . Denote  $T_{\min}$  by the minimum resistance tree. The remainder of the proof is divided into the following four claims. Due to the space limit, we omit the details here.

**Claim 1:**  $\chi((s^0, s^0) \Rightarrow (s^1, s^1)) = m_i + \Psi_i(s^1, s^0) - (u_i(s^1) - \Delta_i(s^1, s^0))$  where  $s^0 \neq s^1$  and the transition  $s^0 \rightarrow s^1$  is feasible with sensor  $i$  as the unilateral deviator.

**Claim 2:** All the edges  $((s, s), (s', s'))$  in  $T_{\min}$  must consist of only one deviator; i.e.,  $s_i \neq s'_i$  and  $s_{-i} = s'_{-i}$  for some  $i \in V$ .

**Claim 3:** Given any edge  $((s, s), (s', s'))$  in  $T_{\min}$ , denote by  $i$  the unilateral deviator between  $s$  and  $s'$ . Then the transition  $s_i \rightarrow s'_i$  is feasible.

**Claim 4:** Let  $h_v$  be the root of  $T_{\min}$ . Then,  $h_v \in \text{diag}(S^*)$ .

**Proof of Proposition 4.1:** It follows from Claim 4 that the state  $h_v \in \text{diag}(S^*)$  has minimum stochastic potential. Then Proposition 4.1 is a direct result of Theorem 3.1. ■

### B. Convergence analysis of the IAL Algorithm

We are now ready to show the convergence in probability of the ISL Algorithm by combining Proposition 4.1 and Proposition 3.1.

**Theorem 4.1:** Consider the Markov chain  $\{\mathcal{P}_t\}$  induced by the ISL Algorithm for the game  $\Gamma_{\text{cov}}$ . Then it holds that  $\lim_{t \rightarrow \infty} \mathbb{P}(z(t) \in \text{diag}(S^*)) = 1$ .

*Proof:* Denote by  $P^{\epsilon(t)}$  the transition matrix of  $\{\mathcal{P}_t\}$ . It is obvious that (C2) in Proposition 3.1 holds. Consider the feasible transition  $z^1 \rightarrow z^2$  with unilateral deviator  $i$ . The corresponding probability is given by

$$P_{z^1 z^2}^{\epsilon(t)} = \begin{cases} \eta_1, & s_i^2 \in \mathcal{F}(a_i^1) \setminus \{s_i^0, s_i^1\}, \\ \eta_2, & s_i^2 = s_i^1, \\ \eta_3, & s_i^2 = s_i^0, \end{cases}$$

where

$$\eta_1 := \frac{\epsilon(t)^{m_i}}{N|\mathcal{F}(a_i^1) \setminus \{s_i^0, s_i^1\}|}, \quad \eta_2 := \frac{1 - \epsilon(t)^{m_i}}{N(1 + \epsilon(t)^{\rho_i(s^0, s^1)})},$$

$$\eta_3 := \frac{(1 - \epsilon(t)^{m_i}) \times \epsilon(t)^{\rho_i(s^0, s^1)}}{N(1 + \epsilon(t)^{\rho_i(s^0, s^1)})}.$$

It is clear that (C3) in Proposition 3.1 holds. We now proceed to verify (C1) in Proposition 3.1 by using Theorem V.3.2 in [11]. Observe that  $|\mathcal{F}(a_i^1)| \leq 5|\mathcal{C}|$ . Since  $\epsilon(t)$  is strictly decreasing, there is  $t_0 \geq 1$  such that  $t_0$  is the first time when  $1 - \epsilon(t)^{m_i} \geq \epsilon(t)^{m_i}$ .

Observe that for all  $t \geq 1$ , it holds that

$$\eta_1 \geq \frac{\epsilon(t)^{m_i}}{N(5|\mathcal{C}| - 1)} \geq \frac{\epsilon(t)^{m_i + m^*}}{N(5|\mathcal{C}| - 1)}.$$

Denote  $b := u_i(s^1) - \Delta_i(s^1, s^0)$  and  $a := u_i(s^0) - \Delta_i(s^0, s^1)$ . Then  $\rho_i(s^0, s^1) = b - a$ . Since  $b - a \leq m^*$ , then for  $t \geq t_0$  it holds that

$$\eta_2 = \frac{1 - \epsilon(t)^{m_i}}{N(1 + \epsilon(t)^{b-a})} = \frac{(1 - \epsilon(t)^{m_i})\epsilon(t)^{\max\{a, b\} - b}}{N(\epsilon(t)^{\max\{a, b\} - b} + \epsilon(t)^{\max\{a, b\} - a})}$$

$$\geq \frac{\epsilon(t)^{m_i} \epsilon(t)^{\max\{a, b\} - b}}{2N} \geq \frac{\epsilon(t)^{m_i + m^*}}{N(5|\mathcal{C}| - 1)}.$$

Similarly, for  $t \geq t_0$ , it holds that

$$\eta_3 = \frac{(1 - \epsilon(t)^{m_i})\epsilon(t)^{\max\{a, b\} - a}}{N(\epsilon(t)^{\max\{a, b\} - b} + \epsilon(t)^{\max\{a, b\} - a})} \geq \frac{\epsilon(t)^{m_i + m^*}}{N(5|\mathcal{C}| - 1)}.$$

Since  $m_i \in (2m^*, Km^*]$ , for all  $i \in V$  and  $Km^* > 1$ , then for any feasible transition  $z^1 \rightarrow z^2$  with  $z^1 \neq z^2$ , it holds

$$P_{z^1 z^2}^{\epsilon(t)} \geq \frac{\epsilon(t)^{(K+1)m^*}}{N(5|\mathcal{C}| - 1)}$$

for all  $t \geq t_0$ . Furthermore, for all  $t \geq t_0$  and all  $z^1 \in \text{diag}(\mathcal{A})$ , we have that:

$$P_{z^1 z^1}^{\epsilon(t)} = 1 - \frac{1}{N} \sum_{i=1}^N \epsilon(t)^{m_i} = \frac{1}{N} \sum_{i=1}^N (1 - \epsilon(t)^{m_i})$$

$$\geq \frac{1}{N} \sum_{i=1}^N \epsilon(t)^{m_i} \geq \frac{\epsilon(t)^{(K+1)m^*}}{N(5|\mathcal{C}| - 1)}.$$

Let  $P(m, m)$  be the identity matrix, and  $P(m, n) := \prod_{t=m}^{n-1} P^{\epsilon(t)}$ ,  $0 \leq m < n$ . Pick  $z \in \mathcal{B}$  and let  $u_z \in \mathcal{B}$  be such that  $P_{u_z z}(t, t + D + 1) = \min_{x \in \mathcal{B}} P_{xz}(t, t + D + 1)$ . Consequently, it follows that for all  $t \geq t_0$ ,

$$\min_{x \in \mathcal{B}} P_{xz}(t, t + D + 1)$$

$$= \sum_{i_1 \in \mathcal{B}} \dots \sum_{i_D \in \mathcal{B}} P_{u_z i_1}^{\epsilon(t)} \dots P_{i_{D-1} i_D}^{\epsilon(t+D-1)} P_{i_D z}^{\epsilon(t+D)}$$

$$\geq P_{u_z i_1}^{\epsilon(t)} \dots P_{i_{D-1} i_D}^{\epsilon(t+D-1)} P_{i_D z}^{\epsilon(t+D)}$$

$$\geq \left( \frac{\epsilon(t)}{N(5|\mathcal{C}| - 1)} \right)^{(D+1)(K+1)m^*}.$$

Hence, we obtain

$$\begin{aligned}
& 1 - \lambda(P(t, t + D + 1)) \\
&= \min_{x, y \in \mathcal{B}} \sum_{z \in \mathcal{B}} \min\{P_{xz}(t, t + D + 1), P_{yz}(t, t + D + 1)\} \\
&\geq \sum_{z \in \mathcal{B}} \min_{x \in \mathcal{B}} P_{xz}(t, t + D + 1) \\
&\geq \sum_{z \in \mathcal{B}} P_{uz}(t, t + D + 1) \\
&\geq |\mathcal{B}| \left( \frac{\epsilon(t)}{N(5|\mathcal{C}| - 1)} \right)^{(D+1)(K+1)m^*}.
\end{aligned}$$

Choose  $k_\ell := (D + 1)\ell$  and let  $\ell_0$  be the smallest integer such that  $(D + 1)\ell_0 \geq t_0$ . Then it holds that

$$\begin{aligned}
& \sum_{\ell=0}^{\infty} (1 - \lambda(P(k_\ell, k_{\ell+1}))) \geq \sum_{\ell=\ell_0}^{\infty} (1 - \lambda(P(k_\ell, k_{\ell+1}))) \\
&\geq \sum_{\ell=\ell_0}^{\infty} |\mathcal{B}| \left( \frac{\epsilon((D + 1)\ell)}{N(5|\mathcal{C}| - 1)} \right)^{(D+1)(K+1)m^*} \\
&= \frac{|\mathcal{B}|}{(N(5|\mathcal{C}| - 1))^{(D+1)(K+1)m^*}} \sum_{\ell=\ell_0}^{\infty} \frac{1}{(D + 1)^\ell} = \infty.
\end{aligned}$$

Hence, the weak ergodicity of  $\{\mathcal{P}_t\}$  follows from Theorem V.3.2 in [11]. The strong ergodicity of  $\{\mathcal{P}_t\}$  follows directly from Proposition 3.1. It follows from Theorem V.4.3 in [11] that the limiting distribution is  $\mu^* = \lim_{t \rightarrow \infty} \mu^t$ . Note that  $\lim_{t \rightarrow \infty} \mu^t = \lim_{t \rightarrow \infty} \mu(\epsilon(t)) = \mu(0)$  and Proposition 4.1 shows that the support of  $\mu(0)$  is contained in the set  $\text{diag}(S^*)$ . Hence, the support of  $\mu^*$  is in  $\text{diag}(S^*)$ . ■

**Remark 4.1:** Compared with the payoff-based learning algorithms in [16], the IAL algorithm optimizes the sum of all local utility functions instead of the potential function [16]. Furthermore, the algorithms in [16] converge to the set of global optima of the potential function with sufficiently large probability by choosing a sufficiently small exploration rate in advance, and the induced evolution is a time-homogeneous Markov chain. In contrast, our IAL algorithm employs a diminishing exploration rate. This leads to the evolution of the IAL algorithm being a time-inhomogeneous Markov chain and a stronger convergence property of reaching the set  $S^*$  in probability.

## V. CONCLUSION

We have formulated a coverage optimization problem as a multi-player game. An asynchronous distributed learning algorithm has been proposed for this coverage game and shown to asymptotically converge to the set of global optima of the coverage performance metric in probability.

## REFERENCES

[1] I.F. Akyildiz, T. Melodia, and K. Chowdhury. Wireless multimedia sensor networks: a survey. *IEEE Wireless Communications Magazine*, 14(6):32–39, 2007.

[2] T. Alpcan, T. Basar, and S. Dey. A power control game based on outage probabilities for multicell wireless data networks. *IEEE Transactions on Wireless Communications*, 5(4):890–899, 2006.

[3] F. Bullo, J. Cortés, and S. Martínez. *Distributed Control of Robotic Networks*. Applied Mathematics Series. Princeton University Press, 2009. Available at <http://www.coordinationbook.info>.

[4] K.Y. Chow, K.S. Lui, and E.Y. Lam. Maximizing angle coverage in visual sensor networks. In *Proc. of IEEE International Conference on Communications*, pages 3516–3521, June 2007.

[5] J. Cortés, S. Martínez, and F. Bullo. Spatially-distributed coverage optimization and control with limited-range interactions. *ESAIM: Control, Optimisation & Calculus of Variations*, 11:691–719, 2005.

[6] M. Freidlin and A. Wentzell. *Random perturbations of dynamical systems*. New York: Springer Verlag, 1984.

[7] D. Fudenberg and J. Tirole. *Game theory*. The MIT press, 1991.

[8] A. Ganguli, J. Cortés, and F. Bullo. Visibility-based multi-agent deployment in orthogonal environments. In *American Control Conference*, pages 3426–3431, New York, July 2007.

[9] A. Ganguli, J. Cortés, and F. Bullo. Multirobot rendezvous with visibility sensors in nonconvex environments. *IEEE Transactions on Robotics*, 25(2):340–352, 2009.

[10] E. Hörster and R. Lienhart. On the optimal placement of multiple visual sensors. In *Proc. of the third ACM international workshop on Video surveillance and sensor networks*, pages 111–120, 2006.

[11] D. Isaacson and R. Madsen. *Markov chains*. Wiley, 1976.

[12] A. Kwok and S. Martínez. Deployment algorithms for a power-constrained mobile sensor network. *International Journal of Robust and Nonlinear Control*, 2009. In press.

[13] K. Laventall and J. Cortés. Coverage control by multi-robot networks with limited-range anisotropic sensory. *International Journal of Control*, 2009. To appear.

[14] J. R. Marden, G. Arslan, and J. S. Shamma. Connections between cooperative control and potential games. *IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics*, 2008. submitted.

[15] J. R. Marden and A. Wierman. Distributed welfare games. *Operations Research*, 2008. submitted.

[16] J.R. Marden and J.S. Shamma. Revisiting log-linear learning : asynchrony, completeness and payoff-based implementation. *Games and Economic Behavior*, 2008. submitted.

[17] J.R. Marden, H.P. Young, G. Arslan, and J.S. Shamma. Payoff based dynamics for multi-player weakly acyclic games. *SIAM Journal on Control and Optimization*, 2006. submitted.

[18] C.B. Margi, V. Petkov, K. Obraczka, and R. Manduchi. Characterizing energy consumption in a visual sensor network testbed. In *Proc. of 2nd International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities*, pages 332–339, March 2006.

[19] R. A. Murphey. Target-based weapon target assignment problems. In P. M. Pardalos and L. S. Pitsoulis, editors, *Nonlinear Assignment Problems: Algorithms and Applications*, pages 39–53. Kluwer Academic Publishers, 1999.

[20] J. O’Rourke. *Art Gallery Theorems and Algorithms*. Oxford University Press, 1987.

[21] R.W. Rosenthal. A class of games possessing pure strategy nash equilibria. *International Journal of Game Theory*, 2:65–67, 1973.

[22] T. Roughgarden. *Selfish routing and the price of anarchy*. MIT press, 2005.

[23] T. C. Shermer. Recent results in art galleries. *Proceedings of the IEEE*, 80(9):1384–1399, 1992.

[24] S. Soro and W. B. Heinzelman. On the coverage problem in video based wireless sensor networks. In *Proc. of 2nd International Conference on Broadband Networks*, pages 932–939, 2005.

[25] A. Tang and L. Andrew. Game theory for heterogeneous flow control. In *Proc. of 42nd Annual Conference on Information Sciences and Systems*, pages 52–56, December 2008.

[26] J. Urrutia. Art gallery and illumination problems. In J. R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 973–1027. North-Holland, 2000.

[27] H.P. Young. The evolution of conventions. *Econometrica*, 61:57–84, January 1993.

[28] M. Zhu and S. Martínez. Distributed coverage games for mobile visual sensors (I): Reaching the set of Nash equilibria. In *Proc. of the 48th IEEE Conf. on Decision and Control and 28th Chinese Control Conference*, Shanghai, China, December 2009. To appear.