

# A Decentralized Ergodic Control Approach to the Multi-Agent Pursuit-Evasion Problem

Ian Abraham and Todd D. Murphey

**Abstract**—In this work, we present ergodic control as a method for solving the pursuit-evasion problem for nonlinear continuous systems. We show that adopting an ergodic control policy the pursuit-evasion problem obtains a Nash equilibrium. Ergodic control is then formulated for a decentralized network of robotic agents. It is shown that the decentralized ergodic control problem is equivalent to solving the centralized ergodic control problem when the network of agents has reached consensus over the past exploration of each agent. Examples are presented for the application of ergodic theory for exploring information-based distributions, localization of multiple targets, and estimation of spatial terrain. Each example emphasizes the ease with which ergodic control can be decentralized and the benefits of a decentralized network of agents actively exploring.

## I. INTRODUCTION

Within the field of autonomous mobile robots, active exploration has been a fundamental area of research. Specifically, exploration and mapping of unknown environments has led to the development of various SLAM-like algorithms [1]–[3]. Recent work in active exploration has utilized the concept of optimally sensing based on the particular sensory modalities available to the robot [4]–[8]. This desire to encode exploratory behavior in robotics has largely been inspired by biology [9], [10]. In particular, the need for unique exploratory behavior subject to constraints imposed by sensory modalities like visual and tactile has led to advances in the use of information theory for active exploration [4], [6], [11], [12]. A clear extension to much of the previously mentioned work is application in multi-agent robotic networks.

Within the field of active exploration distributed robot networks have been shown to improve the efficiency and span of exploration of mobile robots [13]–[15]. Moreover, the application of distributed robotic networks take into consideration the specifications of the task and effectively use the underlying network connection for improved performance [15]–[18]. Interestingly, many of the underlying centralized version of the active exploration methods are shown to be preserved with the use of network consensus. Existing work in applications of ergodic theory in active exploration (or ergodic exploration for short) has been formulated in a centralized multi-agent system for linear dynamics [19]. In addition, work has been done on synthesizing controllers for multi-robot systems using ergodic processes [20]. However, it has

yet to be shown how such a network would benefit from network consensus or whether it can be extended to nonlinear dynamical systems.

In this work, we contribute a formulation of an ergodic controller for active exploration. Specifically, we formulate the controller for nonlinear dynamical systems using sequential action control [21]. The formulation is then extended to a centralized connected robotic network. We then formulate a decentralized ergodic controller and show that under consensus, the objective of the decentralized ergodic controller is equivalent to the objective of the centralized ergodic controller. Numerical examples are presented that show the versatility in ergodic control when applied to exploration, localization, and mapping problems. In particular, these examples emphasize the ease with which ergodic control can be extended to variations of problems within distributed mobile robotics.

The paper is outlined as: Section II formulates the ergodic optimal control problem and then presents ergodic control in a centralized and decentralized manner. Section III discusses the network chosen in this work and its effect on ergodic control. Section IV extends the ergodic control formulation for problems in exploration, estimation and localization. Last, sections V and VI present simulation results and conclusion of the application of a connected decentralized ergodic controller.

## II. ERGODIC CONTROL

In this section, we define the variants of the ergodic control problem. We first start with a formulation of ergodic control for a nonlinear system through a similar process as done with sequential action control [21]. Following the derivation of ergodic control, we modify the specification of the ergodic controller for control of a centralized multi-agent system. We then show in the next section a derivation of a decentralized ergodic controller for a single agent in the multi-agent system and show that under consensus of the network the optimal control solution to the decentralized problem is equivalent to the optimal control solution of the centralized ergodic control problem.

### A. Canonical Ergodic Control

The ergodic exploration problem is defined for a control-affine dynamical systems of the form

$$\dot{x} = f(x, u) = g(x) + h(x)u. \quad (1)$$

Ergodic control is formulated as the control solution that minimizes the ergodic metric

$$\mathcal{E}(x(t)) = \sum_{k \in \mathbb{Z}} (\phi_k - c_k)^2 \quad (2)$$

Authors are with the Neuroscience and Robotics Laboratory (NxR) at the Department of Mechanical Engineering<sup>1</sup>, Northwestern University, 2145 Sheridan Road Evanston, IL 60208 USA.

Email: i-abr@u.northwestern.edu,  
t-murphey@northwestern.edu

where

$$\phi_k = \int_X \Phi(x) F_k(x) dx \quad (3)$$

and

$$c_k = \frac{1}{T} \int_{t_0}^{t_0+T} F_k(x(t)) dt \quad (4)$$

for some time  $t_0$  to  $t_0 + T$ . Here,  $\Phi(x)$  is the underlying spatial distribution (we will sometimes refer to this as the target distribution since it is being used as a reference for the control system) and  $F_k(x)$  is the basis function<sup>1</sup>. Agent actions  $x(t)$  are considered to be “ergodic” with respect to a target distribution  $\Phi(x)$  when the time averaged statistics of the robots actions are equivalent to the statistics of the target distribution.

We define the objective of the ergodic control problem as

$$\begin{aligned} & \underset{u}{\text{minimize}} \quad \mathcal{E}(x(t)) \\ & \text{subject to} \quad \dot{x}(t) = f(x(t), u(t)). \end{aligned}$$

Instead of minimizing directly over control  $u$ , we consider the first order sensitivity to control application time  $\lambda$ . This is found by taking the derivative of  $\mathcal{E}$  with respect to infinitesimal control application time  $\lambda$  defining the mode insertion gradient as

$$\begin{aligned} \frac{d\mathcal{E}}{d\lambda} &= \rho(s)^T (f_2(s, s) - f_1(s)) \\ \text{subject to} \quad \dot{\rho} &= - \sum_{k \in \mathcal{Z}} (\phi_k - c_k) \frac{\partial F_k(x)}{\partial x} - \frac{\partial f^T}{\partial x} \rho. \end{aligned} \quad (5)$$

In the definition of  $\frac{d\mathcal{E}}{d\lambda}$

$$f_1(t) = f(x(t), u_1(t))$$

and

$$f_2(t, \tau) = f(x(t), u_2(\tau))$$

are the systems’ state  $f$  subject to the default control  $u_1$  and optimal control  $u_2$ . Following the formulation of sequential action control (SAC) [21] we define an auxiliary objective function  $J_2$  as

$$J_2 = \frac{d\mathcal{E}}{d\lambda} + \frac{1}{2} \|u_2 - u_1\|_R^2$$

whose minimum occurs when the mode insertion gradient is most negative, regularized by the change in the default control  $u_1$  with  $R$  as a diagonal positive definite matrix. Since  $J_2$  is convex in  $u_2$ , taking a first-order optimality approach to solve for an optimal  $u_2$  gives the closed form solution

$$u^*(t) = -R^{-1} h(x)^T \rho(t) + u_1(t). \quad (6)$$

The time of control application  $\lambda$  that significantly reduces the ergodic cost is obtained via

$$\lambda^* = \underset{\lambda}{\text{argmin}} \quad \frac{d\mathcal{E}(\lambda)}{d\lambda}, \quad (7)$$

<sup>1</sup>In this work, the cosine basis function is used, however, any choice of basis function  $F_k$  can be used.

typically approximated using a line search. Thus, at each sampling instance  $t_{curr}$ , the ergodic controller forward predicts the system trajectory based on the sequence of control actions  $u_1$ . The sensitivity to that trajectory is calculated via the mode insertion gradient (5) and then used to create the closed form solution of control actions that reduce the ergodic metric. The time of application that significantly reduces the ergodic cost is then calculated and the controller returns  $u^*(\lambda)$  which is added to the schedule of control actions  $u_1$ .

### B. Centralized Ergodic Control

In a similar formulation as the canonical ergodic control problem, we define the centralized ergodic control problem for a network of robot agents  $x_i(t)$  subject to dynamical constraints of the form

$$\dot{x}_i(t) = f_i(x_i, u_i) = g_i(x_i) + h_i(x_i)u_i.$$

We can define the ergodic metric as before in a centralized manner as

$$\mathcal{E}(x(t)) = \sum_{k \in \mathcal{Z}} (\phi_k - c_k(x(t)))^2 \quad (8)$$

where

$$x = \begin{bmatrix} x_i \\ \vdots \\ x_N \end{bmatrix}$$

is the stacked centralized representation of all robot agent  $x_i(t) \forall i \in \{1, \dots, N\}$ . For a network of  $N$  agents, the time averaged spatial statistical representation of the collective robot network,  $c_k$ , is defined as

$$c_k = \frac{1}{T} \int_{t_0}^{t_0+T} \frac{1}{N} \sum_i F_k(x_i(t)) = \frac{1}{T} \int_{t_0}^{t_0+T} \tilde{F}_k(x(t)) \quad (9)$$

where  $\tilde{F}_k(x(t)) = \frac{1}{N} \sum_j F_k(x_j(t))$ . Here, the mode insertion gradient is given as

$$\frac{d\mathcal{E}}{d\lambda} = \rho(s)^T (f(x(s), u_2(s)) - f(x(s), u_1(s)))$$

where

$$\rho(t) = \begin{bmatrix} \rho_i \\ \vdots \\ \rho_N \end{bmatrix}, f = \begin{bmatrix} f_i \\ \vdots \\ f_N \end{bmatrix}.$$

The solution to the costate equation  $\rho$  is obtained via the backwards differential equation

$$\dot{\rho} = -q \sum_{k \in \mathcal{Z}} (\phi_k - c_k) \frac{\partial \tilde{F}_k(x)}{\partial x} - \frac{\partial f^T}{\partial x} \rho$$

which is the exactly the same as (5). Assuming that each robot’s state  $x_i(t)$  is independent of one another, we see that

$$\frac{\partial \tilde{F}_k(x)}{\partial x} = \begin{bmatrix} \frac{\partial F_k(x_i)}{\partial x_i} \\ \vdots \\ \frac{\partial F_k(x_N)}{\partial x_N} \end{bmatrix}$$

and

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f_i}{\partial x_i} & 0 & \dots & 0 \\ 0 & \frac{\partial f_{i+1}}{\partial x_{i+1}} & & \\ \vdots & & \ddots & \\ 0 & & & \frac{\partial f_N}{\partial x_N} \end{bmatrix}.$$

As done previously, we define a secondary cost function

$$J_2 = \frac{d\mathcal{E}}{d\lambda} + \frac{1}{2}\|u_2 - u_1\|_R^2. \quad (10)$$

where  $R$  is a positive-definite diagonal matrix with elements  $R_i$ . Taking the derivative of (10) with respect to control  $u_2$  and setting the resulting expression equal to zero gives

$$u^*(t) = -R^{-1}h(x)^T \rho(t) + u_1(t). \quad (11)$$

If we expand (11) we get

$$\begin{bmatrix} u_i^*(t) \\ \vdots \\ u_N^*(t) \end{bmatrix} = -R^{-1} \begin{bmatrix} h_i(x_i) & & \\ & \ddots & \\ & & h_N(x_N) \end{bmatrix}^T \begin{bmatrix} \rho_i(t) \\ \vdots \\ \rho_N(t) \end{bmatrix} + \begin{bmatrix} u_{1,i}(t) \\ \vdots \\ u_{1,N}(t) \end{bmatrix}.$$

Since  $h(x)$  is a block diagonal matrix, the transpose of a block diagonal matrix is the transpose of each element on the diagonal. With  $R$  as a positive definite diagonal matrix whose diagonal elements  $R_i$  are unique to each agent, we find the solution that minimizes (10) with respect to the control  $u_2$  is

$$\begin{bmatrix} u_i^*(t) \\ \vdots \\ u_N^*(t) \end{bmatrix} = \begin{bmatrix} -R_i^{-1}h_i(x_i)^T \rho_i(t) + u_{1,i}(t) \\ \vdots \\ -R_N^{-1}h_N(x_N)^T \rho_N(t) + u_{1,N}(t) \end{bmatrix} \quad (12)$$

which is the stacked representation of the control solution that reduces the ergodic metric for each individual agent. As done before, finding the time of control application that significantly reduces the ergodic metric for all agents is found via

$$\lambda^* = \operatorname{argmin}_{\lambda} \frac{d\mathcal{E}(\lambda)}{d\lambda}.$$

The rest of the controller follows the same formulation as the canonical ergodic control problem.

### C. Decentralized Ergodic Control

The derivation of the decentralized ergodic control problem first assumes that there exists an underlying connected network for which each agent can communicate with one another. Assume that the connections can be represented via a graph Laplacian  $L$  [22]. A consensus matrix  $P$  can be generated from the graph Laplacian such that  $P$  is both row and column stochastic [22], [23]. In the decentralized case, the ergodic metric is defined as

$$\mathcal{E}_i(x_i(t)) = \sum_{k \in \mathbb{Z}} (\phi_k - \sum_j P_{ij} c_{k,j})^2 \quad (13)$$

for each agent  $i$  with neighbors  $j$ . The consensus matrix  $P$  is used to transmit information about where each agent has explored (and is going to explore) to neighboring agent. As a result of sharing this information, each agent is able to synthesize a control solution that is sensitive to what the

network of agents has explored. Hence, consensus improves exploration efficiency for the overall network.

Following the derivation of the centralized ergodic controller, the mode insertion gradient for an individual agent is then,

$$\frac{d\mathcal{E}_i}{d\lambda} = \rho_i(s)(f_i(x_i(s), u_{2,i}(s)) - f_i(x_i(s), u_{1,i}(s)))$$

with the costate equation obtained from the backwards-differential equation

$$\dot{\rho}_i = - \sum_{k \in \mathbb{Z}} (\phi_k - \sum_j P_{ij} c_{k,j}) \frac{\partial F_{k,i}(x_i)}{\partial x_i} - \frac{\partial f_i}{\partial x_i}^T \rho_i.$$

The solution to the secondary cost function for an individual agent

$$J_{2,i} = \frac{d\mathcal{E}_i}{d\lambda} + \frac{1}{2}\|u_{2,i} - u_{1,i}\|_{R_i}^2$$

is

$$u_i^*(t) = -R_i^{-1}h_i(x_i)^T \rho_i(t) + u_{1,i}(t) \quad (14)$$

which is equivalent to the individual agent's control solution as the stacked representation of the centralized case (12). Individual time of control application  $\lambda$  can be found with

$$\lambda^* = \operatorname{argmin}_{\lambda} \frac{d\mathcal{E}_i(\lambda)}{d\lambda}.$$

**Theorem 1.** For a connected network of agents with control-affine dynamics (1), the decentralized ergodic control problem (13) is equivalent to the centralized ergodic control problem (8) when a consensus on the  $c_{k,i}$  values is reached.

*Proof:* Assume there exists a consensus matrix  $P$  that is row and column stochastic and described a connected graph. The operation  $\sum_j P_{ij} c_{k,j}$  is equivalent to taking an average of the local  $c_k$  values for each neighboring agent  $j$ . Thus, we can rewrite the consensus on  $c_k$  as

$$\lim_{t_k \rightarrow \infty} \sum_j P_{ij}^{t_k} c_{k,j} = \frac{1}{N} \sum_j \frac{1}{T} \int_{t_0}^{t_0+T} F_k(x_j(t)) dt. \quad (15)$$

where  $t_k$  is the number of times that the  $P_{i,j} c_{k,j}$  values have been iterated upon. The summation over all neighboring agents can be brought into the integral to give

$$\begin{aligned} \lim_{t_k \rightarrow \infty} \sum_j P_{ij}^{t_k} c_{k,j} &= \frac{1}{T} \int_{t_0}^{t_0+T} \frac{1}{N} \sum_j F_k(x_j(t)) dt \\ &= \frac{1}{T} \int_{t_0}^{t_0+T} \tilde{F}_k(x(t)) dt \end{aligned} \quad (16)$$

which when  $t_k \rightarrow \infty$ , a consensus is reached on where each agent has explored. Thus for any connected graph, the consensus is equivalent to (9).

Regardless of whether consensus amongst agents is reached, the ergodic measure for an individual agent will still be reduced asymptotically as  $t \rightarrow \infty$ . Consider that there exists a control action  $u$  that minimizes the ergodic cost (13). Then the mode insertion gradient (5) is  $\leq 0$  for some  $u^*(\lambda)$  which defines the rate of reduction of the cost at time  $\lambda$ . As a result,

any partial consensus will be included in the sensitivity of the ergodic metric with respect to neighboring agents' actions. Thus, consensus only impacts the rate of convergence of the ergodic metric and is dependent on the network connectivity.

### III. NETWORK AND CONSENSUS

In the previous section, we described the effects of consensus on the decentralized ergodic controller. In particular, the rate of convergence for both minimization of the ergodic metric and consensus on the  $c_k$  values of each agent depend on the connectivity of the graph itself. In this section, we discuss the formulation of the consensus matrix  $P$  that can be used for the decentralized ergodic controller.

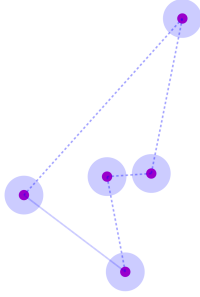


Figure 1. A ring network connection defined for a set of  $n$  vertices with edges  $(i, i + 1) \cup (n, 1)$  for  $i \in \{1, \dots, n\}$  is used for localization and estimation examples. The dotted lines represent the connection between each agent. Each agent is represented by the circle.

For the decentralized ergodic controller, any consensus matrix can be designed so long as the consensus matrix is row and column stochastic. This implies that the resulting consensus matrix must approximate the average of all the agents'  $c_k$  values. In an ideal situation where the graph is fully connected, consensus of the  $c_k$  values will occur rapidly. However, as shown in the last section, even partial consensus will improve the rate of convergence of the ergodic metric. In a later discussion in the results, convergence of the ergodic measure is presented with a 1, 2, and 3 agent network with a complete graph topology. For the remainder of this work, a ring graph (see Fig. 1) is used as an arbitrary choice to show that the decentralized ergodic controller eventually reaches a consensus so long as the graph is connected. Future work will involve examining various consensus matrices and their effects on the rate of convergence of the ergodic metric.

### IV. ESTIMATION AND LOCALIZATION

In this section we present a formulation of the decentralized ergodic control for estimation and localization problems. In general, problems of exploration for distributed information will take the same form as was developed initially in section III where the spatial statistical distribution  $\Phi$  is given.

#### A. Estimation

Typically estimation problems are formulated as a minimization of the sum of square errors of a set of measurement data  $y_i$  taken at state  $x_i \in \mathbb{R}^n$  and a prediction of the underlying measurement model that is likely to represent the data  $\phi(\theta, x)$  where  $\theta$  is some parameter that is to be estimated. Formally, the least-squares problem is described as

$$\min_{\theta} \sum_i^N (y_i - \phi(\theta, x))^2 \quad (17)$$

Solutions to this problem exist in abundance [24]–[26]. There exists an equal abundance of work in estimation for distributed networks [27]–[29]. In this work, we utilize a distributed Kalman filter [27], however any estimator which provides a statistical representation  $p(\theta)$  over the set of parameters  $\theta$  is sufficient for use in ergodic control. The only significant change in the ergodic control formulation is the target distribution  $\Phi(x)$ .

For an estimation problem, we utilize the expected information density (EID) of the estimator. The EID is first formulated with the definition of the Fisher information matrix given as

$$I(\theta, x) = \frac{\partial \phi(\theta, x)}{\partial \theta}^T \Sigma^{-1} \frac{\partial \phi(\theta, x)}{\partial \theta} \quad (18)$$

where  $\Sigma^{-1}$  is the measurement covariance. The EID is then the expectation of information given the certainty of the estimated parameters,  $p(\theta)$ . That is, we define the EID as

$$\text{EID}(x) = \eta \int_{\theta} I(\theta, x) p(\theta) d\theta \quad (19)$$

where  $\eta$  is a normalization factor. For multi-parameter estimation, the determinant of the EID is defined as our target distribution  $\Phi(x)$  which implies that each individual agent in the distributed network will have their own  $\phi_k$  values given by

$$\phi_{k,i} = \int_X \det[\text{EID}_i(x)] F_k(x) dx. \quad (20)$$

However, consensus on each  $\phi_{k,i}$  is not necessary as we are using estimators that form a consensus on the parameters  $\theta$  and their statistical parameters [27]. Thus, the only modification to the distributed ergodic control formulation is the need to recalculate  $\phi_k$  whenever a new estimate of  $\theta$  is provided. In this work, we use the example of mapping the terrain of a square domain region. Thus, measurements are considered to be local topological height.

#### B. Target Localization

Target localization is treated in the same manner as an estimation problem. Specifically, we are estimating the location of targets in the environment. In this work, we examine target localization in an environment filled with obstacles. A benefit of the formulation of ergodic control is the ease with which the objective function can be formulated. We assume that each robot has a map of the environment  $\Gamma(x)$  that penalizes

the robot if it is in collision with the obstacle. The primary objective function for each agent then becomes

$$\mathcal{E}_i(x_i(t)) = q_1 \sum_{k \in \mathbb{Z}} (\phi_k - \sum_j P_{ij} c_{k,j})^2 + q_2 \Gamma(x) \quad (21)$$

where the mode insertion gradient is

$$\frac{d\mathcal{E}_i}{d\lambda} = \rho_i(s)(f_i(x_i(s), u_{2,i}(s)) - f_i(x_i(s), u_{1,i}(s)))$$

with the costate solved by the backwards differential equation

$$\dot{\rho}_i = -2q_1 \sum_{k \in \mathbb{Z}} (\phi_k - \sum_j P_{ij} c_{k,j}) \frac{\partial F_{k,i}(x_i)}{\partial x_i} - q_2 \frac{\partial \Gamma(x_i)}{\partial x_i} - \frac{\partial f_i}{\partial x_i}^T \rho_i.$$

Thus, during forward trajectory prediction, the sensitivity will include future collisions that may occur.

Here, measurements  $y_i$  are taken as global position measurements of the target relative to the world frame. Each robot will then have a sensor with a limited circular range  $r$ . Initially, none of the agents know the location of any of the targets, thus the EID is considered to be a uniform distribution.

The results for exploration, estimation, and target localization are shown in the following section.

## V. RESULTS

In this section, results are presented for three applications of decentralized ergodic control. The first example discusses the reduction of the ergodic measure during information-based exploration of a unknown distribution. In particular, we observe the effects of introducing agents in a complete graph network (defined as a graph with  $n$  vertices connected to all other vertices) and collaborative behavior amongst agents. The second example illustrates the improvement in area coverage for localizing targets in a crowded environment. The last example illustrates how the decentralized ergodic controller can be used for Fisher Information-based estimation of terrain topology.

### A. Information-Based Coverage

Information-based exploration is presented with an example comparing coverage of a distribution using a 1, 2, and 3 agent network. The task for each agent is to maintain coverage of a distribution (see Fig. 2) that is proportional to the information available (given by the heat map). This example is presented to illustrate the collaborative behavior of decentralized ergodic control when under consensus. Specifically, this collaborative behavior can be seen in the trajectory traces of the 2 and 3 agent network. Each agent covers one of the peaks of the distribution while the other spans the remaining peaks. The agents then trade off this behavior to get an improved reduction of the ergodic measure. The benefit of using a multi-agent distributed ergodic system is shown in the ergodicity measure in Fig. 2 where the rate of reduction in the ergodic metric increases with added agents.

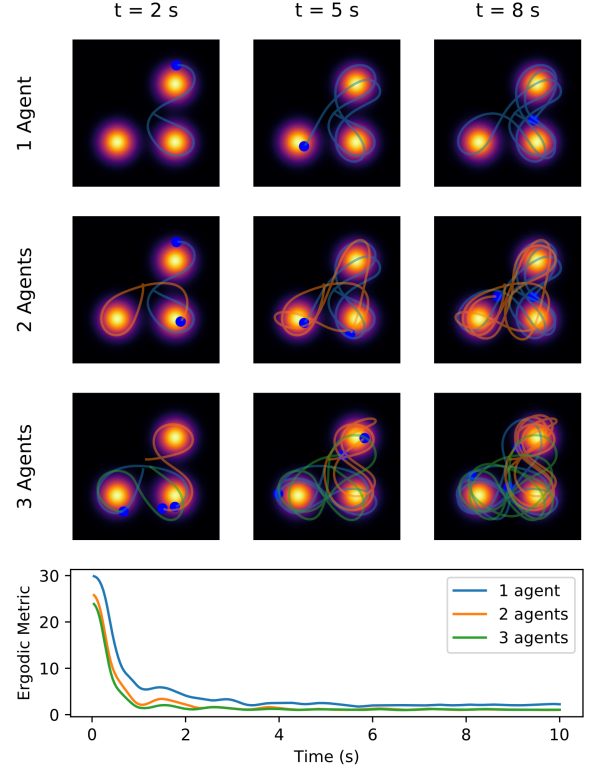


Figure 2. Information based exploration is shown with 1, 2, and 3 agents using ergodic control. Adding an additional agent automatically improves the coverage and reduction of the ergodic metric. Collaborative behavior is seen with the 2 and 3 agent systems where a trade-off occurs allowing a single agent to spend more time at a peak while the other agents span across the remaining information peaks.

### B. Localization

Results for localization are presented with the task to localize 5 targets within an environment with obstacles (see Fig. 3). Targets are shown as a red dot and initialized in the same way for a network of 1 agent and 5 agents. Each time an agent is localized, a blue circle is drawn at the time of localization from the initial starting point (Fig. 3). Each agent is not allowed to traverse the grey obstacles in the environment. Results show a clear benefit of utilizing consensus within ergodic exploration. Specifically, each individual agent explores their own particular subset of the environment. Therefore, each individual agent travels less which improves the overall energetic cost of moving while localizing all the targets at a much faster rate than with a single agent. The longer travel times are clearly noted in the trajectory traces of the single robot agent localizing the targets. Although the single agent requires longer time to localize all the targets, the ergodic control still successfully synthesizes actions that enable a single agent to span a large enough region to eventually localize all the targets.

The analysis of the localization example is furthered with a detailed observation of the ergodic measure. Figure 4 shows the convergence of the ergodic metric for the single agent, the collective 5 agent network, and the individual measures of

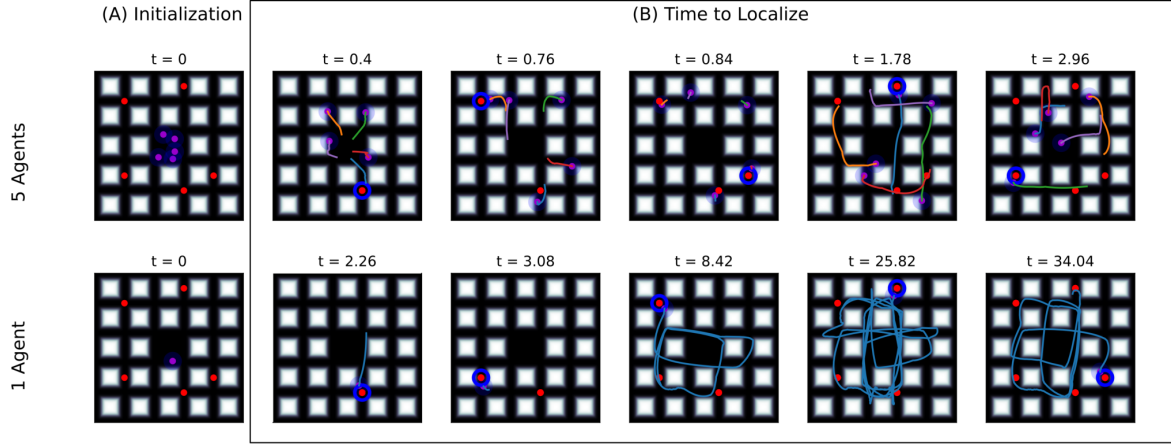


Figure 3. Localization of 5 targets in an environment with obstacles is shown for a single agent and a 5 agent network shown in Fig. 1. Each time frame shown indicates when a new agent was discovered in the environment. (A) The starting frame  $t = 0$  shows the starting location of each agent. Note that the 5 agent network distributes exploration amongst the agents which results in finding all 5 targets within 3 seconds. Interestingly, ergodic exploration of 1 agent eventually results in the localization of all 5 agents given enough time to explore the environment. A monte-carlo simulation of the network localizing sequentially randomly generated targets is provided in the link: <https://vimeo.com/209261679/0d65c1a9ac>.

ergodicity for each agent in the 5 agent network. With more agents, the ergodic measure reduces rapidly and is maintained in a non zero equilibrium value (this is due to the measure never being able to reach zero). The individual robotic agents ergodicity measures show similarity with the single agent network. Thus, regardless if consensus is reached, each agent will reduce the ergodic measure at their own particular rate. As consensus is reached, the rate of decent of the ergodic metric increases. When information about where each agent has explored is transferred through the network, the individual efforts result in an overall reduction in the ergodic metric.

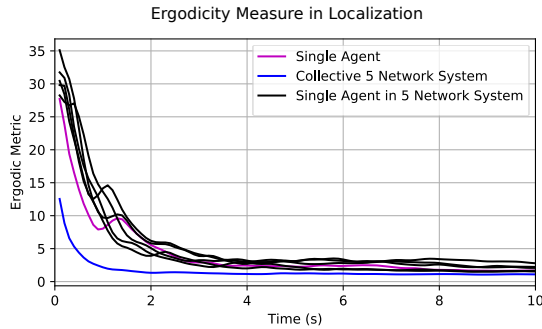


Figure 4. Ergodicity measures for a single agent localizing 5 targets versus a 5 agent network localizing 5 targets. Each individual agent's ergodic measure is shown for the 5 agent network. Compared with the single agent, most robots individually have similar ergodic measures. As a connected network utilizing consensus, the ergodic measure is shown to be significantly less. This is due to shared information between agents on where each agent has been, thus each agent explores in a manner that is best for the collective group.

### C. Estimation

Results for the problem of estimation is presented for a network of 5 autonomous aircraft that is tasked to estimate the

terrain of a square grid. The motion of each agent is subject to the dynamics given by

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} u \cos(\theta) \cos(\psi) \\ u \cos(\theta) \sin(\psi) \\ u \sin(\theta) \\ v \\ w \end{bmatrix}$$

where  $(u, v, w)$  are the control inputs and  $u > 0$  such that the aircraft must always be in motion. Each individual agent then takes measurements of the terrain height with which they are currently above<sup>2</sup>. A distributed Kalman filter is used to estimate the terrain height based on the assumption that the terrain can be described as a network of radial-basis functions (RBFs) [30]. RBFs are typically used to described a continuous surface using a discrete set of radial-basis functions. In this application, the RBFs are used as an estimator which can be used to generate an expected information density. As each agent acquires data, the distributed Kalman filter updates the terrain estimate and then forms a consensus on the estimate based on the network of the agents. The distributed ergodic control then synthesizes control actions that reduce the ergodic metric with respect to the expected information density.

The ground truth terrain is shown in Fig. 5 as well as the estimate for the 5 agent network after 10 second simulation time. For the allotted time, the estimate is shown to converge closely to the ground truth terrain. Although the graph connectivity is not that of a complete graph (a ring graph is used Fig. 1), the agent's coverage is still shown to grasp a significant amount of the allotted space for which the aircraft is allowed to travel.

<sup>2</sup>Each agent has an additional penalty to maintain a specific height above the terrain which prevents collision of the terrain.



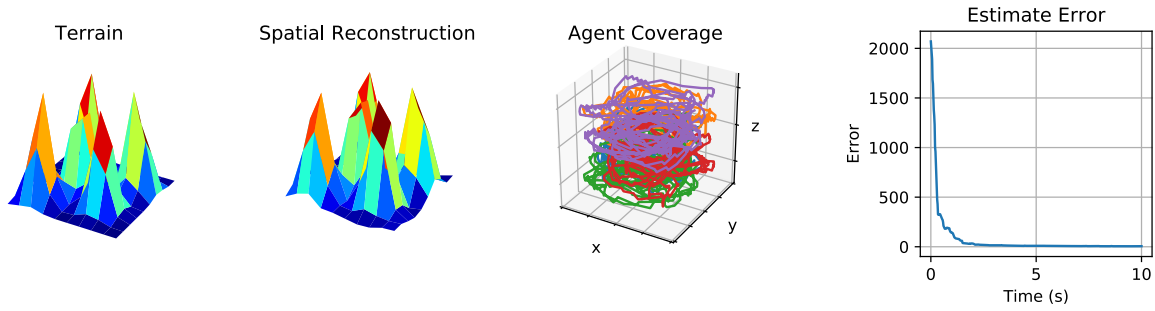


Figure 5. Terrain estimation is shown with a 5 agent aircraft network connected as shown in Fig. 1. Aircrafts take height measurements at each sampling rate and reconstruct the terrain based on a radial-basis-function network. Agent coverage is shown to span the whole space. Estimate error is shown to provide a numerical measure of the accuracy of the spatial reconstruction.

## VI. CONCLUSION

In this work, ergodic control is formulated as a means to synthesize actions that exhibit exploratory coverage behavior subject to a target statistical distribution. The formulation is then extended into a centralized distributed control problem for active exploration. Using graph theory and distributed optimization, ergodic control is reformulated in a completely distributed manner where information about past exploration is shared amongst agents. The sharing of past exploration provides a consensus on what has been explored and the sensitivity to the next action for each individual agent that reduces the ergodic metric. Various examples that utilize variants of ergodic control problem formulation are used to show improved convergence with additional agents, faster localization time and coverage, and rapid estimation of terrain.

In future work, we will examine the effects of the consensus matrix on the distributed ergodic control and augmentations to the controller to maintain network connectivity for time-varying graph topology.

## REFERENCES

- [1] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping SLAM: Part II," *IEEE Robotics & Automation Magazine*, vol. 13, no. 3, pp. 108–117, 2006.
- [2] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: Part I," *IEEE robotics & automation magazine*, vol. 13, no. 2, pp. 99–110, 2006.
- [3] H. Choset and K. Nagatani, "Topological simultaneous localization and mapping (slam): toward exact localization without explicit localization," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 2, pp. 125–137, 2001.
- [4] I. Abraham, A. Prabhakar, M. J. Hartmann, and T. D. Murphey, "Ergodic exploration using binary sensing for nonparametric shape estimation," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 827–834, 2017.
- [5] G. De La Torre, K. Flaßkamp, A. Prabhakar, and T. D. Murphey, "Ergodic exploration with stochastic sensor dynamics," in *American Control Conference (ACC)*, 2016, 2016, pp. 2971–2976.
- [6] L. M. Miller, Y. Silverman, M. A. MacIver, and T. D. Murphey, "Ergodic exploration of distributed information," *IEEE Transactions on Robotics*, vol. 32, no. 1, pp. 36–52, 2016.
- [7] Y. Yang, H. Zhang, Z.-H. Lin, Y. S. Zhou, Q. Jing, Y. Su, J. Yang, J. Chen, C. Hu, and Z. L. Wang, "Human skin based triboelectric nanogenerators for harvesting biomechanical energy and as self-powered active tactile sensor system," *ACS nano*, vol. 7, no. 10, pp. 9213–9222, 2013.
- [8] R. S. Dahiya, G. Metta, M. Valle, and G. Sandini, "Tactile sensing: From humans to humanoids," *IEEE Transactions on Robotics*, vol. 26, no. 1, pp. 1–20, Feb 2010.
- [9] F. Iida and A. J. Ijspeert, "Biologically inspired robotics," in *Springer Handbook of Robotics*. Springer, 2016, pp. 2015–2034.
- [10] L. Zhu, L. Zeng, X. Chen, X. Luo, and X. Li, "A bioinspired touching sensor for amphibious mobile robots," *Advanced Robotics*, vol. 29, no. 22, pp. 1437–1452, 2015.
- [11] L. Mihaylova, T. Lefebvre, H. Bruyninckx, K. Gadeyne, and J. De Schutter, "A comparison of decision making criteria and optimization methods for active robotic sensing," in *International Conference on Numerical Methods and Applications*. Springer, 2002, pp. 316–324.
- [12] A. Ryan and J. K. Hedrick, "Particle filter based information-theoretic active sensing," *Robotics and Autonomous Systems*, vol. 58, no. 5, pp. 574–584, 2010.
- [13] D. Carmel and S. Markovitch, "Exploration strategies for model-based learning in multi-agent systems: Exploration strategies," *Autonomous Agents and Multi-agent systems*, vol. 2, no. 2, pp. 141–172, 1999.
- [14] G. Dudek, M. R. Jenkin, E. Milios, and D. Wilkes, "A taxonomy for multi-agent robotics," *Autonomous Robots*, vol. 3, no. 4, pp. 375–397, 1996.
- [15] C. Manss, D. Shutin, T. Wiedemann, A. Viseras, and J. Mueller, "Decentralized multi-agent entropy-driven exploration under sparsity constraints," in *Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)*, 2016, pp. 143–147.
- [16] A. Viseras Ruiz, M. Angermann, I. Wieser, M. Frassl, and J. Mueller, "Efficient multi-agent exploration with Gaussian processes," 2014.
- [17] M. A. Khamis and W. Goma, "Adaptive multi-objective reinforcement learning with hybrid exploration for traffic signal control based on co-operative multi-agent framework," *Engineering Applications of Artificial Intelligence*, vol. 29, pp. 134–151, 2014.
- [18] F. Pei, M. Wu, and S. Zhang, "Distributed SLAM using improved particle filter for mobile robot localization," *The Scientific World Journal*, 2014.
- [19] G. Mathew and I. Mezić, "Metrics for ergodicity and design of ergodic dynamics for multi-agent systems," *Physica D: Nonlinear Phenomena*, vol. 240, no. 4, pp. 432–442, 2011.
- [20] D. A. Shell and M. J. Matarić, "Ergodic dynamics for large-scale distributed robot systems," in *International Conference on Unconventional Computation*. Springer, 2006, pp. 254–266.
- [21] A. R. Ansari and T. D. Murphey, "Sequential action control: closed-form optimal control for nonlinear and nonsmooth systems," *IEEE Transactions on Robotics*, vol. 32, no. 5, pp. 1196–1214, 2016.
- [22] N. Deo, *Graph theory with applications to engineering and computer science*. Courier Dover Publications, 2016.
- [23] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Prentice hall Englewood Cliffs, NJ, 1989, vol. 23.
- [24] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.
- [25] S. J. Julier and J. K. Uhlmann, "New extension of the kalman filter to nonlinear systems," in *AeroSense'97*. International Society for Optics and Photonics, 1997, pp. 182–193.

- [26] E. A. Wan and R. Van Der Merwe, "The unscented kalman filter for nonlinear estimation," in *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000*, 2000, pp. 153–158.
- [27] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, "Distributed kalman filtering based on consensus strategies," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, 2008.
- [28] Z. Liu, Y. Liu, and C. Li, "Distributed sparse recursive least-squares over networks," *IEEE Trans. Signal Processing*, vol. 62, no. 6, pp. 1386–1395, 2014.
- [29] R. Arablouei, S. Werner, Y.-F. Huang, and K. Doğançay, "Distributed least mean-square estimation with partial diffusion," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 472–484, 2014.
- [30] K.-L. Du and M. Swamy, "Radial basis function networks," in *Neural Networks and Statistical Learning*. Springer, 2014, pp. 299–335.