

Ergodic Coverage In Constrained Environments Using Stochastic Trajectory Optimization

Elif Ayvali, Hadi Salman and Howie Choset

Abstract—In search and surveillance applications in robotics, it is intuitive to spatially distribute robot trajectories with respect to the probability of locating targets in the domain. Ergodic coverage is one such approach to trajectory planning in which a robot is directed such that the percentage of time spent in a region is in proportion to the probability of locating targets in that region. In this work, we extend the ergodic coverage algorithm to robots operating in constrained environments and present a formulation that can capture sensor footprint and avoid obstacles and restricted areas in the domain. We demonstrate that our formulation easily extends to coordination of multiple robots equipped with different sensing capabilities to perform ergodic coverage of a domain.

I. INTRODUCTION

Many applications in robotics require efficient strategies to explore a domain for surveillance [1], target localization [2] and estimation tasks [3], [4]. Directing a robot to exhaustively pass over all points in the exploration domain may be time consuming or impractical when there is limited power budget. Therefore, probabilistic approaches were developed to exploit prior information, when it exists, to direct the exploration to take less time [5]–[7]. Some of these approaches incorporate a coverage metric, called ergodicity, as an objective to guide the exploration with respect to a desired spatial distribution of robot trajectories [8], [9]. For example, given a probability density function (pdf) defined over the exploration domain, such as one indicating where survivors might be located in a search and rescue scenario, an ergodic coverage strategy directs the robot to distribute its time searching regions of the domain in proportion to the probability of locating survivors in those regions.

The ergodic coverage algorithm, originally formulated by Mathew and Mezić [8], applies to simple kinematic systems and has a natural extension to centralized multi-robot systems. The formulation, however, treats a robot as point mass, and robot’s sensor footprint is represented as a Dirac delta function for computational tractability. Additionally, previous implementations of the algorithm in [8], [9] do not consider the obstacles in the exploration domain. To exclude the coverage of certain areas, Mathew and Mezić [8] set the value of the desired spatial distribution of trajectories to zero in those areas. However, the robots still visit those areas although the percentage of time spent is small. We have recently employed a potential field based-approach

to extend the formulation in [8], which considers first-order and second-order systems, to avoid circular obstacles in a planar domain [10]. However, generalization of the derivations for other systems and types of obstacles are nontrivial. Constraints corresponding to arbitrary obstacles can be non-smooth, and, as a consequence, gradient-based optimization methods employed in aforementioned works require special differentiation to guarantee convergence of the solution [11].

In this work, we incorporate the metric for ergodicity [8], as an objective in a sampling-based stochastic trajectory optimization framework. The idea behind this framework is to construct a probability distribution over the set of feasible paths and find the optimal trajectory using the cross entropy method [12], [13]. We chose this approach because it generalizes to nonlinear robotic systems operating in constrained environments and can be coupled with other commonly used sampling-based planners such as rapidly-exploring random trees (RRT) and its variants [14] to generate the set of feasible paths [15].

The novelty of the work in this paper lies in the construction of a new ergodic coverage objective that can take into account the robots’ sensor footprint, and in the formulation of the algorithm within a stochastic trajectory optimization framework that lends itself to constraints rising from arbitrary-shaped obstacles. Our formulation also allows a robot with a wide sensor footprint to perform coarse ergodic coverage of the domain, while a robot with narrow sensor footprint performs dense ergodic coverage of the same domain.¹

This paper is organized as follows. Section II provides background on ergodic coverage and stochastic trajectory optimization using the cross entropy method. In Section III, we derive means to incorporate sensor footprint and introduce a new ergodic coverage objective. In Section IV, we provide a comparison with the formulation in [8]. We also demonstrate the flexibility of our formulation in encoding additional tasks. We present an example in which a robot is directed to perform ergodic coverage while having to reach a destination in finite time.

II. BACKGROUND

A. Ergodic Coverage

Ergodic theory is the statistical study of time-averaged behavior of dynamic systems [16]. A system exhibits ergodic

*This work was supported by NRI IIS-1426655 and ONR N00024-13-D-6400 subcontract through APL/JHU.

¹E. Ayvali, H.Salman, H. Choset are with the Robotics Institute at Carnegie Mellon University, Pittsburgh,PA 15213, USA (eayvali@, hadis@andrew, choset@) cmu.edu

¹Source codes are publicly available at <https://github.com/biorobotics/stoec>

dynamics if it uniformly explores all its possible states over time. Mathew and Mezić [8] introduced a metric to quantify the ergodicity of a robot's trajectory with respect to a given probability density function.

Formally, the time-average statistics of a robot's trajectory, $\gamma: (0, t] \rightarrow X$, quantifies the fraction of time spent at a point, $\mathbf{x} \in X$, where $X \subset \mathbb{R}^d$ is a d -dimensional domain. Mathew and Mezić [8] define the time-average statistics of a trajectory at a point \mathbf{x} as

$$\Gamma(\mathbf{x}) = \frac{1}{t} \int_0^t \delta(\mathbf{x} - \gamma(\tau)) d\tau, \quad (1)$$

where δ is the Dirac delta function.

Let $\xi(\mathbf{x})$ be a pdf – also referred as the desired coverage distribution – defined over the domain. The ergodicity of a robot's trajectory with respect to $\xi(\mathbf{x})$ is defined as

$$\Phi(t) = \sum_{k=0}^m \lambda_k |\Gamma_k(t) - \xi_k|^2, \quad (2)$$

where $\lambda_k = \frac{1}{(1+|k|)^s}$ is a coefficient that places higher weights on the lower frequency components and $s = \frac{d+1}{2}$. The $\Gamma_k(t)$ and ξ_k are the Fourier coefficients of the distributions $\Gamma(\mathbf{x})$ and $\xi(\mathbf{x})$ respectively, i.e.²,

$$\begin{aligned} \Gamma_k(t) &= \langle \Gamma, f_k \rangle = \frac{1}{t} \int_0^t f_k(\gamma(\tau)) d\tau \\ \xi_k &= \langle \xi, f_k \rangle = \int_X \xi(\mathbf{x}) f_k(\mathbf{x}) d\mathbf{x} \end{aligned} \quad (3)$$

where $f_k(\mathbf{x}) = \frac{1}{h_k} \prod_{i=1}^m \cos(\frac{k_i \pi}{L_i} x_i)$ is the Fourier basis functions that satisfy Neumann boundary conditions on the domain X , $m \in \mathbb{Z}$ is the number of basis functions, and $\langle \cdot, \cdot \rangle$ is the inner product with respect to the Lebesgue measure. The term h_k is a normalizing factor.

The goal of ergodic coverage is to generate optimal controls $\mathbf{u}^*(t)$ for a robot, whose dynamics is described by a function $\dot{\mathbf{q}}(t) = g(\mathbf{q}(t), \mathbf{u}(t))$ such that

$$\begin{aligned} \mathbf{u}^*(t) &= \arg \min_{\mathbf{u}} \Phi(\mathbf{q}(t), \mathbf{u}(t), t), \\ \text{subject to } \dot{\mathbf{q}}(t) &= g(\mathbf{q}(t), \mathbf{u}(t)) \\ \|\mathbf{u}(t)\| &\leq u_{max}, \end{aligned} \quad (4)$$

where $\mathbf{q} \in Q$ is the state space and $\mathbf{u} \in U$ denotes the set of controls.

Mathew and Mezić [8] mainly consider first-order, $\dot{\mathbf{q}}(t) = \mathbf{u}(t)$, and second-order systems, $\ddot{\mathbf{q}}(t) = \mathbf{u}(t)$. They derive closed-form solutions that approximate the solution to Eq. (4) by discretizing the exploration time and solving for the optimal control input that maximizes the rate of decrease of Eq. (4) at each time-step. Miller and Murphey [9] use a projection-based trajectory optimization method that solves a first-order approximation of Eq. (2) at each iteration using linear quadratic regulator techniques [17].

²If there are N robots, the time average statistics and the corresponding Fourier coefficients are defined as $\Gamma(\mathbf{x}) = \frac{1}{Nt} \sum_{j=1}^N \int_0^t \delta(\mathbf{x} - \gamma_j(\tau)) d\tau$ and $\Gamma_k(t) = \frac{1}{Nt} \sum_{j=1}^N \int_0^t f_k(\gamma_j(\tau)) d\tau$, respectively.

B. Cross-Entropy Trajectory Optimization

Consider a robot whose dynamics is described by the function $g: Q \times U \rightarrow TX$, such that

$$\dot{\mathbf{q}}(t) = g(\mathbf{q}(t), \mathbf{u}(t)) \quad (5)$$

In addition, the robot is subject to constraints such as actuator bounds and obstacles in the environment. These constraints are expressed through the vector of inequalities

$$F(\mathbf{q}(t)) \geq 0, \quad (6)$$

The goal is to compute the optimal controls \mathbf{u}^* over a time horizon $t \in (0, t_f]$ that minimize a cost function such that

$$\begin{aligned} \mathbf{u}^*(t) &= \arg \min_{\mathbf{u}} \int_0^{t_f} C(\mathbf{u}(t), \mathbf{q}(t)) dt, \\ \text{subject to } \dot{\mathbf{q}}(t) &= g(\mathbf{q}(t), \mathbf{u}(t)), \\ F(\mathbf{q}(t)) &\geq 0, \\ \mathbf{q}(0) &= \mathbf{q}_0, \end{aligned} \quad (7)$$

where \mathbf{q}_0 is the initial state of the robot and $C: U \times Q \rightarrow \mathbb{R}$ is a given cost function.

1) *Trajectory Parameterization*: Following the notation in [15], a trajectory defined by the controls and states over the time interval $[0, T]$ is denoted by the function $\pi: [0, T] \rightarrow U \times Q$, i.e. $\pi(t) = (\mathbf{u}(t), \mathbf{q}(t))$ for all $t \in [0, T]$. The space of all trajectories originating at \mathbf{q}_0 and satisfying Eq. (5) is given by

$$P = \{ \pi: t \in [0, T] \rightarrow (\mathbf{u}(t), \mathbf{q}(t)) \mid \dot{\mathbf{q}}(t) = g(\mathbf{q}(t), \mathbf{u}(t)), \mathbf{q}(0) = \mathbf{q}_0, T > 0. \} \quad (8)$$

Let us consider a finite-dimensional parameterization of trajectories in terms of vectors $\mathbf{z} \in Z$ where $Z \subset \mathbb{R}^{n_z}$ is the parameter space. Assuming that the parameterization is given by a function $\varphi: Z \rightarrow P$ according to

$$\pi = \varphi(\mathbf{z}). \quad (9)$$

The (\mathbf{u}, \mathbf{q}) tuples along a trajectory parameterized by \mathbf{z} are written as $\pi(t) = \varphi(\mathbf{z}, t)$. One choice of parameterization is to use motion primitives defined as $\mathbf{z} = (\mathbf{u}_1, \tau_1, \dots, \mathbf{u}_j, \tau_j)$ where each u_i , for $1 \leq i \leq j$, is a constant control input applied for duration τ_i . For differentially flat systems [18], a sequence of states, $\mathbf{z} = (\mathbf{q}_1, \dots, \mathbf{q}_j)$ can also be used. Another option is to use RRT or probabilistic roadmaps (PRM) to sample trajectories as demonstrated in [15].

Now, we can define a cost function, $J: Z \rightarrow \mathbb{R}$, in terms of the trajectory parameters as

$$J(\mathbf{z}) = \int_0^T C(\varphi(\mathbf{z}, t)) dt \quad (10)$$

Eq. (7) can be restated as finding the optimal $(\mathbf{u}^*, \mathbf{q}^*) = \varphi(\mathbf{z}^*)$ such that

$$\mathbf{z}^* = \arg \min_{\mathbf{z} \in Z_{con}} J(\mathbf{z}). \quad (11)$$

where the constrained parameter space $Z_{con} \subset Z$ is the set of parameters that satisfy the boundary conditions and constraints in Eq. (7).

2) *Cross-Entropy Method*: In this work, we employ the cross entropy (CE) method to optimize the parameters of the trajectory. There are other sampling-based global optimization methods such as Bayesian optimization [19], simulated annealing [20], and other variants of stochastic optimization [21] that can also be used to optimize parameterized trajectories. We use the CE method because it utilizes importance sampling to efficiently compute trajectories that have lower costs after few iterations of the algorithm, and it has been shown to perform well for trajectory optimization of nonlinear dynamic systems [15].

The CE method treats the optimization in Eq. (11) as an estimation problem of rare-event probabilities. The rare event of interest in this work is to find a parameter z whose cost $J(z)$ is very close to the cost of an optimal parameter z^* . It is assumed that the parameter $z \in Z$ is sampled from a Gaussian mixture model defined as

$$p(z; v) = \sum_{k=1}^K \frac{w_k}{\sqrt{(2\pi)^{n_z} |\Sigma_k|}} e^{-\frac{1}{2}(z-\mu_k)^T \sigma_k^{-1}(z-\mu_k)} \quad (12)$$

where $v = (\mu_1, \Sigma_1, \dots, \mu_K, \Sigma_K, w_1, \dots, w_K)$ corresponds to K mixture components with means μ_k , covariance matrices Σ_k , and weights w_k , where $\sum_{k=1}^K w_k = 1$.

The CE method involves an iterative procedure where each iteration has two steps: (i) select a set of parameterized trajectories from $p(z; v)$ using importance sampling [13] and evaluate the cost function $J(z)$, (ii) use a subset of *elite* trajectories³ and update v using expectation maximization [22]. After a finite number of iterations $p(z; v)$ approaches to a delta distribution, thus the sampled trajectories remains unchanged. For implementation details, the reader is referred to [23].

III. STOCHASTIC TRAJECTORY OPTIMIZATION FOR ERGODIC COVERAGE (STOEC)

In this section, we introduce a formulation that optimizes an ergodic coverage objective using the stochastic trajectory optimization algorithm presented in Section II-B. In this formulation, we pose ergodic coverage as a constrained optimization problem subject to constraints associated with obstacle avoidance and with the motion model of the robot. We start by redefining the time average statistics of a robot's trajectory to capture the robot's sensor footprint. We then demonstrate that obstacle avoidance can be easily incorporated into this framework by giving penalties to the sampled trajectories that collide with arbitrarily-shaped obstacles or that cross restricted regions.

Let us start with two key observations. First, Eq. (1) assumes that a robot's trajectory is composed of Dirac delta functions. In other words, it is assumed that the sensor footprint is negligible compared to the size of the domain. This assumption is adopted to efficiently calculate the Fourier coefficients of the distributions given by Eq. (3). In many practical applications, the robots can be equipped with sensors of various types (camera, lidar, ultrasonic, etc.), which

³A fraction of the sampled trajectories with the best costs form an elite set. See [13] for details.

are hardly captured by the Dirac delta function. Having control over the sensor footprint opens up interesting and useful strategies for coverage: a robot equipped with a wide sensor footprint should be able to perform a coarse ergodic coverage of the domain compared to a robot equipped with a narrow sensor footprint. Similarly, the optimal ergodic-coverage trajectory of a ground vehicle equipped with a forward looking ultrasonic sensor would be different than an aerial vehicle equipped with a downward-looking sensor.

Second, as also stated in [9], the motivation in using the norm on the Fourier coefficients is to obtain an objective function that is differentiable with respect to the trajectory. When a sampling-based approach is used for trajectory optimization, there is no need for a differentiable objective function. There are many other measures used in information theory and statistics to compare two distributions. Here, we consider Kullback-Leibler (KL) divergence, which measures the relative entropy between two distributions,

$$D_{KL}(\Gamma|\xi) = \int_X \Gamma(\mathbf{x}) \log \left(\frac{\Gamma(\mathbf{x})}{\xi(\mathbf{x})} \right) d\mathbf{x}. \quad (13)$$

KL divergence can encode an ergodic coverage objective without resorting to the spectral decomposition of the desired coverage distribution, whose accuracy is limited by the number of basis functions used. Additionally, we can explicitly measure the ergodicity of a trajectory without resorting to simplifications in the representation of a robot's sensor footprint.

In our formulation, the optimal control problem posed in Eq. (11) can be solved by defining the cost function $J(\mathbf{z})$ to be

$$J(\mathbf{z}) = D_{KL}(\Gamma|\xi) \quad (14)$$

where z is sampled from a Gaussian mixture model defined in Eq. (12).

A. Sensor Footprints

In this paper, we consider two sensor footprints: (1) Gaussian sensor footprint, and (2) Forward-looking beam sensor footprint. Assuming there are N robots, let $f_j(\cdot)$ be the sensor footprint of the j -th robot. The time-average statistic, $\Gamma(\mathbf{x})$, can be defined as,

$$\Gamma(\mathbf{x}) = \eta \sum_{j=1}^N \int_0^t f_j(\mathbf{x} - \gamma_j(\tau)) d\tau, \quad (15)$$

where γ_j is the trajectory of the j -th robot, and η is a normalizing factor such that $\Gamma(\mathbf{x})$ is a pdf. A Gaussian sensor footprint is described as,

$$f_j(\mathbf{x} - \gamma_j(\tau)) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \gamma_j(\tau))^T \Sigma^{-1}(\mathbf{x} - \gamma_j(\tau))\right) \quad (16)$$

whereas the sensor footprint of a beam model is defined by a radius, r , and angle of view, β , similar to an ultrasonic sensor as shown in Fig. 2. The sensor footprint of the beam

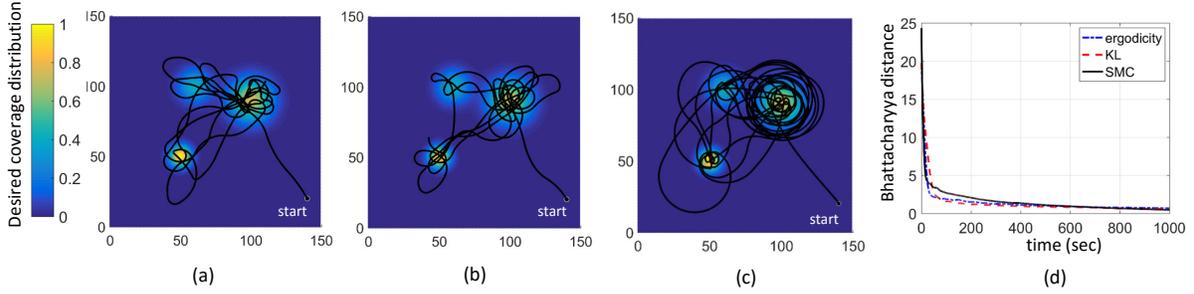


Fig. 1. Ergodic coverage results for three different implementation: (a) Ergodic-STOEC: minimizing Eq.(2), (b) KL-STOEC: minimizing Eq.(13), (c) SMC implementation in Mathew *et. al* [24]. Ergodic coverage performance is assessed through plotting the Bhattacharyya distance between the coverage distribution $\xi(\mathbf{x})$ and the time-average statistics distribution $\Gamma(\mathbf{x})$ in (d) for the different implementations. Notice that (a) and (b) have similar decay rates, while the decay rate of (c) is smaller.

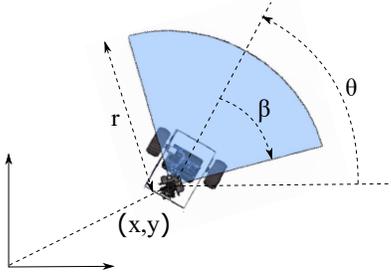


Fig. 2. Forward looking beam sensor model (e.g. ultrasonic sensor): an example of a sensor footprint of the robot. The variables (x, y, θ) represent the state of the robot, whereas r and β represent the bearing measurement limits of the sensor.

model is parameterized as

$$f_j(\mathbf{x} - \gamma_j(\tau)) = \begin{cases} 1, & \text{if } |\mathbf{x} - \gamma_j(\tau)| \leq r \text{ and} \\ & \theta_\tau - \frac{\beta}{2} \leq \arctan\left(\frac{(\mathbf{x} - \gamma_j(\tau))_2}{(\mathbf{x} - \gamma_j(\tau))_1}\right) \leq \theta_\tau + \frac{\beta}{2} \\ 0, & \text{otherwise.} \end{cases} \quad (17)$$

where θ_τ is the heading of the robot at time τ , $(\mathbf{x} - \gamma_j(\tau))_i$ is the i -th element of the vector $(\mathbf{x} - \gamma_j(\tau))$.

B. Obstacle Avoidance

Suppose that the domain X contains p obstacles denoted by $O_1, \dots, O_p \subset X$. We assume that the robot at state \mathbf{q} is occupying a region $A(\mathbf{q}) \subset X$. Borrowing the notation in [15], let the function $\mathbf{prox}(A_1, A_2)$ return the closest Euclidean distance between two sets $A_{1,2} \subset X$. This function returns a negative value if the two sets intersect. Therefore, for an agent to avoid the obstacles O_1, \dots, O_p , we impose a constraint of the form of Eq. (6) expressed as,

$$F(\mathbf{q}(t)) = \min_i \mathbf{prox}(A(\mathbf{q}(t)), O_i), \quad \forall t \in [0, \infty). \quad (18)$$

The sampled trajectories that does not satisfy the constraints of Eq. (6) are penalized.

C. Example: Dubins Car

The trajectory optimization framework can be applied to any robot whose dynamics are given as (5) given that we can parameterize the space of all trajectories satisfying (5) using motion primitives. We will consider the Dubins car model

whose motion is restricted to a plane. The state space for such a model is $Q = SE(2)$ with state $\mathbf{q} = (x, y, \theta)$, where (x, y) are the Cartesian coordinates of the robot, and θ is the orientation of the robot in the plane. The dynamics of a dubins car is defined by,

$$\dot{x} = u \cos \theta, \quad \dot{y} = u \sin \theta, \quad \dot{\theta} = v \quad (19)$$

where $v \in [v_{min}, v_{max}]$ is a controlled forward velocity, and $w \in [w_{min}, w_{max}]$ is a controlled turning rate.

We can represent a trajectory satisfying (19) as a set of connected motion primitives consisting of either straight lines with constant velocity v or arcs of radius v/w . We define a primitive by a constant controls (v, w) . The duration of each primitive is constant and $\tau > 0$. We parameterize the trajectory of the robot using m primitives, and this finite dimensional parameterization is represented by a vector $\mathbf{z} \in \mathbb{R}^{2m}$ such that,

$$\mathbf{z} = (v_1, w_1, \dots, v_m, w_m) \quad (20)$$

The j^{th} primitive ends at time $t_j = j\tau$ for $j = 1, \dots, m$. For any time $t \in [t_j, t_{j+1}]$, the parameterization $\varphi_z = (v, w, x, y, \theta)$ is given by,

$$\begin{aligned} v(t) &= v_{j+1} \\ w(t) &= w_{j+1} \\ \theta(t) &= \theta_j + w_{j+1} \Delta t_j \\ x(t) &= \begin{cases} x(t_j) + \frac{v_{j+1}}{w_{j+1}} (\sin(\theta(t)) - \sin \theta_j), & \text{if } w_{j+1} \neq 0 \\ x(t_j) + v_{j+1} \Delta t_j \cos \theta_j, & \text{otherwise} \end{cases} \\ y(t) &= \begin{cases} y(t_j) + \frac{v_{j+1}}{w_{j+1}} (\cos \theta_j - \cos(\theta(t))), & \text{if } w_{j+1} \neq 0 \\ y(t_j) + v_{j+1} \Delta t_j \sin \theta_j, & \text{otherwise} \end{cases} \end{aligned} \quad (21)$$

where $\theta_j = \theta(t_j)$ and $\Delta t_j = t - t_j$.

IV. RESULTS

In this section, we first compare our algorithm with the implementation of Mathew *et. al* [24] in an unconstrained environment. In Section IV-B, we demonstrate that our formulation can capture coordination of multiple robots equipped with different sensing capabilities. It is also straightforward to encode additional tasks besides ergodic

coverage. In Section IV-C, we present an example in which a robot is directed to reach a destination in a desired time while performing ergodic coverage.

A. Comparisons

We first compare the coverage performance of three different implementations of the ergodic coverage algorithm: (i) We generate a trajectory that minimizes the rate of change of the metric for ergodicity given by Eq. (2) as presented in Mathew *et. al* [24], which will be referred to as spectral-multiscale coverage (SMC) algorithm. (ii) We use STOEC to optimize the trajectory such that Eq. (2) is minimized. This implementation will be referred as Ergodic-STOEC. (iii) We use STOEC to optimize the trajectory such that Eq. (13) is minimized. This will be referred as KL-STOEC.

We perform numerical simulations for each of the aforementioned implementations. We define $\xi(\mathbf{x})$, as a Gaussian mixture model⁴ as shown in Fig. 1. For visualization, $\xi(\mathbf{x})$ is normalized such that the maximum value of the desired coverage distribution takes the value of 1. The initial position of the robot is randomly selected. The robot is assumed to have bounded inputs $v \in [0.1, 5]$ m/s and $w \in [-0.2, 0.2]$ rad/s. For SMC, the total simulation time is $T = 1000$ sec, and the time step is 0.1 sec. For STOEC, we run 20 stages of the algorithm using 5 primitives with a receding horizon of $T = 50$ sec. A single stage of the algorithm corresponds to generating a trajectory over a finite horizon of T seconds by optimizing Eq. (10). Therefore, the full trajectory is 1000 sec. STOEC is seeded with 40 sampled-trajectories.

The results are shown in Fig. 1. Since the metric for ergodicity and KL divergence both measure similarity between two distributions, we need to choose another metric to compare $\Gamma(\mathbf{x})$ of the trajectory optimized by KL-STOEC and $\xi(\mathbf{x})$, and to compare $\Gamma(\mathbf{x})$ of the trajectory optimized by Ergodic-STOEC and $\xi(\mathbf{x})$. To assess the coverage performance of the algorithms, we use Bhattacharyya distance that measures the similarity of two probability distributions. The Bhattacharyya distance is measured as,

$$D_B(\Gamma, \xi) = -\ln \left(\int_{\mathcal{X}} \sqrt{\Gamma(\mathbf{x})\xi(\mathbf{x})} d\mathbf{x} \right). \quad (22)$$

Table 1 shows the computation time⁵ for each algorithm for the example in Fig.1. We should mention that the computational time of STOEC varies depending on the number of trajectory samples that are seeded to the algorithm and the number of primitives. With around 40 sampled trajectories, the algorithm converges to a solution more quickly than SMC. It is also important to note that STOEC plans over a finite time horizon as opposed to the greedy approach in SMC. The metric for ergodicity, given by Eq. (2), gives higher weights to the low frequency components. Therefore, the algorithm favors large scale exploration of the coverage distribution. In the accompanying video of Fig. 1, it can be noticed that the metric for ergodicity given by Eq. (2) results

⁴There are no restrictions on the choice of representation for $\xi(\mathbf{x})$.

⁵The code runs on MATLAB 2016a in Windows 10 on a laptop with i7 CPU and 8 GB RAM.

TABLE I
COMPARISON OF THE COMPUTATION TIME

Method	Computing 1 stage (50 sec)	Computing the full trajectory (1000 sec)
KL-STOEC	0.5	10.0
Ergodic-STOEC	1.2	24.0
SMC	1.7	35.0

in a trajectory that frequently crosses from one region to another, whereas Eq. (13) results in a coverage behavior that gradually expands from one region to another.

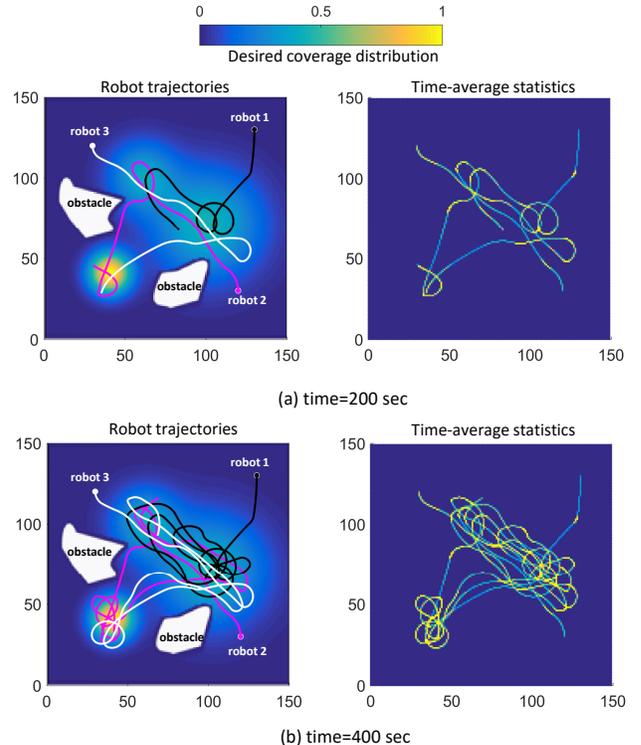


Fig. 3. Ergodic coverage results for three robots with Dirac delta sensor footprint. The left figures show the robots trajectories, and the right figures show the corresponding time-average statistics. The results are shown for two different instants of the KL-STOEC.

B. Sensor Footprints

In this section, we demonstrate that our framework can capture multi-agent systems with different sensing capabilities. It is important to note that the formulation of the ergodic coverage algorithm assumes that robots have perfect communication, thus have access to the trajectory plans of other robots. The robot trajectories are sequentially optimized by taking into account the time-average statistics of every robot's trajectory. Trajectory optimization also assumes deterministic motion models.

Fig. 3 shows the ergodic coverage of a desired distribution using KL-STOEC assuming a Dirac delta sensor footprint. We run 10 stages of the algorithm using 5 primitives with a receding horizon of $T = 40$ sec. The robot is assumed to

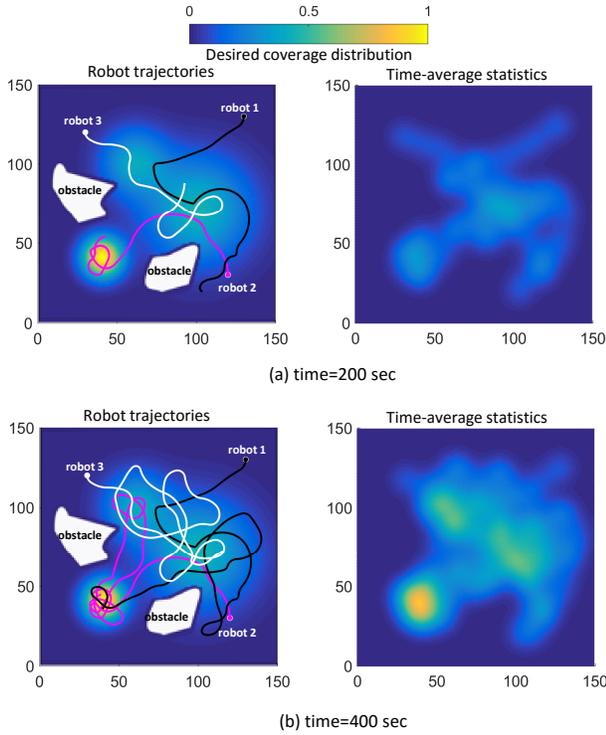


Fig. 4. Ergodic coverage results for three robots with wide Gaussian sensor footprint. The left figures show the robots trajectories, and the right figures show the corresponding time-average statistics. The results are shown for two different instants of the KL-STOEC. Large variance Gaussian sensor footprint

have bounded inputs $v \in [0.1, 5]$ m/s and $w \in [-0.2, 0.2]$ rad/s. Although the coverage focuses on the regions where the measure of the desired coverage distribution, ξ , is greater than 0, the algorithm needs a much longer time for time-average statistics to converge to the desired distribution. Fig. 4 demonstrates results with a wide Gaussian sensor footprint using the same parameters. Fig. 5 shows the errors measured using the Bhattacharyya distance. The computational time for these two scenarios in Fig. 3 and Fig. 4 are 0.7 sec/stage and 1.7 sec/stage, respectively. Fig. 6 shows another example where there are three robots equipped with different sensing capabilities. The sensor footprint of each robot is shown in Fig. 6 (c).

Note that, in these examples, the robots can continue to perform ergodic coverage of the domain indefinitely. That is, once the time-average statistics converge to the desired distribution, any new motion will result in a difference between the two distributions, $\Gamma(\mathbf{x})$ and $\xi(\mathbf{x})$. Thus, the robots will continue covering the domain such that time average-statistics again converges to the desired distribution (see the accompanying video). This is a very useful trait for surveillance applications, where the robots need to continuously monitor a region while spending more time in regions that have high values marked by the desired coverage distributions.

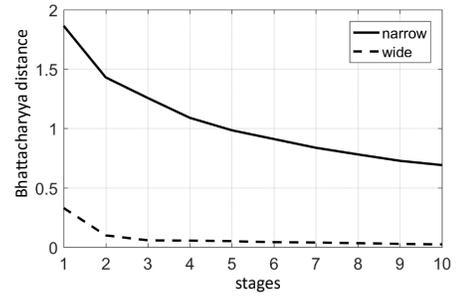


Fig. 5. Ergodic coverage performance comparison for the robots with narrow vs. wide Gaussian footprints.

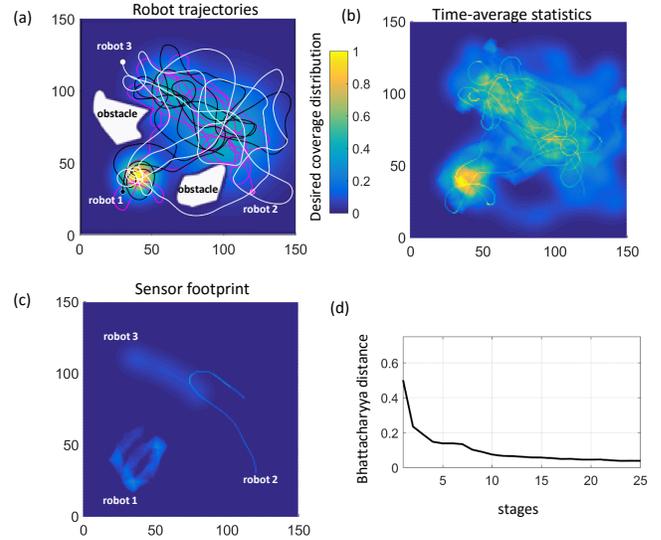


Fig. 6. Ergodic coverage results for three robots equipped with a forward-looking beam sensor footprint, a Dirac delta sensor footprint, and a wide Gaussian sensor footprint, respectively. The resulting trajectories of the robots using KL-STOEC implementation are shown in (a) for 25 concatenated $T = 40$ sec segments. The time-average statistics is shown in (b). The different footprints of the three robots are shown in (c). Finally, (d) shows the Bhattacharyya distance between the coverage distribution $\xi(\mathbf{x})$ and the time-average statistics distribution $\Gamma(\mathbf{x})$ as a function of the number of stages.

C. Point-To-Point Planning

Our algorithm can be extended to encode additional tasks beside ergodic coverage. We present an example in which the robot reaches a desired state in a given time while performing ergodic coverage along the way. This task is easily performed by adding a term to the objective function, Eq. (13), that penalizes the error between the state of the robot at the end of the horizon, $\mathbf{q}(T) = 500$ sec, and the desired state of the robot \mathbf{q}_d . Thus, the cost function is defined as,

$$C(\mathbf{z}, \mathbf{q}) = D_{KL}(\Gamma \|\xi) + \alpha \|\mathbf{q}(T) - \mathbf{q}_d\|_2. \quad (23)$$

where α is a weighting parameter, chosen heuristically, which specifies how *hard* the constraint of reaching a desired state at time T is. As α increases, the robot is guaranteed to reach its final desired state. We simulate an example of such a scenario in Fig. 7. The robot is assumed to have bounded inputs $v \in [0.1, 5]$ m/s and $w \in [-0.2, 0.2]$ rad/s.

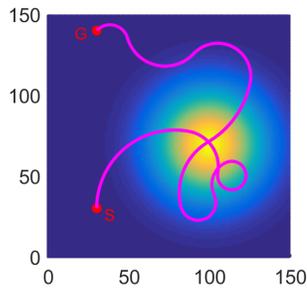


Fig. 7. The robot starts at **S**:start and is supposed to reach point **G**:goal in 500 seconds using a maximum of 10 primitives.

How to systematically choose the parameter α is an open question and is left as a future work.

V. CONCLUSIONS

This work extends the ergodic coverage algorithm presented in [8] to robots operating in constrained environments. The sampling-based trajectory optimization presented in this work allows obstacle avoidance by penalizing sampled trajectories that collide with arbitrarily-shaped obstacles or that cross restricted regions. We demonstrated that Kullback-Leibler divergence can also be used to encode an ergodic coverage objective without resorting to spectral decomposition of the desired coverage distribution, whose accuracy is limited by the number of basis functions used. Compared to [8], which is formulated as a first-order iterative optimization, our formulation allows for trajectory optimization over a receding horizon. We believe, our formulation will be of interest to the wider robotics community because it captures sensor footprints, avoids obstacles, and can be applied to nonlinear dynamic systems [15]. Future work will focus on decentralization of the ergodic coverage algorithm to address multi-agent systems with limited communication.

ACKNOWLEDGMENT

The authors would like to thank Marin Kobilarov for useful discussions and assistance in the implementation of the cross entropy method.

REFERENCES

- [1] L. Caffarelli, V. Crespi, G. Cybenko, I. Gamba, and D. Rus, "Stochastic distributed algorithms for target surveillance," in *Intelligent Systems Design and Applications*. Springer, 2003, pp. 137–148.
- [2] R. R. Murphy, S. Tadokoro, D. Nardi, A. Jacoff, P. Fiorini, H. Choset, and A. M. Erkmen, "Search and rescue robotics," in *Springer Handbook of Robotics*. Springer, 2008, pp. 1151–1173.
- [3] M. Dunbabin and L. Marques, "Robots for environmental monitoring: Significant advancements and applications," *IEEE Robotics & Automation Magazine*, vol. 19, no. 1, pp. 24–39, 2012.
- [4] R. N. Smith, M. Schwager, S. L. Smith, B. H. Jones, D. Rus, and G. S. Sukhatme, "Persistent ocean monitoring with underwater gliders: Adapting sampling resolution," *Journal of Field Robotics*, vol. 28, no. 5, pp. 714–741, 2011.
- [5] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. MIT press, 2005.
- [6] E. Acar, H. Choset, Y. Zhang, and M. Schervish, "Path planning for robotic demining: Robust sensor-based coverage of unstructured environments and probabilistic methods," *The Int. Journal of Robotics Research*, vol. 22, no. 7 - 8, pp. 441 – 466, July 2003.

- [7] F. Bourgault, T. Furukawa, and H. F. Durrant-Whyte, "Optimal search for a lost target in a Bayesian world," in *Field and Service Robotics*. Springer, 2003, pp. 209–222.
- [8] 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.
- [9] L. M. Miller and T. D. Murphey, "Trajectory optimization for continuous ergodic exploration," in *American Control Conf. (ACC), 2013*. IEEE, 2013, pp. 4196–4201.
- [10] H. Salman, E. Ayvali, and H. Choset, "Multi-agent ergodic coverage with obstacle avoidance," in *International Conference on Automated Planning and Scheduling (ICAPS), Accepted, 2017*.
- [11] F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth analysis and control theory*. Springer Science & Business Media, 2008, vol. 178.
- [12] M. Kobilarov, "Cross-entropy randomized motion planning," in *Proceedings of Robotics: Science and Systems*, Los Angeles, CA, USA, June 2011.
- [13] R. Y. Rubinstein and D. P. Kroese, *The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation and machine learning*. Springer Science & Business Media, 2013.
- [14] S. Karaman, M. R. Walter, A. Perez, E. Frazzoli, and S. Teller, "Any-time motion planning using the RRT," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1478–1483.
- [15] M. Kobilarov, "Cross-entropy motion planning," *The International Journal of Robotics Research*, vol. 31, no. 7, pp. 855–871, 2012.
- [16] K. E. Petersen, *Ergodic theory*. Cambridge University Press, 1989, vol. 2.
- [17] J. Hauser, "A projection operator approach to the optimization of trajectory functionals," *IFAC Proceedings Volumes*, vol. 35, no. 1, pp. 377–382, 2002.
- [18] R. M. Murray, M. Rathinam, and W. Sluis, "Differential flatness of mechanical control systems: A catalog of prototype systems," in *ASME international mechanical engineering congress and exposition*. Citeseer, 1995.
- [19] J. Snoek, H. Larochelle, and R. P. Adams, "Practical bayesian optimization of machine learning algorithms," in *Advances in neural information processing systems*, 2012, pp. 2951–2959.
- [20] P. J. Van Laarhoven and E. H. Aarts, "Simulated annealing," in *Simulated Annealing: Theory and Applications*. Springer, 1987, pp. 7–15.
- [21] A. Zhigljavsky and A. Žilinskas, *Stochastic global optimization*. Springer Science & Business Media, 2007, vol. 9.
- [22] G. McLachlan and D. Peel, *Finite mixture models*. John Wiley & Sons, 2004.
- [23] P.-T. De Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A tutorial on the cross-entropy method," *Annals of operations research*, vol. 134, no. 1, pp. 19–67, 2005.
- [24] G. Mathew, S. Kannan, A. Surana, S. Bajekal, and K. R. Chevva, "Experimental implementation of spectral multiscale coverage and search algorithms for autonomous uavs," in *AIAA Guidance, Navigation, and Control (GNC) Conference*, 2013, p. 5182.