

Trajectory planning for autonomous nonholonomic vehicles for optimal monitoring of spatial phenomena

Sisi Song, Abel Rodriguez, and Mircea Teodorescu

Abstract—This paper considers optimal trajectory planning for autonomous nonholonomic vehicles used in investigating environmental phenomena. In particular, we present an algorithm that generates locally optimal trajectories to find the global maximum of the underlying environmental field. Our algorithm uses Gaussian process priors to estimate the unknown field and the notion of expected improvement to develop an objective function for optimal planning. Monte Carlo simulations focusing on two-dimensional spatial fields show the advantage of our algorithm at finding the global maximum over existing methods.

I. INTRODUCTION

Autonomous vehicles are becoming the platform of choice for large-scale monitoring of spatial and spatio-temporal phenomena, replacing human-piloted ones. This switch is happening in good part due to their low cost and the increasing accuracy and dependability of off-the-shelf sensors and parts. In this context, automated trajectory planning (i.e., the determination of the path and velocity of the vehicle so that it can accomplish its mission) is a critical but difficult task, particularly in the case of uncertain environments (where trajectories cannot be pre-planned in advance but need to be determined online) and nonholonomic vehicles (in which vehicle constraints can seriously limit the number of trajectories available) [1].

Common applications of trajectory planning involve either tracking a target [2], [3] or traveling between two or more pre-specified locations [4]. In these applications, the goal is usually to achieve

the task either in the minimal amount of time or using the least amount of fuel/energy [5]. In contrast, this work focuses on trajectory planning for “optimal monitoring” of an unknown spatio-temporal field (e.g., concentration of a gas plume released by a natural process such as a volcano, or the level of soil humidity), which is observed subject to measurement error. The meaning of “optimal monitoring” can vary according to the specific application and might correspond to minimizing the variance associated with the field reconstruction [6], [7], identifying its maximum/minimum [8], [9], or identifying and following a given contour level of the field [10].

The focus of this paper is on identifying the maximum of the field, which might be of interest on its own or as a proxy for other measurements of interest (e.g., in the case of a gas plume where the source of the gas is unknown, the maximum of the field can be used as a proxy for the source). Our approach uses a statistical model based on Gaussian process priors to learn the unknown spatial field. Gaussian process priors are widely used in the spatial statistics literature [11], [12] as well as in the context of computer experiments to create a surrogate for computationally expensive computer code [13]. Gaussian process priors have also been used for the optimal design of mobile sensor networks [14], [15]. Our approach differs from this work in that we are interested in trajectory planning rather than the locations of vehicles at discrete and possibly distant times, and in that we focus on identifying the maximum of the field rather than on minimizing the uncertainty associated with its reconstruction. We have recently discussed the use of a similar model in the context of waypoint selection [16]. However, unlike the techniques described here, the approach in [16] requires a small training sample to guide waypoint selection and, more importantly, using such an approach as

This work was supported by awards DMS-1322216 and NASA/UARC NAS2-03144/TO.030.18.MD.D.

Sisi Song and Abel Rodriguez are with the Department of Applied Mathematics and Statistics, University of California Santa Cruz, Santa Cruz, CA 95064, USA song@soe.ucsc.edu, abel@soe.ucsc.edu

Mircea Teodorescu is with the Department of Computer Engineering, University of California Santa Cruz, Santa Cruz, CA 95064, USA mteodorescu@soe.ucsc.edu

part of trajectory planning is usually not practical for nonholonomic vehicles because it would often require them to follow infeasible or inefficient paths. Because of our focus on identifying the maximum of the underlying field generating the data, the optimality criterion used to construct the objective function in our planning algorithm is the expected improvement [17], [18], which has also been widely applied in the context of computer experiments [19], [20]. The expected improvement is particularly appealing because it automatically incorporates a trade-off between exploration and exploitation of the unknown field. However, the trajectory planning algorithm we describe can be easily extended to address other operational goals (such as the optimal reconstruction and optimal boundary tracking problems mentioned above) by using appropriate utility functions.

The remainder of the paper is organized as follows. Section II poses the problem and introduces our solution based on the expected improvement function. In Section III we detail our method to reconstruct the field, whose estimates are used in the expected improvement function. Section IV provides an algorithm for optimal trajectory planning. Section V presents the results of our Monte Carlo simulations, which show that even in multimodal phenomena, our algorithm is capable of identifying the global maximum. Section VI discusses extensions of the model and future work.

II. PROBLEM FORMULATION

Let $\mathbb{R}$ denote the real numbers, $\mathbb{R}_{\geq 0}$ be the nonnegative real numbers, $\mathbb{R}^d$ be the d dimensional Euclidean space, and $\mathcal{S}$ be a connected region in $\mathbb{R}^d$ that represents the space on which we are interested in monitoring the environmental field. Our main focus in this paper is on monitoring spatial phenomena that can be considered to not evolve over the time scale of interest (such as remotely monitoring soil humidity levels), so we assume throughout that $d = 2$. However, extending the model to three-dimensional time-varying settings (such as monitoring the movement of a gas plume) is conceptually straightforward (see Section VI).

In the sequel we denote $f : \mathcal{S} \rightarrow \mathbb{R}$ the unknown scalar field that we are interested in monitoring and consider an autonomous vehicle that follows

a nonholonomic trajectory $\mathbf{q} : \mathbb{R}_{\geq 0} \rightarrow \mathcal{S}$. For example, for monitoring soil humidity levels we could consider aerial vehicles, and we might require that the curvature of $\mathbf{q}$ does not exceed the maximum turning rate of the vehicle and that the speed should not be lower than the minimum flight speed. While moving, the vehicle takes noisy measurements $y_1, y_2, \dots$ of the unknown field f with sampling frequency $1/\Delta$ so that measurements are taken at locations $\mathbf{s}_1, \mathbf{s}_2, \dots$, where $\mathbf{s}_i = \mathbf{q}(i\Delta)$.

We are interested in developing algorithms that use the information about f provided by the observations $y_1, y_2, \dots$ to guide the design of the globally optimal trajectory. Because this information is acquired sequentially as the vehicle explores the region $\mathcal{S}$, we approach this problem by designing locally optimal trajectories that are revised every few steps to incorporate the most recent information collected by the vehicle. These locally optimal trajectories are then stitched together to generate $\mathbf{q}(t)$. Trajectory planning thus becomes a problem of sequential experimental design for which Bayesian design methods [21] have important advantages over classical tools. For example, Bayesian approaches to experimental design naturally deal with the exploration/exploitation trade-off [22] and, under mild conditions, their results are not influenced by the stopping rule used [23].

In the context of our problem, applying a Bayesian approach to experimental design involves the maximization of an objective function $\tilde{U}(\mathbf{q})$,

$$\tilde{U}(\mathbf{q}) = \int U(\mathbf{q}, f) p(f \mid \text{Data}) df, \quad (1)$$

where $U(\mathbf{q}, f)$ is a utility function that depends on the trajectory $\mathbf{q}$ and the underlying field f , and $p(f \mid \text{Data})$ is the posterior probability density associated with f , which captures our knowledge about the field, including that from the observations. The previous expressions implicitly assume that we treat f as a stochastic process over $\mathbb{R}$, so that for any point $\mathbf{s}$ the value of $f(\mathbf{s})$ is a random variable, and we can talk about quantities such as its expected value and its variance.

We use the improvement function as the starting point to construct our objective function. Given a finite sample of measurements $y_1, \dots, y_n$ taken at locations $\mathbf{s}_1, \dots, \mathbf{s}_n$, the expected improvement associated with a new point $\mathbf{s}_*$ is given by

$$\mathbb{E}(\max\{f(\mathbf{s}_*) - y_{\max}^n, 0\}) = \int_{y_{\max}^n}^{\infty} \{f(\mathbf{s}_*) - y_{\max}^n\} p(f(\mathbf{s}_*) | \text{Data}) df(\mathbf{s}_*), \quad (2)$$

where $y_{\max}^n = \max_{i \leq n} \{y_i\}$ is the maximum value observed so far. When considering sampling at $\tilde{n}$ new points that reside on a trajectory $\mathbf{q}$, (2) can be extended in a few different ways. In this paper we consider the maximum expected improvement over the next $\tilde{n}$ measurements, defining $\tilde{U}(\mathbf{q})$ as

$$\max_{i=1, \dots, \tilde{n}} \mathbb{E}(\max\{f(\mathbf{q}(\{n+i\}\Delta)) - y_{\max}^n, 0\}). \quad (3)$$

Another possible definition for the objective function, based on the average expected improvement, is $\frac{1}{\tilde{n}} \sum_{i=1}^{\tilde{n}} \mathbb{E}(\max\{f(\mathbf{q}(\{n+i\}\Delta)) - y_{\max}^n, 0\})$, which in our experience, tends to be more conservative in terms of exploration.

As we will discuss in Section III-C, the expected improvement automatically trades off exploration and exploitation of $\mathcal{S}$. However, when little information about the structure of f is available, directly optimizing (3) can potentially lead to trajectories that take the vehicle outside $\mathcal{S}$. To address this issue, the objective function we use in our analysis is a slightly modified version of (3) that incorporates a penalty that discourages the vehicle from going far away from $\mathcal{S}$ [24]. In particular, if a measurement y_{n+i} is taken at a location $\mathbf{s}_{n+i}$ that is far from $\mathcal{S}$, we penalize the associated expected improvement with a term that is proportional to the distance between $\mathbf{s}_{n+i}$ and the barycenter of $\mathcal{S}$, $\mathbf{s}_o$. This results in the reward function

$$R(\mathbf{q}) = \tilde{U}(\mathbf{q}) - \alpha \max_{i=1, \dots, \tilde{n}} \|\mathbf{q}(\{n+i\}\Delta) - \mathbf{s}_o\| \mathbf{1}_{\mathcal{S}_B^c}(\mathbf{q}(\{n+i\}\Delta)), \quad (4)$$

where $\alpha \geq 0$ is a tuning parameter, $\mathbf{1}_{\mathcal{S}_B^c}$ is the indicator function of the set $\mathcal{S}_B^c$, which is the complement of the smallest ball completely enclosing $\mathcal{S}$. When $\alpha = 0$ there is no penalty for exploring far away from $\mathcal{S}_B$, while as α increases the incentive to explore outside of $\mathcal{S}_B$ decreases. Our goal is to construct a trajectory $\mathbf{q}$ consisting of segments $\tilde{\mathbf{q}}_1, \tilde{\mathbf{q}}_2, \dots$ that maximize (4) sequentially.

III. FIELD ESTIMATION

Implementation of the strategy described in the previous section requires that we describe a Bayesian model for the unknown function f (i.e., a likelihood for the observations $y_1, y_2, \dots$

conditional on the unknown function f , a prior distribution for f that reflects our assumptions about the form of the underlying field before any data is collected, and hyperprior distributions on key parameters of the prior distribution for f). This section describes such a model and derives the form of the expected improvement in (2) under the corresponding posterior distribution.

A. Gaussian process model

We assume that, given the field f , the observations collected by our vehicle are independent and follow an additive model $y_i = f(\mathbf{s}_i) + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ represents the measurement noise. Furthermore, we model f as following a Gaussian process with mean function $m(\mathbf{s})$ and covariance function $\gamma(\mathbf{s}, \mathbf{s}')$, denoted by $f \sim \mathcal{GP}(m, \gamma(\mathbf{s}, \mathbf{s}'))$. This implies that for any finite set of locations $\{\mathbf{s}_1, \dots, \mathbf{s}_n\}$, the joint distribution of $(f(\mathbf{s}_1), \dots, f(\mathbf{s}_n))^T$ is normal $\mathcal{N}(\mathbf{m}, \mathbf{\Gamma})$, where $\mathbf{m} = (m(\mathbf{s}_1), \dots, m(\mathbf{s}_n))^T$ and $[\mathbf{\Gamma}]_{ij} = \gamma(\mathbf{s}_i, \mathbf{s}_j)$ for $i, j = 1, 2, \dots, n$.

In the sequel we assume that no prior information is available about the shape of f and thus assume a constant prior mean function $m(\mathbf{s}) = \beta$. This assumption is not central to our approach; if prior information is available (either from measurements obtained from other sources, such as satellites or fixed sensors, or from a computer model of the underlying phenomena), it can easily be incorporated by altering the form of m . Furthermore, in this paper we assume that f is stationary and isotropic and employ an exponential covariance function, $\gamma(\mathbf{s}, \mathbf{s}') = \tau^2 \exp\{-\|\mathbf{s} - \mathbf{s}'\|/\lambda\}$, where τ^2 is the process variance and $\lambda > 0$ is the length scale. Again, the assumption of an exponential covariance function is not key, and the approach can be modified to accommodate more general stationary families such as the power exponential or Matérn families (which provide additional parameters that can be used to control the differentiability of f), or even covariance functions that can capture nonstationarity.

Note that, if the values of the parameters β, σ^2, τ^2 and λ are known, then for any $\mathbf{s}_* \in \mathcal{S}$, the posterior distribution of $f(\mathbf{s}_*)$ given n observations $\mathbf{y} = (y_1, \dots, y_n)$ taken at locations $\mathbf{s} = (\mathbf{s}_1, \dots, \mathbf{s}_n)$ implied by the Gaussian process model is normal

with mean

$$\mu(\mathbf{s}_*) = \beta + \gamma(\mathbf{s}_*, \mathbf{s})^\top [\mathbf{\Gamma} + \sigma^2 \mathbf{I}]^{-1} (\mathbf{y} - \beta \mathbf{1}) \quad (5)$$

and variance

$$\varsigma^2(\mathbf{s}_*) = \tau^2 - \gamma(\mathbf{s}_*, \mathbf{s})^\top [\mathbf{\Gamma} + \sigma^2 \mathbf{I}]^{-1} \gamma(\mathbf{s}_*, \mathbf{s}), \quad (6)$$

where $\mathbf{I}$ is the identity matrix, $\mathbf{1}$ is a vector of ones, and $\gamma(\mathbf{s}_*, \mathbf{s}) = (\gamma(\mathbf{s}_*, \mathbf{s}_1), \dots, \gamma(\mathbf{s}_*, \mathbf{s}_n))^\top$. Hence, under squared error loss, $\mu(\mathbf{s}_*)$ provides the optimal estimate of the true value of the field f at location $\mathbf{s}_*$.

B. Optimal selection of hyperparameters

The quality of the reconstruction provided by the Gaussian process model just described strongly depends on the values of the hyperparameters β , σ^2 , τ^2 and λ . Hence, rather than fixing their values, we treat them as unknown and estimate them from the data using a quasi-empirical Bayes approach [25], [26]. More specifically, we set β , σ^2 , τ^2 , and λ to the values that maximize their joint marginal posterior distribution $p(\beta, \sigma^2, \tau^2, \lambda \mid \text{Data}) = \int p(f, \beta, \sigma^2, \tau^2, \lambda \mid \text{Data}) \, df$.

Standard applications of empirical Bayes procedures assume a flat prior for the hyperparameters, i.e., $p(\beta, \sigma^2, \tau^2, \lambda) \propto 1$. However, because flat priors are improper and there are potential identifiability issues associated with the pair (τ^2, λ) , we instead assume that the hyperparameters are independent a priori and let $p(\beta) \propto 1$, $\sigma^2 \sim \text{Inverse-Gamma}(a_1, b_1)$, $\tau^2 \sim \text{Inverse-Gamma}(a_2, b_2)$, and $\lambda \sim \text{Gamma}(a_3, b_3)$, where $a_1, b_1, a_2, b_2, a_3, b_3$ are fixed to reflect the general scale of the data. Under these priors, the logarithm of the joint posterior distribution is

$$\begin{aligned} \log p(\beta, \sigma^2, \tau^2, \lambda \mid \text{Data}) = & c - \frac{1}{2} \log |\mathbf{\Gamma} + \sigma^2 \mathbf{I}| \\ & - \frac{1}{2} (\mathbf{y} - \beta \mathbf{1})^\top [\mathbf{\Gamma} + \sigma^2 \mathbf{I}]^{-1} (\mathbf{y} - \beta \mathbf{1}) - (a_1 + 1) \log \sigma^2 \\ & - \frac{b_1}{\sigma^2} - (a_2 + 1) \log \tau^2 - \frac{b_2}{\tau^2} + (a_3 - 1) \log \lambda - b_3 \lambda, \end{aligned} \quad (7)$$

where c is a constant. The optimal values for the hyperparameters are the $\beta, \sigma^2, \tau^2, \lambda$ that jointly maximize $\log p(\beta, \sigma^2, \tau^2, \lambda \mid \text{Data})$, which are found numerically using the Nelder-Mead algorithm [27]. These values are then substituted into (5) and (6) for posterior inference about f .

We use this empirical Bayes approach for estimating the hyperparameters rather than more common Markov chain Monte Carlo (MCMC) algorithms [28] because we are interested in fast algorithms that can be used in real time applications. Since the trajectory is constructed sequentially as observations are collected, our estimate of the hyperparameters also needs to be updated online, precluding the use of computationally intensive algorithms such as MCMC.

C. The expected improvement under the Gaussian process prior

One advantage of working with Gaussian process priors for the unknown field f is that the computation of the expected improvement is relatively simple. Indeed, under our model the expected improvement (2) at any point $\mathbf{s}_* \in \mathcal{S}$ reduces to

$$\begin{aligned} \mathbb{E}(\max\{f(\mathbf{s}_*) - y_{\max}^n, 0\}) = \\ (\mu(\mathbf{s}_*) - y_{\max}^n) \Phi(z(\mathbf{s}_*)) + \varsigma(\mathbf{s}_*) \phi(z(\mathbf{s}_*)), \end{aligned} \quad (8)$$

where $z(\mathbf{s}_*) = \frac{\mu(\mathbf{s}_*) - y_{\max}^n}{\varsigma(\mathbf{s}_*)}$, and ϕ and Φ are, respectively, the density and the cumulative distribution function of the standard Gaussian distribution.

The structure of (8) highlights how the expected improvement trades off exploration and exploitation. For locations $\mathbf{s}_*$ where $\mu(\mathbf{s}_*)$ is large and $\varsigma(\mathbf{s}_*)$ is small the first term of (8) dominates, whereas for locations where $\mu(\mathbf{s}_*)$ is small and $\varsigma(\mathbf{s}_*)$ is large the second term dominates. Hence, we can think of the first term as the exploitation term, which favors points where the Gaussian process estimation suggests that we can improve with a high probability, while the second term can be interpreted as an exploration term that favors visiting locations where the uncertainty of the value of the field is large.

IV. OPTIMAL TRAJECTORY PLANNING ALGORITHM

As mentioned in Section II, our approach to constructing the optimal trajectory $\mathbf{q}$ involves sequentially computing locally optimal trajectories $\tilde{\mathbf{q}}_1, \tilde{\mathbf{q}}_2, \dots$ as batches of new observations become available. These local trajectory segments are stitched together by enforcing continuity in the position and the velocity fields at their endpoints.

Let k be the number of locations on a local trajectory. Our algorithm selects the local trajectories by looking $pk\Delta$ time units ahead and maximizing the reward in (4) associated with the potential pk new locations along each possible trajectory that satisfies the nonholonomic constraints given by the vehicle kinematics model. The vehicle then follows this locally optimal trajectory for $k\Delta$ time units, at which point the information from the most recent k measurements is incorporated into the posterior distribution for f , and the process is repeated to compute the next locally optimal segment. In other words, the vehicle plans an optimal segment over $pk\Delta$ time, follows it for $k\Delta$ time, and then plans the next optimal segment. The values of Δ and k are decided by the user on the basis of factors such as the sampling frequency of the sensors, the average speed of the vehicle, and the size of the region to be studied. $p \geq 1$ is a tuning parameter for the planning horizon. Larger values of p generate less myopic trajectory segments at the cost of longer computation time.

More formally, given $n = (j-1)k$ total observations collected during the first $j-1$ locally optimal trajectory segments, the next local trajectory $\tilde{\mathbf{q}}_j$ is obtained by maximizing (4) with $\tilde{n} = pk$ and subject to the constraints

$$\begin{aligned} c_1(\tilde{\mathbf{q}}_j) &= \tilde{\mathbf{q}}_j(n\Delta) - \tilde{\mathbf{q}}_{j-1}(n\Delta) = \mathbf{0} \\ c_2(\tilde{\mathbf{q}}_j) &= \tilde{\mathbf{q}}'_j(n\Delta) - \tilde{\mathbf{q}}'_{j-1}(n\Delta) = \mathbf{0} \\ c_3(\tilde{\mathbf{q}}_j) &= \mathbf{g}(\tilde{\mathbf{q}}'_j(t), \tilde{\mathbf{q}}''_j(t), t) \leq \mathbf{0} \end{aligned} \quad (9)$$

for all $t \in [n\Delta, (n+pk)\Delta]$. In the previous expressions, $c_1(\tilde{\mathbf{q}}_j)$ and $c_2(\tilde{\mathbf{q}}_j)$ represent the constraints on the continuity of the position and velocity fields, while $c_3(\tilde{\mathbf{q}}_j)$ represents a vector of nonholonomic constraints. The procedure just described is then applied iteratively until the stopping rule is met (e.g., fixed amount of time or the maximum expected improvement over $\mathcal{S}$ falls below a predetermined threshold). The general structure of the algorithm is outlined in Algorithm 1.

A. Constant speed circular arcs

In practice, we simplify the optimization process by parametrizing $\tilde{\mathbf{q}}_1, \tilde{\mathbf{q}}_2, \dots$ in terms of a small set of parameters $\xi_1, \xi_2, \dots$, thereby transforming the trajectory planning problem into a low-dimensional nonlinear optimization problem. As

Algorithm 1 Optimal trajectory algorithm

Input: $\mathcal{S}$, out-of-bound penalty α , vehicle's initial location and heading angle, prior distributions for hyperparameters $(\beta, \sigma^2, \tau^2, \lambda)$, length of local trajectory k , sampling rate Δ , planning horizon parameter p , nonholonomic constraints $c_3(\cdot)$, and stopping rule.

Output: Optimal trajectory $\mathbf{q}$ that finds the maximum over $\mathcal{S}$.

- 1: Initialize $j = 1$, $n = k$, and fix $\tilde{n} = pk$.
- 2: Initialize $\mathbf{q}$ by setting $\tilde{\mathbf{q}}_1$ as a straight-line trajectory (along an initial heading angle) from vehicle's initial location.
- 3: Initialize $\mathbf{y}$ with the measurements taken at $\tilde{\mathbf{q}}_1(\Delta), \dots, \tilde{\mathbf{q}}_1(k\Delta)$.
- 4: **while** stopping rule is not met **do**
- 5: Set $j = j + 1$.
- 6: Estimate $(\beta, \sigma^2, \tau^2, \lambda)$ by maximizing (7).
- 7: Obtain $\tilde{\mathbf{q}}_j$ by maximizing $R(\tilde{\mathbf{q}}_j)$ in (4) subject to (9).
- 8: Add $\tilde{\mathbf{q}}_j(\{n+i\}\Delta)$, $i = 1, \dots, k$ to $\mathbf{q}$.
- 9: Sample at $\tilde{\mathbf{q}}_j(\{n+i\}\Delta)$, $i = 1, \dots, k$.
- 10: Add new measurements to $\mathbf{y}$.
- 11: Set $n = n + k$.
- 12: **end while**

an example, consider designing a trajectory for an aerial vehicle that needs to maintain a minimum flight speed $\nu_{\min}$ and whose turning radius cannot be smaller than a constant $1/\kappa_{\max}$. A simple choice for the local trajectories is to use constant-speed circular arcs parameterized as

$$\tilde{\mathbf{q}}_j(t) = \begin{bmatrix} \xi_{j,1} \cos(\xi_{j,2}t + \xi_{j,3}) + \xi_{j,4} \\ \xi_{j,1} \sin(\xi_{j,2}t + \xi_{j,3}) + \xi_{j,5} \end{bmatrix}, \quad (10)$$

where $\xi_{j,1}\xi_{j,2} = \nu^* > \nu_{\min}$. (Note that a straight trajectory arises as a special case when the radius of curvature goes to infinity.) An advantage of using these circular trajectories is that, after continuity on the position and velocity fields are enforced, finding the optimal trajectory reduces to finding the optimal radius $\xi_{j,1}$ for the circle subject to the constraint $|\xi_{j,1}| \geq 1/\kappa_{\max}$. This is a simple univariate maximization problem that can be solved using a numerical method such as Brent's method [29]. Fig. 1 shows a sample trajectory parametrized by circular arcs as described above.

B. Variable speed circular arcs

For more generality, the constant speed circular arc parametrization can be extended to allow the aerial vehicle to fly at speeds in a range $\nu \in [\nu_{\min}, \nu_{\max}]$. In particular, we consider a parametrization of the form

$$\tilde{\mathbf{q}}_j(t) = \begin{bmatrix} \xi_{j,1} \cos\left(\frac{1}{2}\frac{\xi_{j,2}}{\xi_{j,1}}t^2 + \xi_{j,3}t + \xi_{j,4}\right) + \xi_{j,5} \\ \xi_{j,1} \sin\left(\frac{1}{2}\frac{\xi_{j,2}}{\xi_{j,1}}t^2 + \xi_{j,3}t + \xi_{j,4}\right) + \xi_{j,6} \end{bmatrix}, \quad (11)$$

which is a specific solution of the Dubin's vehicle system $q'_{j,1}(t) = \nu(t) \cos \theta_j(t)$, $q'_{j,2}(t) =$

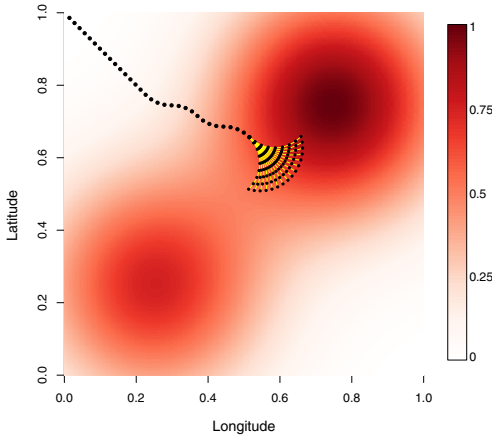


Fig. 1. Vehicle following trajectory composed straight of lines and circular arcs. Large (small) black dots show where the vehicle has taken (will possibly take) measurements. Yellow curves sweep over an area of feasible trajectories for the next $pk\Delta$ time period.

$\nu(t) \sin \theta_j(t)$, $\theta'_j(t) = u_j(t)$, where the speed $\nu(t)$ is continuous for all t , and the angular velocity control has the form $u_j(t) = \frac{\xi_{j,2}}{\xi_{j,1}}t + \xi_{j,3}$ and is piecewise continuous during each iteration of the algorithm. Note that $\xi_{j,2}$ is a linear acceleration term and (11) reduces to (10) when $\xi_{j,2} = 0$. With this parametrization, $\xi_{j,3}$ and $\xi_{j,4}$ are set by enforcing continuity of velocity, and $\xi_{j,5}$ and $\xi_{j,6}$ are set by enforcing continuity of position. This leaves optimizing over $\xi_{j,1}$ and $\xi_{j,2}$ subject to the constraints $|\xi_{j,1}| \geq 1/\kappa_{\max}$ and $\nu_{\min} \leq \|\tilde{\mathbf{q}}'(k\{j-1\}\Delta)\| + \xi_{j,2}t \leq \nu_{\max}$ for all $t \in [(j-1)k\Delta, (j+p-1)k\Delta]$, where $\|\tilde{\mathbf{q}}'(k\{j-1\}\Delta)\|$ is the speed of the vehicle at the end of the previous segment. We solve this bivariate maximization problem using a quasi-Newton method with box constraints [30].

V. SIMULATIONS

In this section we use simulations to investigate the performance of Algorithm 1 in two different scenarios. Our first experiment compares the performance of the constant speed and variable speed circular arc parametrizations and that of the biased random walk strategy presented in [9] (with a mean free path bias of 25%) in the context of a bimodal field whose global maximum is obstructed by a local maximum. Our second experiment focuses on a trimodal field and investigates the performance of our constant speed algorithm against that of a popular “lawnmower” trajectory that exhaustively explores the region of interest.

A. Constant and variable speed algorithms versus biased random walk

In this first experiment data is generated from a true field of the form

$$f_1(\mathbf{s}) = \exp \left\{ -10 \left[(s_1 - 0.75)^2 + (s_2 - 0.75)^2 \right] \right\} + 0.75 \exp \left\{ -10 \left[(s_1 - 0.25)^2 + (s_2 - 0.25)^2 \right] \right\}$$

over $\mathcal{S} = [0, 1]^2$ (see Fig. 2). The global maximum value of 1.0052 occurs at $(0.7473, 0.7473)$, and there is a local maximum at $(0.2549, 0.2549)$ with value 0.7571. The true observational noise associated with the measurements in this simulated system is set as $\sigma_T = 0.04$.

When applying our algorithm to this simulated system, we set out-of-bounds penalty $\alpha = 10^{-10}$, sampling period $\Delta = 1$, local trajectory length $k = 4$, planning horizon factor $p = 2$, maximum curvature allowed $\kappa_{\max} = 10$, constant speed $\nu^* = 0.04$, and a variable speed in the range $\nu \in [0.01, 0.04]$. The hyperpriors for σ^2 and τ^2 are based on rough estimates of the variance of the first n observations, with $a_1 = a_2 = 2$ and $b_1 = b_2 = \frac{1}{2n} \left\{ \sum_{i=1}^n y_i^2 - \frac{1}{n} \left(\sum_{i=1}^n y_i \right)^2 \right\}$. This choice reflects the assumption that, a priori, we expect about half the variability of the data to be due to measurement noise and about half to come from the variability in f . Furthermore, for the prior on λ we take $a_3 = 2$ and $b_3 = -\sqrt{2} \log(0.05)$, reflecting the assumption that, a priori, the correlation between measurements taken at the two farthest points on $\mathcal{S}$ is around 0.05. The algorithm is run for 200

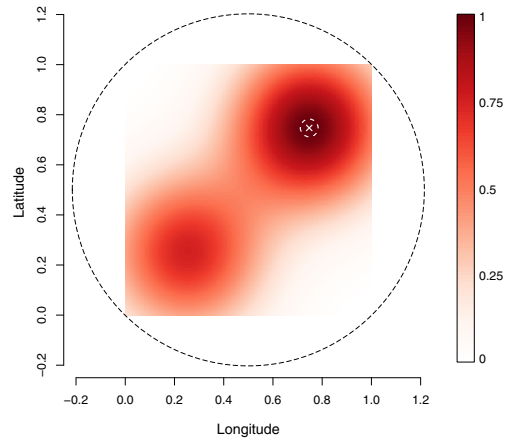


Fig. 2. True field f_1 . The white $\times$ marks the location of the global maximum of f_1 . The white circle has radius $0.025\sqrt{2}$. The outer black circle circumscribes $[0, 1]^2$ and is the boundary of $\mathcal{S}_B$.

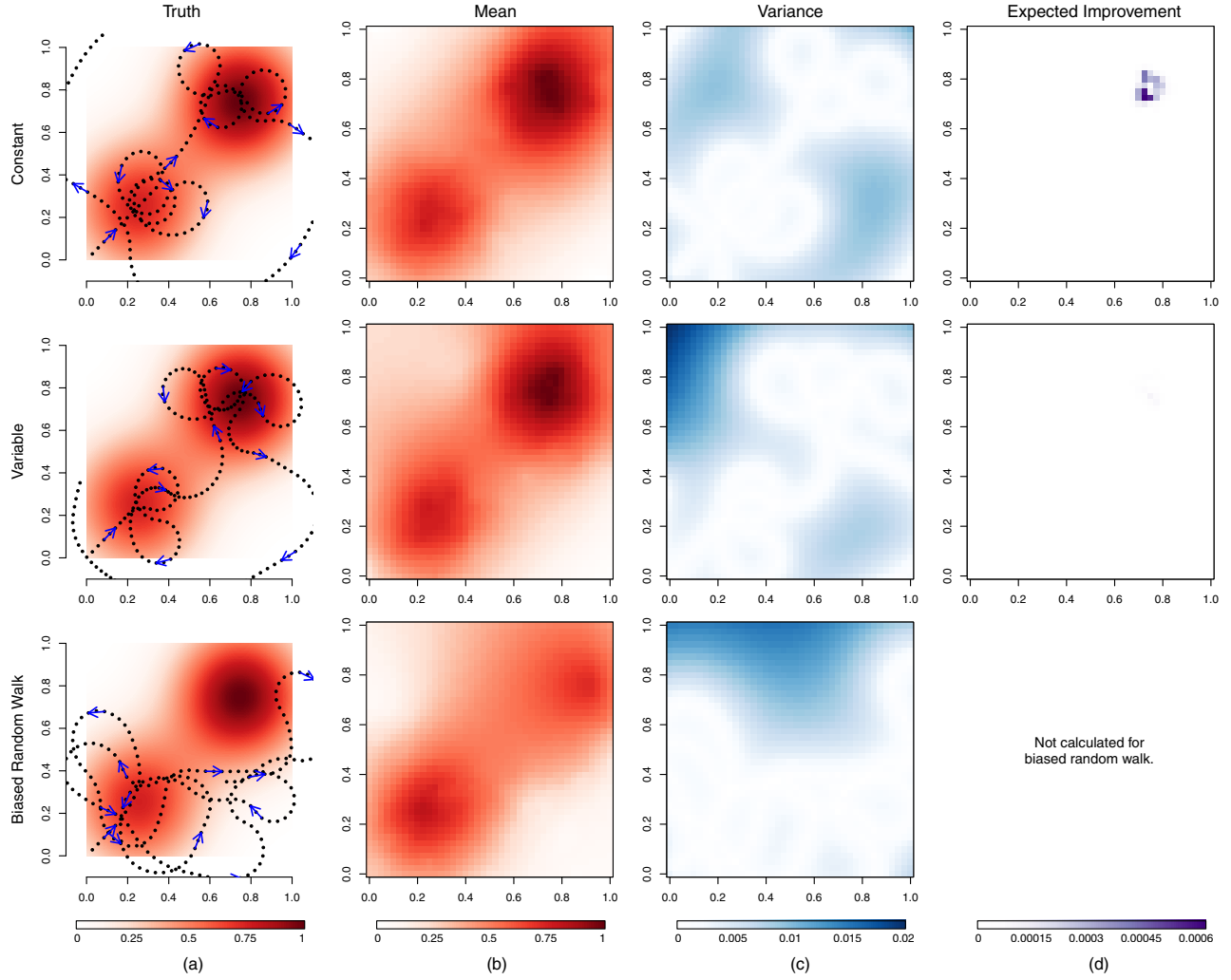


Fig. 3. For each of constant speed circular arcs, variable speed circular arcs, and biased random walk trajectories, (a) true field with snapshot of the optimal trajectory at $t = 224$; (b) the estimate of the field at $t = 224$ over a discretized 41×41 grid; (c) the variance of the estimate over the grid; (d) expected improvement evaluated over the grid.

iterations, equivalent to 800Δ time units, for a total of 800 measurements.

To evaluate the performance of our method, we run 200 simulations of each of the constant speed and variable speed cases in addition to 200 biased random walk trajectories. In every simulation the vehicle starts at the origin with a heading angle of $\frac{\pi}{4}$. We assume that no data is available at the beginning of the simulation and instruct the vehicle to go straight for $k\Delta$ time units, collecting k measurements that are used to initialize our algorithm. The code for these simulations is available from the first author by request. Fig. 3 shows examples of trajectories obtained for one of these simulations after 224 measurements have been taken (first column), as well as the posterior

mean of f_1 (which provides the current optimal reconstruction, second column), the variance of posterior distribution (which provides an estimate of the uncertainty associated with that reconstruction, third column) and the expected improvement (fourth column). We see that the trajectories generated by our constant and variable speed algorithms find and circle a few times the local maximum and then head toward and find the global maximum. By this point in time, the posterior means of f_1 computed from the constant and variable speed trajectories capture most of the features of the true field. However, the biased random walk trajectory so far has stayed mostly in the lower half of $\mathcal{S}$ and has not gotten near enough to the global maximum

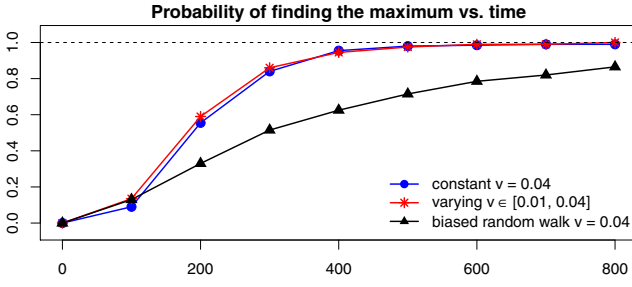


Fig. 4. The proportion of our simulations that have located the global maximum as a function of time.

to estimate the field or the location of the global maximum accurately.

A summary of the results for all simulations under the first scenario is presented in Fig. 4, which shows the probability that the algorithm has found the true global maximum as a function of flight time. For the purpose of this plot we assume that the global maximum has been accurately found if the location of the maximum of the posterior mean is inside an ε -disk centered at the true global maximum's location, and preset $\varepsilon = 0.025\sqrt{2}$. In both constant and variable speed cases, our algorithm seems to be able to find the global maximum with high probability more quickly than the biased random walk strategy. Moreover, the curves for constant and variable speed are very close. To illustrate why, we show in Fig. 5 a representative path for the optimal speed resulting from the variable speed parametrization (11). Note that the vehicle tends to fly at maximum speed when it moves *away* from the currently estimated maximum and it tends to slow down when it approaches and flies over a maximum. The result that the constant and variable speed algorithms seem to perform very similarly and that the two-dimensional optimization associated with

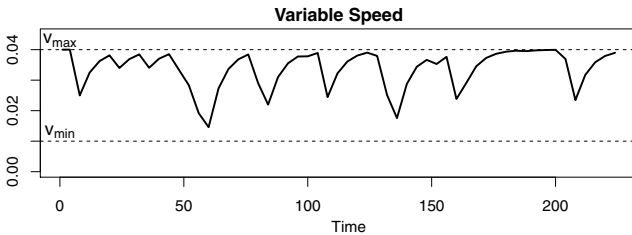


Fig. 5. An example of the speed path under variable speed parametrization in our first simulation study.

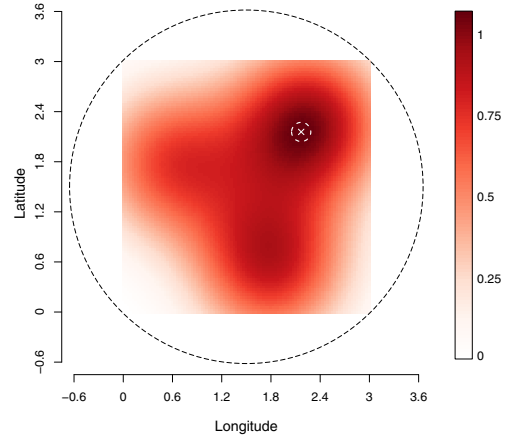


Fig. 6. True field f_2 . The white $\times$ marks the location of the maximum of f_2 . The white circle has radius $0.075\sqrt{2}$. The outer black circle circumscribes $[0, 3]^2$ and is the boundary of $\mathcal{S}_B$.

the variable speed is computationally expensive suggest that the constant speed algorithm is better suited for online trajectory generation for a physical system.

B. Constant speed algorithm versus lawnmower

In our second experiment data is generated from a trimodal true field

$$f_2(\mathbf{s}) = \exp \left\{ -\frac{[(s_1 - 2.25)^2 + (s_2 - 2.25)^2]}{0.9} \right\} + 0.85 \exp \left\{ -\frac{[(s_1 - 1.8)^2 + (s_2 - 0.6)^2]}{0.9} \right\} + 0.7 \exp \left\{ -\frac{[(s_1 - 0.6)^2 + (s_2 - 1.8)^2]}{0.9} \right\}$$

over $\mathcal{S} = [0, 3]^2$ (see Fig. 6). Of the three modes, the global maximum occurs at $(2.1685, 2.1575)$ with value 1.0721. In this simulation we compare Algorithm 1 with constant speed circular arc parametrization against the “lawnmower” trajectory that exhaustively explores the region of interest. As in our first experiment, we set $\sigma_T = 0.04$, $\alpha = 10^{-10}$, $\Delta = 1$, $k = 4$, $p = 2$, $\kappa_{\max} = 10$, $\nu^* = 0.04$, and use the same reasoning for specifying the hyperprior distributions of σ^2 , τ^2 , and λ . The algorithm is run for a total of 1224Δ time units, which is the amount of time needed for the “lawnmower” trajectory to cover the region at the same speed and maximum curvature constraint as our constant speed algorithm.

Fig. 7 shows a sample trajectory of the constant speed algorithm obtained for one of these simulations after 1224 measurements. Results from this experiment show that, without a local mode close

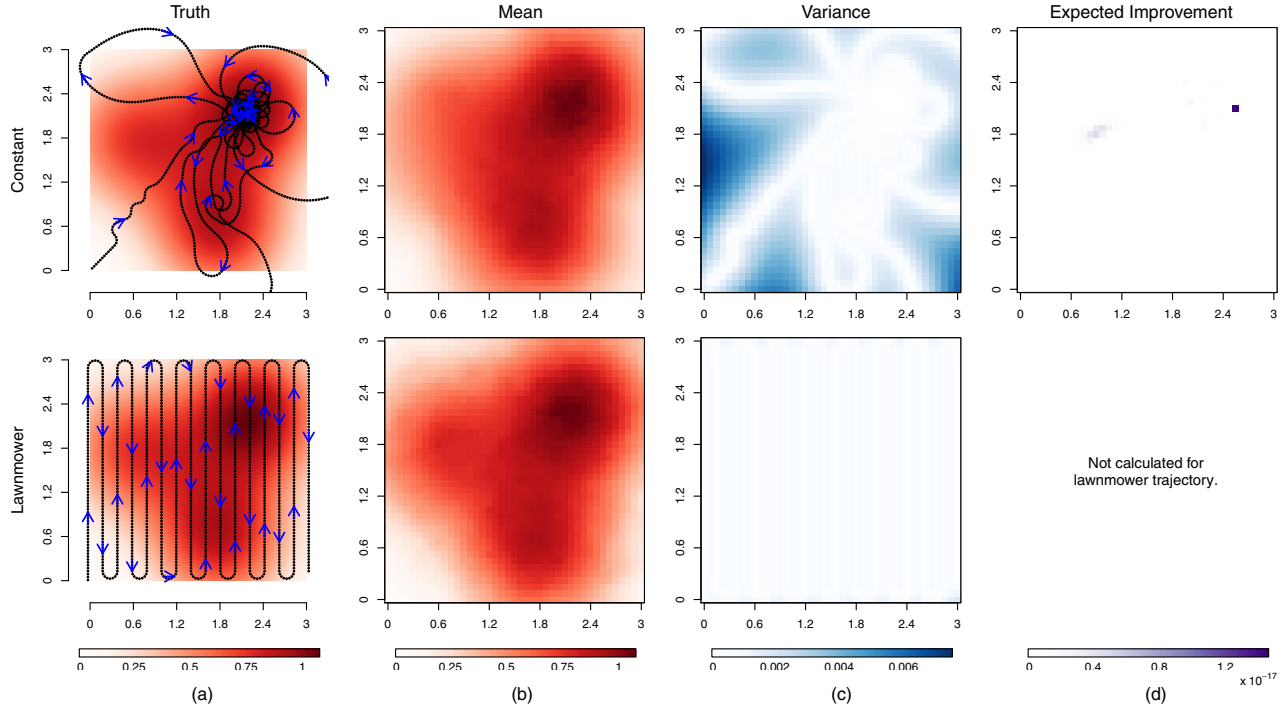


Fig. 7. For constant speed circular arcs and lawnmower trajectories, (a) true field with snapshot of the optimal trajectory at $t = 1224$; (b) the estimate of the field at $t = 1224$ over a discretized 41×41 grid; (c) the variance of the estimate over the grid; (d) expected improvement evaluated over the grid.

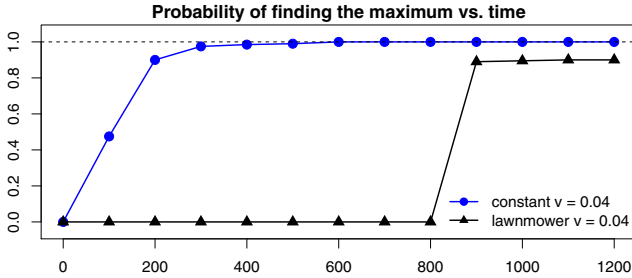


Fig. 8. Proportion of our simulations (out of 200) that have located the global maximum as a function of time.

to the origin obstructing the global maximum, our algorithm generates optimal trajectories that head towards the global maximum very quickly from the beginning. Furthermore, because the global maximum is far away from the vehicle's starting position, the lawnmower trajectory takes much longer to find it, as plotted in Fig. 8. The one advantage of the lawnmower trajectory is that its computed variance for the posterior distribution of f_2 is, on average over the region, lower than that of our algorithm's. This is due to the even measurements over the region, and so the lawnmower trajectory is perhaps better suited for different applications

of optimal monitoring of a spatio-temporal field. On a similar note, the variance around the global maximum is much lower for our algorithm because the optimally computed trajectories converge toward the global maximum and sample around there often, giving more accurate estimates of the maximum's location.

VI. CONCLUSIONS

We have developed a method for trajectory planning to find the maximum of a spatial field that seems to outperform both an exhaustive search of the space as well as search algorithms based on biased random walks. Extending our method to three dimensional and/or time varying environmental processes can be done by increasing the dimension of s and q and modifying the covariance function of the Gaussian process to reflect all dimensions. Similarly, extending to nonstationary processes can be done through nonstationary covariance functions [31]. Because the vehicle has an up-to-date reconstruction of the field, other operational goals (e.g., finding the boundary of a gas plume) can be achieved using the same basic framework

with replacement of the expected improvement by a different objective function. Another direction for future work is to involve multiple vehicles by looking into additional penalty terms in the reward function and additional constraints in the sequential optimization of the reward function.

Although the results from these and similar experiments are encouraging, it is important to acknowledge some shortcomings we noted during our simulations. Firstly, although in all of our experiments we seemed to be able to eventually find the global mode, the myopic nature of our algorithm can be an issue (especially if Δ , k , or p is too small), as the vehicle may stay close to a local mode for a long time. Also, our focus on circular trajectory segments can be a problem because the vehicle may get trapped flying in circles around the perceived mode for a while. A third direction for future work explores more general parametrizations of the trajectory segments.

REFERENCES

- [1] J. Laumond, S. Sekhavat, and F. Lamiroux, "Guidelines in nonholonomic motion planning for mobile robots," in *Robot Motion Planning and Control*, ser. Lecture Notes in Control and Information Sciences. Springer, 1998, vol. 229, pp. 1–53.
- [2] J. Lee, R. Huang, A. Vaughn, X. Xiao, J. K. Hedrick, M. Zennaro, and R. Sengupta, "Strategies of path-planning for a UAV to track a ground vehicle," in *AINS Conference*, 2003.
- [3] S.-O. Lee, Y.-J. Cho, M. Hwang-Bo, B.-J. You, and S.-R. Oh, "A stable target-tracking control for unicycle mobile robots," in *IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, vol. 3, 2000, pp. 1822–1827.
- [4] E. Anderson, R. Beard, and T. McLain, "Real-time dynamic trajectory smoothing for unmanned air vehicles," *IEEE Trans. Control Syst. Technol.*, vol. 13, no. 3, pp. 471–477, May 2005.
- [5] K. Savla, E. Frazzoli, and F. Bullo, "Traveling salesperson problems for the dubins vehicle," *IEEE Trans. Autom. Control*, vol. 53, no. 6, pp. 1378–1391, July 2008.
- [6] N. Yilmaz, C. Evangelinos, P. F. J. Lermusiaux, and N. Patrikalakis, "Path planning of autonomous underwater vehicles for adaptive sampling using mixed integer linear programming," *IEEE J. Ocean. Eng.*, vol. 33, no. 4, pp. 522–537, Oct 2008.
- [7] N. Leonard, D. Paley, F. Lekien, R. Sepulchre, D. Fratantoni, and R. Davis, "Collective motion, sensor networks, and ocean sampling," *Proc. IEEE*, vol. 95, no. 1, pp. 48–74, Jan 2007.
- [8] J. Cochran and M. Krstic, "Nonholonomic source seeking with tuning of angular velocity," *IEEE Trans. Autom. Control*, vol. 54, no. 4, pp. 717–731, April 2009.
- [9] A. Dhariwal, G. Sukhatme, and A. Requicha, "Bacterium-inspired robots for environmental monitoring," in *IEEE Int. Conf. Robotics and Automation*, vol. 2, 2004, pp. 1436–1443.
- [10] F. Zhang, E. Fiorelli, and N. Leonard, "Exploring scalar fields using multiple sensor platforms: Tracking level curves," in *IEEE Conf. Decision and Control*, Dec 2007, pp. 3579–3584.
- [11] N. Cressie, *Statistics for Spatial Data*. Wiley, 1993.
- [12] S. Banerjee, A. E. Gelfand, and B. P. Carlin, *Hierarchical Modeling and Analysis for Spatial Data*. CRC Press, 2004.
- [13] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn, "Design and analysis of computer experiments," *Statistical Science*, pp. 409–423, 1989.
- [14] R. Graham and J. Cortés, "Cooperative adaptive sampling of random fields with partially known covariance," *Int. J. Robust and Nonlinear Control*, vol. 22, no. 5, pp. 504–534, 2012.
- [15] Y. Xu and J. Choi, "Adaptive sampling for learning gaussian processes using mobile sensor networks," *Sensors*, vol. 11, no. 3, pp. 3051–3066, 2011.
- [16] S. Hening, P. Regueiro, A. Rodriguez, M. Teodorescu, N. Nguyen, and C. Ippolito, "Bayesian modeling for decentralized uav control and task allocation," in *AIAA Science and Technology Forum*, Jan 2015.
- [17] T. Wagner, M. Emmerich, A. Deutz, and W. Ponweiser, "On expected-improvement criteria for model-based multi-objective optimization," in *Parallel Problem Solving from Nature*, ser. Lecture Notes in Computer Science. Springer, 2010, vol. 6238, pp. 718–727.
- [18] E. Vazquez and J. Bect, "Convergence properties of the expected improvement algorithm with fixed mean and covariance functions," *Journal of Statistical Planning and Inference*, vol. 140, no. 11, pp. 3088–3095, 2010.
- [19] M. Frean and P. Boyle, "Using Gaussian processes to optimize expensive functions," in *AI 2008: Advances in Artificial Intelligence*. Springer, 2008, pp. 258–267.
- [20] R. Benassi, J. Bect, and E. Vazquez, "Robust Gaussian process-based global optimization using a fully Bayesian expected improvement criterion," in *Learning and Intelligent Optimization*. Springer, 2011, pp. 176–190.
- [21] K. Chaloner and I. Verdinelli, "Bayesian experimental design: A review," *Statistical Science*, vol. 10, no. 3, pp. 273–304, 1995.
- [22] N. Vlassis, M. Ghavamzadeh, S. Mannor, and P. Poupart, "Bayesian reinforcement learning," in *Reinforcement Learning*, ser. Adaptation, Learning, and Optimization. Springer, 2012, vol. 12, pp. 359–386.
- [23] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer, 1993.
- [24] E. Rimon and D. Koditschek, "Exact robot navigation using artificial potential functions," *IEEE Trans. Robot. Autom.*, vol. 8, no. 5, pp. 501–518, Oct 1992.
- [25] H. Robbins, "The empirical Bayes approach to statistical decision problems," *The Annals of Mathematical Statistics*, pp. 1–20, 1964.
- [26] C. N. Morris, "Parametric empirical Bayes inference: theory and applications," *Journal of the American Statistical Association*, vol. 78, no. 381, pp. 47–55, 1983.
- [27] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965.
- [28] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. Springer, 2005.
- [29] R. P. Brent, *Algorithms for Minimization Without Derivatives*. Courier Corporation, 1973.
- [30] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM Journal on Scientific Computing*, vol. 16, no. 5, pp. 1190–1208, 1995.
- [31] C. Paciorek, "Nonstationary gaussian processes for regression and spatial modelling," Ph.D. dissertation, Carnegie Mellon University, Pittsburgh, Pennsylvania, 2003.