On Trajectory Optimization for Active Sensing in Gaussian Process Models

Jerome Le Ny and George J. Pappas

Abstract— We consider the problem of optimizing the trajectory of a mobile sensor with perfect localization whose task is to estimate a stochastic, perhaps multidimensional field modeling the environment. When the estimator is the Kalman filter, and for certain classes of objective functions capturing the informativeness of the sensor paths, the sensor trajectory optimization problem is a deterministic optimal control problem. This estimation problem arises in many applications besides the field estimation problem, such as active mapping with mobile robots. The main difficulties in solving this problem are computational, since the Gaussian process of interest is usually high dimensional. We review some recent work on this problem and propose a suboptimal non-greedy trajectory optimization scheme with a manageable computational cost, at least in static field models based on sparse graphical models.

I. INTRODUCTION

Mobile robots with sensing capabilities are rapidly becoming a critical asset for Intelligence, Surveillance, and Reconnaissance (ISR) missions [1] and environmental monitoring [2]. Improving their sensing performance requires the development of path planning algorithms to design informative paths. In turn, defining this sensing performance requires a model of the underlying phenomenon of interest. For a number of environmental monitoring applications, such as ocean sampling [2], processing meteorological data [3] or geostatistical data [4], environment models based on collections of Gaussian random variables have long been popular, in part due to the computational tractability of the associated inference problems [5]. In robotics, probabilistic models have also been proposed to create a *stochastic map* of the environment of a robot [6]. Gaussian processes in particular can be used to build this stochastic map as well as model the uncertainty about the robot state and its sensor measurements, and form the foundation of most approaches to the Simultaneous Localization and Mapping (SLAM) problem [7] [8].

Once a stochastic environment model has been chosen, it becomes important to design trajectories for mobile sensing platforms allowing them to take measurements that provide as much information as possible about the hidden parameters of the model. This design problem is close to certain problems studied under the names of optimal design of experiments, active learning, or sensor management [9]–[11]. The main complication introduced by mobile robotic networks that is mostly absent from the classical sensor management

problems, is that one can only indirectly control the sensor configurations by modifying the robot state. Concretely, whereas in radar management problems for example one can assume that a sensor can switch instantaneously between different targets of interest [12], for mobile robots we need to design a path in the environment that is feasible for the robot, a problem that in itself is the object of significant research [13]. Nonetheless, the trajectory optimization problem to improve the estimate of a stochastic map (potentially coupled with the robot localization problem to obtain the *active SLAM* problem) has received significant attention in recent years [2], [10], [14]–[36]. We review the literature that is closest to our problem formulation in later sections.

This paper is organized as follows. Section II describes a general model of the environment based on Gaussian Markov random fields. Such an environment model typically involves a large number of variables. We review some important facts about the consequence of natural *conditional* independence assumptions about these variables on the sparsity pattern of the information matrix of the field (the inverse of the covariance matrix). Section III introduces the mobile sensor model and a class of objective functions which allows the formulation of the sensor trajectory optimization problem to be formulated as a *deterministic* optimal control problem. This feature depends critically on the properties of the Kalman filter and has been exploited since the early work on sensor management problems [37]. In section IV, we describe possible approaches to the trajectory optimization problem. We illustrate the drawbacks of the greedy heuristic, which is the most frequently used, in numerical simulations presented in section IV-E. To improve on this heuristic, we note that the deterministic nature of the optimal control problem allows us to solve the problem using a value iteration scheme progressing forward, which has critical computational advantages in such high-dimensional state spaces. We propose a heuristic to reduce the dimension of the space over which we execute the dynamic programming algorithm to the dimension of the robot state space. We provide some analysis of this heuristic and briefly show its performance in the numerical simulations.

II. ENVIRONMENT MODELING

A. Gaussian Markov Random Fields

In this paper, we model the environment in which the robot (also described hereafter as the mobile sensor) evolves as a time-varying Gaussian Markov random field (GMRF) $\{\phi_t(s)\}_{s \in D}, t = 0, 1, \ldots, t_f$, where t_f is the time horizon, and D is a discrete set. Hence for each t, ϕ_t is a Gaussian

vector, and the components $\phi_t(s)$ for $s \in D$ are also allowed to be vector valued. Typically, the partition specified by the set D is related to certain conditional independence assumptions made in the model, as explained below. In the sensing applications mentioned in the introduction, s typically represents a position on a spatial grid, whereas in a SLAM problem it can be the index of an obstacle or "feature" in the environment. In the latter case $\phi_t(s)$ could then be the position and orientation in the plane of a moving-obstacle at time t, a 3-dimensional vector.

In some fields of spatial statistics, for example geostatistics, weather prediction or ocean modeling, models using Gaussian processes (GPs) with D a continuous space arise frequently, see e.g. [4], [38], [39], and have been used in the context of active sensing with mobile robots [2], [22], [24], [28], [35]. However, the model is usually reduced to a finite-dimensional model similar to ours by considering a discrete set of locations of interest, including a discrete set a measurements along the robot trajectories. Moreover, even if the basic environment model is specified in terms of a continuous space GP, there are known techniques to approximate it by a discrete GMRF [40], [41]. A potential drawback of the approach based on continuous-space GPs however is that they are specified using covariance functions, whereas information matrices (the inverses of covariance matrices) have often useful properties for modeling (using the graphical model formalism [5]), numerical analysis [41], [42] and sensor fusion in multi-sensor systems [43]. Note finally that some methods in spatial statistics which have been used in the context of robotic sensing networks, such as Kriging [2], [4], [24], [38], do not make any Gaussian assumption about the field but require the estimator to be linear in the measurements. In this case the filtering equations are identical to the ones used in this paper for the optimal filter with the Gaussian assumption (Kalman filter).

The dynamics of the environment model are assumed to be linear

$$\phi_{t+1} = A_t \phi_t + w_t, \qquad (1)$$

where A_t is a matrix of appropriate dimensions, and the noise w_t is Gaussian with covariance matrix $E(w_t w_t') = W_t\, \delta(t - t')$. We assume here that the mean of $[\phi_t(s)]_{s \in D}$ for t fixed is known and taken without loss of generality to be 0. Indeed adding a known deterministic process to the model (1) has no influence on the path planning problem *if the performance measure is intended to only reflect the estimate uncertainty*, as will be the case in this paper. The initial covariance matrix $\Sigma_0 := \mathrm{Var}(\phi_0)$ is known.

B. Basic Properties of GMRFs

Fixing the time $t = 0$, the vector $\{\phi_0(s)\}_{s \in D}$ is jointly Gaussian $\phi_0 \sim N(0, \Sigma_0)$. Assuming that Σ_0 is nonsingular, we can work instead with the *information matrix* $\Omega_0 = \Sigma_0^{-1}$. We drop the subscript indicating time $t = 0$ in this section, so that we write ϕ, Σ, Ω. An important fact, which does not play a role in the formulation of the problem below but only in its numerical solution, is that problem modeling

using GMRFs tends to involve information matrices Ω that are *sparse* (whereas Σ is not) if the interactions between features of the environment are local. The origin of this assumption is based on the following facts. We capture the structure of the Gaussian density $p(\phi) \propto \exp\{-\frac{1}{2}\phi^T \Omega \phi\}$ by an *undirected graphical model*. This is an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, i.e., with set of vertices $\mathcal{V}$ and set of edges $\mathcal{E}$, with $|\mathcal{V}| = |D|$ and where vertex s is associated to the subvector $\phi(s)$. For a GMRF this graph captures certain *conditional* independence assumptions, namely, conditional independence is associated with graph separation. If R, S, T are subsets of $\mathcal{V}$ and S separates R and T (i.e., there is no path between R and T that does not pass through S), then $\phi(R)$ and $\phi(T)$ are conditionally independent given $\phi(S)$ (here $\phi(A)$ for $A \subset \mathcal{V}$ denotes the vector $\{\phi(s)\}_{s \in A}$). The neighborhood of a node s of the graph is defined as $\mathcal{N}_s = \{t \in \mathcal{V} | (s,t) \in \mathcal{E}\}$. Hence we have immediately the Markov property $p(\phi(s)|\phi(\mathcal{V} \setminus \{s\})) = p(\phi(s)|\phi(\mathcal{N}_s))$. We use the notation $\phi(-A) := \phi(\mathcal{V} \setminus \{A\})$ for $A \subset V$, and simplify to $\phi(-s) := \phi(-\{s\})$ if $A = \{s\}$. So more generally, if we let $\mathcal{N}_A = \bigcup_{s \in A}(\mathcal{N}(s) \setminus A)$ be the neighborhood of the subset A of nodes, then $p(\phi(A)|\phi(-A)) = p(\phi(A)|\phi(\mathcal{N}_A))$. Hence we can use the graph formalism to specify the conditional independence relationships of a set of Gaussian random variables. Moreover, these relationships are reflected in the sparsity pattern of the *information* matrix Ω (see e.g. [44], [5, prop. 2.1]). To see why, consider for simplicity the case where $\phi(s), \phi(t)$ are scalars. Then the conditional canonical correlation coefficient of $\phi(s)$ and $\phi(t)$, conditioned on their local neighborhood $\phi(\mathcal{N}_{\{s,t\}})$ is

$$
\begin{aligned}
\rho_{st|\mathcal{N}_{\{s,t\}}} &= \frac{\mathrm{cov}[\phi(s), \phi(t)|\phi(\mathcal{N}_{\{s,t\}})]}{\sqrt{\mathrm{Var}[\phi(s)|\phi(\mathcal{N}_{\{s,t\}})]\mathrm{Var}[\phi(t)|\phi(\mathcal{N}_{\{s,t\}})]}} \\
&= \frac{-\Omega^{st}}{\sqrt{\Omega^{ss}\Omega^{tt}}}, \qquad (2)
\end{aligned}
$$

where in general we denote by Ω^{st} the block (s,t) of the matrix Ω (and Ω^{st} is a scalar for $\phi(s)$ and $\phi(t)$ scalars). A similar formula holds for multidimensional variables $\phi(s), \phi(t)$ [5]. If $(s,t) \notin \mathcal{E}$ then $\mathcal{N}_{\{s,t\}}$ separates s and t, which therefore must be conditionally independent, and this is equivalent to $\Omega^{st} = 0$ by (2). Hence conditional independence assumptions, typically made based on the local nature of certain environmental phenomena, translate directly into a sparse information matrix.

Because ultimately the numerical procedures for inference and trajectory optimization involve computations with the covariance or information matrix, sparsity patterns in these matrices are important. The fact that *covariance matrices are not sparse in general* can be seen with the simple example of a one-dimensional Gaussian autoregressive process $\phi(0) = 0$, $\phi(s+1) = \lambda\phi(s) + w(s)$, where the $w(s)$ are independent and identically distributed standard normal random variables and $|\lambda| < 1$. It is desirable that such a simple model captures some notion of local interactions in the one-dimensional field. Indeed the associated graphical model is just a line graph and so $\Omega^{ij} = 0$ for $|i - j| \geq 2$, i.e., the information

matrix is tridiagonal. However, the covariance matrix is full, namely we have $\Sigma^{ij} = \frac{1}{1-\lambda^2}\lambda^{|i-j|}$. One can try to truncate this covariance matrix [24], [35], [45], but the probabilistic interpretation of this approach is not clear, and one must then be careful to maintain positive-definiteness of the sparse version of the covariance matrix.

Finally, this section has focused on the case of static fields, and numerical examples are provided in section IV-E. An important research direction is to understand how to maintain sparse models for time-varying fields [42].

III. PERFORMANCE CRITERION FOR TRAJECTORY OPTIMIZATION

The (possibly time-varying) Gaussian stochastic map described in the previous section is a model of the environment in which our mobile measurement system (or robot) evolves. This measurement system is a controlled dynamical system $x_{t+1} = f_t(x_t, u_t)$, where the state x_t is assumed perfectly known in this paper. Potentially, the state x_t can be partitioned to represent the state of several robots. There is a control signal $u_0, \ldots, u_{t_f-1}$ available to steer the vehicle. The robot can take measurements $y_{1:t_f} := [y_1^T, \ldots, y_{t_f}^T]^T$, and we assume a measurement model which is *linear in the state of the environment*: $y_t = C_t(x_t)\phi_t + v_t$, where the noise v_t is Gaussian with covariance matrix $E[v_t v_{t'}] = V_t(x_t)\delta(t - t')$. Here the matrices $C(x_t)$ and $V(x_t)$ can depend nonlinearly on the state of the robot, and typically reflect the fact that the robot can sense only the part of the map that is close to its position. The random vector ϕ_t given a set $y_{1:t}$ of these linear measurements is still Gaussian with information matrix $\Omega_t = \Sigma_t^{-1}$, which can be computed recursively using the information filter

$$\bar{\Omega}_t = (A_{t-1}\,\Omega_{t-1}^{-1}\,A_{t-1}^T + W_{t-1})^{-1} \quad (3)$$

$$\Omega_t = \bar{\Omega}_t + C_t(x_t)^T V_t(x_t)^{-1} C_t(x_t). \quad (4)$$

Note that $\bar{\Omega}_t$ is the conditional information matrix of ϕ_t given the measurements $y_{1:t-1}$. The conditional mean of ϕ_t given $y_{1:t}$ can also be computed recursively via the Kalman filter equations but will not play a role in this paper.

We would like to maximize some performance measure reflecting the quality of the measurements taken by the mobile sensor. There are a variety of design criterions proposed in the literature on optimal experiment design [10] and for concreteness we choose one of them which is theoretically appealing. We wish to maximize the *mutual information* [46] between the field state and the measurements. Recall that the entropy of a normal random vector of dimension n, with normal distribution $N(\mu, \Sigma)$ with mean μ and covariance matrix Σ, is given by

$$H(N(\mu, \Sigma)) = \frac{n}{2}\ln(2\pi e) + \frac{1}{2}\ln|\Sigma|, \quad (5)$$

where $|\Sigma|$ denotes the determinant of Σ (hence entropy is independent of the mean of the Gaussian). Also, if x is Gaussian with distribution $N(\mu, \Sigma)$, $\Omega = \Sigma^{-1}$ and $z =$

$Cx + v$ with v Gaussian $N(0, V)$ independent of x, we can write the mutual information between x and z as

$$I(x; z) = \frac{1}{2}\ln\frac{|\Omega + C^T V^{-1} C|}{|\Omega|}. \quad (6)$$

Then maximizing the mutual information between the field state over time and the measurements along a trajectory of the robot is maximizing $I(\phi_{1:t_f}; y_{1:t_f})$, with $\phi_{1:t_f} := [\phi_1^T, \ldots, \phi_{t_f}^T]^T$. We could more generally replace ϕ_t by $\phi_t(L)$, for $L \subset D$ a subset of the field locations. By our Markov assumptions and the chain rule of mutual information, we can write (see e.g. [11, p.43]) $I(\phi_{1:t_f}; y_{1:t_f}) = \sum_{t=1}^{t_f} I(\phi_t; y_t | y_{1:t-1})$. Now using (5) or (6) we get

$$I(\phi_t; y_t | y_{1:t-1}) = H(\phi_t | y_{1:t-1}) - H(\phi_t | y_{1:t})$$
$$= \frac{1}{2}\left(\ln|\bar{\Omega}_t + C_t(x_t)^T V_t^{-1}(x_t) C_t(x_t)| - \ln|\bar{\Omega}_t|\right).$$

Hence

$$I(\phi_{1:t_f}; y_{1:t_f}) = \frac{1}{2}\sum_{t=1}^{t_f}\left(\ln|\bar{\Omega}_t + C_t^T V_t^{-1} C_t| - \ln|\bar{\Omega}_t|\right). \quad (7)$$

In particular if the field is static ($A_t = I, W_t = 0$ for all t), we have $\bar{\Omega}_t = \Omega_{t-1}$ and so

$$I(\phi_{1:t_f}; y_{1:t_f}) = I(\phi_0; y_{1:t_f}) = \frac{1}{2}(\ln|\Omega_{t_f}| - \ln|\Omega_0|).$$

Other measures of information can be used, such as the trace of the covariance matrix [10], [17]. More generally, we can try to optimize an objective function of the form

$$\max J(\Omega_0, x_0) = \sum_{t=0}^{t_f-1} r_t(\Omega_t, x_t, u_t) + R(\Omega_{t_f}, x_{t_f}). \quad (8)$$

over all control policies $u_0, \ldots, u_{t_f-1}$. Here $r_t(\Omega, x, u)$ and $R(\Omega, x)$ are immediate and terminal rewards respectively.

IV. CLOSING THE LOOP

A. Active Mapping

We can now consider the active mapping problem with perfect localization. Most of the related literature in robotics focuses on the filtering problem in the presence of measurements nonlinear in the state of the map [42], and on the robot localization problem in the presence of uncertain dynamics, which are not considered here. The design of the control inputs is usually not considered and assumed to be provided by an operator. Even for a robot with perfect localization and linear measurements however, the design of a control policy which optimizes an information criterion such as (7) is a challenging task. Given an initial information matrix Ω_0 and an initial state for the robot x_0, the optimization problem to solve is of the form

$$\max_\pi J(\Omega_0, x_0) = \sum_{t=0}^{t_f-1} r_t(\Omega_t, x_t, u_t) + R(\Omega_{t_f}, x_{t_f}) \quad (9)$$

$$\text{subject to } \Omega_t = (A_{t-1}\,\Omega_{t-1}^{-1}\,A_{t-1}^T + W_{t-1})^{-1}$$
$$+ C_t(x_t)^T V_t(x_t)^{-1} C_t(x_t) \quad (10)$$

$$x_t = f_t(x_{t-1}, u_{t-1}), \Omega(0) = \Omega_0, x(0) = x_0, \quad (11)$$

where the optimization is over all control policies $\pi = u_1, \ldots, u_{t_f-1}$. Note that (9) is a *deterministic* optimal control problem. This is a consequence of our assumption that the stochastic map dynamics are linear Gaussian, that the measurements are linear in ϕ_t with Gaussian noise, and of the fact that our objective does not depend on the actual estimated value of the field but only on the information (or covariance) matrix. This fact has been exploited since the early work on sensor management [9], [37]. Note that in robotics, except for the SLAM problem, maps are often based on "occupancy grids" [8, chap. 9], where the presence of obstacles is modeled as binary random variables. In this framework, active mapping requires solving a *stochastic* control problem (say, using POMDP techniques [47]), which is potentially more difficult.

The continuous time problem corresponding to the mapping problem (9) is

$$\max_{u(\cdot)} J(\Omega_0, x_0) = \int_{t=0}^{t_f} r(t, \Omega, x, u)dt$$
$$+ R(\Omega(t_f), x(t_f)) \quad (12)$$

subject to

$$\dot{\Omega} = -\Omega A - A^T \Omega - \Omega W \Omega + C(x)^T V^{-1}(x)C(x), \quad (13)$$
$$\dot{x} = f(t, x, u), \Omega(0) = \Omega_0, x(0) = x_0,$$

where the continuous-time Riccati equation (13) replaces the discrete-time version (10).

B. The Greedy Policy

In order to solve the active mapping problem (9) (or its continuous time version), the greedy policy is often used, see e.g. [16], [18], [23], [25], [35]. The greedy policy consists in choosing in state (Ω, x) at time t the control u_t which maximizes the immediate reward $r_t(\Omega, x, u)$ in (9). For the objective (7), the greedy policy therefore is chosen as

$$u_t(\Omega, x) \in \arg\max_{u \in U_t(x)} \left(\ln |\bar{\Omega} + C_{t+1}^T(u)V_{t+1}^{-1}(u)C_{t+1}(u)| \right.$$
$$\left. - \ln |\bar{\Omega}| \right)$$

where $\bar{\Omega} = (A_{t-1}\Omega^{-1}A_{t-1}^T + W_{t-1})^{-1}$ and by a slight abuse of notation $C_t(u) := C_t(f_t(x, u))$ and $V_t(u) := V_t(f_t(x, u))$. The set $U_t(x)$ denotes the controls available at time t when the robot is in state x. Grocholsky [16] uses a continuous-time version of the greedy policy for this objective function. Other heuristic approaches have been proposed. For example in [2], instead of using the greedy policy, the sensing vehicles are constrained to move along ellipses whose parameters are optimized.

C. A Solution using Forward Value Iteration

Sim and Roy [17] have discussed the benefits of developing better policies than the greedy policy for the active SLAM problem, and the same is true for the simpler mapping problem (9) with perfect robot localization. In [29], the author proposes to use the maximum principle to solve the continuous time problem (12) (using the formulation

in terms of covariances). This requires solving a two-point boundary value problem of the size of the joint state (Ω, x), an approach which is unlikely to scale to realistic field models. Moreover, the procedure is not guaranteed a priori to converge to a global optimum. In [34], [36] the authors parametrize the control history using a piecewise linear function and optimize the parameters by solving a nonlinear program. Note that optimal control approaches for the simplified version of this problem which arises in sensor management has a long history, see e.g. [9].

In general, the main bottleneck for computations is the large dimension of the state Ω_t in (9). Suppose that ϕ_t is a d dimensional vector, then $\Omega_t \in \mathbb{R}^N$, $N = \frac{d(d+1)}{2}$. In discrete time, the simplest approach to solving (9) is to use dynamic programming and in particular backward value iteration. This would require discretizing this N dimensional state space, which is not practical for most purposes. For example, considering a scalar field over a relatively small grid of 50×50 points, we have $d = 2500$, $N > 3 \cdot 10^6$ and if each real valued entry of the information matrix is quantized over n bits, the size of the discrete state space used for the computations is then $(2^n)^N$! Even if Ω_t turns out to be sparse and belongs in fact to a subspace of dimension of order d (i.e., in section II-B each node in the graph has a small number of neighbors), we cannot hope to handle realistic problems by general numerical methods simply relying on discretization. However, we can take advantage of the fact that (12) is a deterministic problem to solve it instead using an alternative version of value iteration which progresses forward in time [13], [48], and which is not available to solve stochastic control problems. Alternatively, this forward version of value iteration could also be replaced by general purpose shortest path algorithms [48]. This approach has in fact been used together with additional heuristics in the active SLAM problem [17] and in a problem complementary to (9) which considers the robot localization problem in a perfectly known environment [49], [50]. As soon as robot localization is introduced however, the problem becomes dependent on the actual measurement values obtained by the sensors, and we face a stochastic optimal control problem. In [17], [49], [50], the problem is approximated and reduced to a deterministic problem by neglecting the deviations of the vehicle from its mean trajectory for the purpose of planning. It is important therefore to develop first good solution methods for the simpler problem (9), which can later serve as base heuristics for more complicated scenarios, in particular involving uncertain robot dynamics.

The main advantage of the forward version of value iteration for this problem is that we can build the set of reachable information matrices progressively starting from the initial state, and consider only this reachable subspace, instead of discretizing the space of all information matrices as the backward version of the algorithm would require. The algorithm constructs iteratively a graph containing the reachable space. We start with the initial node corresponding to the initial state (x_0, Ω_0). From this node, there are arcs for each possible control values u_0, which lead to nodes

corresponding to the state values (x_1, Ω_1) reachable at time $t = 1$. The graph can be organized in stages, with one stage for each time period. Starting from each node corresponding to a reachable value of (x_t, Ω_t), there is one arc for each control value u_t leading to a node of stage $t+1$ corresponding to the value (x_{t+1}, Ω_{t+1}) obtained from the dynamics (10), (11). Let us denote these dynamics more succinctly by the difference equation $(\Omega_t, x_t) = F_t(\Omega_{t-1}, x_{t-1}, u_{t-1})$. Let S_t be the set of states (Ω, x) reachable at time t from (Ω_0, x_0) by some control policy. Also, let $U_t(x_t)$ be the set of available controls at time t when the robot is in state x_t, and let $U_t = \cup_{x \in S_t} U_t(x)$. Next, define for all $(\Omega, x) \in S_t$

$$F_t^-(\Omega, x) = \{(\hat{\Omega}, \hat{x}, \hat{u}) \in S_{t-1} \times U_{t-1} | F_t(\hat{\Omega}, \hat{x}, \hat{u}) = (\Omega, x)\},$$

the set of state-control pairs which can reach (Ω, x). These sets can be built while building the graph above. The forward value iteration algorithm computing the optimal value of $J(\Omega_0, x_0)$ is described in algorithm 1.

Algorithm 1 Forward Value Iteration

$V_0(\Omega_0, x_0) = 0$
for all $u \in U_0(x_0)$ **do**
  $S_1 \leftarrow S_1 \cup \{F_0(\Omega_0, x_0, u_0)\}$
end for
for all $1 \leq t \leq t_f - 1$ **do**
  for all $(\Omega, x) \in S_t$ **do**
    for all $u \in U_t(x)$ **do**
      $S_{t+1} \leftarrow S_{t+1} \cup \{F_t(\Omega, x, u)\}$
    end for
    $V_t(\Omega, x) = \max_{(\hat{x}, \hat{\Omega}, \hat{u}) \in F_t^-(\Omega, x)} \{r_{t-1}(\hat{\Omega}, \hat{x}, \hat{u}) + V_{t-1}(\hat{\Omega}, \hat{x})\}$
  end for
end for
for all $(\Omega, x) \in S_{t_f}$ **do**
  $V_{t_f}(\Omega, x) = \max_{(\hat{x}, \hat{\Omega}, \hat{u}) \in F_{t_f}^-(\Omega, x)} \{r_{t_f - 1}(\hat{\Omega}, \hat{x}, \hat{u}) + R(\Omega, x) + V_{t_f - 1}(\hat{\Omega}, \hat{x})\}$
end for
$J(\Omega_0, x_0) = \max_{(\Omega, x) \in S_{t_f}} V_{t_f}(\Omega, x).$

D. Heuristic Pruning in the Forward Value Iteration Algorithm

Even though the forward value iteration algorithm avoids considering the part of the state space that is not reachable, its complexity is still too high for most problems and all but very short time horizons. Indeed, the conditional information matrix Ω_t obtained at period t by a robot reaching state x_t depends essentially on the whole path $x_0, x_1, \ldots, x_t$, not just on the value of x_t. Hence the number of nodes (Ω, x) at stage t grows essentially as the number of feasible paths of length t in the state space of the robot. We can reduce the complexity of the algorithm by considering the following heuristic, which reintroduces a degree of myopia in the algorithm. At stage t of the forward value iteration algorithm, for a given value of the robot state $x_t = \tilde{x}$, we discard all state values

$(\Omega, \tilde{x}) \in S_t$ except the one which achieves the maximum of $V_t(\Omega, \tilde{x})$ over the different information matrices Ω which can be obtained by the different robot paths terminating in state $\tilde{x}$ at time t. This way, we ensure that the number of nodes in the graph at stage t remains bounded by the number of reachable vehicle state values x_t. Denote by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ the graph built by the complete forward value iteration algorithm. That is, the nodes $\mathcal{V}$ at stage t correspond to the reachable values of the state (Ω_t, x_t), and the arcs $\mathcal{E}$ to the available control actions. Let $\hat{\mathcal{G}} = (\hat{\mathcal{V}}, \hat{\mathcal{E}})$ be the pruned graph. A node in $\hat{\mathcal{G}}$ at stage t can be represented simply by the value of the robot state x_t.

In the rest of this section, we discuss the performance of the pruning heuristic in more details. Consider a path $\mathcal{P} = (\Omega_0, x_0), (\tilde{\Omega}_1, \tilde{x}_1), \ldots, (\tilde{\Omega}_{t_f}, \tilde{x}_{t_f})$ in the graph $\mathcal{G}$. We let $t_0 = 0$ and define recursively the *blocking times* for this path to be,

$$t_k = \min\{t : t_f \geq t > t_{k-1} \text{ and } (\tilde{x}_{t-1}, \tilde{x}_t) \notin \hat{\mathcal{E}}\}, \quad (14)$$

for all $k \geq 1$ for which the set is nonempty. That is, t_k is the kth time where we find that there is no edge from $\tilde{x}_{t-1}$ to $\tilde{x}_t$ in $\hat{\mathcal{G}}$. This can happen when there is another state $(\Omega'_t, \tilde{x}_t)$ in $\mathcal{G}$ such that $V_t(\Omega'_t, \tilde{x}_t) > V_t(\tilde{\Omega}_t, \tilde{x}_t)$ so that the state $(\tilde{\Omega}_t, \tilde{x}_t)$ is removed from $\hat{\mathcal{G}}$. Let m be the maximum index k for which t_k is defined, and note that $m \leq t_f$.

Let us now consider the specific objective function (7), i.e., we wish to maximize $I(\phi_{1:t_f}; y_{1:t_f})$. We define the following notation. First, the objective function is completely determined by the trajectory of the robot $x_{0:t_f} := x_0, x_1, \ldots, x_{t_f}$. For simplicity of notation, we write $I(x_1, \ldots, x_{t_f}) := I(\phi_{1:t_f}; y_{1:t_f})$, for the measurements $y_{1:t_f}$ obtained along the trajectory $x_{0:t_f}$. Similarly, we write for the conditional mutual information $I(x_{t+1:t'}|x_{1:t}) := I(\phi_{1:t_f}; y_{t+1:t'}|y_{1:t})$ Next, consider an optimal vehicle trajectory $x_0, x_1^*, \ldots x_{t_f}^*$, i.e., maximizing the objective function (7). Let $t_1, \ldots, t_m$ be the blocking times (14) for this optimal trajectory. We have, by definition of the first blocking time and by the construction of $\hat{\mathcal{G}}$

$$I(x_1^*, \ldots, x_{t_1-1}^*, x_{t_1}^*) \leq I(x_1^{(1)}, \ldots, x_{t_1-1}^{(1)}, x_{t_1}^*), \quad (15)$$

where $x_1^{(1)}, \ldots, x_{t_1-1}^{(1)}, x_{t_1}^*$ is a feasible path through the graph $\hat{\mathcal{G}}$, which terminates in the same vehicle state as the optimal path at time t_1. Next,

$$I(x_{1:t_1-1}^{(1)}, x_{t_1}^*, x_{t_1+1:t_2}^*) \leq I(x_1^{(2)}, \ldots, x_{t_2-1}^{(2)}, x_{t_2}^*),$$

for some path $x_1^{(2)}, \ldots, x_{t_2-1}^{(2)}, x_{t_2}^*$ feasible in $\hat{\mathcal{G}}$. That is, by the chain rule for mutual information,

$$I(x_{1:t_1-1}^{(1)}, x_{t_1}^*) + I(x_{t_1+1:t_2}^*|x_{1:t_1-1}^{(1)}, x_{t_1}^*) \leq I(x_{1:t_2-1}^{(2)}, x_{t_2}^*).$$

Similarly, for $k \leq m - 1$,

$$I(x_{1:t_k-1}^{(k)}, x_{t_k}^*) + I(x_{t_k+1:t_{k+1}}^*|x_{1:t_k-1}^{(k)}, x_{t_k}^*)$$
$$\leq I(x_{1:t_{k+1}-1}^{(k+1)}, x_{t_{k+1}}^*), \quad (16)$$

and finally

$$I(x_{1:t_m-1}^{(m)}, x_{t_m}^*) + I(x_{t_m+1:t_f}^*|x_{1:t_m-1}^{(m)}, x_{t_m}^*)$$
$$\leq I(x_{1:t_f-1}^{(m+1)}, x_{t_f}^*). \qquad (17)$$

Here for each $1 \leq k \leq m+1$, the path $x_1^{(k)}, \ldots, x_{t_k-1}^{(k)}, x_{t_k}^*$, with the convention $t_{m+1} = t_f$, is feasible in $\hat{\mathcal{G}}$ and terminates at a state of the optimal vehicle trajectory. Now define, for $1 \leq k \leq m$,

$$\Delta^{(k)} = I(x_{t_k+1:t_{k+1}}^*|x_{1:t_k-1}^*, x_{t_k}^*) - I(x_{t_k+1:t_{k+1}}^*|x_{1:t_k-1}^{(k)}, x_{t_k}^*)$$

Let $x_0, \hat{x}_1, \ldots \hat{x}_{t_f}$ be the path returned by the pruned forward value iteration algorithm, and $\hat{I}$ be the corresponding value for the mutual information. Then we have

Proposition 1: $\hat{I} \geq I^* - \sum_{k=1}^{m} \Delta^{(k)}$.

Proof: Recall our convention $t_{m+1} = t_f$. By the chain rule for mutual information, we have

$$I^* = I(x_{1:t_1}^*) + \sum_{k=1}^{m} I(x_{t_k+1:t_{k+1}}^*|x_{1:t_k}^*)$$

$$= I(x_{1:t_1}^*) + \sum_{k=1}^{m} \Delta^{(k)} + \sum_{k=1}^{m} I(x_{t_k+1:t_{k+1}}^*|x_{1:t_k-1}^{(k)}, x_{t_k}^*)$$

$$\overset{(a)}{\leq} I(x_{1:t_f-1}^{(m+1)}, x_{t_f}^*) + \sum_{k=1}^{m} \Delta^{(k)}$$

$$\overset{(b)}{\leq} \hat{I} + \sum_{k=1}^{m} \Delta^{(k)}.$$

Inequality (a) is obtained by summing (15), (16) and (17). Inequality (b) holds because the path $x_0, x_{1:t_f-1}^{(m+1)}, x_{t_f}^*$ is a path in the graph $\hat{\mathcal{G}}$ and $\hat{I}$ is the maximal value of mutual information which can be obtained over such paths. ∎

Although proposition 1 does not provide an approximation guarantee for the pruned value iteration algorithm, it sheds some light on the conditions under which we should expect the heuristic to perform well. Essentially, the term $\Delta^{(k)}$ remains small if changing the subpath $x_{1:t_k}^*$ of the optimal trajectory by the subpath $x_{1:t_k-1}^{(1)}, x_{t_k}^*$ feasible in $\hat{\mathcal{G}}$ does not impact too much the value of the subsequent measurements on the optimal trajectory for $t_k + 1 \leq t \leq t_{k+1}$. Obtaining better performance bounds is the subject of current research.

E. Numerical Simulations

We briefly present some numerical simulations illustrating the performance of the pruned value iteration algorithm with respect to the greedy heuristic. The examples considered involve static fields, and the sites D correspond to a regular spatial grid. Fig. 1 and 2 show examples of trajectories obtained using the greedy heuristic and the pruned forward value iteration algorithm. The vehicle can take noisy measurements of the field values at its current position and its nearest neighbors, and can move at each period to its neighboring positions on the grid, including the diagonals in the two-dimensional case. In both cases, we see an important drawback of the greedy policy, which remains trapped in

the first local region of relatively high variance encountered, and fails to see that there are perhaps even more interesting regions which it should explore instead during the limited available time.
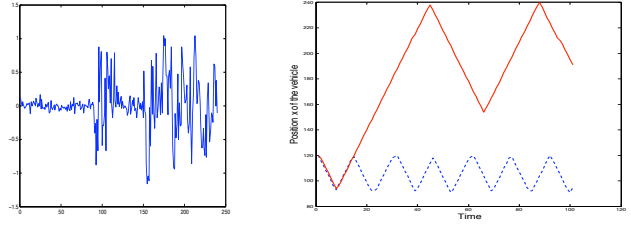
(a) One-dimensional field sample.

(b) Trajectories for the greedy (dashed line) and pruned value iteration heuristic. We represent the position of the vehicle on the line, as a function of time.

Fig. 1. One-dimensional Example. The greedy heuristic remains in the left region of high variance, even though traversing the middle region of low variance would allow the vehicle to access the region on the right which turns out to be more rewarding.
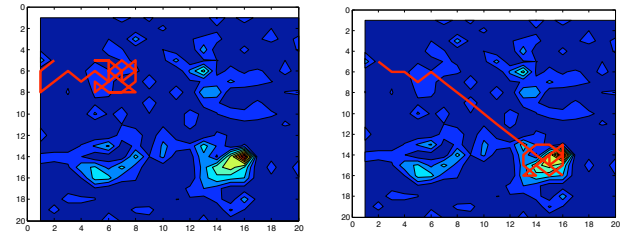
(a) Trajectory for the greedy policy.

(b) Trajectory for the pruned value iteration heuristic.

Fig. 2. Two-dimensional example for a 20×20 grid. The region in the lower right corner is of highest variance and most rewarding for the sensor. However, the greedy policy again remains trapped in the first region of higher variance encountered, for the total duration of the mission.

V. CONCLUSION AND FUTURE WORKS

We have discussed the problem of informative path planning in Gaussian fields. This problem is currently motivated by the deployment of mobile sensors, and by the interest in developing robots which map their environment autonomously. We have delineated essentially the simplest problem in active sensing with mobile robots, for which the trajectory optimization problem is deterministic. In this problem, the robots are perfectly localized, the measurements are linear in the state of the environment, and the objective function does not depend on the actual values of the sensor measurements, but only on their variance. Even for this problem, computing an optimal trajectory is difficult for realistic environment models, and more work in needed to develop compact environment representations and design suboptimal trajectories with performance approximation guarantees. Advances on this problem should be of interest for more complex applications, such as active SLAM, by providing sound base policies for the development of high-performance heuristics.

VI. Acknowledgements

The authors thank Vijay Kumar for interesting discussions during the preparation of this paper.

References

[1] UAV roadmap, "Unmanned aircraft systems roadmap 2005-2030," Office of the Secretary of Defense, Tech. Rep., 2005. [Online]. Available: http://www.acq.osd.mil/usd/Roadmap%20Final2.pdf

[2] N. Leonard, D. Paley, F. Lekien, R. Sepulchre, and D. Fratantoni, "Collective motion, sensor networks and ocean sampling," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 48–74, 2007.

[3] A. Willsky, M. Bello, D. Castanon, B. Levy, and G. Verghese, "Combining and updating of local estimates and regional maps along sets of one-dimensional tracks," *IEEE Transactions on Automatic Control*, vol. 27, no. 4, pp. 799–813, August 1982.

[4] N. Cressie, *Statistics for Spatial Data*. Wiley, 1993.

[5] E. Sudderth, "Embedded trees: Estimation of gaussian processes on graphs with cycles," Master's thesis, Massachusetts Institute of Technology, 2002.

[6] R. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," in *Autonomous Robot Vehicles*, I. Cox and G. Wilfon, Eds. New York: Springer Verlag, 1990, pp. 167–193.

[7] M. G. Dissanayake, P. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (SLAM) problem," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 3, pp. 229–241, June 2001.

[8] S. Thrun, W. Burgard, and S. Fox, *Probabilistic Robotics*. MIT Press, 2005.

[9] M. Athans, "On the determination of optimal costly measurement strategies for linear stochastic systems," *Automatica*, vol. 8, pp. 397–412, 1972.

[10] D. Uciński, *Optimal Measurement Methods for Distributed Parameter System Identification*. CRC Press, 2005.

[11] J. Williams, "Information theoretic sensor management," Ph.D. dissertation, Massachusetts Institute of Technology, February 2007.

[12] V. Krishnamurthy and R. Evans, "Hidden Markov model multiarm bandits: a methodology for beam scheduling in multitarget tracking," *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 2893 – 2908, December 2001.

[13] S. LaValle, *Planning Algorithms*. Cambridge University Press, 2006.

[14] B. Yamauchi, A. Schultz, and W. Adams, "Mobile robot exploration and map building with continuous localization," in *Proceedings of the International Conference on Robotics and Automation*, Leuven, Belgium, May 1998, pp. 3715–3720.

[15] K. Murphy, "Bayesian map learning in dynamic environments," in *Advances in Neural Information Processing Systems (NIPS)*, 1999.

[16] B. Grocholsky, "Information theoretic control of multiple sensor platforms," Ph.D. dissertation, The University of Sidney, 2002.

[17] R. Sim and N. Roy, "Global A-optimal robot exploration in SLAM," *International Conference of Robotics and Automation*, 2005.

[18] T. Chung, J. Burdick, and R. Murray, "A decentralized motion coordination strategy for dynamic target tracking," in *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, 2006.

[19] V. Gupta, T. Chung, B. Hassibi, and R. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage," *Automatica*, vol. 42, no. 2, pp. 251–260, 2006.

[20] C. Leung, S. Huang, N. Kwok, and G. Dissanayake, "Planning under uncertainty using model predictive control for information gathering," *Robotics ad Autonomous Systems*, 2006.

[21] C. Leung, S. Huang, and G. Dissanayake, "Active SLAM using model predictive control and attractor based exploration," in *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006.

[22] A. Singh, A. Krause, C. Guestrin, W. Kaiser, and M. Batalin, "Efficient planning of informative paths for multiple robots," Carnegie-Mellon University, Tech. Rep., 2006.

[23] D. Popa, M. Mysorewala, and F. Lewis, "EKF-based adaptive sampling with mobile robotic sensor nodes," in *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006.

[24] J. Cortés, "Distributed kriged Kalman filter for spatial estimation," *Submitted to the IEEE Transactions on Automatic Control*, 2007.

[25] I. Hussein, "A Kalman filter-based control strategy for dynamic coverage control," in *Proceedings of the 2007 American Control Conference*, New York City, July 2007.

[26] R. Martinez-Cantin, N. de Freitas, A. Doucet, and J. Castellanos, "Active policy learning for robot planning and exploration under uncertainty," in *Robotics: Science and Systems (RSS)*, 2007.

[27] B. Zhang and G. Sukhatme, "Adaptive sampling for estimating a scalar field using a robotic boat and a sensor network," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2007.

[28] J. Choi, J. Lee, and S. Oh, "Swarm intelligence for achieving the global maximum using spatio-temporal Gaussian processes," in *Proceedings of the American Control Conference*, 2008.

[29] I. Hussein, "Kalman filtering with optimal sensor motion planning," in *Proceedings of the American Control Conference*, 2008.

[30] J. Le Ny, "Performance optimization for unmanned vehicle systems," Ph.D. dissertation, Massachusetts Institute of Technology, 2008.

[31] T. Kollar and N. Roy, "Trajectory optimization using reinforcement learning for map exploration," *The International Journal of Robotics Research*, vol. 27, no. 2, pp. 175–196, 2008.

[32] ——, "Efficient optimization of information-theoretic exploration in SLAM," in *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, Chicago, IL, 2008, pp. 1369–1375.

[33] F. Amigoni, "Experimental evaluation of some exploration strategies for mobile robots," in *2008 IEEE International Conference on Robotics and Automation*, 2008.

[34] H.-L. Choi and J. How, "Continuous motion planning for information forecast," in *Proceedings of the 47th Conference on Decision and Control*, 2008.

[35] R. Graham and J. Cortés, "A cooperative deployment strategy for optimal sampling in spatiotemporal estimation," in *Proceedings of the Conference on Decision and Control*, 2008.

[36] H.-L. Choi, "Adaptive sampling and forecasting with mobile sensor networks," Ph.D. dissertation, Massachusetts Institute of Technology, 2009.

[37] L. Meier, J. Perschon, and R. Dressler, "Optimal control of measurement systems," *IEEE Transactions on Automatic Control*, vol. 12, no. 5, pp. 528–536, 1967.

[38] A. Bennett, *Inverse Modeling of the Ocean and Atmosphere*. Cambridge University Press, 2002.

[39] E. Kalnay, *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge University Press, 2003.

[40] H. Rue and H. Tjelmeland, "Fitting Gaussian Markov random fields to Gaussian fields," *Scandinavian Journal of Statistics*, vol. 29, pp. 31–49, 2002.

[41] H. Rue and L. Held, *Gaussian Markov Random Fields: Theory and Applications*. Chapman & Hall/CRC, 2005.

[42] S. Thrun, Y. Liu, D. Koller, A. Ng, Z. Ghahramani, and H. Durrant-Whyte, "Simultaneous localization and mapping with sparse extended information filters," *The International Journal of Robotics Research*, vol. 23, no. 7-8, pp. 693–716, July-August 2004.

[43] J. Manyika and H. Durrant-Whyte, *Data Fusion and Sensor Management: A Decentralized Information-Theoretic Approach*. Ellis Horwood, 1994.

[44] S. Lauritzen, *Graphical Models*. Oxford University Press, 1996.

[45] R. Furrer, M. Genton, and D. Nychka, "Covariance tapering for interpolation of large spatial datasets," *Journal of Computational and Graphical Statistics*, vol. 15, no. 3, 2006.

[46] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, 1991.

[47] E. Sondik, "The optimal control of partially observable Markov processes," Ph.D. dissertation, Stanford University, 1971.

[48] D. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Athena Scientific, 2005.

[49] A. Censi, D. Calisi, A. De Luca, and G. Oriolo, "A Bayesian framework for optimal motion planning with uncertainty," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Pasadena, CA, May 2008.

[50] R. He, S. Prentice, and N. Roy, "Planning in information space for a quadrotor helicopter in a GPS-denied environment," in *Proceedings of the 2008 International Conference on Robotics and Automation*, 2008.