

# Filter-Based Stochastic Abstractions for Constrained Planning with Limited Sensing

Hasan A. Poonawala and Ufuk Topcu

**Abstract**—As the complexity of the specifications that must be met by a system increases, hierarchical control protocols that merge control and planning decisions at multiple levels of abstraction become necessary. For such hierarchical reasoning, a suitable finite-state abstraction for dynamical systems evolving over continuous state spaces may be needed. The implementation of existing controllers derived using a finite-state abstraction often require that the current continuous state be known exactly, in order to guarantee that the required transitions in the finite-state abstraction occur. When the measurements are partial or noisy, the true state is unknown, and these controllers cannot be implemented. We propose an abstraction that can be used to overcome the uncertainty in the state resulting from imperfect measurement, at the cost of providing only probabilistic guarantees. The abstraction is based on the filter used to maintain an estimate of the true state. We show how the abstraction can be used to create a time-varying policy which maximizes the minimum probability that a target discrete state is reached in finite time from any initial state.

## I. INTRODUCTION

Motion control problems can sometimes be solved effectively by designing continuous feedback control laws [1]. More often, the complexity of the desired motion or system dynamics leads to the adoption of a hierarchical scheme. In such a scheme, a planning algorithm determines a suitable high-level plan and a continuous controller is designed to implement this plan.

The application of planning algorithms for motion planning over continuous spaces often rely on a discretization of the space into a finite (or at least a countable) set of states. The discretization may be based on regular grids [2], random sampling [2], or the satisfaction of logical propositions [3]. Furthermore, suitable transitions must be defined between the discrete states such that they appropriately capture the behaviour of the original continuous system [4]. The discretization and transitions together form an abstraction of the continuous system. Various decision-making algorithms can be applied to the abstracted system, to generate a high-level plan. Once a suitable path has been planned in the discretized space, the continuous control is tasked with achieving the desired transitions between discrete states.

When the uncertainty in the state due to partial and/or noisy measurements is significant, these hierarchical plan-

ning methods must often be modified in order to account for the resulting uncertainty in the state. In the case of planning and control for autonomous rendezvous and docking of spacecraft [5], the limited control authority of the spacecraft and noisy sensing make the problem highly challenging. The control goal is more complicated than a regulation or tracking problem, since it involves reaching a given convex region in finite time. These challenges motivate us to develop planning methods that account for state uncertainty due to measurement noise, in the presence of constrained inputs.

The most common approach for modifying hierarchical planning methods in order to account for noise is to assume that the estimation and control aspects of the problem can be treated separately. For certain systems, such separation is justified due to the existence of a separation principle [1], [6] guaranteeing that the state and estimate will independently converge to their desired values. When such a principle does not exist, this approach may still be reasonable when the estimation error can be reduced arbitrarily fast through design of estimator parameters (e.g. high-gain observers [6]). Alternatively, the system may be shown to be stable despite large errors in the estimate of the state. These large errors may lead to poor control performance or prevent the control objective from being achieved during the initial time period. As the estimation error decrease over time, eventually the control performance will improve or the control objective will be achieved.

The separation between estimation and control may not always be relied upon in many control problems. When the inputs are constrained, or when the control goals need to be achieved in finite time, the planning methods used may need to explicitly account for the uncertainty in the state due to noisy measurement at the instant of implementing the derived control policy. Alternate plans or control strategies may be implemented depending on the amount of information or uncertainty in the system. Solution methods for partially observable Markov decision processes (POMDPs) [7] explicitly account for the uncertainty in the state induced by imperfect measurement at the time of choosing actions. The derived optimal control policy is able to handle the trade-off between needing to improve the quality of information and achieving the control objective.

The work in [8] proposes a finite-state abstraction for dynamical systems that yields a finite-state POMDP. The resulting planning problem is computationally expensive, as noted by the authors. We therefore wish to avoid the POMDP formulation if possible. Work by [9] develops controllers over finite-state abstractions for the case when

This work is supported by grants from AFRL (# FA9453-15-1-0317) and ONR (# N00014-13-1-0778).

H. A. Poonawala is a Postdoctoral Researcher at the Institute for Computational Engineering and Science, University of Texas at Austin, Austin, TX 78712, USA [hasanp@utexas.edu](mailto:hasanp@utexas.edu)

Ufuk Topcu is with the Department of Aerospace Engineering, University of Texas at Austin, Austin, TX 78712, USA [utopcu@utexas.edu](mailto:utopcu@utexas.edu)

no measurement noise exists. The controller is modified to be robust to measurement noise, however the robustness depends directly on the initial uncertainty, and the evolution of the covariance of the estimates is not exploited in the method.

We propose a method of abstraction of a continuous dynamical system with noisy measurements that enables probabilistic-reachability-based methods [10] to be used to predict the probability that a discrete goal state is reached from any other discrete state in the abstraction. These reachability probabilities are valid when the uncertainty in the state is constant. Since the estimation scheme will reduce the uncertainty over time, the policy that maximizes the probability of successfully transitioning between two states at the initial time may be different from that at a later time. We also propose a planning method that uses the predicted levels of uncertainty at future times in order to select the policy at those future times. The result is a time-varying policy which maximizes the lower bound on the probability with which the final discrete state is reached from the initial state. Due to the design of the abstraction and the planning method, a continuous control exists which achieves this lower bound on the reachability probability at least.

## II. PROBLEM STATEMENT

We now discuss a number of preliminary concepts and notation as a prelude to the statement of the main problem.

### A. System Model

Consider the discrete Linear Time Invariant (LTI) system

$$q(k+1) = A_d q(k) + B_d u(k) + w \quad (1a)$$

$$y(k) = C q(k) + v \quad (1b)$$

where  $q(k) \in Q \subset \mathbb{R}^n$  is the state at the  $k^{\text{th}}$  time instant,  $u(k) \in U \subset \mathbb{R}^p$  is the control at this instant,  $y(k)$  is the measurement,  $v$  is a zero-mean Gaussian noise term with covariance matrix  $\Sigma_v$ ,  $w$  is a zero-mean Gaussian noise term with covariance matrix  $\Sigma_w$ ,  $A_d \in \mathbb{R}^{n \times n}$  is the system matrix,  $B_d \in \mathbb{R}^{n \times p}$  is the input matrix, and  $C \in \mathbb{R}^{m \times n}$  is the sensing matrix. The dimension  $m$  of the range space of  $C$  may be less than  $n$ , implying that the observations are partial. Note that we will assume  $p = n$  in this work, implying that the system is fully actuated. A controllable underactuated discrete time linear system can be converted to a fully actuated system by grouping multiple time steps together and appropriately redefining the control and measurement signals.

### B. State Space Partitions

A common choice for partitioning a continuous space is to partition it into convex polytopes. A convex polytope in  $\mathbb{R}^n$  is a convex set  $\{x \in \mathbb{R}^n \mid a_i^T x \leq b_i, \}$  where  $a_i \in \mathbb{R}^n$  and  $b_i \in \mathbb{R}$  for  $i \in \{1, 2, \dots, m\}$ . These  $m$  constraints can be represented by a matrix  $L \in \mathbb{R}^{m \times n}$  and  $w \in \mathbb{R}^m$  as  $Lx \leq w$ . The advantage of such sets is that their images under affine transformations remain as convex polytopes. Furthermore, several algorithms exist for computing probability bounds related to distributions over such sets [11].

### C. Markov Decision Process

A Markov decision process (MDP) [10] is a tuple  $(S, A, P)$  where

- $S$  is a set of states,
- $A$  is a set of labels (or actions), and
- $P$  is a probability transition function  $P: S \times A \rightarrow \text{Distr}(S)$ .

where  $\text{Distr}(S)$  is the space of probability distributions on the set  $S$ . The MDP is assumed to evolve in discrete time steps. At a given discrete time  $k \in \mathbb{N}$ , an action  $a \in A$  is chosen when the current state is  $s \in S$ . The state at the next time step  $k+1$  is a random variable described by the distribution  $P(s, a)$ . We can denote the transition from  $s_i \in S$  to  $s_j \in S$  due to action  $a$  as  $s_i \xrightarrow{a} s_j$ . Thus,  $P(s_i, a)(s_j)$  is the probability with which this transition will occur, given that the initial state is  $s_i$  and action  $a \in A$  is chosen.

An execution of an MDP is a sequence of states  $s_0, s_1, s_2, \dots$  under a sequence of actions  $a_0, a_1, \dots$  such that  $P(s_i, a_i)(s_{i+1}) > 0$ . A policy  $\pi: S \rightarrow A$  is a map that assigns each state in  $S$  to an action  $A$ . When the policy is fixed, the MDP reduces to a Markov chain (MC). An execution of an MDP under policy  $\pi$  is a sequence of states  $s_0, s_1, s_2, \dots$  such that  $P(s_i, \pi(s_i))(s_{i+1}) > 0$ .

Let  $s \models \diamond g$  denote the logical condition that an execution  $s_0, s_1, s_2, \dots$  exists such that  $s_0 = s$  and there exists  $i \geq 0$  such that  $s_i = g$ . The probability with which this will occur in an execution beginning at  $s$  is denoted as  $Pr(s \models \diamond g)$ . The probability that  $i \leq N$  is denoted by  $Pr_N(s \models \diamond g)$ . If the execution is determined by a policy  $\pi$ , then we denote the same probability as  $Pr_N^\pi(s \models \diamond g)$ .

### D. Recursive Estimation

Given a partially observable dynamical (control) system, a belief is a probability distribution representing the estimate of the true state of the system. The belief space  $\mathcal{B}$  is the set of all possible beliefs regarding the true state. We assume that there exists a recursive estimator  $F: \mathcal{B} \times U \times Y \rightarrow \mathcal{B}$  for such a system, which maps the current belief  $b_k$ , control action  $u_k$  and resulting measurement  $y_{k+1}$  into a new belief  $b_{k+1}$ . In other words, the new belief is  $b_{k+1} = F(b_k, u_k, y_{k+1})$ .

The measurement  $y$  which will be received is unknown, but can be represented as a random variable that depends on the (future) true state and noise. This measurement is used to update the future belief of the system. The current belief  $b$ , action  $u$ , and the model of the dynamical system together can be used to generate a distribution over the possible expected value of  $F(b, u, \cdot)$ , since  $y$  is unknown but is This distribution will be denoted as  $p_F(b, u)$ . In this paper, we will represent a belief by its mean and covariance.

### E. Problem Statement

Consider a system with state space  $Q \subset \mathbb{R}^n$  and dynamics given by (1). Let  $F$  be a recursive filter designed for the system (1). Let  $Q_f$  be a connected subset of  $Q$ , and  $N \in \mathbb{N}$  be a desired final time. Let  $q(k)$  be the unique solution of (1a) given a control signal  $u(k)$ , where  $k \geq 0$ . Given an initial

probability distribution  $b_0$ , derive a control  $u(k) \subset U \forall k \in [0, N]$  which maximizes the probability that  $q(N) \in Q_f$ .

### III. FILTER-BASED ABSTRACTION

Given an MDP, techniques from model checking [10] can be used to determine whether a certain state can be reached from a set of initial states or not. These techniques have been modified in order to obtain a method to generate controllers that can guarantee that a goal state will be reached from a given initial state. Furthermore, reachability while avoiding certain unsafe states can also be achieved. In order to extend these methods to continuous state systems, finite-state abstraction methods have been proposed [3], [12], [13] which capture the relevant behavior of the continuous system in the form of an MDP or transition system [10].

However, these methods assume that the true state is known when implementing the control policy derived from the abstraction. When the state is uncertain, it is not clear which action should actually be used. Even if the correct action is selected in a discrete state, it is not clear whether the corresponding continuous control action when applied to the estimated continuous state will result in the true continuous state achieving the intended transition. Thus, the guarantees available for the case in which the complete true state is measured precisely will not necessarily extend to the case in which the state is uncertain due to noisy measurement, even when the uncertainty is low.

A common method for dealing with uncertainty in the state is to derive control policies by solving planning problems in the belief space [7], [14], [15]. The control is selected based on the current belief, instead of the (unknown) current true state. For example, value iteration algorithms used to solve for the optimal policy in MDPs are extended to POMDPs by solving for a value function defined on the belief space [7]. In [14], [15], the state space is continuous, and hence the planning method in [7] cannot be applied. These planning methods use an extended Kalman filter (EKF) to maintain a belief (represented by a Gaussian distribution) which is updated whenever measurements are received. The EKF update equations define the dynamics of the belief, which we refer to as the filter dynamics.

We are motivated to define a finite-abstraction of the *belief* space using the filter dynamics in order to derive controllers that can guarantee reachability. However, given that the measurements are random, we can only guarantee probabilistic reachability. The planning methods in [14], [15] can maximize the probability of reaching a desired belief. However, these methods rely on predicting the future beliefs obtained resulting from a given initial belief and a sequence of actions. This is done by assuming that the measurement (which is random) in a state will equal its expected value. This is known as the assumption of maximum likelihood measurement. Due to this ability to predict the evolution of the beliefs deterministically, these methods are similar to model predictive control techniques.

In order to depend on the probabilistic guarantees we seek to obtain, we need to relax the assumption of maximum

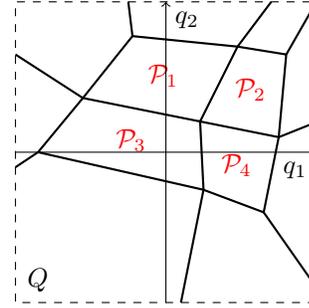


Fig. 1. Partition of the continuous state space  $Q \subset \mathbb{R}^2$ . Each polytope is an element of  $\mathcal{P}$ .

likelihood estimation. Indeed, the use of this assumption in deriving a control policy may result in unreliable performance in practice [16]. The authors in [16] overcome this limitation by picking a nominal trajectory and deriving a locally-optimal controller that is robust to the randomness of the measurement. The benefit of this approach is that the state space does not need to be discretized, while the disadvantage is that the control policy needs to be computed online once the initial belief is specified. We believe that a judicious discretization of the belief space is worth the increased complexity in designing a controller, in order to be able to compute control strategies offline. We now describe the derivation of a finite-state abstraction of the belief space and its dynamics.

The true state  $q$  is treated as a random variable  $\hat{Q}$  and is represented as a belief  $b$ . The expected value of the belief is  $\mu = E[\hat{Q}] \in \mathbb{R}^n$  and the covariance is  $\Sigma = E[(\hat{Q} - \mu)(\hat{Q} - \mu)^T] \in \mathbb{R}_{++}^{n \times n}$ , where  $\mathbb{R}_{++}^{n \times n}$  is the space of positive semidefinite matrices of rank  $n$ . We will represent the belief  $b$  as the pair  $(\mu, \Sigma)$ .

We can partition  $Q$  into a set of polytopes  $\mathcal{P}$  as depicted in Figure 1. A key aspect of the method we propose is that this partition  $\mathcal{P}$  also serves as a partition of the subspace of  $\mathcal{B}$  corresponding to the mean of the distributions. Let the map  $T: Q \rightarrow \mathcal{P}$  assign a vector  $x \in \mathbb{R}^n$  to one of the elements of  $\mathcal{P}$ . Given this partition, if the current belief is  $b = (\mu, \Sigma)$ , then  $\mu = E[\hat{Q}]$  belongs to  $T(\mu) \in \mathcal{P}$ . Starting from this belief, an action and measurement pair  $(u, y)$  in the continuous space will result in a new belief  $\bar{b} = (\bar{\mu}, \bar{\Sigma})$ , based on the filter map  $F$ .

When only the control  $u$  is known, the future belief  $\bar{b} = (\bar{\mu}, \bar{\Sigma})$  is a random variable since it depends on the measurement which is random. We denote its probability distribution function (PDF) by  $p(\bar{b}|b, u)$ , which is a joint PDF over  $(\bar{\mu}, \bar{\Sigma})$ . We marginalize it with respect to the random variable  $\bar{\Sigma}$  in order to obtain  $p(\bar{\mu}|b, u)$ , which is the PDF of the expected value of the future belief given the current belief  $b$  and control  $u$ . We denote the mean and covariance of  $p(\bar{\mu}|b, u)$  by  $\tilde{\mu}$  and  $\tilde{\Sigma}$  respectively.

The action is chosen such that  $E[p(\bar{\mu}|b, u)] \in T^{-1}(\mathcal{P}_j)$ . Once the measurement is received,  $\bar{b}$  may be such that  $\bar{\mu}$  belongs to a partition  $\mathcal{P}_k \neq \mathcal{P}_j$ . The probability that  $\tilde{\mu} \in \mathcal{P}_k$  depends on  $\tilde{\Sigma}$ . Thus, given a discrete state and action,

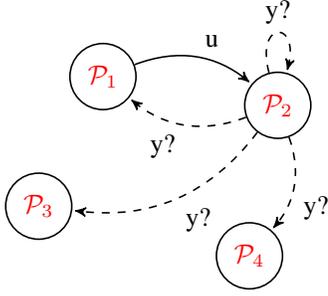


Fig. 2. The filter update can be viewed as a two step process. The action  $u$  from  $\mathcal{P}_1$  defines a transition such that  $T(E[p_F(b, u)]) = \mathcal{P}_2$ , which can be viewed as a deterministic transition. The measurement  $y$  results in a new belief. Since  $y$  is a random variable, the effect of action  $u$  can be viewed as a probabilistic transition from the state  $\mathcal{P}_2$  (and thus from  $\mathcal{P}_1$ ) to other states.

we obtain a probability distribution on the resulting discrete state. This process is depicted in Figure 2.

We can define an abstraction of the dynamics of the mean (expected value) of the belief  $b$  as an MDP  $\mathcal{M} = (S, A, P_{\tilde{\Sigma}_k})$ , where

- $S = \{s_1, s_2, \dots, s_{|\mathcal{P}|}\} \cup s_a$  is a finite set states, one for each partition in  $\mathcal{P}$  and an absorbing state  $s_a$  representing  $\mathbb{R}^n \setminus Q$ . If  $\mathbb{R}^n \setminus Q = \emptyset$  then  $s_a = \emptyset$ ,
- $A = \{a_1, a_2, \dots, a_{|\mathcal{P}|}\}$  is a set of actions,
- $P_{\tilde{\Sigma}_k}: S \times A \rightarrow \text{Distr}(S)$  is a probabilistic transition function.

Here, the subscript  $\tilde{\Sigma}_k$  indicates that the transition probabilities depend on the time-varying matrix  $\tilde{\Sigma}_k \in \mathbb{R}_{++}^{n \times n}$ , where  $k$  denotes the time instant. For each state  $s_i \in S$ , we can define a fixed point  $d_i \in T^{-1}(s_i)$ . The point  $d_i$  is a target point in the partition corresponding to  $s_i$ . If for all  $\mu \in T^{-1}(s_i)$ , there exists  $d_j \in T^{-1}(s_j)$  and a control  $u \in U$  such that  $\tilde{\mu} = d_j$ , then the action  $a_j$  is said to exist in state  $s_i$ . The action  $a_j \in A$  in any state represents an attempt to induce a transition from that state to the state  $s_j \in S$ . This definition of actions is one of many alternatives when defining an abstraction  $\mathcal{M}$ .

As mentioned earlier, the existence of an action  $a_j$  that ensures that  $E[p(\tilde{\mu}|b, u)] \in T^{-1}(s_j)$  does not necessarily imply that  $E[F(b, u, y)] \in T^{-1}(s_j)$ . Furthermore, it is possible that the updated belief has an expected value in  $\mathbb{R}^n \setminus Q$  if the latter is non-empty, and so we include the absorbing state  $s_a$  in  $S$ .

For an action  $a_l \in A$ , the transition probabilities are given by

$$P_{\tilde{\Sigma}}(s_i, a_l, s_j) = \begin{cases} \int_{x \in T^{-1}(s_j)} p(\tilde{\mu}|b, u)(x) dx, & l = j \text{ and } s_i \xrightarrow{a_j} s_j \text{ exists,} \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

where the definition of  $P_{\tilde{\Sigma}}$  depends only on  $\tilde{\Sigma}$  since we require that  $\tilde{\mu} = E[p(\tilde{\mu}|b, u)]$  is identical for all  $b = (\mu, \Sigma)$ . This is achieved by choosing an appropriate control  $u$  for each  $\mu \in T^{-1}(s_i)$ . By definition of action  $a_l$ , such a control  $u$  exists for each such  $\mu$ .

In general, it may not be possible to compute the transition probabilities in (2) exactly. However, efficient techniques to bound the transition probabilities may exist. Thus, in practice we may need to resort to an uncertain MDP [17], [18], in which the transition probabilities belong to known sets. In the next section, we will describe how to derive a control policy for the abstraction that maximizes the probability of reaching a goal discrete state.

#### IV. PLANNING USING FILTER-BASED ABSTRACTIONS

The continuous control objective consists of maximizing the probability that a target set  $Q_f$  is reached from an initial state (or belief) within some finite time  $N$ . Once we have formulated an abstraction  $\mathcal{M} = (S, A, P_{\tilde{\Sigma}_k})$  that describes the filter dynamics, maximizing the probability of reaching  $Q_f$  is equivalent to maximizing  $Pr_N(s \models \diamond g)$  where  $s \in S$  represents the partition of  $Q$  that contains the initial belief, and  $g \in S$  represents  $Q_f$ . Thus, the control objective in the abstraction is to design a planning algorithm which provides a policy  $\pi$  that maximizes  $Pr_N^\pi(s \models \diamond g)$ . When the transition probabilities in  $\mathcal{M}$  are not known exactly, then we can only compute a lower bound on  $Pr_N^\pi(s \models \diamond g)$ . In this case, the control objective is to maximize this lower bound. Let the lower bound be denoted by  $\underline{Pr}_N^\pi(s \models \diamond g)$ .

Let the space of policies for  $\mathcal{M}$  be  $\Pi$ . Then, the policy  $\pi^* \in \Pi$  that maximizes the lower bound on the probability of reaching  $g$  from  $s$  in  $N$  steps is

$$\pi^* = \arg \max_{\pi \in \Pi} \underline{Pr}_N^\pi(s \models \diamond g) \quad (3)$$

and the reachability probability is

$$\underline{Pr}_N^{max}(s \models \diamond g) = \underline{Pr}_N^{\pi^*}(s \models \diamond g). \quad (4)$$

We can compute  $\pi^*$  by extending a value-iteration method [18], [19] for the solution of parametric MDPs. In [18], the method focuses on the case when each state has only two successor states, and the transition probabilities are fixed. In the abstraction  $\mathcal{M}$ , each state has multiple successors, and the transition probabilities vary due to their dependence on  $\tilde{\Sigma}_k$ .

In order to deal with the case where the covariance changes, we capture the change by creating a larger MDP  $\mathcal{M}_p$  from  $N + 1$  copies of  $S$ , one for each time step  $k \in \{0, \dots, N\}$ . The transitions in  $\mathcal{M}_p$  are defined such that the transition probability from a state  $s$  in the  $k^{\text{th}}$  copy of  $S$  to a state  $s'$  in the  $(k + 1)^{\text{th}}$  copy of  $S$  reflects the covariance  $\tilde{\Sigma}_k$  at the  $k^{\text{th}}$  time step. The method from [18] can then be applied to  $\mathcal{M}_p$ .

#### V. LINEAR SYSTEMS WITH KALMAN FILTERING

In this section, we describe the details of obtaining a filter-based abstraction of a discrete-time linear system driven by Gaussian measurement noise, assuming that the state space  $Q$  has been partitioned into a finite set of convex polytopes  $\mathcal{P}$ , and a Kalman filter [20] is used to update the belief.

### A. Kalman Filter Equations

Let the current estimate of the true state  $q(k)$  at time step  $k$  be represented by a Gaussian distribution with mean  $\mu_k$  and variance  $\Sigma_k$ . Assume that a control  $u_k$  is selected. According to the The Kalman filter equations, the predicted mean and covariance are

$$\mu_{k+1}^{pred} = A_d \mu_k + B_d u_k \quad (5a)$$

$$\Sigma_{k+1}^{pred} = A_d \Sigma_k A_d^T + \Sigma_w \quad (5b)$$

where  $\Sigma_w$  is the covariance of the process noise. The innovation  $\tilde{y}_{k+1} = y_{k+1} - C \mu_{k+1}^{pred}$  is the difference between the measurement received and the expected measurement. The optimal Kalman gain  $K_{k+1}$  is

$$K_{k+1} = \Sigma_{k+1}^{pred} C^T \left( C \Sigma_{k+1}^{pred} C^T + \Sigma_v \right)^{-1}. \quad (6)$$

where  $\Sigma_v$  is the covariance of the measurement noise. The updated mean and covariance is then

$$\mu_{k+1} = \mu_{k+1}^{pred} + K_{k+1} \tilde{y}_{k+1} \quad (7a)$$

$$\Sigma_{k+1} = (I - K_{k+1} C) \Sigma_{k+1}^{pred} \quad (7b)$$

where the updated covariance turns out to be independent of the measurement received.

The innovation  $\tilde{y}$  is a Gaussian random variable  $\tilde{y}_{k+1} \sim \mathcal{N} \left( 0, C \Sigma_{k+1}^{pred} C^T + \Sigma_v \right)$ . Thus, given  $\mu_{k+1}^{pred}$  but not  $y_{k+1}$ , the mean  $\mu_{k+1}$  is a Gaussian random variable with mean  $\mu_{k+1}^{pred}$  and covariance

$$\tilde{\Sigma}_k = K_{k+1} \left( C \Sigma_{k+1}^{pred} C^T + \Sigma_v \right) K_{k+1}^T. \quad (8)$$

### B. Constructing the Abstraction

The abstraction of the filter dynamics is given by  $\mathcal{M} = (S, A, P_{\tilde{\Sigma}_k})$  as mentioned in Section III. Given the finite partition of  $Q$  into convex polytopes  $\mathcal{P}$ , we can define the states as  $S = \{s_1, s_2, \dots, s_{|\mathcal{P}|}\} \cup s_a$ , where  $s_i = \mathcal{P}_i$  and  $s_a$  is an absorbing state that represents  $\mathbb{R}^n \setminus Q$ . The set of possible actions in a state is  $A = \{a_1, a_2, \dots, a_{|\mathcal{P}|}\}$ . The implementation of action  $a_j$  in state  $s_i$  as a continuous control input involves the selection of a control  $u \in U$  given  $\mu = E[\hat{Q}] \in T^{-1}(s_i)$  such that the resulting predicted mean of the belief  $\mu^{pred} = A_d \mu + B u$  is equal to the target point  $d_j \in \mathcal{P}_j$  associated with discrete state  $s_j$ .

### C. Determining Transition Probabilities

Given a belief  $b = (\mu, \Sigma)$ , the PDF of the expected value of the future belief under action  $u$  is a Gaussian distribution with mean  $\tilde{\mu}$  and covariance  $R_{K_{\tilde{y}}}$  given by the right hand sides of (5a) and (8) respectively. We use an implementation of the algorithm in [21] to compute the right hand sides of (5a). The method provides estimates of the numerical error in the result, which are used to provide upper and lower bounds on the true transition probability.

Given an initial covariance  $\Sigma_0$  representing the uncertainty in the estimate of the true state, the covariance  $\Sigma_k$  at subsequent time steps  $k \in \mathbb{N}$  can be computed using (7b) repeatedly. The transition probabilities used in the value iteration [18] at the  $k^{\text{th}}$  time step are computed using the covariance matrix  $\tilde{\Sigma}_k$ , given by (8).

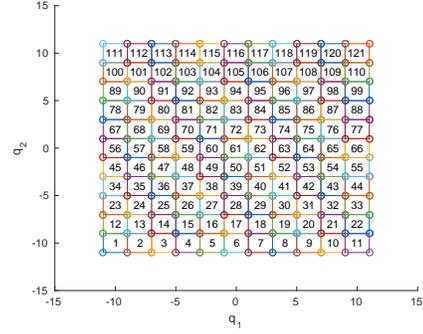


Fig. 3. The plane is partitioned into 122 regions, depicted here as 121 squares in a  $11 \times 11$  grid. The region outside this grid is the 122<sup>nd</sup> state, which is treated as an absorbing state. Each unit square - marked by a colored box - represents a single discrete state, and is identified by the number located inside this box. The state numbered 61 contains the origin at its centre.

## VI. SIMULATION

In this section, we show the results of computing a policy that maximizes a lower bound on the probability of reaching the origin for a discretized double integrator system. This system has a 2-dimensional state space, with discretized double integrator dynamics

$$q(k+1) = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix} q(k) + \begin{bmatrix} \frac{\tau^2}{2} \\ \tau \end{bmatrix} u(k) + w \quad (9a)$$

$$y(k) = \begin{bmatrix} 0 & 1 \end{bmatrix} q(k) + v, \quad (9b)$$

where  $q(k) \in \mathbb{R}^2$  is the state at discrete time  $k$ . The control  $u(k) \in \mathbb{R}$  belongs to the control region  $U = [-5, 5]$ . The terms  $w$  and  $v$  represent process noise and measurement noise respectively, which are modeled as gaussian random variables with mean 0 and variance  $\Sigma_v = 0.3$  and  $\Sigma_w = 0.3$  respectively. The sampling period of the discretization is  $\tau$ , which has value  $1 \text{ sec}$  in the results presented here.

The partition  $\mathcal{P}$  of  $\mathbb{R}^2$  consists of 121 squares of side 2 units, arranged in a  $11 \times 11$  grid, as seen in Figure 3. The centre of this grid is the origin. The initial covariance is

$$\Sigma_0 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad (10)$$

The planning horizon is  $N = 30$ . The objective is to find a time-varying policy which maximizes the probability of reaching the polytope containing the origin (state  $s_{61}$ ) within these 30 time steps.

The results of the policy computation can be seen in Figure 4. The value iteration procedure described in Section IV yields the maximum lower bound on  $Pr_N^{\pi}(s \in \diamond g)$  over all policies, for each time step  $k$ , where  $1 \leq k \leq 30$ . The probability of reaching the goal state depends on both the uncertainty of the state estimate (represented by  $\Sigma_k$ ), and the number of time steps remaining, i.e.,  $N - k$ . As  $k$  increases, the Kalman filter reduces the uncertainty in the state estimate. However, due to the limited control authority, the goal may not be reached from some states when  $N - k$  is small. This

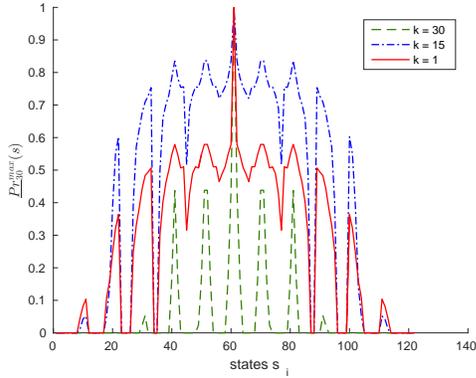


Fig. 4. This figure shows the computed lower bound  $P_{T_N}^{max}(s)$  on the probability of reaching the goal state  $s_{61}$  from the state  $s_i$ , at different time steps  $k$ , over a planning horizon  $N = 30$  time steps. For small  $k$ , the trace of  $\Sigma_k$  is large, and therefore the probability of reaching the goal state is not very high (red line). As estimation proceeds, the uncertainty decreases, therefore the lower bound on the probability of reaching the goal states increases. This can be seen in the plot for  $k = 15$  (blue dashed and dotted line). The probability increases from  $k = 1$  to  $k = 15$  despite the decrease in the remaining time  $N - k$ . For  $k$  close to  $N = 30$ , the probability of reaching the goal state from most states is small or even zero, since the goal state cannot be reached from most states in a small number of steps (green dashed line).

tradeoff is captured by the computed lower bounds, as seen in Figure 4.

As an illustrative example, we provide the computed optimal policy for state  $s_{57}$ . The policy for the remaining states are omitted due to space constraints. The optimal policy  $\pi_k^*(s_{57})$  is

$$\pi_k^*(s_{57}) = \begin{cases} a_{69} & , \text{ if } k = 1 \text{ or } k = 2, \\ a_{57} & , \text{ if } k = 30, \\ a_{81} & , \text{ otherwise.} \end{cases} \quad (11)$$

The choice of action  $a_{69}$  at  $k = 1$  in state  $s_{57}$  would result in reaching the goal at  $k = 4$ , via  $s_{69}$  and  $s_{62}$ , assuming the intended transitions are achieved. In contrast, choosing action  $a_{81}$  at some time instant  $k' > 2$  results in reaching the goal at time step  $k' + 2$ , via  $s_{81}$ . In other words, the method determines that a longer path has a higher probability of reaching the goal when the uncertainty is high, but chooses a shorter path when the uncertainty reduces. Thus, the proposed method takes into account the sensing and control limitations of the system while planning over finite horizons.

## VII. CONCLUSION

We proposed a filter-based abstraction for dynamical systems with noisy and partial measurement. We also presented a planning method which used this abstraction to compute an optimal control policy in the discrete state space of the abstraction. The optimal policy maximizes the lower bound of the probability of reaching a goal state in finite time from any initial state. The derived control policy can then be implemented by a continuous control signal which will achieve this bound. We demonstrated the planning method on a simple continuous control system.

## VIII. ACKNOWLEDGEMENT

The authors would like to thank Nils Jansen for his helpful discussions related to parametric MDPs.

## REFERENCES

- [1] M. W. Spong, S. Hutchinson, and M. Vidyasagar, *Robot modeling and control*. Hoboken (N.J.): John Wiley & Sons, 2006. [Online]. Available: <http://opac.inria.fr/record=b1119287>
- [2] S. M. LaValle, "Planning algorithms," 2004.
- [3] T. Wongpiromsarn, U. Topcu, and R. M. Murray, "Receding horizon temporal logic planning for dynamical systems," in *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, Dec 2009, pp. 5997–6004.
- [4] R. Alur, T. A. Henzinger, G. Lafferriere, and G. J. Pappas, "Discrete abstractions of hybrid systems," *Proceedings of the IEEE*, vol. 88, no. 7, pp. 971–984, July 2000.
- [5] C. Jewison and S. Erwin, "A spacecraft benchmark problem for hybrid control and estimation," in *Decision and Control, IEEE Conference on*, 2016.
- [6] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.
- [7] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *ARTIFICIAL INTELLIGENCE*, vol. 101, pp. 99–134, 1998.
- [8] K. Lesser and M. Oishi, "Finite state approximation for verification of partially observable stochastic hybrid systems," in *Proceedings of the 18th International Conference on Hybrid Systems: Computation and Control*, ser. HSCC '15. New York, NY, USA: ACM, 2015, pp. 159–168. [Online]. Available: <http://doi.acm.org/10.1145/2728606.2728632>
- [9] S. Haesaert, P. M. J. V. den Hof, and A. Abate, "Observer-based correct-by-design controller synthesis," *CoRR*, vol. abs/1509.03427, 2015. [Online]. Available: <http://arxiv.org/abs/1509.03427>
- [10] C. Baier and J.-P. Katoen, *Principles of Model Checking (Representation and Mind Series)*. The MIT Press, 2008.
- [11] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, March 2004.
- [12] O. Mickelin, N. Ozay, and R. M. Murray, "Synthesis of correct-by-construction control protocols for hybrid systems using partial state information," in *American Control Conference (ACC), 2014*, June 2014, pp. 2305–2311.
- [13] J. Liu, N. Ozay, U. Topcu, and R. M. Murray, "Synthesis of reactive switching protocols from temporal logic specifications," *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1771–1785, July 2013.
- [14] A. Bry and N. Roy, "Rapidly-exploring random belief trees for motion planning under uncertainty," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, May 2011, pp. 723–730.
- [15] S. Prentice and N. Roy, *Robotics Research: The 13th International Symposium ISRR*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, ch. The Belief Roadmap: Efficient Planning in Linear POMDPs by Factoring the Covariance, pp. 293–305.
- [16] J. van den Berg, S. Patil, and R. Alterovitz, "Motion planning under uncertainty using iterative local optimization in belief space," *The International Journal of Robotics Research*, vol. 31, no. 11, pp. 1263–1278, 2012.
- [17] E. M. Wolff, U. Topcu, and R. M. Murray, "Robust control of uncertain markov decision processes with temporal logic specifications," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, Dec 2012, pp. 3372–3379.
- [18] T. Quatmann, C. Dehnert, N. Jansen, S. Junges, and J. Katoen, "Parameter synthesis for Markov models: Faster than ever," in *ATVA*.
- [19] C. Dehnert, S. Junges, N. Jansen, F. Corzilius, M. Volk, H. Bruinjtes, J. Katoen, and E. Ábrahám, "Prophesy: A probabilistic parameter synthesis tool," in *CAV (I)*, ser. Lecture Notes in Computer Science, vol. 9206. Springer, 2015, pp. 214–231.
- [20] R. E. Kalman, "'a new approach to linear filtering and prediction problems,'" *Transactions of the ASME—Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.
- [21] J. P. Cunningham, P. Hennig, and S. Lacoste-Julien, "Gaussian probabilities and expectation propagation," *arXiv:1111.6832 [stat.ML]*.