Controller Synthesis for Probabilistic Safety Specifications using Observers [⋆]

Kendra Lesser [*] **Alessandro Abate** [*]

[*] *Department of Computer Science, University of Oxford*
e-mail: {kendra.lesser,alessandro.abate}@cs.ox.ac.uk

Abstract: We present a method for the correct-by-design synthesis of controllers that maximize the safety probability of partially observable stochastic systems. Given a stochastic system with outputs that are corrupted by Gaussian measurement noise, we construct a stochastically contracting observer that produces estimates of the internal state of the system. The contractivity guarantees that the distance between the internal state and the estimate produced by the observer remains bounded, and we can treat the observer as a fully observable abstraction of the original system. For the bounded-horizon probabilistic safety objective, we can synthesize a control law for the observer using a modified safe region according to the bound on the distance above. The control law applied to the original system guarantees that the safety objective is met with some given probability. We showcase the approach on a temperature control problem using a Kalman filter as the observer for a linear stochastic model.

Keywords: partially observable stochastic processes; state observers; probabilistic safety; controller synthesis; contraction theory; abstractions; temperature control.

1. INTRODUCTION

Safety critical systems, such as aircraft, satellites, and electricity grids, have a prohibitively high cost of failure. Controllers for these systems must therefore be designed with guarantees to uphold rigorous safety requirements. Many of these systems further rely on sensors to measure their state and their environment. Sensors, however, are often prone to noise, and cannot capture the entire state of the system. It is therefore of interest to develop formal methods for provably-safe controller synthesis that take this lack of information into account. In particular, our focus is on partially observable stochastic systems, which comprise states evolving with some uncertainty and an observation process. The observation process quantifies which states are accessible, and if their measurements are corrupted by noise. A controller then chooses actions based only on the observation process.

To date, little attention has been given to safety verification of partially observable systems. Results on formal verification are few, and either produce feasible controllers without optimality guarantees (Giro and Rabe (2012)), or are restricted to uncontrolled systems (cf. Zhang et al. (2005)). In a fully observable setting, reachability analysis is often used to asses whether the state of the system will remain within a desired (safe) region of the state space over a given time horizon, particularly in the context of hybrid systems (see Tomlin et al. (2003), Prandini and Hu (2006), Abate et al. (2008)). Verma and Del Vecchio (2012) examine deterministic hybrid systems with hidden modes, and Ghaemi and Del Vecchio (2014) uncertain systems with incomplete information on a partial order, but reachability analysis of a partially observable stochastic hybrid system has been approached only recently by Ding et al. (2013) and Lesser and Oishi (2014).

To the best of our knowledge the only existing approach to actually *computing* safety probabilities and controllers for partially observable stochastic hybrid systems is in Lesser and Oishi (2015). This approach models the partially observable system and safety objective as a partially observable Markov decision process (POMDP) with a multiplicative cost function, and employs approximate POMDP optimization techniques. However, the method suffers from scalability issues, and leads to encoding the synthesized controller into a set of functions that may be difficult to implement in practice.

We propose an alternative and simple approach to synthesize provably safe controllers of partially observable stochastic systems. We first generate an observer to produce a state estimate, then synthesize a controller and compute corresponding safety probabilities for the state estimate, rather than for the actual state of the system. The use of an observer, whose state is known completely, mitigates some of the implementation issues that arise from the POMDP approach of Lesser and Oishi (2015).

The state estimate, however, cannot be treated as the true state of the system. There is a risk that the state estimate is kept within the safe region while the true state is not, which may lead to an (undesirable) overestimate of the probability of safety. We therefore treat the model of the state estimate as an *abstraction* of the actual system, and draw upon existing work that employs abstractions to formally verify properties of complex systems (see, e.g., Tabuada (2009) for an overview of verification using symbolic abstractions). To our knowledge, these techniques have never been used for partially observable systems.

[⋆] This work is supported in part by the European Commission IAPP project AMBI 324432, and by the John Fell Oxford University Press (OUP) Research Fund.

A well-known quantitative notion of model abstraction is that of approximate bisimulation, introduced for stochastic systems in Julius and Pappas (2009). An abstract model approximately bisimulates a given concrete model if the models' outputs remain within a bounded distance of each other, either in expected value or in probability. Abate (2009) uses the notion of stochastic contractivity (introduced in Pham et al. (2009)) to show the existence of an approximate bisimulation between uncontrolled diffusion processes, and Zamani et al. (2014) use incremental input-to-state stability (originally introduced in Angeli (2002) for non-probabilistic models) to generate approximate bisimulations for controlled diffusion processes. The contractivity or stability properties of the system guarantee the existence of a Lyapunov-like function over the two outputs, which is integral to bounding the distance between trajectories of the two models.

The contribution of this paper is an indirect extension of the above notions of formal abstraction via approximate probabilistic bisimulation. Although we do not employ the notion of bisimulation explicitly, we design an observer and show that, if it is stochastically contractive, a Lyapunov-like function exists that bounds the expected value of the distance between the trajectories. The observer can therefore be thought of as a formal abstraction of the original system that is, however, fully observable. As such, we show that safety properties can be verified over the observer with classical techniques, employing a subset of the safe set of interest that is defined according to the maximum expected distance between trajectories. Further, a controller is synthesized over the observer, and shown to guarantee a certain probability of safety when applied directly to the original system. We conclude by showcasing the approach on a linear temperature control problem using a Kalman filter as the observer.

2. PRELIMINARIES

2.1 Notations

We denote the expected value of random quantities with $\mathbb{E}$; when a probability measure or an expected value is induced by a control policy π (to be defined later), we write $\mathbb{P}^\pi$ and $\mathbb{E}^\pi$, respectively. A random variable $x \sim \mathcal{N}(\mu, \Sigma)$ follows a Gaussian distribution with mean μ and covariance Σ. The complement of a set B is written B^c, and $\mathbb{P}[B^c] = 1 - \mathbb{P}[B]$. For a compact set $B \subset \mathbb{R}^n$, the boundary of B is denoted ∂B.

For a space $\mathcal{X}$, $\mathcal{X}^n = \mathcal{X} \times \ldots \times \mathcal{X}$ is the n-times product space of $\mathcal{X}$. The state of a system at discrete time step n is denoted x_n, and the sequence of states $(x_0, x_1, \ldots, x_n)$ is abbreviated as $x_{0:n}$. We use $\|\cdot\|$ to denote the Euclidean vector norm, and if a different norm is required, i.e. $\|\cdot\|_\infty$, then it is explicitly stated. For a matrix A, the norm $\|x\|_A = x^T A x$. The trace of a matrix A is written $\mathrm{tr}(A)$, and the maximum eigenvalue of A is denoted $\lambda_{\max}(A)$.

2.2 Problem Formulation

We consider the following partially observable discrete time controlled stochastic system

$$\begin{aligned} x_{n+1} &= f(x_n, u_n) + \sigma(x_n) w_n, \qquad x_0 = \xi, \\ y_n &= h(x_n) + v_n, \end{aligned} \tag{1}$$

with internal state $x_n \in \mathcal{X}$, output $h(x_n) \in \mathcal{Y}$, noisy output measurement $y_n \in \mathcal{Y}$, and initial condition ξ (to be further characterised shortly). The function f is assumed to be continuous and differentiable, $f \in C^1$. The control input $u_n \in \mathcal{U}$ is assumed known (perfectly measured without noise), but the controller only has access to the noisy output y_n. The state space $\mathcal{X}$ is assumed equal to $\mathbb{R}^m$, and output space $\mathcal{Y}$ equal to $\mathbb{R}^l$, with $l \leq m$. The control space $\mathcal{U}$ is assumed bounded.

The noise terms w_n and v_n are each independent and identically distributed Gaussian random variables, with $w_n \sim \mathcal{N}(0, \mathcal{W})$ and $v_n \sim \mathcal{N}(0, \mathcal{V})$. The initial condition ξ is also an independent Gaussian random variable, with $\xi \sim \mathcal{N}(\mu_0, \Sigma_0)$. We therefore do not know the starting state x_0 exactly, but only know that it initializes according to a Gaussian distribution.

When the internal state of the model is unknown, it can be estimated by constructing an *observer*. An observer is a model designed to produce an estimate of the internal state of the system through knowledge of the dynamics and output measurements, and often takes the form

$$\hat{x}_{n+1} = f(\hat{x}_n, u_n) + L_n(y_n - h(\hat{x}_n)), \tag{2}$$

with L_n the so-called observer gain. The state $\hat{x}_n \in \hat{\mathcal{X}} = \mathbb{R}^m$ is an estimate of x_n. To be practically useful, the observer should produce an accurate state estimate that converges quickly in time to the internal state.

The goal is to construct an observer for system (1), and use that observer to verify safety properties over the internal state trajectory $x_{0:N}$. We want to determine the probability, starting from a given initial condition ξ, that the state trajectory $x_{0:N} \in \mathbb{R}^{mN}$ remains within some compact safe region $K \subset \mathbb{R}^m$ for all time steps $n = 0, \ldots, N$, with $N < \infty$:

$$p^N_{\text{safe}}(\xi, \pi) = \mathbb{P}^\pi[x_{0:N} \in K \mid \xi], \tag{3}$$

where π denotes a policy that generates the sequence of control inputs $u_{0:N-1}$. We specifically consider only Markov policies that map the current state estimate to a control input, formally defined as follows.

Definition 1. A Markov policy π for the observer-based control problem is a sequence of functions $\pi = (\pi_0, \ldots, \pi_{N-1})$ that map the current state estimate to the space of control inputs: $\pi_n : \hat{\mathcal{X}} \to \mathcal{U}$, for all $n = 0, \ldots, N-1$. The set of all such policies is denoted by Π.

While in (3) the policy π is assumed to be given, in general we may want to find the policy that maximizes the safety probability, namely

$$\pi^* = \arg\max_{\pi \in \Pi} p^N_{\text{safe}}(\xi, \pi). \tag{4}$$

We plan to synthesize a control policy for the observer in (2) that, when applied online to (1), guarantees that the internal state remains within K with a probability at least equal to $1 - \alpha$.

Problem 1. Given a partially observable system (1), a compact safe region $K \subset \mathbb{R}^m$, and a design parameter $0 \leq \alpha < 1$, we would like to design an observer (2) such that we can

1. Generate a lower bound to safety probability (3) by finding an equivalent safety probability for the observer.

(2) Synthesize a control policy for the observer that guarantees that the internal state of the model remains within K with probability at least $1-\alpha$.

The evolution of the system in (1) under an observer-based control policy would proceed as follows. At the initial time 0, state x_0 is generated according to $x_0 = \xi \sim \mathcal{N}(\mu_0, \Sigma_0)$, and the observer is initialized as $\hat{x}_0 = \mu_0$. The control input u_0 is selected as $u_0 = \pi_0(\hat{x}_0)$. At time $n-1$, the observation y_{n-1} is recorded. At time n, $\hat{x}_n$ is generated according to (2), and the control input u_n is selected as $u_n = \pi_n(\hat{x}_n)$. The observation y_n is then recorded, and state x_n evolves to x_{n+1} according to (1).

2.3 Stochastic Contraction Theory

A deterministic dynamical system is contracting in some region $\mathcal{O} \subseteq \mathbb{R}^m$ if all trajectories that start within a ball centered around a given trajectory lying inside $\mathcal{O}$ for all time, remain within that ball and converge to the given trajectory. It is distinct from the usual notions of stability that describe convergence to an equilibrium point or to a nominal trajectory, instead describing convergence of trajectories to *each other*. The following definition is extended from Lohmiller and Slotine (1998) for autonomous systems.

Definition 2. The nonlinear controlled discrete-time system

$$x_{n+1} = f(x_n, u_n).$$

is said to be contracting with respect to the uniformly positive definite metric $M_n = \Theta_n^T \Theta_n$ if there exists a constant $0 < \beta < 1$ such that

$$\sup_{u \in \mathcal{U}} \lambda_{\max}\left(F_n^T F_n\right) \leq \beta,$$

for all $x \in \mathcal{X}$ and for all n, with generalized Jacobian $F_n = \Theta_{n+1} \frac{\partial f}{\partial x} \Theta_n^{-1}$.

If the Jacobian $\frac{\partial f}{\partial x}$ is a function of u, then the system may be contracting for a specific value of u, or for all possible $u \in \mathcal{U}$, depending on the problem of interest. In our case, we are interested in contractivity for all $u \in \mathcal{U}$, since we are searching for an optimal control input. It may also be the case that the system is contracting for all $u \in \overline{\mathcal{U}} \subset \mathcal{U}$, in which case we could search for control inputs only over the subset $\overline{\mathcal{U}}$.

The above definition can be extended to stochastic systems, where we are instead concerned with convergence properties over the expected values of the trajectories. As presented in Pham (2008), a stochastic system is *stochastically contracting* if its noiseless dynamics are contracting, and the impact of the noise is bounded.

Definition 3. A stochastic control system of the form

$$x_{n+1} = f(x_n, u_n) + \sigma(x_n) w_n, \tag{5}$$

with control input u_n and noise input $w_n \sim \mathcal{N}(0, \mathcal{W})$, is stochastically contracting in the metric M_n if

(1) The noiseless dynamics $f(x_n, u_n)$ are contracting according to Definition 2 with metric M_n, and
(2) There exists a finite constant C such that

$$\mathrm{tr}(\sigma(x)^T M_n \sigma(x) \mathcal{W}) \leq C$$

for all $x \in \mathcal{X}$, and for all $n \in [0, N]$.

Contraction theory (and the related concept of incremental stability, see Abate (2009)) have been used in the design and analysis of observers for non-probabilistic systems (cf. Sontag and Wang (1997)) or systems with noisy observations (cf. Pham et al. (2009)).

3. SAFETY VERIFICATION OVER ABSTRACTIONS

As mentioned in the Introduction, the notion of approximate probabilistic bisimulation is often used to generate abstractions over stochastic, continuous space models. Two systems are approximately probabilistically bisimilar if control inputs can be chosen for each system such that the outputs of each system (which lie in the same space) remain bounded in probability or expected value, i.e. $\mathbb{E}[\|h^1(x_n^1) - h^2(x_n^2)\|^2] \leq \kappa$ where h^1 and h^2 are the output mappings of two systems indexed by 1 and 2, respectively. In previous work the output is not noisy, and verification is considered directly over the output rather than the internal state.

In particular, Abate (2009) and Julius and Pappas (2009) introduce the notion of a probabilistic bisimulation function $\phi(x^1, x^2)$, which is a supermartingale that bounds the distance between outputs, $\phi(x^1, x^2) \geq \|h^1(x_n^1) - h^2(x_n^2)\|^2$ for all $x^1 \in \mathbb{R}^{m_1}$, $x^2 \in \mathbb{R}^{m_2}$. In discrete time, a supermartingale is a non-increasing random process satisfying

$$\mathbb{E}\left[\phi(x_{n+1}^1, x_{n+1}^2) \mid \phi(x_n^1, x_n^2)\right] \leq \phi(x_n^1, x_n^2). \tag{6}$$

Because the distance between outputs is bounded above by a supermartingale, we may use the following known inequality:

$$\mathbb{P}[\sup_{0 \leq n \leq \infty} \|h^1(x_n^1) - h^2(x_n^2)\|^2 > \epsilon \mid x_0^1, x_0^2]$$
$$\leq \mathbb{P}[\sup_{0 \leq n \leq \infty} \phi(x_n^1, x_n^2) > \epsilon \mid x_0^1, x_0^2] \leq \frac{\phi(x_0^1, x_0^2)}{\epsilon}. \tag{7}$$

If system 2 represents a simplified abstraction of system 1, we can construct a set $K_\epsilon \subset K$, which is the set K with boundaries deflated by size ϵ (K_ϵ is the set K minus the ϵ-neighborhood of the boundary ∂K of K), and determine the probability that x_n^2 remains inside K_ϵ for all n. The inequality (7) then gives a bound on the probability that $x_n^1 \in K$ for all n, as a function of the probability that $x_n^2 \in K_\epsilon$ for all n and of the quantity $\frac{\phi(x_0^1, x_0^2)}{\epsilon}$.

While we will utilize supermartingale properties to analyze partially observable systems, the above approach is not directly applicable in our context because a) we consider outputs with additive noise; and b) we do not wish to verify properties over the noisy output, but rather over the internal states of the original system.

4. OBSERVER DESIGN AND CONTROLLER SYNTHESIS

To overcome the discrepancy between verifying properties over the internal state versus the noisy output, we design an observer and treat it as the abstraction of the concrete model. The state estimate produced by the observer is available for controller synthesis. The control input synthesized over the "abstraction" $\hat{x}_n$ is then directly applied to the concrete model x_n.

4.1 A Bound on the Distance Between Internal States and State Estimates

To guarantee an upper bound to the distance $\mathbb{E}\left[\|x_n - \hat{x}_n\|\right]$, we must have an observer that is stochastically contracting. An observer defined according to (2) is stochastically contracting in the metric M_n if there exists $\beta < 1$ and $C_2 < \infty$ such that

$$\sup_{u,x,n} \lambda_{\max}\left(\Theta_{n+1}\left(\frac{\partial f}{\partial x} - L_n \frac{\partial h}{\partial x}\right)\Theta_n^{-1}\right) \le \beta, \qquad (8)$$

$$\max_n \operatorname{tr}\left(L_n^T M_{n+1} L_n \mathcal{V}\right) \le C_2. \qquad (9)$$

We will also make the additional assumption that

$$\sup_{x,n} \operatorname{tr}\left(\sigma(x)^T \sigma(x) \mathcal{W}\right) \le C_1, \qquad (10)$$

for $C_1 < \infty$, which will be necessary when comparing trajectories $x_{0:N}$ and $\hat{x}_{0:N}$.

For system (1) without measurement noise v_n nor process noise w_n, a contracting observer guarantees convergence of $\hat{x}_n$ to x_n. In the presence of noise, however, we cannot have complete convergence, but rather convergence up to fluctuations because of noise.

To clarify, given the distinct and independent noise processes $v_{0:n}$ and $w_{0:n}$ driving the trajectories $\hat{x}_{0:n}$ and $x_{0:n}$, respectively, and considering the joint process $(x, \hat{x})$, we have the following theorem, extended from Pham (2008).

Theorem 4. For a discrete-time observer (2) that is stochastically contracting in the metric M_n, i.e. that satisfies (8)-(9), and for dynamics (1) with process noise on the internal state satisfying (10), it holds that

$$\mathbb{E}[\|x_n - \hat{x}_n\|^2_{M_n}] \le \beta^n \mathbb{E}[\|\xi - \hat{\xi}\|^2_{M_0}] + \frac{C_1 + C_2}{1-\beta} \qquad (11)$$

for all $n \ge 0$ and for the same control sequence $u_{0:N}$ applied to both $x_{0:N}$ and $\hat{x}_{0:N}$.

The proof can be shown by deriving a Lyapunov-like function $V_n(x_n, \hat{x}_n) = \|x_n - \hat{x}_n\|^2_{M_n}$ over the composed system $(x, \hat{x})$, for which we can state the following.

Theorem 5. The Lyapunov function $V_n(x, \hat{x}) = \|x - \hat{x}\|^2_{M_n}$ over the composed system $(x, \hat{x})$ with dynamics (1) and (2) using the same sequence of control inputs $u_{0:N}$, and under assumptions (8)-(10), satisfies

$$\mathbb{E}[V_{n+1}(x_{n+1}, \hat{x}_{n+1}) \mid V_n(x_n, \hat{x}_n)] \le \beta V_n(x_n, \hat{x}_n) + C,$$

with $C = C_1 + C_2$, for all $n \ge 0$.

Theorem 5 gives a bound on the distance between x_n and $\hat{x}_n$ at any time n, but is not a supermartingale, as in (6), because of the presence of the constant C. Therefore, the bisimulation approach of, e.g., Julius and Pappas (2009) does not apply. However, because we are only interested in *finite time* properties of (1), we can apply inequalities related to supermartingales, as given in (Kushner, 1967, p. 86). We then get the following theorem.

Theorem 6. For observer trajectory $\hat{x}_{0:N}$ with dynamics (2), initialized by $\hat{\xi} = \mu_0$ (Dirac probability distribution centered at μ_0), and true state trajectory $x_{0:N}$ with dynamics (1) initialized by a random variable $\xi \sim \mathcal{N}(\mu_0, \Sigma_0)$, and under assumptions (8) - (10), for any $\epsilon > 0$, policy $\pi \in \Pi$, and time horizon $0 \le N < \infty$, it follows that

$$\mathbb{P}^\pi[\sup_{0\le n\le N} \|x_n - \hat{x}_n\| > \epsilon] \le$$

$$\begin{cases} 1 - \left(1 - \dfrac{\mathbb{E}[\|\xi - \hat{\xi}\|^2]}{\epsilon^2}\right)\left(1 - \dfrac{C}{\epsilon^2}\right)^N, & \text{for } \epsilon \ge \sqrt{\dfrac{C}{1-\beta}}, \\ \dfrac{\mathbb{E}[\|\xi - \hat{\xi}\|^2]\beta^N}{\epsilon^2} + \dfrac{(1-\beta^N)C}{\epsilon^2(1-\beta)}, & \text{for } \epsilon < \sqrt{\dfrac{C}{1-\beta}}. \end{cases}$$

4.2 Correct-by-Design Control using Observer

Based on Theorem 6, we can proceed in the same manner as Julius and Pappas (2009) and Abate (2009) to synthesize a correct-by-design controller, and to generate a lower bound to the safety probability (3).

More precisely, given a parameter ϵ, we can find α such that $P[\sup_{0\le n\le N} \|x_n - \hat{x}_n\| > \epsilon] \le \alpha$. Alternately, given a desired α, we can iterate over possible $\epsilon > 0$ until we find the minimal ϵ for which $\mathbb{P}[\sup_{0\le n\le N} \|x_n - \hat{x}_n\| > \epsilon] \le \alpha$. However ϵ is selected, we then define the set K_ϵ as

$$K_\epsilon = \{x \in \mathbb{R}^m : x \in K \cap \|x - \overline{x}\| > \epsilon, \forall \overline{x} \in \partial K\}. \qquad (12)$$

A controller is then synthesized using existing methods for fully observable stochastic systems (either as done later through dynamic programming, or through an additional finite abstraction step), which maximizes the probability that the state estimate $\hat{x}$ remains within K_ϵ. Applying the synthesized controller to x, we can conclude the following.

Theorem 7. For any $\epsilon > 0$, compact safe sets $K \in \mathbb{R}^m$ and $K_\epsilon \in \mathbb{R}^m$, and given that $\mathbb{P}^\pi[\sup_{0\le n\le N} \|x_n - \hat{x}_n\| > \epsilon] \le \alpha$ for all $\pi \in \Pi$, it follows that

$$\mathbb{P}^{\pi^*}[x_{0:N} \in K \mid \xi] \ge \mathbb{P}^{\pi^*}[\hat{x}_{0:N} \in K_\epsilon \mid \hat{\xi}] - \alpha, \qquad (13)$$

with $\pi^* = \arg\sup_{\pi\in\Pi} \mathbb{P}^\pi[\hat{x}_{0:N} \in K_\epsilon \mid \hat{\xi}]$.

In summary, we have shown that we can solve a conservative safety problem over the observer $\hat{x}$ to generate π^*, as well as $\mathbb{P}^{\pi^*}[\hat{x}_{0:N} \in K_\epsilon \mid \hat{\xi}]$, and apply the optimal policy π^* to the original system online by additionally employing the state estimate at each time step. The closed-loop state trajectory $x_{0:N}$ is then guaranteed to remain within K under policy π^*, with probability at least $\mathbb{P}^{\pi^*}[\hat{x}_{0:N} \in K_\epsilon \mid \hat{\xi}] - \alpha$.

In practice, there is a risk that the bound obtained for α is quite large, requiring ϵ to increase in size to reduce that of α. This could then render the set K_ϵ small, so that the probability of $\hat{x}$ remaining within K_ϵ also becomes small. This is an issue we are currently exploring, but believe there may be a way to characterize and optimize a tradeoff between the size of ϵ and the size of α.

5. CASE STUDY: TEMPERATURE REGULATION

We apply the observer-based control method to a two-room heating example with continuous state dynamics. The state $x_n = [x_n^1, x_n^2]^T$, with x_n^i the temperature in degrees Celsius of room i at time n, so that $\mathcal{X} = \mathbb{R}^2$. The control input $u \in \mathcal{U} = \{0, 1, 2\}$ is a command that programs the heater to heat room one ($u = 1$), room two ($u = 2$), or shut off ($u = 0$). The effect of the input is in $q(u) \in \mathbb{Z}^2$, for which the i^{th} element of $q(u)$ is 1 if $u = i$, and 0 otherwise. The dynamics of the temperature in the two rooms, taken from Abate et al. (2008), are:

$$\begin{bmatrix} x^1_{n+1} \\ x^2_{n+1} \end{bmatrix} = \begin{bmatrix} 0.9613 & 0.022 \\ 0.022 & 0.9613 \end{bmatrix} \begin{bmatrix} x^1_n \\ x^2_n \end{bmatrix} + \begin{bmatrix} 0.8 & 0 \\ 0 & 0.9333 \end{bmatrix} q(u_n) + \begin{bmatrix} 0.1002 \\ 0.1002 \end{bmatrix} + w_n = Ax_n + Bq(u_n) + c + w_n$$

where the state x_n is subject to the additive Gaussian noise $w_n \sim \mathcal{N}(0, [0.025, 0; 0, 0.025])$. Further, x_n is unknown to the controller, and only a noisy observation of the temperature in the first room is available,

$$y_n = [1 \ 0]\, x_n + v_n, \tag{14}$$

with $\mathcal{Y} = \mathbb{R}$ and $v_n \sim \mathcal{N}(0, 0.1)$. The state at time 0 is initialized according to a Gaussian distribution with $x_0 \sim \mathcal{N}(\mu_0, \Sigma_0)$, $\Sigma_0 = [0.02, 0; 0, 0.02]$.

Because the dynamics are linear with additive Gaussian noise, the optimal state estimate is generated by a Kalman filter. The observer dynamics use the steady state Kalman gain matrix L, which in this case is equal to $L = [0.3759, 0.089]^T$,

$$\begin{aligned} \hat{x}_{n+1} = {} & A\hat{x}_n + Bq(u_n) + c + L\,[H(Ax_n + Bq(u_n) + c \\ & + v_n) + w_{n+1} - H(A\hat{x}_n + Bq(u_n) + c)] \\ = {} & (I - LH)A\hat{x}_n + Bq(u_n) + c \\ & + LHAx_n + LHv_n + Lw_{n+1}. \end{aligned}$$

Note that the observer is of the form $\hat{x}_{n+1} = A\hat{x}_n + Bu_n + c + L(y_{n+1} - C\hat{x}_{n+1})$, which is distinct from (2) because we assume a constant gain L and use the observation at time $n+1$ rather than at time n. However, all of the above results still apply.

The state estimate is a function of x_n, which is a Gaussian random variable when conditioned on $y_{0:n}$, with mean $\hat{x}_n$ and covariance Σ_n, calculated according to the Kalman filter. Therefore the probability distribution of the state estimate $\hat{x}_{n+1}$, conditioned on $\hat{x}_n$, is also Gaussian, with mean $A\hat{x}_n + Bq(u_n) + c$ and covariance $A^T H^T L^T P_n LHA + H^T L^T \mathcal{V} LH + L^T \mathcal{W} L$. The state estimate $\hat{x}_n$ can be treated as an independent Markov process with control input u_n, whose transition kernel follows a Gaussian distribution initialized by $\hat{x}_0 = \mu_0$ (μ_0 being the mean of the initial distribution of x_0).

We are interested in controlling the temperature to remain within bounds $K = [17.5, 22] \times [17.5, 22]$. The Kalman observer is contracting in the identity metric $M_n = I$, with contraction rate $\beta = 0.923$, and we can use Theorem 6 to construct the set K_ϵ. Further, the constant C_1 equals 0.0349, and C_2 equals 0.0144. We can then solve a dynamic program over the fully observable estimator model using safe set K_ϵ (see Abate et al. (2008)), to generate a control policy and corresponding safety estimates for the system, according to Theorem 7. Note that we do not consider computation error arising from the dynamic program (which requires discretization of the space $\hat{\mathcal{X}}$), although we could incorporate such an analysis through the work of, e.g., Soudjani and Abate (2013).

We can choose to set either ϵ or α as a design parameter, and then calculate the other value according to Theorem 6. For example, if we desire $\epsilon = 1$, then α must equal 0.254. Alternately, if we desire $\alpha = 0.1$, then $\epsilon = 1.66$. We again point out the trade-off between setting α small (which is desirable), which might cause ϵ to be large and thus likely reduce the probability that $\hat{x}_{0:N} \in K_\epsilon$, versus setting α large (undesirable) which might allow ϵ to be small and likely increase the probability that $\hat{x}_{0:N} \in K_\epsilon$. For this example, ϵ is set to 1.5, and $\alpha = 0.12$, making the set $K_\epsilon = [19, 20.5] \times [19, 20.5]$. Fig. 1 shows the estimated probability that x_n remains within K for $N = 5$ time steps (which is a lower bound to the actual probability) given $\mu_0 = (\hat{x}^1_0, \hat{x}^2_0)$. This is the probability that $\hat{x}_n$ remains in K_ϵ for N time steps minus α, as per Theorem 7.

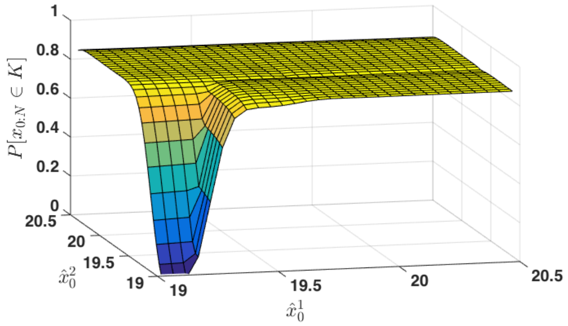

Fig. 1. Probability that $x_n \in K$ for all $n \in [0, 5]$, given a range of initial means $\mu_0 = (\hat{x}^1_0, \hat{x}^2_0) \in K_\epsilon = [19, 20.5] \times [19, 20.5]$. The values are obtained from the probability $\mathbb{P}[\hat{x}_{0:N} \in K_\epsilon]$, minus the calculated bound $\alpha = 0.12$ on the probability that the distance between x and $\hat{x}$ exceeds $\epsilon = 1.5$.

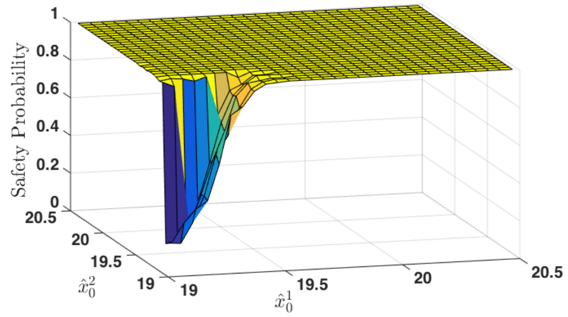

Fig. 2. Simulated probability that $x_n \in K$, $\mathbb{P}[x_{0:N} \in K]$, using the control policy computed from the dynamic program over $\hat{x}$ over varying $\mu_0 = (\hat{x}^1_0, \hat{x}^2_0)$, obtained empirically as the percentage of safe trajectories out of 200 trials.

We also tested the performance of the controller constructed by the dynamic program over the state estimate. We performed 200 simulations for each initial $\hat{x}_0$, and generated a sample trajectory of the state and observations. The state estimate is constructed online using the Kalman filter algorithm, and the appropriate control input is selected using the look-up table generated by the dynamic program. The ratio of successful runs (where the state x_n remains within K for 5 time steps) is presented for varying $\hat{x}_0$ in Fig. 2. Note that the controller performs quite well, and the probability of success according to the simulation is higher than the computed lower bound, as expected.

Fig. 3, however, shows that the maximum distance $\sup_n \|x_n - \hat{x}_n\|$ is not negligible, and the conservativeness in K_ϵ is indeed necessary. Were we to treat $\hat{x}$ as the

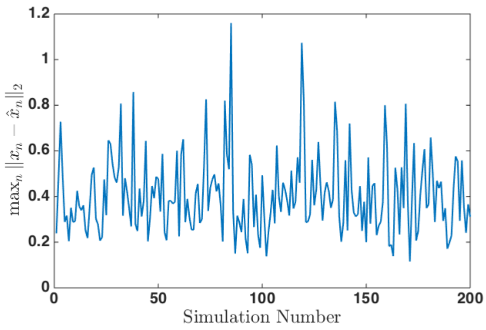

Fig. 3. Maximum Euclidean distance between x and $\hat{x}$ over $N = 5$ time steps, for each of 200 simulations for a random starting value $\hat{x}_0 \in [19, 20.5] \times [19, 20.5]$.

true state and simply generate the probability that $\hat{x}$ remains within K rather than K_ϵ, our controller would not be guaranteed correct to a certain probability, and the associated probabilities of safety would be overestimates.

6. CONCLUSIONS

We have discussed the use of a stochastically contracting observer as an abstraction for a partially observable stochastic system. The expected value of the distance between the state estimate, produced by the observer, and the actual state, remains bounded and converges to a constant value over time, which allows us to bound the probability that the distance between the internal state and the state estimate exceeds some given threshold. We have then used this threshold along with the fully observable observer dynamics to solve an equivalent probabilistic safety problem defined over a subset of the given safe set. As a result, we have been able to synthesize a correct-by-design controller for the original, partially observable system. This approach enables the use of known techniques to synthesize a controller that is a function of the state estimate, which we have access to, rather than using methods for partially observable systems, which are known to be more computationally demanding.

REFERENCES

Abate, A. (2009). A contractivity approach for probabilistic bisimulation of diffusion processes. In *IEEE Conference on Decision and Control*, 2230 – 2235.

Abate, A. (2014). Approximation metrics based on probabilistic bisimulations for general state-space Markov processes: a survey. *Electronic Notes in Theoretical Computer Sciences*, 297, 3–25.

Abate, A., Prandini, M., Lygeros, J., and Sastry, S. (2008). Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. *Automatica*, 44(11), 2724–2734.

Angeli, D. (2002). A Lyapunov approach to incremental stability properties. *IEEE Transactions on Automatic Control*, 47(3), 410–421.

Ding, J., Abate, A., and Tomlin, C. (2013). Optimal control of partially observable discrete time stochastic hybrid systems for safety specifications. In *American Control Conference*, 6231–6236.

Ghaemi, R. and Del Vecchio, D. (2014). Control for safety specifications of systems with imperfect information on a partial order. *IEEE Transactions on Automatic Control*, 59(4).

Giro, S. and Rabe, M.N. (2012). Verification of partial-information probabilistic systems using counterexample-guided refinements. In *Proc. on Automated Technology for Verification and Analysis*, LNCS, 333–348. Springer.

Julius, A.A. and Pappas, G.J. (2009). Approximations of stochastic hybrid systems. *IEEE Transactions on Automatic Control*, 54(6), 1193 – 1203.

Kushner, H.J. (1967). *Stochastic stability and control.* Academic Press, Inc.

Lesser, K. and Oishi, M. (2015). Finite state approximation for verification of partially observable stochastic hybrid systems. In *Hybrid Systems: Computation and Control.*

Lesser, K. and Oishi, M. (2014). Reachability for partially observable discrete time stochastic hybrid systems. *Automatica*, 50(8), 1989–1998.

Lohmiller, W. and Slotine, J.J.E. (1998). On contraction analysis for non-linear systems. *Automatica*, 34(6), 683–696.

Mitchell, I. and Templeton, J. (2005). A toolbox of Hamilton-Jacobi solvers for analysis of nondeterministic continuous and hybrid systems. In *Hybrid Systems: Computation and Control*, 480–494.

Pham, Q.C. (2008). Analysis of discrete and hybrid stochastic systems by nonlinear contraction theory. In *Proceedings of the 10th International Conference on Control, Automation, Robotics, and Vision.*

Pham, Q.C., Tabareau, N., and Slotine, J.J. (2009). A contraction theory approach to stochastic incremental stability. *IEEE Transactions on Automatic Control*, 54(4), 816–820.

Prandini, M. and Hu, J. (2006). *Stochastic Reachability: Theoretical Foundations and Numerical Approximation*, 107–139. Lecture Notes in Control and Information Sciences. Springer Verlag.

Sontag, E. and Wang, Y. (1997). Output-to-state stability and detectability of nonlinear systems. *Systems and Control Letters*, 29, 279–290.

Soudjani, S. and Abate, A. (2013). Adaptive and sequential gridding procedures for the abstraction and verification of stochastic processes. *SIAM Journal on Applied Dynamical Systems*, 12(2), 921–956.

Tabuada, P. (2009). *Verification and Control of Hybrid Systems: A Symbolic Approach.* Springer.

Tomlin, C., Mitchell, I., Bayen, A., and Oishi, M. (2003). Computational techniques for the verification and control of hybrid systems. In *Proceedings of the IEEE*, volume 91, 986–1001.

Verma, R. and Del Vecchio, D. (2012). Safety control of hidden mode hybrid systems. *IEEE Transactions on Automatic Control*, 57(1), 62–77.

Zamani, M., Esfahani, M., Majumdar, R., Abate, A., and Lygeros, J. (2014). Symbolic control of stochastic systems via approximately bisimilar finite abstractions. *IEEE Transactions on Automatic Control*, 59(12), 3135–3150.

Zhang, L., Hermanns, H., and Jansen, D.N. (2005). Logic and model checking for hidden Markov models. In *Proc. on Formal Techniques for Networked and Distributed Systems*, 98–112. Springer.