



State-based confidence bounds for data-driven stochastic reachability using Hilbert space embeddings[☆]

Adam J. Thorpe, Kendric R. Ortiz, Meeko M.K. Oishi^{*}

Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM, USA

ARTICLE INFO

Article history:

Received 30 September 2020

Received in revised form 6 September 2021

Accepted 8 December 2021

Available online 17 January 2022

Keywords:

Machine learning

Stochastic systems

Reachability

Safety

Optimal control

ABSTRACT

In this paper, we compute finite sample bounds for data-driven approximations of the solution to stochastic reachability problems. Our approach uses a nonparametric technique known as kernel distribution embeddings, and provides probabilistic assurances of safety for stochastic systems in a model-free manner. By implicitly embedding the stochastic kernel of a Markov control process in a reproducing kernel Hilbert space, we can approximate the safety probabilities for stochastic systems with arbitrary stochastic disturbances as simple matrix operations and inner products. We present finite sample bounds for point-based approximations of the safety probabilities through construction of probabilistic confidence bounds that are state- and input-dependent. One advantage of this approach is that the bounds are responsive to non-uniformly sampled data, meaning that tighter bounds are feasible in regions of the state- and input-space with more observations. We numerically evaluate the approach, and demonstrate its efficacy on a neural network-controlled pendulum system.

© 2021 Elsevier Ltd. All rights reserved.

1. Introduction

In expensive, high risk, or safety-critical systems, tools for verification are important for ensuring correctness before testing, implementation, or deployment. As autonomy grows in prevalence, and systems continue to grow in size and complexity, there is a need to extend such tools to accommodate learning-enabled components in autonomous systems that resist traditional modeling. Stochastic reachability is an established tool for model-based verification and probabilistic safety (Abate, Prandini, Lygeros, & Sastry, 2008; Summers & Lygeros, 2010) that has been applied to safety-critical problems in a variety of domains, including spacecraft rendezvous and docking (Lesser, Oishi, & Scott Erwin, 2013), robotics and path planning (Malone, Chiang, Lesser, Oishi, & Tapia, 2017), and vehicle control (Kariotoglou, Kamgarpour, Summers, & Lygeros, 2017). Safety refers to the ability of trajectories of the system to respect known constraints on the state space, with at least a desired likelihood, despite bounded control authority. In this paper, we characterize confidence bounds on a data-driven approach for stochastic reachability, enabling rigorous assurances of safety in a model-free manner.

Model-based stochastic reachability has received considerable attention, with methods and tools specific to Markov decision processes (Cauchi & Abate, 2019; Soudjani, Gevaerts, & Abate, 2015), polynomial dynamical systems (Prajna & Jadbabaie, 2004; Prajna, Jadbabaie, & Pappas, 2007), and linear dynamical systems with log-concave disturbances (Vinod, Gleason, & Oishi, 2019; Vinod, HomChaudhuri, & Oishi, 2017; Vinod & Oishi, 2021), as well as the development of various benchmarks for comparisons (Abate et al., 2019, 2020, 2018). Additional work has been done in the model checking community to develop tools for probabilistic, bounded-time reachability problems on parametric stochastic systems which provide interval-based assurances on the safety probabilities (Shmarov et al., 2020; Shmarov & Zuliani, 2015). However, far less work has been done on data-driven stochastic reachability. Methods for statistical verification (Roohi, Wang, West, Dullerud, & Viswanathan, 2017; Wang, Roohi, West, Viswanathan, & Dullerud, 2019; Zarei, Wang, & Pajic, 2020), forward reachable set estimation (Lew & Pavone, 2020; Thorpe, Ortiz, et al., 2021), and data-driven interval-based methods based in adaptive sampling and Gaussian processes (Devonport & Arca, 2020) have been explored, with application to powertrains, bipedal robots, and learning-enabled cyber-physical systems.

We focus in particular on a framework for stochastic reachability that is based in dynamic programming (Abate et al., 2008; Summers & Lygeros, 2010). In contrast to model checking approaches, methods based on this framework are amenable to simultaneous controller synthesis, and the underlying theory accommodates a wide range of dynamical systems that can be

[☆] The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Alessandro Abate under the direction of Editor Ian R. Petersen.

^{*} Corresponding author.

E-mail addresses: ajthor@unm.edu (A.J. Thorpe), kendric@unm.edu (K.R. Ortiz), oishi@unm.edu (M.M.K. Oishi).

captured as a Markov decision process with both discrete- and continuous-valued variables. This approach is general, in that it encompasses formulations for the terminal-hitting time problem (Abate et al., 2008; Summers & Lygeros, 2010), as well as for the first-hitting time problem, which is far more computationally complex (Summers & Lygeros, 2010). We employ a point-based formulation, in which we seek to find the likelihood of maintaining desirable time-varying state constraints over a finite time horizon, from a known initial condition. In contrast to set-based methods (Devonport & Arcak, 2020; Gleason, Vinod, & Oishi, 2017; Gleason, Vinod, & Oishi, 2021), we do not seek to compute the stochastic reachable set, the largest set of states for which the desired likelihood of safety can be maintained.

We draw in particular on tools based in reproducing kernel Hilbert spaces (RKHS), a family of machine learning techniques which enable an implicit approximation of the underlying stochastic kernel. In our previous work (Thorpe & Oishi, 2019), we integrated the RKHS framework with dynamic programming-based approaches for stochastic reachability, to facilitate a data-driven approach to stochastic reachability (Thorpe & Oishi, 2019; Thorpe, Sivaramakrishnan, et al., 2021). Key to the utility of this approach are bounds on the approximation error associated with the RKHS methods. In Thorpe and Oishi (2019), we described the existence of an upper bound on the asymptotic convergence rate, which can be used to infer the existence of a confidence bound on the stochastic reachability probability. However, these bounds are difficult to compute, and hence are not practicable for ascertaining the quality of the result. In this paper, we present *computable confidence bounds* that are state- and input-dependent. The benefit of this approach is that it enables higher accuracy bounds in regions of the state- and input-space for which more training data exists.

Kernel methods have long been used in probability and statistics, (Berlinet & Thomas-Agnan, 2004; Parzen, 1961), and more recently applied to Markov models (Grünewälder, Lever, Baldassarre, Pontil, & Gretton, 2012b), partially observable systems (Nishiyama, Boularias, Gretton, & Fukumizu, 2012; Song, Boots, Siddiqi, Gordon, & Smola, 2010a), and policy synthesis (Lever & Stafford, 2015) in the context of optimal control. Finite sample bounds for conditional distribution embeddings in these contexts are presented in Grünewälder et al. (2012, 2012b), Kanagawa and Fukumizu (2014), Song et al. (2010a), Song, Gretton, and Guestrin (2010b), Song, Huang, Smola, and Fukumizu (2009), which show that the estimated expectation converges in probability to the true expected value with a known, asymptotic convergence rate. The main drawback of kernel-based approaches is the tradeoff that can occur between computability and accuracy. Higher volumes of data are required to ensure an accurate result, which inevitably increases the computational burden associated with inversion of a matrix whose size is proportional to the sample size (Muandet, Fukumizu, Sriperumbudur, Schölkopf, et al., 2017; Rahimi & Recht, 2008). Fortunately, several methods have been developed to significantly reduce the computational overhead (Le, Sarló, Smola, et al., 2013; Rahimi & Recht, 2008; Thorpe, Sivaramakrishnan, et al., 2021).

In this paper, we construct finite sample bounds for kernel embedding-based computation of the stochastic reachability probability, that are specific to the stochastic reachability problem. Our main contribution is the construction of computable state- and input-based upper and lower bounds, obtained via concentration inequalities (McDiarmid, 1989) and tools from statistical learning theory (Vapnik, 1998). Our proposed approach accommodates the fact that the gathered data may not be available uniformly through the state and input space. That is, the bound is dependent upon the sample set from which the kernel embedding is inferred.

The paper is organized as follows. In Section 2, we formulate the problem and provide relevant background information. In Section 3, we derive the finite sample bounds. We discuss implications of the proposed bounds in Section 4, and the problem of parameter selection. In Section 5, we numerically validate the bounds on a stochastic chain of integrators, and demonstrate its utility on a nonlinear pendulum system and a nonlinear cart-pole system with black-box neural network controllers. Concluding remarks are provided in Section 6.

2. Preliminaries

We employ the following notational conventions. Let E be an arbitrary nonempty space. The indicator function $\mathbf{1}_A : E \rightarrow \{0, 1\}$ of $A \subseteq E$ is defined such that $\mathbf{1}_A(x) = 1$ if $x \in A$, and $\mathbf{1}_A(x) = 0$ if $x \notin A$. Let $\mathcal{E}$ denote the σ -algebra on E . If E is a topological space (Çinlar, 2011), the σ -algebra generated by the set of all open subsets of E is called the Borel σ -algebra, denoted by $\mathcal{B}(E)$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ denote a probability space, where $\mathcal{F}$ is the σ -algebra on Ω and $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ is a *probability measure* on the measurable space $(\Omega, \mathcal{F})$. A measurable function $X : \Omega \rightarrow E$ is called a *random variable* taking values in $(E, \mathcal{E})$. The image of $\mathbb{P}$ under X , $\mathbb{P}(X^{-1}A)$, $A \in \mathcal{E}$ is called the *distribution* of X . Let T be an arbitrary set, and for each $t \in T$, let X_t be a random variable. The collection of random variables $\{X_t : t \in T\}$ on $(\Omega, \mathcal{F})$ is a *stochastic process*.

2.1. System model & stochastic reachability problem

Consider a discrete-time stochastic dynamical system described by a Markov control process as defined in Summers and Lygeros (2010).

Definition 1 (Markov Control Process). A Markov control process $\mathcal{H}$ is defined as a 3-tuple, $\mathcal{H} = (\mathcal{X}, \mathcal{U}, Q)$, consisting of: a Borel space $\mathcal{X} \subseteq \mathbb{R}^n$ representing the state space, a compact Borel space $\mathcal{U} \subset \mathbb{R}^m$ representing the control space, and $Q : \mathcal{B}(\mathcal{X}) \times \mathcal{X} \times \mathcal{U} \rightarrow [0, 1]$, a Borel-measurable stochastic kernel which assigns a probability measure $Q(\cdot | x, u)$ to every $x \in \mathcal{X}$ and $u \in \mathcal{U}$ on the Borel space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$.

The system evolves from an initial condition $x_0 \in \mathcal{X}$ over a finite time horizon $t = 0, 1, \dots, N$, $N < \infty$. The control inputs are chosen according to a Markov control policy $\pi = \{\pi_0, \pi_1, \dots, \pi_{N-1}\}$, which is a sequence of universally-measurable (Bertsekas & Shreve, 1978) maps $\pi_i : \mathcal{X} \rightarrow \mathcal{U}$, $i = 0, 1, \dots, N - 1$.

For simplicity, we assume that the policy π is stationary, meaning $\pi(x) = \pi_0(x) = \pi_1(x) = \dots = \pi_{N-1}(x)$ for all $x \in \mathcal{X}$. This is a simplifying assumption for the purpose of analysis, and the extension to non-stationary policies is trivial.

We consider a probabilistic reach/avoid problem known as the *terminal-hitting time problem* where the objective is to determine the likelihood that the system starting at an initial condition $x_0 \in \mathcal{X}$ will remain in a pre-defined safe set and reach a target set at the final time instant. Let $\mathcal{K}, \mathcal{T} \in \mathcal{B}(\mathcal{X})$, denote the safe set and target set, respectively. From Summers and Lygeros (2010), the *terminal-hitting time safety probability* $r_{x_0}^\pi(\mathcal{K}, \mathcal{T})$ is defined as the probability that a system $\mathcal{H}$ following a Markov policy π from the initial condition x_0 will reach the target set $\mathcal{T}$ at time N while remaining within the safe set $\mathcal{K}$ for all $t = 0, 1, \dots, N - 1$,

$$r_{x_0}^\pi(\mathcal{K}, \mathcal{T}) = \mathbb{E}_{x_0}^\pi \left[\left(\prod_{i=0}^{N-1} \mathbf{1}_{\mathcal{K}}(x_i) \right) \mathbf{1}_{\mathcal{T}}(x_N) \right]. \quad (1)$$

For a fixed Markov policy π , we define the value functions $V_t^\pi : \mathcal{X} \rightarrow [0, 1]$, $t = 0, \dots, N$, via *backward recursion*:

$$V_N^\pi(x) = \mathbf{1}_{\mathcal{T}}(x), \quad (2)$$

$$V_t^\pi(x) = \mathbf{1}_{\mathcal{K}}(x) \int_{\mathcal{X}} V_{t+1}^\pi(y) Q(dy | x, \pi(x)). \quad (3)$$

Then according to [Summers and Lygeros \(2010\)](#), we have that $V_0^\pi(x_0) = r_{x_0}^\pi(\mathcal{K}, \mathcal{T})$ for every $x_0 \in \mathcal{X}$.

We consider the case where the stochastic kernel Q is unknown, which means the integral in (3) becomes intractable.

Assumption 2. The stochastic kernel Q is unknown.

Instead, we presume that a finite sample $\mathcal{S}$ of $M \in \mathbb{N}$ observations is available, taken independently and identically distributed (i.i.d.) from the stochastic kernel Q ,

$$\mathcal{S} = \{(y_i, x_i, u_i)\}_{i=1}^M, \quad (4)$$

where x_i are taken i.i.d. from a probability measure with support $\mathcal{X}$, $u_i = \pi(x_i)$, and $y_i \sim Q(\cdot | x_i, u_i)$.

2.2. Conditional distribution embeddings

Because Q is assumed to be unknown, we cannot compute the safety probabilities $r_{x_0}^\pi(\mathcal{K}, \mathcal{T})$ directly. Thus, we seek to approximate the safety probabilities by numerically approximating the integral with respect to $Q(\cdot | x, \pi(x))$ in (3) using the sample $\mathcal{S}$. In order to approximate the integral of V_{t+1}^π in (3), we choose to represent the integral operator with respect to $Q(\cdot | x, \pi(x))$ as an element in a high-dimensional space of functions known as a reproducing kernel Hilbert space (RKHS).

Define a positive-definite ([Christmann & Steinwart, 2008](#), Definition 4.15) function $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ known as a kernel function and let $\mathcal{H}$ be a Hilbert space of functions of the form $\mathcal{X} \rightarrow \mathbb{R}$ equipped with an inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and the induced norm $\|\cdot\|_{\mathcal{H}}$.

Definition 3 (RKHS). Let E be an arbitrary space. A Hilbert space $\mathcal{H}$ is a reproducing kernel Hilbert space (RKHS) if there exists a positive-definite kernel function $k : E \times E \rightarrow \mathbb{R}$ that satisfies the following properties ([Aronszajn, 1950](#)):

1. $k(x, \cdot) \in \mathcal{H}$ for all $x \in E$, and
2. $f(x) = \langle f, k(x, \cdot) \rangle_{\mathcal{H}}$ for all $f \in \mathcal{H}$ and $x \in E$.

where the second property is known as the *reproducing property*.

Remark 4. Conversely, by the Moore–Aronszajn theorem ([Aronszajn, 1950](#)), for any positive-definite kernel k , there exists a unique RKHS with k as its reproducing kernel, where $\mathcal{H}$ is defined as the closure of the linear span of kernel functions, i.e. $\mathcal{H} = \overline{\text{span}\{k(x, \cdot) | x \in \mathcal{X}\}}$. In short, this means that by defining a reproducing kernel, we obtain a corresponding RKHS. See, e.g., [Berlinet and Thomas-Agnan \(2004\)](#), [Christmann and Steinwart \(2008\)](#), [Schölkopf and Smola \(2001\)](#) for more information on reproducing kernel Hilbert spaces.

The reproducing property is central to our approach, since it allows us to evaluate any function in the RKHS as a Hilbert space inner product. By embedding the integral operator with respect to $Q(\cdot | x, \pi(x))$ as an element in the RKHS, we can use the reproducing property to evaluate the integral in (3).

For the measurable space $\mathcal{X}$, define the kernel function $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ with the associated RKHS $\mathcal{H}$. We impose a mild simplifying assumption for the purpose of analysis and assume that the kernel k is bounded above by a real number $\rho < \infty$, where $\sup_{x \in \mathcal{X}} (k(x, x))^{1/2} \leq \rho < \infty$. Such a kernel is called a bounded kernel ([Christmann & Steinwart, 2008](#), §4.3). This assumption (along with the measurability of k) ensures that we can represent the integral operator as an element in the RKHS.

Let $\mathcal{P}$ denote the set of probability measures on $\mathcal{X}$ which are densities of $Y \in \mathcal{X}$ conditioned on $(X, U) \in \mathcal{X} \times \mathcal{U}$ (of which

the probability measures defined by the stochastic kernel Q are a part). For any probability measure $Q(\cdot | x, u) \in \mathcal{P}$, if the following necessary and sufficient condition $\mathbb{E}_{Y \sim Q(\cdot | x, u)}[k(Y, Y)] < \infty$ is satisfied ([Sriperumbudur, Gretton, Fukumizu, Schölkopf, & Lanckriet, 2010](#)) (which is satisfied if k is bounded and measurable on $\mathcal{X}$), there exists an element $m(x, u) \in \mathcal{H}$ called the *conditional distribution embedding* ([Song et al., 2009](#)), defined as:

$$m : \mathcal{P} \rightarrow \mathcal{H},$$

$$Q(\cdot | x, u) \mapsto m(x, u) := \int_{\mathcal{X}} k(y, \cdot) Q(dy | x, u). \quad (5)$$

Using this representation, we can embed the integral with respect to $Q(\cdot | x, u)$ as an element in the RKHS $\mathcal{H}$. Furthermore, if the kernel is *characteristic* ([Sriperumbudur, Fukumizu, & Lanckriet, 2011](#)), then the embedding is unique, meaning the embedding captures all statistical information of the underlying distribution, and no information is lost by this representation.

As shown in [Thorpe and Oishi \(2019\)](#), we can use conditional distribution embeddings to solve the stochastic reachability problem. By the reproducing property, we can evaluate the integral of any function $f \in \mathcal{H}$ with respect to $Q(\cdot | x, \pi(x))$ as a Hilbert space inner product with the embedding $m(x, \pi(x))$. Thus, assuming the value functions V_t^π , $t = 1, \dots, N$, are in $\mathcal{H}$, we can evaluate the integral in (3) as:

$$\begin{aligned} \langle V_{t+1}^\pi, m(x, \pi(x)) \rangle_{\mathcal{H}} \\ = \left\langle V_{t+1}^\pi, \int_{\mathcal{X}} k(y, \cdot) Q(dy | x, \pi(x)) \right\rangle_{\mathcal{H}} \end{aligned} \quad (6)$$

$$= \int_{\mathcal{X}} \langle V_{t+1}^\pi, k(y, \cdot) \rangle_{\mathcal{H}} Q(dy | x, \pi(x)) \quad (7)$$

$$= \int_{\mathcal{X}} V_{t+1}^\pi(y) Q(dy | x, \pi(x)). \quad (8)$$

Using the embedding $m(x, \pi(x))$, we can substitute the integral expression in the stochastic reachability backward recursion (3) with a Hilbert space inner product in order to compute the safety probabilities $r_{x_0}^\pi(\mathcal{K}, \mathcal{T})$. This reduces the evaluation of potentially expensive integrals to a simple linear operation in Hilbert space.

2.3. Empirically approximating the value functions

In practice, we do not have access to the true embedding $m(x, \pi(x))$ since the stochastic kernel Q is unknown, meaning we cannot compute (3) directly. Thus, we compute an empirical approximation of the embedding $m(x, \pi(x))$ using a finite sample $\mathcal{S}$ as in (4) of size $M \in \mathbb{N}$ collected i.i.d. from Q .

Following [Micchelli and Pontil \(2005\)](#), using a sample $\mathcal{S}$, we can compute an empirical estimate $\hat{m}(x, u)$ of an embedding $m(x, u)$ as the solution to the following regularized least-squares problem ([Grünewälder et al., 2012](#); [Micchelli & Pontil, 2005](#)),

$$\min_{\hat{m}} \frac{1}{M} \sum_{i=1}^M \|k(y_i, \cdot) - \hat{m}(x_i, u_i)\|_{\mathcal{H}}^2 + \lambda \|\hat{m}\|_{\Gamma}^2, \quad (9)$$

where Γ is a vector-valued RKHS ([Micchelli & Pontil, 2005](#)) and $\lambda > 0$ is the regularization parameter. By the representer theorem ([Micchelli & Pontil, 2005](#)), the solution to (9) is unique and has the following form:

$$\hat{m} = \sum_{i=1}^M \alpha_i k(x_i, \cdot) l(u_i, \cdot), \quad (10)$$

where $\alpha \in \mathbb{R}^M$ is a vector of real-valued coefficients and $l : \mathcal{U} \times \mathcal{U} \rightarrow \mathbb{R}$ is a reproducing kernel function over $\mathcal{U}$. For simplicity, we assume that the kernel l is bounded and the bound ρ of the kernel k is also a bound for l . By substituting (10) into (9) and

taking the derivative with respect to α , we obtain the following closed-form solution,

$$\hat{m}(x, u) = \Phi^\top W \Psi k(x, \cdot) l(u, \cdot), \quad (11)$$

where Φ and Ψ are known as *feature vectors*, with elements $\Phi_i = k(y_i, \cdot)$ and $\Psi_i = k(x_i, \cdot) l(u_i, \cdot)$, respectively, and $W = (\Psi \Psi^\top + \lambda MI)^{-1}$. For simplicity, let

$$\beta(x, u) := W \Psi k(x, \cdot) l(u, \cdot) \quad (12)$$

be a vector of real-valued coefficients that depends on the value of the conditioning variables x and u , such that $\hat{m}(x, u) = \Phi^\top \beta(x, u)$.

As shown in Thorpe and Oishi (2019), using an estimate $\hat{m}(x, \pi(x))$ of $m(x, \pi(x))$, we can approximate the integral of the value functions with respect to $Q(\cdot | x, \pi(x))$ in (3) as an inner product,

$$\begin{aligned} \langle V_{t+1}^\pi, \hat{m}(x, \pi(x)) \rangle_{\mathcal{H}} &= \mathbf{V}_{t+1}^{\pi^\top} \beta(x, \pi(x)) \\ &\approx \int_{\mathcal{X}} V_{t+1}^\pi(y) Q(dy | x, \pi(x)), \end{aligned} \quad (13)$$

where $\mathbf{V}_{t+1}^\pi = [V_{t+1}^\pi(y_1), \dots, V_{t+1}^\pi(y_M)]^\top$. Thus, using the estimate $\hat{m}(x, \pi(x))$, we can recursively approximate and substitute the value functions in the stochastic reachability backward recursion in order to approximate the safety probabilities. We summarize this procedure as Lemma 5.

Lemma 5 (Approximate Backward Recursion, Thorpe & Oishi, 2019). Let π be a fixed Markov policy. Define the approximate value functions $V_t^\pi : \mathcal{X} \rightarrow [0, 1]$, $t = 0, \dots, N$ by the backward recursion:

$$\bar{V}_N^\pi(x) = V_N^\pi(x), \quad (14)$$

$$\bar{V}_t^\pi(x) = \mathbf{1}_{\mathcal{K}}(x) \langle \bar{V}_{t+1}^\pi, \hat{m}_{Y|x, \pi(x)} \rangle_{\mathcal{H}}. \quad (15)$$

Then $r_{x_0}^\pi(\mathcal{K}, \mathcal{T}) \approx \bar{V}_0^\pi(x_0)$.

Note that since the estimate $\hat{m}(x, \pi(x))$ is conditioned on a particular value of $x \in \mathcal{X}$, by approximating the safety probability $r_{x_0}^\pi(\mathcal{K}, \mathcal{T})$ using Lemma 5, we obtain a point-based approximation of the safety probability at a particular value of $x_0 \in \mathcal{X}$. Lemma 5 provides a model-free approach to approximate the stochastic reachability probability, and can easily be extended to solve related problems, including the first-hitting time problem (Summers & Lygeros, 2010) and the multiplicative and maximal cost stochastic reachability problems in Abate et al. (2008).

3. Finite sample bounds

The difficulty in finding bounds on the stochastic reachability probability stems from the underlying structure of the conditional distribution embedding estimate. Unlike the embedding for a marginal distribution (Smola, Gretton, Song, & Schölkopf, 2007), which has uniform coefficients $1/M$, the conditional distribution embedding has non-uniform coefficients $\beta(x, u)$ (12) which depend upon the value of the conditioning variables. This complicates the application of existing mathematical techniques from statistical learning theory.

It is worth noting that in our case, we directly bound the variation of the estimator at a particular value of the conditioning variables x and u . This means the bounds we derive provide a localized result, which can be used to assess the quality of the approximation at a particular point. This is not a significant limitation in the context of Lemma 5, since in the case of the stochastic reachability backward recursion, we seek to evaluate the safety probability at a single point.

In order to determine a bound on the quality of the approximation obtained using Lemma 5, we seek a bound on the difference between the expectation of the value function and its empirical counterpart.

3.1. Worst-case difference between the true and the empirical expectation

Assume that $V_t^\pi \in \mathcal{H}$, $t = 1, \dots, N$, and assume that for all $f \in \mathcal{H}$, $f \in [0, 1]$ and $\|f\|_{\mathcal{H}} \leq 1$. We begin by upper bounding the deviation of the empirical expectation computed using an estimate $\hat{m}(x, u)$ from the true expectation for any value function V_t^π . In other words, we seek a state-based bound $B(x, u) \in \mathbb{R}$ such that for any value function V_t^π , $t = 1, \dots, N$,

$$|\mathbb{E}_{Y \sim Q(\cdot | x, u)}[V_t^\pi(Y)] - \mathbf{V}_t^{\pi^\top} \beta(x, u)| \leq B(x, u). \quad (16)$$

For simplicity of notation, let $\mathbb{E}f := \mathbb{E}_{Y \sim Q(\cdot | x, u)}[f(Y)]$.

We can uniformly bound the difference between the value function expectation and the empirical expectation computed using an estimate $\hat{m}(x, u)$. Note that for any value function $V_t^\pi \in \mathcal{H}$, not necessarily attaining the supremum,

$$|\mathbb{E}V_t^\pi - \mathbf{V}_t^{\pi^\top} \beta(x, u)| \leq \sup_{\|h\|_{\mathcal{H}} \leq 1} |\mathbb{E}h - \mathbf{h}^\top \beta(x, u)|, \quad (17)$$

where h is some function in $\mathcal{H}$, $\mathbf{h} = [h(y_1), \dots, h(y_M)]^\top$, and $\beta(x, u)$ is defined as in (12). This means that in the worst case, the deviation of the estimated value function expectation from the true value function expectation is less than or equal to the deviation of a function $h^* \in \mathcal{H}$ which satisfies the supremum.

We then bound the right-hand side of (17) using McDiarmid's inequality. In simple terms, McDiarmid's inequality states that if the empirical estimate $\mathbf{h}^\top \beta(x, u)$ computed using $\mathcal{S}$ has bounded variation when a single observation in the sample $\mathcal{S}$ is changed, then the deviation of the estimate from the true expectation is bounded by some quantity that depends on the variation bound.

Definition 6 (Bounded Differences Condition). Given coefficients $c_i \geq 0$, $i = 1, \dots, M$, a function $f : E^M \rightarrow \mathbb{R}$ satisfies the bounded differences condition if

$$\sup_{\substack{x_1, \dots, x_M \\ x'_i \in E}} |f(x_1, \dots, x_M) - f(x_1, \dots, x'_i, \dots, x_M)| \leq c_i \quad (18)$$

for every $i = 1, \dots, M$.

Lemma 7 (McDiarmid's Inequality, McDiarmid, 1989). Let $X = \{X_1, \dots, X_M\}$ be independent random variables taking values in a set E , and assume that the function $f : E^M \rightarrow \mathbb{R}$ satisfies the bounded differences condition (Definition 6). Then for every $\varepsilon > 0$,

$$\Pr(|f(X) - \mathbb{E}[f(X)]| \geq \varepsilon) \leq \exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^M c_i^2}\right). \quad (19)$$

Alternatively, Lemma 7 implies that, given a small probability $\delta/2 \in (0, 1)$, then with probability $1 - \delta/2$, the deviation of the function f from the expectation $\mathbb{E}[f(X)]$ is bounded by:

$$f(X) - \mathbb{E}[f(X)] \leq \sqrt{\frac{MC^2 \log(2/\delta)}{2}}, \quad (20)$$

where $C \geq c_i$ for all $i = 1, \dots, M$.

In order to determine the bound on (17), we seek some constant C that satisfies the bounded differences condition. However, the effect of changing an individual observation in the empirical expectation term in (17) is non-trivial, since changing a single observation affects all elements of the coefficient vector $\beta(x, u)$.

Therefore, in order to determine C , we make use of a well-known result in linear algebra known as the push-through identity (Bernstein, 2009, Fact 2.16.16).

Lemma 8 (Push-Through Identity, Bernstein, 2009, Fact 2.16.16). Let $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{m \times n}$, and assume that $AB + I$ is non-singular. Then $BA + I$ is non-singular and

$$(AB + I)^{-1}A = A(BA + I)^{-1}. \quad (21)$$

We now prove the following theorem:

Theorem 9. *The variation of $\mathbf{h}^\top \beta(x, u)$ by changing a single observation is at most $\rho/(\lambda M)$.*

Proof. For some $h \in \mathcal{H}$, let $\mathbf{h}^\top \beta(x, u)$ be defined as in (17), such that $\beta(x, u) = W\Psi k(x, \cdot)l(u, \cdot)$, where Ψ is a feature vector with elements $\Psi_i = k(x_i, \cdot)l(u_i, \cdot)$, $W = (\Psi\Psi^\top + \lambda MI)^{-1}$, and the kernels k and l are bounded by some constant $\rho < \infty$.

Using Lemma 8, we can write $W\Psi$ as:

$$W\Psi = (\Psi\Psi^\top + \lambda MI)^{-1}\Psi = \Psi(\Psi^\top\Psi + \lambda M)^{-1}, \quad (22)$$

where we note that $(\Psi^\top\Psi + \lambda M)^{-1}$ is scalar. Using the identity in (22), we can rewrite the estimate $\mathbf{h}^\top \beta(x, u)$ as:

$$\mathbf{h}^\top W\Psi k(x, \cdot)l(u, \cdot) = \frac{1}{\Psi^\top\Psi + \lambda M} \mathbf{h}^\top \Psi k(x, \cdot)l(u, \cdot). \quad (23)$$

Since $h \in [0, 1]$ by assumption, we have from (23) that the variation for changing a single observation is at most $\rho/(\lambda M)$, which proves the result. $\square$

Remark 10. Note that the result provided by Theorem 9 is strong, and is verified by the result presented in Bousquet and Elisseeff (2002, Theorem 22) (with a scaling factor of 1 instead of $1/2$, which comes from the worst-case bound).

From Theorem 9, we have that $C = \rho/(\lambda M)$ satisfies the bounded differences condition (Definition 6). Continuing from (17), using McDiarmid's inequality (19), we have that given $\delta/2 \in (0, 1)$, with probability $1 - \delta/2$,

$$\sup_{\|h\|_{\mathcal{H}} \leq 1} |\mathbb{E}h - \mathbf{h}^\top \beta(x, u)| \leq \mathbb{E}_S \left[\sup_{\|h\|_{\mathcal{H}} \leq 1} |\mathbb{E}h - \mathbf{h}^\top \beta(x, u)| \right] + \sqrt{\frac{\rho^2 \log(2/\delta)}{2\lambda^2 M}}. \quad (24)$$

We now have an expression for the worst-case difference between the true and empirical stochastic reachability probabilities. However, because the expectation on the right-hand side of (24) relies upon the true expectation of h , it is not directly computable.

3.2. Removing reliance upon the true expectation

In order to enable computability of the bound in (24), we bound the first term on the right-hand side of (24) via symmetrization and then utilize the properties of the RKHS to bound the expectation of the worst-case empirical estimate. In particular, we make use of the reproducing property and the definition of the dual norm for Hilbert spaces, which we present here adapted from Rudin (1991, Theorem 4.3).

Definition 11 (Dual Norm). Let $\mathcal{H}$ be a Hilbert space. For any $f, g \in \mathcal{H}$, $\|f\|_{\mathcal{H}} = \sup_{\|g\|_{\mathcal{H}} \leq 1} |\langle f, g \rangle_{\mathcal{H}}|$.

In addition, we rely upon a special type of random variable known as a Rademacher variable (cf. Bartlett & Mendelson, 2002), which is a uniform random variable taking values in $\{-1, 1\}$.

Definition 12 (Rademacher Variable, Bartlett & Mendelson, 2002). A random variable σ is called a *Rademacher variable* if it is independent uniform, such that $\Pr(\sigma = 1) = \Pr(\sigma = -1) = 1/2$.

We now prove the following lemma:

Lemma 13. *Let $h \in \mathcal{H}$ and $\mathbf{h}^\top \beta(x, u)$ be defined as in (17). Given $\delta/2 \in (0, 1)$, then with probability $1 - \delta/2$,*

$$\mathbb{E}_S \left[\sup_{\|h\|_{\mathcal{H}} \leq 1} |\mathbb{E}h - \mathbf{h}^\top \beta(x, u)| \right]$$

$$\leq 2\sqrt{\text{tr}(\beta^\top \Phi \Phi^\top \beta)} + 2\sqrt{\frac{\rho^2 \log(2/\delta)}{2\lambda^2 M}}. \quad (25)$$

Proof. We begin by bounding the first term on the right-hand side of (24) via symmetrization (Vapnik, 1998). Let $\tilde{S}$ be a *ghost sample*, that is, an independent copy of S that is drawn from the same sampling distribution as S (Vapnik, 1998). We replace the expectation in the first term on the right-hand side of (24) with a second empirical estimate computed using $\tilde{S}$, to obtain:

$$\mathbb{E}_S \left[\sup_{\|h\|_{\mathcal{H}} \leq 1} |\mathbb{E}h - \mathbf{h}^\top \beta(x, u)| \right] \leq \mathbb{E}_S \left[\sup_{\|h\|_{\mathcal{H}} \leq 1} |\mathbb{E}_{\tilde{S}}[\tilde{\mathbf{h}}^\top \tilde{\beta}(x, u) - \mathbf{h}^\top \beta(x, u)]| \right] \quad (26)$$

$$\leq \mathbb{E}_{S\tilde{S}} \left[\sup_{\|h\|_{\mathcal{H}} \leq 1} |\tilde{\mathbf{h}}^\top \tilde{\beta}(x, u) - \mathbf{h}^\top \beta(x, u)| \right], \quad (27)$$

where $\tilde{\beta}(x, u)$ is computed using $\tilde{S}$ as in (12) and $\tilde{\mathbf{h}} = [h(\tilde{y}_1), \dots, h(\tilde{y}_M)]^\top$. Eq. (26) follows almost surely from the properties of conditional expectations and (27) follows by the convexity of the supremum.

We next exploit the symmetry of the empirical distributions to upper bound (27). Let σ be Rademacher variables with $\sigma_i \in \{-1, 1\}$, and let $\beta := \text{diag}(\beta(x, u))$.

Since the distribution of the difference in empirical expectations $\tilde{\mathbf{h}}^\top \tilde{\beta}(x, u) - \mathbf{h}^\top \beta(x, u)$ is symmetric around 0, which follows since $\mathbf{f}^\top \beta(x, u) \in [0, 1]$ for all $f \in \mathcal{H}$, we see that $\tilde{\mathbf{h}}^\top \tilde{\beta}\sigma - \mathbf{h}^\top \beta\sigma$ has the same distribution. In effect, the Rademacher variables randomly exchange observations in S and $\tilde{S}$ with probability $1/2$. When we take the expectation over σ , the expectations of the empirical estimates computed using S and $\tilde{S}$ are the same. Using this fact, we obtain the following:

$$\mathbb{E}_{S\tilde{S}} \left[\sup_{\|h\|_{\mathcal{H}} \leq 1} |\tilde{\mathbf{h}}^\top \tilde{\beta}(x, u) - \mathbf{h}^\top \beta(x, u)| \right] = \mathbb{E}_{S\tilde{S}\sigma} \left[\sup_{\|h\|_{\mathcal{H}} \leq 1} |\tilde{\mathbf{h}}^\top \tilde{\beta}\sigma - \mathbf{h}^\top \beta\sigma| \right] \quad (28)$$

$$\leq 2\mathbb{E}_{S\sigma} \left[\sup_{\|h\|_{\mathcal{H}} \leq 1} |\mathbf{h}^\top \beta\sigma| \right]. \quad (29)$$

Then by applying the reproducing property (Definition 3) and the definition of the dual norm for Hilbert spaces (Definition 11), we remove the dependence on h :

$$2\mathbb{E}_{S\sigma} \left[\sup_{\|h\|_{\mathcal{H}} \leq 1} |\mathbf{h}^\top \beta\sigma| \right] = 2\mathbb{E}_{S\sigma} \left[\sup_{\|h\|_{\mathcal{H}} \leq 1} |\langle h, \Phi^\top \beta\sigma \rangle_{\mathcal{H}}| \right] \quad (30)$$

$$= 2\mathbb{E}_{S\sigma} [\|\Phi^\top \beta\sigma\|_{\mathcal{H}}]. \quad (31)$$

We then utilize the definition of the Hilbert space norm and the concavity of the square root and note that the expectation of the Rademacher variables $\mathbb{E}_\sigma[\sigma_i \sigma_j]$ vanishes except when $i = j$,

$$2\mathbb{E}_{S\sigma} [\|\Phi^\top \beta\sigma\|_{\mathcal{H}}] \leq 2\mathbb{E}_S \left[\left(\mathbb{E}_\sigma [\sigma^\top \beta^\top \Phi \Phi^\top \beta \sigma] \right)^{1/2} \right] \quad (32)$$

$$= 2\mathbb{E}_S \left[\left(\text{tr}(\beta^\top \Phi \Phi^\top \beta) \right)^{1/2} \right]. \quad (33)$$

By bounding the expectation in (33) by McDiarmid's inequality again, we obtain:

$$2\mathbb{E}_S \left[\left(\text{tr}(\beta^\top \Phi \Phi^\top \beta) \right)^{1/2} \right] \leq 2\sqrt{\text{tr}(\beta^\top \Phi \Phi^\top \beta)} + 2\sqrt{\frac{\rho^2 \log(2/\delta)}{2\lambda^2 M}}, \quad (34)$$

which proves the result. $\square$

Continuing from (24), using McDiarmid's inequality with $C = \rho/(\lambda M)$ via Theorem 9 and Lemma 13, we have that given $\delta/2 \in (0, 1)$, with probability $1 - \delta/2$,

$$\sup_{\|h\|_{\mathcal{H}} \leq 1} |\mathbb{E}h - \mathbf{h}^\top \beta(x, u)| \leq 2\sqrt{\text{tr}(\beta^\top \Phi \Phi^\top \beta)} + 3\sqrt{\frac{\rho^2 \log(2/\delta)}{2\lambda^2 M}}. \quad (35)$$

Thus, we have a computable bound for the worst-case difference between the true and empirical expectation of a function $h \in \mathcal{H}$.

Remark 14. We note that even in the worst case, the first term on the right-hand side of (35) is bounded, since $\|\Phi^\top \beta(x, u)\|_{\mathcal{H}} \leq 1$ by assumption. However, in practice, with appropriate kernel selection and choice of regularization parameter λ , this term will often be significantly less than 1.

3.3. Finite sample bound on the safety probability

We now can state the main result, which we present as Theorem 15.

Theorem 15. For any value function V_t^π , given $\delta/2 \in (0, 1)$, with probability $1 - \delta/2$, the difference between the true and empirical expectation of the value functions is bounded by:

$$|\mathbb{E}V_t^\pi - \mathbf{V}_0^{\pi^\top} \beta(x, u)| \leq 2\sqrt{\text{tr}(\beta^\top \Phi \Phi^\top \beta)} + 3\sqrt{\frac{\rho^2 \log(2/\delta)}{2\lambda^2 M}}. \quad (36)$$

Proof. The proof follows from the arguments presented for (35). In (17), we uniformly bound the difference between the true and empirical expectation of the value functions by the worst-case function $h \in \mathcal{H}$. We then use McDiarmid's inequality (Lemma 7) with the bound $C = \rho/(\lambda M)$ satisfying the bounded differences condition (Theorem 9) to obtain (24). Using a ghost sample, the symmetrization argument, and the definition of the dual norm for Hilbert spaces, we then bound the first term on the right-hand side of (24) in order to remove the dependence on h ,

$$\mathbb{E}_S \left[\sup_{\|h\|_{\mathcal{H}} \leq 1} |\mathbb{E}h - \mathbf{h}^\top \beta(x, u)| \right] \leq 2\sqrt{\text{tr}(\beta^\top \Phi \Phi^\top \beta)} + 2\sqrt{\frac{\rho^2 \log(2/\delta)}{2\lambda^2 M}}. \quad (37)$$

We then substitute the bound in (37) into (24) and obtain the result,

$$|\mathbb{E}V_t^\pi - \mathbf{V}_0^{\pi^\top} \beta(x, u)| \leq 2\sqrt{\text{tr}(\beta^\top \Phi \Phi^\top \beta)} + 3\sqrt{\frac{\rho^2 \log(2/\delta)}{2\lambda^2 M}}, \quad (38)$$

which proves (36). $\square$

Let $B(x, u)$ be the bound on the difference between the expected value of the value functions and its empirical counterpart in (36), given by

$$B(x, u) = 2\sqrt{\text{tr}(\beta^\top \Phi \Phi^\top \beta)} + 3\sqrt{\frac{\rho^2 \log(2/\delta)}{2\lambda^2 M}}. \quad (39)$$

This means that for any value function V_t^π , given $\delta/2 \in (0, 1)$, with probability $1 - \delta/2$, that the absolute difference between the actual expectation and the empirical expectation computed using $\hat{m}(x, \pi(x))$ is bounded by

$$-B(x, u) \leq \mathbb{E}V_t^\pi - \mathbf{V}_t^{\pi^\top} \beta(x, u) \leq B(x, u). \quad (40)$$

Thus, by applying this bound to the expectations in the backward recursion, we obtain an overall bound on the approximation of the safety probabilities obtained using Lemma 5. Furthermore, the bound in (39) depends on the value of the conditioning variables x and u , which means the bound can serve as an indication of the quality of the approximation at a particular point.

Note that the bound in (39) applies to the error of a value function at a single time step. This means that the total error on the approximation of the safety probabilities increases linearly with the number of time steps used in the backward recursion in Lemma 5. Following Thorpe and Oishi (2019, Corollary 3), if the error of each value function is bounded by $B(x, u)$ in (39), we obtain an overall bound on the safety probabilities V_0^π of $NB(x, u)$, where N is the time horizon. We summarize this result in the following lemma:

Lemma 16. Given a bound $B(x, u)$, the error in the safety probabilities obtained via Lemma 5 is given by

$$|\mathbb{E}V_0^\pi - \mathbf{V}_0^{\pi^\top} \beta(x, u)| \leq NB(x, u), \quad (41)$$

where N is the time horizon.

Proof. The proof is by induction. If the error of a single value function is $B(x, u)$ as in (39), and the safety probabilities are computed via recursive substitution as in Lemma 5, then at time $t = 0$, the maximum error of V_0^π is $NB(x, u)$.

4. Kernel and parameter selection

The quality of the approximation of the stochastic reachability probability is governed by a number of factors, including the choice of kernel function, parameters associated with the kernel function, the regularization parameter of the least-squares problem (9), and the sample, S . In a realistic setting, we typically do not have explicit control over the sample S or the number of observations in the sample. Thus, the choice of kernel function and the model parameters plays an important role in the quality of the kernel-based approximation.

4.1. Kernel selection

The performance of kernel-based learning algorithms is closely tied to the choice of kernel function and the structure of the RKHS. In essence, the Hilbert space needs to be rich enough to model the set of probability measures underlying the observed data without overfitting. Thus, we frame the problem of kernel selection as a problem of limiting the complexity of the RKHS (Bartlett & Mendelson, 2002; Vapnik, 1998), where complexity in statistical learning literature refers to the ability of a function class to fit random noise. Intuitively, by choosing a function class that lowers the complexity term, we reduce the possibility that our function class $\mathcal{H}$ will overfit the observed data.

We propose a complexity term that is closely related to the Rademacher complexity (Bartlett & Mendelson, 2002) from statistical learning theory, but instead based on the bound presented in Theorem 15 that accommodates non-uniform coefficients $\beta(x, u)$. Define the random variable

$$\hat{\mathcal{C}}(\mathcal{H}) = \mathbb{E}_\sigma \left[\sup_{\|h\|_{\mathcal{H}} \leq 1} |h^\top \beta \sigma| \right]. \quad (42)$$

Then the conditional complexity of $\mathcal{H}$ is defined as $\mathcal{C}(\mathcal{H}) = \mathbb{E}_S[\hat{\mathcal{C}}(\mathcal{H})]$. Using the bound in (33), we have

$$\mathcal{C}(\mathcal{H}) \leq \mathbb{E}_S \left[\left(\text{tr}(\beta^\top \Phi \Phi^\top \beta) \right)^{1/2} \right]. \quad (43)$$

This choice of complexity term is equivalent to the first term on the right-hand side of (39). Since the conditional complexity appears in the bounds presented in Theorem 15, minimizing the complexity term also minimizes the finite sample bounds on the difference in expectations. Thus, we can choose the kernel functions which minimize the complexity term for all (x, u) , effectively minimizing the finite sample bound on the difference in expectations of the value functions.

The kernel should also be chosen to satisfy universality (Sriperumbudur et al., 2011, 2010) (resp. characteristic) and boundedness properties. One common choice of kernel that satisfies these properties is the Gaussian RBF kernel $k(x, x') = \exp(-\|x - x'\|_2^2 / 2\sigma^2)$, where $\sigma > 0$. Universal (Smola et al., 2007; Sriperumbudur et al., 2010) kernels are so-named because they satisfy a universal approximation property and are able to learn any real-valued function arbitrarily well. Because the mapping from the set of all probability measures $\mathcal{P}$ into the RKHS (5) is injective for universal (resp. characteristic, see Sriperumbudur et al., 2011) kernels, this means there is a unique element in the RKHS $\mathcal{H}$ for any $\mathbb{P}, \mathbb{Q} \in \mathcal{P}$, such that $\|m_{\mathbb{P}} - m_{\mathbb{Q}}\|_{\mathcal{H}} = 0$ if and only if $\mathbb{P} = \mathbb{Q}$. This ensures that the conditional distribution embedding admits a unique solution (Sriperumbudur et al., 2010), and that we can distinguish between distributions in Hilbert space. By choosing a bounded kernel function, we ensure that $\rho < \infty$, and we can achieve tighter bounds by selecting a kernel function with small ρ . The Gaussian kernel function, for example, has $\rho = 1$.

Typically, a parameterized kernel which is known to satisfy these properties is chosen, and then kernel parameters are tuned via cross-validation techniques. However, nascent work has posed kernel synthesis for a given sample S as an optimization problem. This approach has been demonstrated via convex optimization (Kim, Zymnis, Magnani, Koh, & Boyd, 2008) and semi-definite programming (Lanckriet, Cristianini, Bartlett, Ghaoui, & Jordan, 2004) for marginal distributions with scalar-valued regression. While this idea is promising, the connection between conditional distribution embeddings and the underlying regression problem has only recently begun to be explored (Grünwälder et al., 2012). Further, the extension from scalar-valued regression to vector-valued regression (Micchelli & Pontil, 2005) that is required by the objective function in (9) is not straightforward. Hence additional work will be needed to evaluate the feasibility of kernel synthesis methods for this problem.

4.2. Parameter selection

Tighter bounds may be possible by identifying a strict upper bound on the elements of $\beta(x, u)$, which in turn are influenced by the parameters ρ , the upper bound on the kernel function, and λ , the least-squares regularization coefficient in (9).

The value of ρ is determined primarily by the choice of kernel and its parameters, and can be tuned by minimizing the complexity term $\mathcal{C}(\mathcal{H})$. The value of λ affects the convergence rate associated with the uniform finite sample bound. The convergence guarantees in Song et al. (2009) and Grünwälder et al. (2012b) typically depend upon λ going to zero as the number of observations increases. See Caponnetto and De Vito (2007), De Vito, Caponnetto, and Rosasco (2005) for a discussion of optimal values of λ .

4.3. Scalability

One significant advantage of using the kernel-based approach in Lemma 5 is that approximating the value function expectation in (3) using conditional distribution embeddings does not scale exponentially as the system dimensionality increases (also known as the curse of dimensionality). This is primarily due to the fact that the system dimensionality (as well as the input dimension) only has an effect on the evaluation of the kernel functions k and l . Instead, the complexity of computing the conditional distribution embedding estimate is generally $\mathcal{O}(M^3)$, which is primarily driven by the matrix inverse W in (12), and scales solely with the sample size M used to construct the embedding estimate. This is demonstrated empirically in Thorpe and Oishi (2019), which shows that the computational complexity increases roughly linearly as the system dimensionality is increased. However, in order to adequately characterize the stochastic kernel of a high-dimensional state space, a large sample size may be required. In effect, this means that the quality of the approximation is in large part governed by the sample used to characterize the region of interest—which further motivates our approach to state-based finite sample bounds.

As the time horizon N increases, the backward recursion in Lemma 5 shows that if the policy π is time-invariant, meaning $\pi_0 = \pi_1 = \dots = \pi_{N-1}$, we do not need to recompute the embedding estimate $\hat{m}(x, \pi(x))$ at every time step. This means we only need to compute the estimate *once* for a given sample S , and indicates that the complexity of Lemma 5 increases linearly as the time horizon increases. In the case of time-varying policies, this of course means that we must recompute the embedding at every time step since the closed-loop dynamics of the system vary with time in accordance with the policy.

5. Examples

We implemented Lemma 5 on a stochastic chain of integrators for the purpose of validation, and on a nonlinear pendulum system (Dutta, Chen, & Sankaranarayanan, 2019) and a closed-loop nonlinear cart-pole benchmark system (Lopez, Musau, Tran, & Johnson, 2019) with black-box neural network controllers to demonstrate the capabilities of the proposed approach. For each problem, we generated a sample S of observations via simulation, and then assumed no knowledge of the system dynamics or the structure of the disturbance for the purposes of computing the stochastic reachability probability. We then computed finite sample bounds via Theorem 15. For all problems, we chose a Gaussian kernel $k(x, x') = \exp(-\|x - x'\|_2^2 / 2\sigma^2)$ with $\sigma = 0.1$, and chose λ using the optimal rate computed in Caponnetto and De Vito (2007).

All computations were done in Matlab, and code to reproduce the analysis and figures is available at: github.com/unm-hscl/ajthor-Automatica2020a.

5.1. Stochastic chain of integrators

We first consider a 2-D stochastic chain of integrators (Vinod & Oishi, 2017), in which the input appears at the second derivative and each element of the state vector is the discretized integral of the element that follows it. The dynamics with sampling time $T = 0.25$ are given by:

$$\mathbf{x}_{k+1} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \mathbf{x}_k + \begin{bmatrix} \frac{T^2}{2} \\ \frac{T}{T} \end{bmatrix} u_k + \mathbf{w}_k \quad (44)$$

where $\mathbf{w}_k$ is an i.i.d. disturbance defined on the probability space $(\mathcal{W}, \mathcal{B}(\mathcal{W}), \Pr_{\mathbf{w}})$. We presume a Gaussian disturbance $\mathbf{w}_k \sim \mathcal{N}$

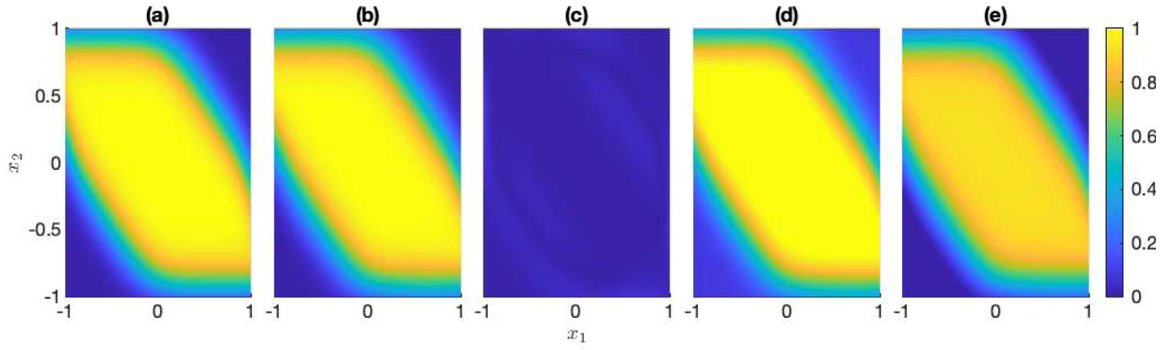


Fig. 1. (a) Safety probabilities at $k = 0$, $N = 5$, for a double integrator computed using dynamic programming. (b) Safety probabilities at $k = 0$ for a double integrator computed using Lemma 5 for $N = 5$. (c) Absolute error $|V_0^\pi(x) - \tilde{V}_0^\pi(x)|$ between the dynamic programming solution and Lemma 5. (d, e) Upper and lower finite sample bounds, respectively, of the safety probabilities of a double integrator system computed using Theorem 15 with $\delta = 0.1$, where (d) is $\tilde{V}_0^\pi(x) + B(x, \pi(x))$ and (e) is $\tilde{V}_0^\pi(x) - B(x, \pi(x))$, where $B(x, \pi(x))$ is computed as in (39).

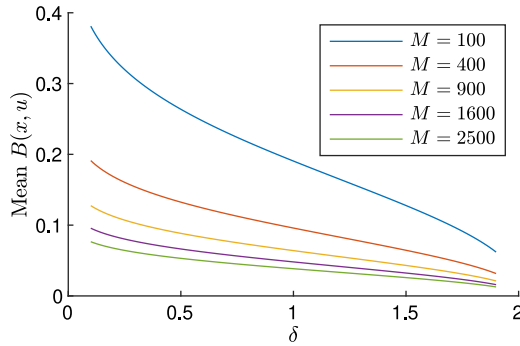


Fig. 2. Figure showing the mean of the finite sample bounds $B(x, \pi(x))$ in the region $[-1, 1] \times [-1, 1]$ for a double integrator system as a function of δ .

$(0, \Sigma)$ with $\Sigma = 0.01I$, a control policy $\pi(x) = 0$, and target and safe sets $\mathcal{T} = [-1, 1]^2$ and $\mathcal{K} = [-1, 1]^2$.

We consider a sample $\mathcal{S}$ of $M = 2500$ observations drawn i.i.d. from Q , a representation of (44) as a Markov control process (Definition 1). The initial conditions $x \in \mathcal{X}$ in the sample were chosen uniformly in the interval $[-1.1, 1.1] \times [-1.1, 1.1]$ in order to ensure that a subset of the initial conditions violates the safety constraints, $\mathcal{K}$ and $\mathcal{T}$. We do this to ensure the “learned” model does not map all initial conditions to a safe set. Otherwise, in the regression, the value function estimate maps all values to 1. The resulting state $y \in \mathcal{X}$ is drawn from $Q(\cdot | x, \pi(x))$ using the dynamics in (44). We then presumed no knowledge of the system dynamics or Q and computed the estimate $\hat{m}(x, \pi(x))$ according to (11), with $\beta(x, \pi(x))$ computed as in (12). Using $\hat{m}(x, \pi(x))$, we then computed the stochastic reachability probability using Lemma 5 for a time horizon of $N = 5$. We compare the stochastic reachability probability computed according to Lemma 5 against the solution via dynamic programming, that presumes the stochastic kernel Q is known. The absolute error $|V_0^\pi(x) - \tilde{V}_0^\pi(x)|$ between the results obtained from Lemma 5 and the dynamic programming solution is shown in Fig. 1(c). As expected, the stochastic reachability probabilities computed using Lemma 5, show low absolute error as compared with the dynamic programming solution, with a maximum absolute error of 0.1158 and a mean absolute error of 0.0122.

We then evaluated the finite sample bounds of the approximation using Theorem 15. We computed the bound $B(x, \pi(x))$ as in (39) with $\delta = 0.1$ to obtain bounds on the safety probabilities. The bounds obtained from Theorem 15 are probabilistic upper and lower bounds, meaning that given any $\delta/2 \in [0, 1]$, with probability $1 - \delta/2$, the approximation from Lemma 5 is bounded

by Theorem 15, which validates the result. Fig. 1(d) shows the upper bound on the safety probabilities, while Fig. 1(e) shows the lower bound.

We can see that the difference between the upper and lower bounds is small, which indicates that the quality of the approximation obtained via Lemma 5 is close to the true solution with high probability. Further, the absolute difference is larger than the error values in Fig. 1(c), as expected, meaning that the computed bound is a reasonable probabilistic upper bound.

We then consider the effect of varying the parameters M and δ in (39) on the finite sample bound computed using Theorem 15. In order to demonstrate the effect of the parameter M , we drew 5 new samples from the stochastic kernel Q , where the number of observations in each sample was chosen to be of size $M \in [100, 2500]$. An estimate was then computed for each sample and the finite sample bounds were computed for each estimate. For each sample of length $M \in [100, 2500]$, we computed the mean of the finite sample bounds in the region $[-1, 1] \times [-1, 1]$ for different values of $\delta \in [0.1, 1.9]$. Fig. 2 shows the mean of $B(x, \pi(x))$ (39), the finite sample bounds computed using Theorem 15 for the five samples of length $M \in [100, 2500]$ as a function of δ .

As expected, we can see that as the size of the samples increases, we obtain tighter probabilistic bounds $B(x, \pi(x))$ in the region $[-1, 1] \times [-1, 1]$ via Theorem 15. Effectively, this means that as the number of observations from the stochastic kernel increases, we obtain a better estimate of the conditional distribution embedding $m(x, \pi(x))$, and thus a better estimate of the safety probabilities via Lemma 5. Also as expected, as the violation threshold decreases (i.e., δ increases), the mean of the probabilistic bound decreases. Fig. 2 also shows that for low values of delta, we obtain higher values of the probabilistic bound $B(x, \pi(x))$. This corresponds to a high desired confidence. Further, we can see in Fig. 2 that the finite sample bounds do not improve appreciably as the sample size increases beyond a certain point.

5.2. Linearized cart-pole system

We then considered a benchmark cart-pole system (Lopez et al., 2019) with a black-box neural network feedback controller. The dynamics for the linearized cart-pole system (Lopez et al., 2019) are given by:

$$\begin{aligned} \ddot{x} &= 0.0043\dot{\theta} - 2.75\theta + 1.94u - 10.95\dot{x} \\ \ddot{\theta} &= 28.58\theta - 0.044\dot{\theta} - 4.44u + 24.92\dot{x} \end{aligned} \quad (45)$$

with state $x = [x, \dot{x}, \theta, \dot{\theta}]^\top \in \mathbb{R}^4$ and control input $u \in \mathbb{R}$. The dynamics are then discretized in time with sampling time $T_s = 0.2$ s. We add an additional Gaussian disturbance $w_k \sim \mathcal{N}(0, \Sigma)$

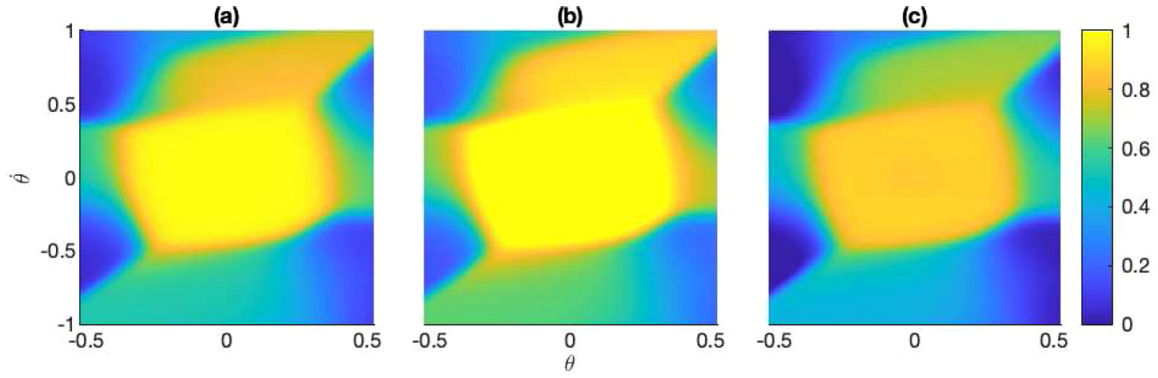


Fig. 3. (a) Safety probabilities for a linearized cart-pole system computed using Lemma 5 for $N = 10$. (b) Upper bound on the safety probabilities computed using Theorem 15. (c) Lower bound on the safety probabilities computed using Theorem 15.

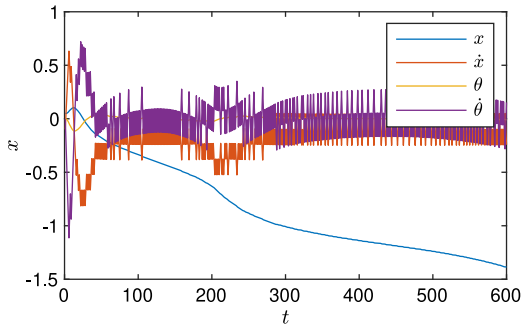


Fig. 4. Sample realization of the nonlinear cart-pole system (48) over 600 time steps.

with $\Sigma = 0.01I$ to the dynamical state equations, which can simulate dynamical uncertainty or minor system perturbations. The control input is computed via a neural network controller (Lopez et al., 2019), which takes the current state and outputs a real number $u \in \mathbb{R}$, which can be interpreted as the input torque.

The benchmark is defined (Lopez et al., 2019) such that the neural network controller must keep the lateral position of the cart x within $[-0.7, 0.7]$, maintain a low cart velocity $\dot{x} \in [-1, 1]$, and keep the pendulum angle θ within $[-\pi/6, \pi/6]$ while the angular velocity $\dot{\theta}$ is unconstrained. We define the safe set $\mathcal{K}$ according to the above constraints, and define the target set $\mathcal{T}$ such that the pendulum angle θ must be within $[-0.05, 0.05]$.

$$\mathcal{K} = \{x \in \mathbb{R}^4 \mid |x_1| \leq 0.7, |x_2| \leq 1, |x_3| \leq \pi/6\} \quad (46)$$

$$\mathcal{T} = \{x \in \mathbb{R}^4 \mid |x_3| \leq 0.05\} \quad (47)$$

We simulated 10 trajectories from initial conditions taken uniformly from the ranges specified above, and extracted a sample $\mathcal{S}$ of $M = 12,234$ observations taken i.i.d. from the stochastic kernel Q , a representation of the dynamics (45) as a Markov control process (Definition 1).

We then computed the safety probabilities for the system over a time horizon $N = 10$ using Lemma 5 to demonstrate the feasibility of the approach. The safety probabilities computed using Lemma 5 at $k = 0$ for $N = 10$ are shown in Fig. 3(a). We can see that the closed-loop system has a high probability of stabilizing the pendulum from an initial condition within the range $\theta \in [-\pi/6, \pi/6]$ and the results also show an underlying symmetry around $\theta = 0$, as expected. We then computed the finite sample bounds on the approximation using Theorem 15 with $\delta = 0.1$ to obtain probabilistic upper and lower bounds on the safety probabilities. Fig. 3(b) shows the probabilistic upper bound on the safety probabilities, while Fig. 3(c) shows the

lower bound. As expected, because we use a high number of observations $M = 12,234$, the bounds computed in Theorem 15 show that with high probability, the solution is close to the true solution.

This means that using the proposed data-driven approach, we can utilize stochastic reachability to analyze the safety properties of a dynamical system with a black-box neural network controller. Similarly, we can expose the underlying structure of the closed-loop system to reveal useful knowledge of the system properties, such as symmetry.

5.3. Nonlinear cart-pole system

We then analyzed a nonlinear cart-pole system with a neural network controller (Lopez et al., 2019), with dynamics given by:

$$\begin{aligned} \ddot{x} &= \frac{u + m_l \omega^2 \sin(\theta)}{m_t} \\ &\quad - \frac{m_l(g \sin(\theta) - \cos(\theta))(\frac{u + m_l \omega^2 \sin(\theta)}{m_t}) \cos(\theta)}{l(\frac{4}{3} - m \frac{\cos^2(\theta)}{m_t}) m_t} \\ \ddot{\theta} &= \frac{g \sin(\theta) - \cos(\theta)(\frac{u + m_l \omega^2 \sin(\theta)}{m_t}) \cos(\theta)}{l(\frac{4}{3} - m \frac{\cos^2(\theta)}{m_t}) m_t} \end{aligned} \quad (48)$$

where $g = 9.8$ is the gravitational constant, the pole mass is $m = 0.1$, half the pole's length is $l = 0.5$, and $m_t = 1.1$ is the total mass. The dynamics are then discretized in time with sampling time $T_s = 0.015$ s. The control input, $u \in \{-10, 10\}$, which affects the lateral position of the cart, is chosen by the neural network controller (Lopez et al., 2019). This means the controller is less “smooth” than the neural network controller for the linearized cart-pole system, because when the pendulum is near vertical, the controller rapidly switches between a high positive and negative control input value. Thus, the velocity components of the system state, $\dot{x}$ and $\dot{\theta}$, will not stabilize to zero, as shown in Fig. 4.

As before, we added a Gaussian disturbance $w_k \sim \mathcal{N}(0, \Sigma)$ with $\Sigma = 0.01I$ to the dynamical equations and represent the system as a Markov control process. The benchmark is defined such that the pole angle θ will remain within $[-\pi/6, \pi/6]$, while the other variables are unconstrained. As such, we define the target set $\mathcal{T}$ such that $\theta \in [-0.05, 0.05]$ as with the linearized cart-pole system, but define the safe set $\mathcal{K}$ to be the entire state space, meaning there is no unsafe region.

$$\mathcal{T} = \{x \in \mathbb{R}^4 \mid |x_3| \leq 0.05\} \quad (49)$$

This allows us to analyze the behavior of the controller to reach a pre-specified objective without enforcing constraints on the system before the terminal time. We then simulated 10 trajectories

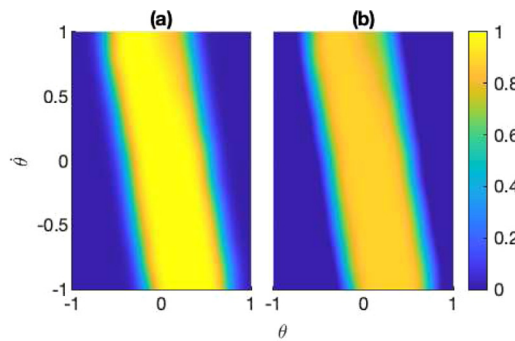


Fig. 5. (a) Safety probabilities of the nonlinear cart-pole system computed using Lemma 5 for a time horizon of $N = 600$. (b) Probabilistic lower bound on the safety probabilities computed using Theorem 15 with $\delta = 0.1$.

from initial conditions sampled uniformly from the ranges specified above, and collected a sample $\mathcal{S}$ of $M = 10,000$ observations. Then, we computed the safety probabilities using Lemma 5 with $N = 600$. The results are shown in Fig. 5(a). We then computed the finite sample bounds using Theorem 15 for $\delta = 0.1$, and plotted the lower bound on the safety probabilities in Fig. 5(b).

As expected, we see that the safety probabilities show that the nonlinear pendulum system is able to stabilize a pendulum starting within a small range of θ . Interestingly, it is revealed in Fig. 5(a) that the controller is not completely symmetric, meaning it has a higher chance of stabilizing the pendulum for positive values of θ than negative values of θ . Fig. 5(b) shows the probabilistic lower bound on safety probabilities computed using Theorem 15 with $\delta = 0.1$.

6. Conclusion

We provided state- and input-based finite sample bounds for the stochastic reachability probability constructed via conditional distribution embeddings. Our approach is based on an application of statistical learning theory, that relates the observed data to the quality of the approximation of the stochastic reachability probability at a given state and input. This approach enables rigorous bounds on model-free stochastic reachability. We validated our approach on a nonlinear dynamical system with a neural net controller, and numerically characterized our approach on the stochastic double integrator.

Acknowledgments

This material is based upon work supported by the National Science Foundation under NSF Grant Number CNS-1836900. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. This research was supported in part by the Laboratory Directed Research and Development program at Sandia National Laboratories, a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. The views expressed in this article do not necessarily represent the views of the U.S. Department of Energy or the United States Government. This work was also supported in part by the University of New Mexico's Vice President for Research Office.

References

- Abate, Alessandro, Blom, Henk, Cauchi, Nathalie, Degiorgio, Kurt, Fränzle, Martin, Hahn, Ernst Moritz, et al. (2019). ARCH-COMP19 category report: Stochastic modelling. *EPiC Series in Computing*, 61, 62–102.
- Abate, Alessandro, Blom, Henk, Cauchi, Nathalie, Delicaris, Joanna, Hartmanns, Arnd, Khaled, Mahmoud, et al. (2020). ARCH-COMP20 category report: Stochastic models. *EPiC Series in Computing*, 74, 76–106.
- Abate, Alessandro, Blom, H. A. P., Cauchi, Nathalie, Haesaert, Sofie, Hartmanns, Arnd, Lesser, Kendra, et al. (2018). ARCH-COMP18 category report: Stochastic modelling. *EPiC Series in Computing*, 54.
- Abate, Alessandro, Prandini, Maria, Lygeros, John, & Sastry, Shankar (2008). Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. *Automatica*, 44(11), 2724–2734.
- Aronszajn, Nachman (1950). Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68(3), 337–404.
- Bartlett, Peter L., & Mendelson, Shahar (2002). Rademacher and Gaussian complexities: Risk bounds and structural results. *Journal of Machine Learning Research*, 3, 463–482.
- Berlinet, Alain, & Thomas-Agnan, Christine (2004). *Reproducing Kernel Hilbert spaces in probability and statistics*. Springer.
- Bernstein, Dennis S. (2009). *Matrix mathematics*. Princeton University Press.
- Bertsekas, Dimitri P., & Shreve, Steven E. (1978). *Stochastic optimal control: the discrete time case*. Elsevier.
- Bousquet, Olivier, & Elisseeff, André (2002). Stability and generalization. *Journal of Machine Learning Research*, 2, 499–526.
- Caponnetto, Andrea, & De Vito, Ernesto (2007). Optimal rates for the regularized least-squares algorithm. *Foundations of Computational Mathematics*, 7(3), 331–368.
- Cauchi, Nathalie, & Abate, Alessandro (2019). StocHy - automated verification and synthesis of stochastic processes. In *Conference on hybrid systems: computation and control* (pp. 258–259).
- Christmann, Andreas, & Steinwart, Ingo (2008). *Support vector machines*. Springer.
- Çınlar, Erhan (2011). *Probability and stochastics*. Springer.
- De Vito, Ernesto, Caponnetto, Andrea, & Rosasco, Lorenzo (2005). Model selection for regularized least-squares algorithm in learning theory. *Foundations of Computational Mathematics*, 5(1), 59–85.
- Devonport, Alex, & Arcak, Murat (2020). Data-driven reachable set computation using adaptive Gaussian process classification and Monte Carlo methods. In *American control conference* (pp. 2629–2634). IEEE.
- Dutta, Souradeep, Chen, Xin, & Sankaranarayanan, Sriram (2019). Reachability analysis for neural feedback systems using regressive polynomial rule inference. In *Conference on hybrid systems: computation and control* (pp. 157–168).
- Gleason, Joseph D., Vinod, Abraham P., & Oishi, Meeko M. K. (2017). Underapproximation of reach-avoid sets for discrete-time stochastic systems via Lagrangian methods. In *Conference on decision and control* (pp. 4283–4290). IEEE.
- Gleason, Joseph D., Vinod, Abraham P., & Oishi, Meeko M. K. (2021). Lagrangian approximations for stochastic reachability of a target tube. *Automatica*, 128, Article 109546.
- Grünewälder, Steffen, Lever, Guy, Baldassarre, Luca, Patterson, Sam, Gretton, Arthur, & Pontil, Massimiliano (2012). Conditional mean embeddings as regressors. In *International conference on machine learning* (pp. 1803–1810).
- Grünewälder, Steffen, Lever, Guy, Baldassarre, Luca, Pontil, Massimiliano, & Gretton, Arthur (2012). Modelling transition dynamics in MDPs with RKHS embeddings. In *International conference on machine learning* (pp. 1603–1610).
- Kanagawa, Motonobu, & Fukumizu, Kenji (2014). Recovering distributions from Gaussian RKHS embeddings. In *Artificial intelligence and statistics* (pp. 457–465). PMLR.
- Kariotoglou, Nikolaos, Kamgarpour, Maryam, Summers, Tyler H., & Lygeros, John (2017). The linear programming approach to reach-avoid problems for Markov decision processes. *Journal of Artificial Intelligence Research*, 60(1), 263–285.
- Kim, Seung-Jean, Zymnis, Argyrios, Magnani, Alessandro, Koh, Kwangmoo, & Boyd, Stephen (2008). Learning the kernel via convex optimization. In *Conference on acoustics, speech and signal processing* (pp. 1997–2000). IEEE.
- Lanckriet, Gert R. G., Cristianini, Nello, Bartlett, Peter, Ghaoui, Laurent El., & Jordan, Michael I. (2004). Learning the kernel matrix with semidefinite programming. *Journal of Machine Learning Research*, 5(Jan), 27–72.
- Le, Quoc, Sarló, Tamás, Smola, Alex, et al. (2013). Fastfood-approximating kernel expansions in loglinear time. In *International conference on machine learning*.
- Lesser, Kendra, Oishi, Meeko, & Scott Erwin, R. (2013). Stochastic reachability for control of spacecraft relative motion. In *Conference on decision and control* (pp. 4705–4712).
- Lever, Guy, & Stafford, Ronnie (2015). Modelling policies in MDPs in reproducing kernel Hilbert space. In *Artificial intelligence and statistics* (pp. 590–598).
- Lew, Thomas, & Pavone, Marco (2020). Sampling-based reachability analysis: A random set theory approach with adversarial sampling. In *Conference on robot learning*.

- Lopez, Diego Manzanar, Musau, Patrick, Tran, Hoang-Dung, & Johnson, Taylor T. (2019). Verification of closed-loop systems with neural network controllers. In *International workshop on applied verification of continuous and hybrid systems*. Vol. 61 (pp. 201–210).
- Malone, N., Chiang, H., Lesser, K., Oishi, M., & Tapia, L. (2017). Conference on hybrid dynamic moving obstacle avoidance using a stochastic reachable set-based potential field. *IEEE Transactions on Robotics*, 33(5), 1124–1138.
- McDiarmid, Colin (1989). On the method of bounded differences. *Surveys in Combinatorics*, 141(1), 148–188.
- Micchelli, Charles A., & Pontil, Massimiliano (2005). On learning vector-valued functions. *Neural Computation*, 17(1), 177–204.
- Muandet, Krikamol, Fukumizu, Kenji, Sriperumbudur, Bharath, Schölkopf, Bernhard, et al. (2017). Kernel mean embedding of distributions: A review and beyond. *Foundations and Trends in Machine Learning*, 10(1–2), 1–141.
- Nishiyama, Yu, Boularias, Abdeslam, Gretton, Arthur, & Fukumizu, Kenji (2012). Hilbert space embeddings of POMDPs. In *Conference on uncertainty in artificial intelligence* (pp. 644–653).
- Parzen, Emanuel (1961). An approach to time series analysis. *The Annals of Mathematical Statistics*, 951–989.
- Prajna, Stephen, & Jadbabaie, Ali (2004). Safety verification of hybrid systems using barrier certificates. In *Conference on hybrid systems: computation and control* (pp. 477–492). Springer.
- Prajna, Stephen, Jadbabaie, Ali, & Pappas, George J. (2007). A framework for worst-case and stochastic safety verification using barrier certificates. *Transactions on Automatic Control*, 52(8), 1415–1428.
- Rahimi, Ali, & Recht, Benjamin (2008). Random features for large-scale kernel machines. In J. Platt, D. Koller, Y. Singer, & S. Roweis (Eds.), *Advances in neural information processing systems*. Vol. 20.
- Roohi, Nima, Wang, Yu, West, Matthew, Dullerud, Geir E., & Viswanathan, Mahesh (2017). Statistical verification of the Toyota powertrain control verification benchmark. In *Conference on hybrid systems: computation and control*. New York, NY, USA (pp. 65–70).
- Rudin, Walter (1991). *Functional analysis*. McGraw-Hill.
- Schölkopf, Bernhard, & Smola, Alexander J. (2001). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT Press.
- Shmarov, Fedor, Soudjani, Sadegh, Paoletti, Nicola, Bartocci, Ezio, Lin, Shan, Smolka, Scott A., et al. (2020). Automated synthesis of safe digital controllers for sampled-data stochastic nonlinear systems. *IEEE Access*, 8, 180825–180843.
- Shmarov, Fedor, & Zuliani, Paolo (2015). Probreach: verified probabilistic delta-reachability for stochastic hybrid systems. In *Conference on hybrid systems: computation and control* (pp. 134–139).
- Smola, Alex, Gretton, Arthur, Song, Le, & Schölkopf, Bernhard (2007). A Hilbert space embedding for distributions. In *International conference on algorithmic learning theory* (pp. 13–31). Springer.
- Song, Le, Boots, Byron, Siddiqi, Sajid M., Gordon, Geoffrey, & Smola, Alex (2010). Hilbert space embeddings of hidden Markov models. In *International conference on machine learning* (pp. 991–998).
- Song, Le, Gretton, Arthur, & Guestrin, Carlos (2010). Nonparametric tree graphical models. In *International conference on artificial intelligence and statistics* (pp. 765–772).
- Song, Le, Huang, Jonathan, Smola, Alex, & Fukumizu, Kenji (2009). Hilbert space embeddings of conditional distributions with applications to dynamical systems. In *International conference on machine learning* (pp. 961–968).
- Soudjani, Sadegh, Esmail Zadeh, Gevaerts, Caspar, & Abate, Alessandro (2015). FAUST2: Formal abstractions of uncountable-State Stochastic processes. In *International conference on tools and algorithms for the construction and analysis of systems* (pp. 272–286). Springer.
- Sriperumbudur, Bharath K., Fukumizu, Kenji, & Lanckriet, Gert RG. (2011). Universality, characteristic kernels and RKHS embedding of measures. *Journal of Machine Learning Research*, 12(7).
- Sriperumbudur, Bharath K., Gretton, Arthur, Fukumizu, Kenji, Schölkopf, Bernhard, & Lanckriet, Gert RG. (2010). Hilbert space embeddings and metrics on probability measures. *Journal of Machine Learning Research*, 11(Apr), 1517–1561.
- Summers, Sean, & Lygeros, John (2010). Verification of discrete time stochastic hybrid systems: A stochastic reach-avoid decision problem. *Automatica*, 46(12), 1951–1961.
- Thorpe, Adam J., & Oishi, Meeko MK. (2019). Model-free stochastic reachability using kernel distribution embeddings. *IEEE Control Systems Letters*.
- Thorpe, Adam J., Ortiz, Kendrick R., & Oishi, Meeko MK. (2021). Learning approximate forward reachable sets using separating kernels. In *Learning for dynamics and control* (pp. 201–212). PMLR.
- Thorpe, Adam J., Sivaramakrishnan, Vignesh, & Oishi, Meeko MK. (2021). Approximate stochastic reachability for high dimensional systems. In *American control conference* (pp. 1287–1293). IEEE.
- Vapnik, Vladimir (1998). *Statistical learning theory*. Wiley.
- Vinod, Abraham P., Gleason, Joseph D., & Oishi, Meeko M. K. (2019). SReachTools: A MATLAB stochastic reachability toolbox. In *Conference on hybrid systems: computation and control*. (pp. 33–38).
- Vinod, A. P., HomChaudhuri, B., & Oishi, M. M. K. (2017). Forward stochastic reachability analysis for uncontrolled linear systems using Fourier transforms. In *Conference on hybrid systems: computation and control* (pp. 35–44).
- Vinod, Abraham P., & Oishi, Meeko MK. (2017). Scalable underapproximation for the stochastic reach-avoid problem for high-dimensional LTI systems using Fourier transforms. *IEEE Control Systems Letters*, 1(2), 316–321.
- Vinod, Abraham P., & Oishi, Meeko M. K. (2021). Probabilistic occupancy via forward stochastic reachability for Markov jump affine systems. *Transactions on Automatic Control*, 66(7), 3068–3083.
- Wang, Yu, Roohi, Nima, West, Matthew, Viswanathan, Mahesh, & Dullerud, Geir E. (2019). Statistical verification of PCTL using antithetic and stratified samples. *Formal Methods in System Design*, 54, 145–163.
- Zarei, Mojtaba, Wang, Yu, & Pajic, Miroslav (2020). Statistical verification of learning-based cyber-physical systems. In *Conference on hybrid systems: computation and control*. New York, NY.



Adam J. Thorpe received the B.S. EE (2017) and M.S. (2019) in the field of Electrical Engineering with a focus in systems and controls from the University of New Mexico. He is currently working toward a Ph.D. degree since 2019. His research interests are in the areas of stochastic optimal control, autonomous systems, and statistical learning, with an emphasis on data-driven control techniques.



Kendrick R. Ortiz received the B.S. EE in 2020 with a focus in Electromagnetics from the University of New Mexico. He is currently working toward a M.S. degree in the field of Electrical Engineering with a focus in systems and controls. His research interests are in the areas of statistical learning, stochastic optimal control, and human-automation systems.



Meeko M.K. Oishi received the Ph.D. (2004) and M.S. (2000) in Mechanical Engineering from Stanford University (Ph.D. minor, Electrical Engineering), and a B.S.E. in Mechanical Engineering from Princeton University (1998). She is a Professor of Electrical and Computer Engineering at the University of New Mexico. Her research interests include human-centric control, stochastic optimal control, and autonomous systems. She previously held a faculty position at the University of British Columbia at Vancouver, and postdoctoral positions at Sandia National Laboratories and at the National Ecological Observatory Network. She was a Visiting Researcher at AFRL Space Vehicles Directorate, and a Science and Technology Policy Fellow at The National Academies. She is the recipient of the NSF CAREER Award and a member of the 2021–2023 DoD Defense Science Study Group.