2021 American Control Conference (ACC)
New Orleans, USA, May 25-28, 2021

Approximate Stochastic Reachability for High Dimensional Systems

Adam J. Thorpe, Vignesh Sivaramakrishnan, Meeko M. K. Oishi

Abstract— **We present a method to compute the stochastic reachability safety probabilities for high-dimensional stochastic dynamical systems. Our approach takes advantage of a nonparametric learning technique known as conditional distribution embeddings to model the stochastic kernel using a data-driven approach. By embedding the dynamics and uncertainty within a reproducing kernel Hilbert space, it becomes possible to compute the safety probabilities for stochastic reachability problems as simple matrix operations and inner products. We employ a convergent approximation technique, random Fourier features, in order to alleviate the increased computational requirements for high-dimensional systems. This technique avoids the curse of dimensionality, and enables the computation of safety probabilities for high-dimensional systems without prior knowledge of the structure of the dynamics or uncertainty. We validate this approach on a double integrator system, and demonstrate its capabilities on a million-dimensional, nonlinear, non-Gaussian, repeated planar quadrotor system.**

I. Introduction

Stochastic reachability is an established verification technique which is used to compute the likelihood that a system will reach a desired state without violating a predefined set of safety constraints. The solutions to stochastic reachability problems are broadly framed in terms of a dynamic program [1], which scales poorly with the system dimensionality. Methods using approximate dynamic programming [2], particle filtering [3, 4], and abstractions [5] have been posed, but are limited to systems of moderate dimensionality. Optimization-based solutions have garnered modest computational tractability via chance constraints [4, 6], sampling methods [7]–[9], and convex optimization with Fourier transforms [10, 11], but are limited to linear dynamical systems and Gaussian or log-concave disturbances.

Recent work in reachability for non-stochastic, linear dynamical systems has accommodated systems with up to a billion dimensions [12]–[14], an unprecendented size. However, comparably scalable solutions for stochastic systems, even with considerable structure in the dynamics and in the uncertainty, remain elusive. Techniques based in reinforcement learning, for example using parameterized representations of the policy, have been applied to the problem of policy evaluation. In many cases, the scalability of existing reinforcement learning techniques that leverage dynamic programming is limited, due to the well-known curse of dimensionality [15].

We propose a model-free method for stochastic reachability analysis of high-dimensional systems using a class of machine learning techniques known as kernel methods. We take advantage of kernel distribution embeddings [16], a nonparametric learning technique that uses observations of the system evolution to construct a model of probability distributions in a Hilbert space of functions. Broadly, these techniques enable efficient computation of expectations by representing integral operators as elements in a reproducing kernel Hilbert space. We apply these techniques to the area of stochastic reachability and present a method to compute the safety probabilities for high-dimensional, nonlinear stochastic systems with arbitrary disturbances. These techniques scale exponentially with the number of samples, meaning that they can suffer from high computational complexity and memory storage requirements. This can be prohibitive for high-dimensional systems, which may require a large number of samples in order to effectively capture the system dynamics. The utility of distribution embeddings for the terminal-hitting time problem has been demonstrated for systems of up to 10,000 dimensions [17], but the jump to a million presents significant computational challenges.

To facilitate stochastic reachability calculations for extremely high-dimensional systems, we couple distribution embeddings with a technique known as random Fourier features (RFF) [18, 19], that uses an empirical Fourier approximation to deal with high-dimensional systems. RFF is a well-known speedup technique for kernel methods to reduce the computational overhead. However, application to stochastic reachability requires additional considerations to quantify the quality of the approximation obtained by this approach. *The main contribution of this paper is incorporation of random Fourier features into a kernel distribution embeddings approach to compute safety probabilities for high-dimensional stochastic systems.*

The paper is outlined as follows. Section II formulates the problem. Section III outlines the theory of conditional distribution embeddings. Section IV applies random Fourier features to the computation of safety probabilities. In section V, we demonstrate our approach on two examples: a stochastic chain of integrators for validation, and a million-dimensional, non-Gaussian, repeated planar quadrotor.

II. Problem Formulation

The following notation is used throughout the paper. For any nonempty space Ω, the indicator $\mathbf{1}_{\mathcal{A}} : \Omega \to \{0, 1\}$ of $\mathcal{A} \subseteq \Omega$ is a function defined such that $\mathbf{1}_{\mathcal{A}}(\omega) = 1$ if $\omega \in$

This material is based upon work supported by the National Science Foundation under NSF Grant Number CMMI-1254990, and CNS-1836900, and in part by the LDRD program at Sandia Nat'l Laboratories, a multimission laboratory managed and operated by Nat'l Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell Int'l, Inc., for the US Dep't of Energy's National Nuclear Security Administration under contract DE-NA-0003525. 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, US Dep't of Energy, or the United States Government.

2021 American Control Conference (ACC) | 978-1-6654-4197-1/21/$31.00 ©2021110.23919/ACC50511.2021.9483404

$\mathcal{A}$, and $\mathbf{1}_{\mathcal{A}}(\omega) = 0$ if $\omega \notin \mathcal{A}$. Let $(\Omega, \mathcal{F}(\Omega), \Pr)$ define a probability space, where $\mathcal{F}(\Omega)$ denotes the σ-algebra relative to Ω, and $\Pr$ is the assigned probability measure. Let $\mathscr{B}(\Omega)$ denote the Borel σ-algebra associated with Ω. Given $i \in \mathbb{N}$ random variables $\boldsymbol{x}_i$, which are measurable functions on $(\Omega, \mathcal{F}(\Omega), \Pr)$, let $\boldsymbol{x} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]^\top$ be a random vector defined on the induced probability space $(\Omega^n, \mathcal{F}(\Omega^n), \Pr_{\boldsymbol{x}})$, where $\Pr_{\boldsymbol{x}}$ is the induced probability measure. A stochastic process is defined as a sequence of random vectors $\{\boldsymbol{x}_k\}_{k=0}^N$, $N \in \mathbb{N}$. For a real, measurable function $\boldsymbol{x}$ on $(\Omega, \mathcal{F}(\Omega), \Pr)$, the Lebesgue integral $\int_\Omega \boldsymbol{x} \Pr$ is denoted by the expectation operator $\mathbb{E}_{\boldsymbol{x}\sim\Pr}[\boldsymbol{x}]$.

A. System Model

We consider a Markov control process $\mathcal{H}$ [1], which is defined as a 3-tuple: $\mathcal{H} = (\mathcal{X}, \mathcal{U}, Q)$, where $\mathcal{X} \subseteq \mathbb{R}^n$ and $\mathcal{U} \subset \mathbb{R}^m$ are Borel spaces representing the state and control spaces, and $Q : \mathscr{B}(\mathcal{X}) \times \mathcal{X} \times \mathcal{U} \to [0, 1]$ is a stochastic kernel, which is a Borel-measurable function that maps a probability measure $Q(\cdot \,|\, x, u)$ to each $x \in \mathcal{X}$ and $u \in \mathcal{U}$ in $(\mathcal{X}, \mathscr{B}(\mathcal{X}))$. The system evolves over a finite time horizon $k \in [0, N]$, where the inputs are chosen from a Markov control policy $\pi = \{\pi_0, \pi_1, \ldots\}$ [15], which is a sequence of universally-measurable maps $\pi_k : \mathcal{X} \to \mathcal{U}$.

We consider the case where the stochastic kernel is unknown, but observations of the system are available. Consider a sample $\mathcal{S} = \{(\bar{x}_i, \bar{u}_i, \bar{y}_i)\}_{i=1}^M$ of size M drawn i.i.d. from Q, such that $\bar{y}_i \sim Q(\cdot \,|\, \bar{x}_i, \bar{u}_i)$ and $\bar{u}_i = \pi(\bar{x}_i)$, where π is a fixed Markov control policy. We denote sample vectors with a bar to differentiate them from time-indexed vectors.

B. First-Hitting Time Problem

Let $\mathcal{K}, \mathcal{T} \in \mathscr{B}(\mathcal{X})$, $\mathcal{T} \subseteq \mathcal{K}$, denote the *safe set* and *target set*, respectively. We define the *first-hitting time safety probability* [1] as the probability that a system $\mathcal{H}$ following a control policy π and starting at the initial condition x_0 will reach a target set $\mathcal{T}$ at some time $j \in [0, N]$ while remaining within the safe set $\mathcal{K}$ for all time $i \in [0, j-1]$.

$$r^\pi_{x_0}(\mathcal{K}, \mathcal{T}) \coloneqq \mathbb{E}^\pi_{x_0}\left[\sum_{j=0}^{N}\left(\prod_{i=0}^{j-1} \mathbf{1}_{\mathcal{K}\setminus\mathcal{T}}(x_i)\right)\mathbf{1}_{\mathcal{T}}(x_j)\right] \tag{1}$$

For a fixed Markov policy π, we define the value functions $V^\pi_k : \mathcal{X} \to [0, 1]$, $k \in [0, N]$, via the *backward recursion*:

$$V^\pi_N(x) = \mathbf{1}_{\mathcal{T}}(x) \tag{2a}$$

$$V^\pi_k(x) = \mathbf{1}_{\mathcal{T}}(x) + \mathbf{1}_{\mathcal{K}\setminus\mathcal{T}}(x)\mathbb{E}_{\boldsymbol{y}\sim Q}[V^\pi_{k+1}(\boldsymbol{y})] \tag{2b}$$

Then, $V^\pi_0(x) = r^\pi_{x_0}(\mathcal{K}, \mathcal{T})$ for every $x_0 \in \mathcal{X}$. In general, computing $r^\pi_{x_0}(\mathcal{K}, \mathcal{T})$ is difficult due to the expectation in (2b). We seek a representation of the stochastic kernel which enables an efficient computation of this expectation.

C. Problem Statement

We consider the following problems:

Problem 1. *Without direct knowledge of Q, use a sample $\mathcal{S}$ of observations taken from Q to compute a kernel-based approximation of* (2b) *that converges in probability.*

Problem 2. *Use RFF to compute an approximation of the kernel that enables efficient computation of* (2b) *for high-dimensional systems.*

The computational efficiencies afforded by RFF transform (2b) and thus (1) into simple matrix operations and inner products, enabling us to handle high-dimensional systems.

III. RKHS Embeddings of Distributions

For any set $\mathcal{X}$, let $\mathscr{H}_{\mathcal{X}}$ denote a Hilbert space of real-valued functions $f : \mathcal{X} \to \mathbb{R}$, with the inner product $\langle \cdot, \cdot \rangle_{\mathscr{H}_{\mathcal{X}}}$. A Hilbert space $\mathscr{H}_{\mathcal{X}}$ is a reproducing kernel Hilbert space (RKHS) if there exists a positive definite [20] kernel function $k_{\mathcal{X}} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ that satisfies the following properties [21]: 1) For any $x, x' \in \mathcal{X}$, $k_{\mathcal{X}}(x, \cdot) : x' \to k_{\mathcal{X}}(x, x')$ is an element of $\mathscr{H}_{\mathcal{X}}$. 2) An element $k_{\mathcal{X}}(x, x')$ of $\mathscr{H}_{\mathcal{X}}$ satisfies the reproducing property such that $\forall f \in \mathscr{H}_{\mathcal{X}}$ and $x \in \mathcal{X}$,

$$f(x) = \langle k_{\mathcal{X}}(x, \cdot), f \rangle_{\mathscr{H}_{\mathcal{X}}} \tag{3a}$$

$$k_{\mathcal{X}}(x, x') = \langle k_{\mathcal{X}}(x, \cdot), k_{\mathcal{X}}(x', \cdot) \rangle_{\mathscr{H}_{\mathcal{X}}} \tag{3b}$$

We define the positive definite kernels $k_{\mathcal{X}} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ and $k_{\mathcal{U}} : \mathcal{U} \times \mathcal{U} \to \mathbb{R}$, and let $\mathscr{H}_{\mathcal{X}}$ and $\mathscr{H}_{\mathcal{U}}$ denote the RKHS induced by $k_{\mathcal{X}}$ and $k_{\mathcal{U}}$, respectively. Further, we define $k_{\mathcal{X}\times\mathcal{U}} : (\mathcal{X} \times \mathcal{U}) \times (\mathcal{X} \times \mathcal{U}) \to \mathbb{R}$, via the tensor product $k_{\mathcal{X}\times\mathcal{U}}((x, u), (x', u')) = k_{\mathcal{X}}(x, x')k_{\mathcal{U}}(u, u')$. Let $\mathscr{H}_{\mathcal{X}\times\mathcal{U}}$ be the associated RKHS. We can also view an element $k_{\mathcal{X}}(x, \cdot) \in \mathscr{H}_{\mathcal{X}}$ as a *feature map* $\phi : \mathcal{X} \to \mathscr{H}_{\mathcal{X}}$, such that $k_{\mathcal{X}}(x, x') = \langle \phi(x), \phi(x') \rangle_{\mathscr{H}_{\mathcal{X}}}$. Intuitively, the feature map can be viewed as a basis function, such that a function $f \in \mathscr{H}_{\mathcal{X}}$ can be represented as a weighted sum $f(x) = \langle w, \phi(x) \rangle$ for some possibly infinite-dimensional weight vector w. However, constructing ϕ and computing the inner product explicitly can be computationally expensive or even impossible, depending on the choice of kernel. Instead, the inner product can be computed using $k_{\mathcal{X}}(x, x')$ directly. This is known as the *kernel trick* [20].

A. Conditional Distribution Embeddings

Because Q is unknown, we cannot compute the expectation in (2b) directly. Instead, we embed the expectation with respect to the stochastic kernel in an RKHS. Using this representation, we can efficiently compute the expectation in Hilbert space, even when the structure of the stochastic kernel is unknown.

For any measurable space $\mathcal{X}$, let $\mathscr{P}$ denote the set of probability distributions on $\mathcal{X}$. For any distribution $\mathbb{P} \in \mathscr{P}$, if the sufficient condition $\mathbb{E}_{\boldsymbol{x}\sim\mathbb{P}}[k_{\mathcal{X}}(\boldsymbol{x}, \boldsymbol{x})] < \infty$ is satisfied [16], there exists an element $m_{\mathbb{P}}$ in the RKHS $\mathscr{H}_{\mathcal{X}}$ called the *kernel distribution embedding*,

$$\begin{aligned} m : \mathscr{P} &\to \mathscr{H}_{\mathcal{X}} \\ \mathbb{P} &\mapsto m_{\mathbb{P}} \coloneqq \int_{\mathcal{X}} k_{\mathcal{X}}(y, \cdot)\mathbb{P}(y)\mathrm{d}y \end{aligned} \tag{4}$$

This representation has several advantages. First, if the kernel function is *universal* [16], the mapping is injective, meaning there is a unique element in the RKHS $\mathscr{H}_{\mathcal{X}}$ for any $\mathbb{P}, \mathbb{Q} \in \mathscr{P}$, such that $\|m_{\mathbb{P}} - m_{\mathbb{Q}}\|_{\mathscr{H}_{\mathcal{X}}} = 0$ if and only

if $\mathbb{P} = \mathbb{Q}$. A popular kernel which satisfies these properties is the Gaussian kernel, $k_{\mathcal{X}}(x, x') = \exp(-\|x - x'\|_2^2/2\sigma^2)$, $\sigma > 0$. Second, using the reproducing property (3a), we can compute the expectation of a function with respect to the distribution $\mathbb{P}$ as an inner product with the embedding.

$$\langle m_{\mathbb{P}}, f\rangle_{\mathscr{H}_{\mathcal{X}}} = \int_{\mathcal{X}} f(y)\mathbb{P}(y)\mathrm{d}y \tag{5}$$

Lastly, because $\mathbb{P}$ is typically unknown, we can compute an efficient estimate of $m_{\mathbb{P}}$. As shown in [22], the estimate for a conditional distribution embedding is the closed-form solution of a regularized least-squares problem.

We consider mapping the stochastic kernel Q into the RKHS $\mathscr{H}_{\mathcal{X}}$ (4) [23]. Its representation in $\mathscr{H}_{\mathcal{X}}$ is given by

$$m_{\boldsymbol{y}|x,u} \coloneqq \int_{\mathcal{X}} k_{\mathcal{X}}(y, \cdot)Q(y \,|\, x, u)\mathrm{d}y \tag{6}$$

Because Q is unknown, we do not have access to $m_{\boldsymbol{y}|x,u}$ directly. Instead, we use a sample $\mathcal{S}$ drawn i.i.d. from Q to compute an estimate $\hat{m}_{\boldsymbol{y}|x,u} \in \mathscr{H}_{\mathcal{X}}$ which can be found by minimizing the following optimization problem:

$$\sum_{i=1}^{M} \|k_{\mathcal{X}}(\bar{y}_i, \cdot) - \hat{m}_{\boldsymbol{y}|\bar{x}_i,\bar{u}_i}\|^2_{\mathscr{H}_{\mathcal{X}}} + \lambda\|\hat{m}_{\boldsymbol{y}|\boldsymbol{x},\boldsymbol{u}}\|^2_{\Gamma} \tag{7}$$

where Γ is a vector-valued RKHS [24] and $\lambda > 0$ is the regularization parameter. According to [22], the solution to (7) is unique and has the following form:

$$\hat{m}_{\boldsymbol{y}|x,u} = \Phi(\Psi\Psi^\top + \lambda MI)^{-1}\Psi k_{\mathcal{X}\times\mathcal{U}}((x, u), \cdot) \tag{8}$$

The vectors Φ and Ψ are known as *feature vectors*, given by

$$\Phi = [k_{\mathcal{X}}(\bar{y}_1, \cdot), \ldots, k_{\mathcal{X}}(\bar{y}_M, \cdot)]^\top \tag{9}$$

$$\Psi = [k_{\mathcal{X}\times\mathcal{U}}((\bar{x}_1, \bar{u}_1), \cdot), \ldots, k_{\mathcal{X}\times\mathcal{U}}((\bar{x}_M, \bar{u}_M), \cdot)]^\top \tag{10}$$

Using the estimate $\hat{m}_{\boldsymbol{y}|x,u}$, we can approximate the expectation $\mathbb{E}_{\boldsymbol{y}\sim Q}[f(\boldsymbol{y})]$ for any $f \in \mathscr{H}_{\mathcal{X}}$ as an inner product.

$$\langle \hat{m}_{\boldsymbol{y}|x,u}, f\rangle_{\mathscr{H}_{\mathcal{X}}} \approx \mathbb{E}_{\boldsymbol{y}\sim Q}[f(\boldsymbol{y})] \tag{11}$$

For simplicity, we can write this as

$$\langle \hat{m}_{\boldsymbol{y}|x,u}, f\rangle_{\mathscr{H}_{\mathcal{X}}} = \boldsymbol{f}^\top\beta(x, u) \tag{12}$$

where $\boldsymbol{f} = [f(\bar{y}_1), \ldots, f(\bar{y}_M)]^\top$ and $\beta(x, u) \in \mathbb{R}^M$ is a vector of coefficients that depends on the value of the conditioning variables $(x, u) \in \mathcal{X} \times \mathcal{U}$.

$$\beta(x, u) = (\Psi\Psi^\top + \lambda MI)^{-1}\Psi k_{\mathcal{X}\times\mathcal{U}}((x, u), \cdot) \tag{13}$$

This means we can approximate the expectation of the value functions $\mathbb{E}_{\boldsymbol{y}\sim Q}[V^{\pi}_{k+1}(\boldsymbol{y})]$ in (2b) *as an inner product with the conditional distribution embedding estimate.*

Computing the estimate typically requires us to compute and store a matrix $G = \Psi\Psi^\top \in \mathbb{R}^{M\times M}$, which is at least $\mathcal{O}(M^2)$. For large sample sizes, the storage and computation of G may be prohibitive. In order to overcome this computational challenge, we compute an approximation of the kernel itself and thus obtain a low-dimensional approximation of G using a technique known as random Fourier features [18].

B. Random Fourier Features

As shown in [18], we can reduce the computational complexity of computing (12) by exploiting Bochner's theorem [25]. This allows us to approximate the inner product in (12) by approximating the Fourier transform of the kernel.

Bochner's Theorem. *[25] A translation-invariant kernel $k_{\mathcal{X}}(x, x') = \varphi(x - x')$ on $\mathcal{X}$ is positive definite if and only if $\varphi(x - x')$ is the Fourier transform of a non-negative Borel measure Λ.*

$$\varphi(x - x') = \int_{\mathcal{X}} \exp(j\omega^\top(x - x'))\Lambda(\omega)\mathrm{d}\omega \tag{14}$$

$$= \int_{\mathcal{X}} \cos(\omega^\top(x - x'))\Lambda(\omega)\mathrm{d}\omega \tag{15}$$

where (15) *follows from the real-valued property of φ.*

Following [18], we construct an estimate of (15) using a sample $\Omega = \{\bar{\omega}_i\}_{i=1}^{D}$ of size D, such that $\bar{\omega}_i$ is drawn i.i.d. from the Borel measure Λ according to $\bar{\omega}_i \sim \Lambda(\cdot)$.

$$k_{\mathcal{X}}(x, x') \approx \frac{1}{D}\sum_{i=1}^{D} \cos(\bar{\omega}_i^\top(x - x')) \tag{16}$$

We define a *random feature map* $z : \mathcal{X} \to \mathbb{R}^D$ such that

$$k_{\mathcal{X}}(x, x') \approx \frac{1}{D}\sum_{i=1}^{D} z_{\bar{\omega}_i}(x)z_{\bar{\omega}_i}(x') =: \langle z(x), z(x')\rangle \tag{17}$$

$$z_{\omega}(x) = \sqrt{2}\cos(\omega^\top x + b) \tag{18}$$

where b is drawn uniformly from $[0, 2\pi]$. Let $\hat{k}_{\mathcal{X}} \approx k_{\mathcal{X}}$ denote the kernel approximation. Using random feature maps to approximate $k_{\mathcal{X}}$ and $k_{\mathcal{U}}$, we define the feature vector Z,

$$Z = [z(\bar{x}_1) \otimes z(\bar{u}_1), \ldots, z(\bar{x}_M) \otimes z(\bar{u}_M)]^\top \tag{19}$$

where $\otimes$ denotes the algebraic tensor product. Using (19), we can approximate (12) as

$$\langle \hat{m}_{\boldsymbol{y}|x,u}, f\rangle_{\mathscr{H}_{\mathcal{X}}} \approx \boldsymbol{f}^\top\gamma(x, u) \tag{20}$$

where $\gamma(x, u) \in \mathbb{R}^M$ is a vector of coefficients computed using the random feature vector Z in (19) (cf. [26]).

$$\gamma(x, u) = (ZZ^\top + \lambda MI)^{-1}Z(z(x) \otimes z(u)) \tag{21}$$

This means we can approximate the expectation of the value function $\mathbb{E}_{\boldsymbol{y}\sim Q}[V^{\pi}_{k+1}(\boldsymbol{y})]$ in (2b) *as an inner product of random feature maps.* Note that the matrix $H = ZZ^\top \in \mathbb{R}^{D\times D}$ has lower dimensionality than G if $D < M$, making it more computationally efficient to compute and store. As remarked in [18], evaluating a function using the kernel trick requires $\mathcal{O}(Md)$ operations, where d is the dimensionality of the data, whereas RFF only requires $\mathcal{O}(D+d)$ operations.

IV. Approximate Stochastic Reachability

With the conditional distribution embedding $m_{\boldsymbol{y}|x,u}$, the value function in (2b) can be written as

$$V^{\pi}_{k}(x) = \mathbf{1}_{\mathcal{T}}(x) + \mathbf{1}_{\mathcal{K}\setminus\mathcal{T}}(x)\langle m_{\boldsymbol{y}|x,u}, V^{\pi}_{k+1}\rangle_{\mathscr{H}_{\mathcal{X}}} \tag{22}$$

With the estimate $\hat{m}_{\boldsymbol{y}|x,u}$ and using the RFF approximation in (20), we obtain the approximation

$$V_k^\pi(x) \approx \mathbf{1}_{\mathcal{T}}(x) + \mathbf{1}_{\mathcal{K}\setminus\mathcal{T}}(x)\langle \hat{m}_{\boldsymbol{y}|x,u}, V_{k+1}^\pi\rangle_{\mathscr{H}_\mathcal{X}} \tag{23}$$

We define the approximate value functions $\overline{V}_k^\pi : \mathcal{X} \to [0,1]$, $k \in [0, N-1]$ via the backward recursion

$$\overline{V}_k^\pi \coloneqq \mathbf{1}_{\mathcal{T}}(x) + \mathbf{1}_{\mathcal{K}\setminus\mathcal{T}}(x)\langle \hat{m}_{\boldsymbol{y}|x,u}, V_{k+1}^\pi\rangle_{\mathscr{H}_\mathcal{X}} \tag{24}$$

where $V_k^\pi(x) \approx \overline{V}_k^\pi$. Let $\overline{V}_N^\pi = V_N^\pi$. Following [17], we can approximate the safety probability in (1) by approximating $\overline{V}_{k+1}^\pi$ and recursively substituting it into (24). This procedure is outlined in Algorithm 1. Using this, we obtain the approximation $r_{x_0}^\pi(\mathcal{K},\mathcal{T}) \approx \overline{V}_0^\pi(x)$.

A. Convergence

We now seek to characterize the quality of the approximation and analyze the conditions for its convergence. First, we analyze the convergence of the estimate $\hat{m}_{\boldsymbol{y}|x,u}$. As shown in [27], the estimate $\hat{m}_{\boldsymbol{y}|x,u}$ converges in probability to $m_{\boldsymbol{y}|x,u}$ at a rate of $\mathcal{O}_p(M^{-1/4})$ if the regularization parameter λ is decreased at a rate of $\mathcal{O}(M^{-1/2})$ (cf. [22]).

Theorem 1. *[27, Theorem 6] Assume $k_\mathcal{X}$ is in the range of $\mathbb{E}_{\boldsymbol{x}}[k_\mathcal{X}(\boldsymbol{x},\cdot)\otimes k_\mathcal{X}(\boldsymbol{x},\cdot)]$, then $\hat{m}_{\boldsymbol{y}|x,u}$ converges to $m_{\boldsymbol{y}|x,u}$ in the RKHS norm at a rate of $\mathcal{O}_p((M\lambda)^{-1/2} + \lambda^{1/2})$.*

This means we have theoretical guarantees of convergence of the embedding $\|m_{\boldsymbol{y}|x,u} - \hat{m}_{\boldsymbol{y}|x,u}\|_{\mathscr{H}_\mathcal{X}} \to 0$ as $M \to \infty$. Thus, for any function $f \in \mathscr{H}_\mathcal{X}$, using Cauchy-Schwarz,

$$\begin{aligned}|\langle m_{\boldsymbol{y}|x,u}, f\rangle_{\mathscr{H}_\mathcal{X}} - \langle \hat{m}_{\boldsymbol{y}|x,u}, f\rangle_{\mathscr{H}_\mathcal{X}}| \\ \leq \|f\|_{\mathscr{H}_\mathcal{X}}\|m_{\boldsymbol{y}|x,u} - \hat{m}_{\boldsymbol{y}|x,u}\|_{\mathscr{H}_\mathcal{X}}\end{aligned} \tag{25}$$

Since $\|m_{\boldsymbol{y}|x,u} - \hat{m}_{\boldsymbol{y}|x,u}\|_{\mathscr{H}_\mathcal{X}}$ converges in probability according to Theorem 1, the approximation of f computed using the estimate $\hat{m}_{\boldsymbol{y}|x,u}$ also converges in probability.

Next, we consider the convergence of the RFF approximation in (20). Convergence rates and finite-sample bounds for RFF in a generalized setting have been explored in [19, 26, 28]. We utilize the results in [28], which presents bounds for RFF in the context of least-squares problems with Tikhonov regularization. According to [28, Theorem 1], the approximation computed via RFF in (20) has an error of $\mathcal{O}_p(M^{-1/2})$ if we choose D according to $\mathcal{O}(M^{1/2}\log M)$ and decrease λ at a rate of $\mathcal{O}(M^{-1/2})$.

We can use this result to show that by properly choosing D, the approximate value functions converge in probability at the rate in Theorem 1. Following [28] and under the assumptions of Theorem 1, we present the following theorem.

Theorem 2. *The approximate value functions $\overline{V}_k^\pi(x)$ converge in probability to $V_k^\pi(x)$ at a rate of $\mathcal{O}_p(M^{-1/4})$ if D is chosen according to $\mathcal{O}(M^{1/2}\log M)$ and λ is decreased at a rate of $\mathcal{O}(M^{-1/2})$.*

The proof follows by combining the convergence rates from [28, Theorem 1] and Theorem 1 to obtain the result. Thus, if $|V_k^\pi(x) - \overline{V}_k^\pi(x)|$ has a probabilistic error bound of $\varepsilon > 0$ at every time $k < N$, the approximation $r_{x_0}^\pi(\mathcal{K},\mathcal{T}) \approx \overline{V}_0^\pi(x)$ computed using Algorithm 1 converges in probability with an error of $N\varepsilon$ [17].

Algorithm 1 Backward Recursion via RFF

Input: sample $\mathcal{S}$, evaluation point x, policy π, horizon N, sample $\Omega = \{\bar{\omega}_i\}_{i=1}^D$ such that $\bar{\omega}_i \sim \Lambda(\cdot)$
Output: value function estimate $\overline{V}_0^\pi(x) \approx r_{x_0}^\pi(\mathcal{K},\mathcal{T})$

$\overline{V}_N^\pi(x) \leftarrow \mathbf{1}_\mathcal{T}(x)$
for $k \leftarrow N-1$ to 0 **do**
  Compute $\gamma(x,\pi_k(x))$ from (21) using $\mathcal{S}$ and Ω
  $\mathcal{Y} \leftarrow [\overline{V}_{k+1}^\pi(\bar{y}_1),\ldots,\overline{V}_{k+1}^\pi(\bar{y}_M)]^\top$
  $\overline{V}_k^\pi(x) \leftarrow \mathbf{1}_\mathcal{T}(x) + \mathbf{1}_{\mathcal{K}\setminus\mathcal{T}}(x)\mathcal{Y}^\top\gamma(x,\pi_k(x))$
end for
Return $\overline{V}_0^\pi(x)$

V. Numerical Results

We implemented Algorithm 1 on a stochastic chain of integrators for the purposes of validation, and on a million-dimensional repeated planar quadrotor example in order to demonstrate the method for high-dimensional systems. We generate observations via simulation, and then presume no knowledge of the dynamics or the structure of the uncertainty for the purposes of computing the safety probability $r_{x_0}^\pi(\mathcal{K},\mathcal{T})$ in (1) using Algorithm 1. For all problems, we used a Gaussian kernel $\exp(-\|x-x'\|_2^2/2\sigma^2)$ with $\sigma = 0.1$, and chose $\lambda = 1$ as the default regularization parameter. The Borel measure Λ that corresponds to the Fourier transform of the Gaussian kernel is a Gaussian distribution of the form $\Lambda(\omega) = \sigma^{-1}\exp(-\sigma^2\|\omega\|_2^2/2)$.

All computations were done in Matlab on a 3.8GHz Intel Xeon CPU with 32 GB RAM. Computation times were obtained using Matlab's Performance Testing Framework. Code to generate all figures is available at *https://github.com/unm-hscl/ajthor-ACC2021*.

A. Stochastic Chain of Integrators

We consider a 2-D stochastic chain of integrators [11], in which the input appears at the 2^{nd} derivative and each element of the state vector is the discretized integral of the element that follows it. The dynamics with sampling time T are given by:

$$\boldsymbol{x}_{k+1} = \begin{bmatrix}1 & T\\ 0 & 1\end{bmatrix}\boldsymbol{x}_k + \begin{bmatrix}\frac{T^2}{2!}\\ T\end{bmatrix}u_k + \boldsymbol{w}_k \tag{26}$$

where $\boldsymbol{w}_k$ is an i.i.d. disturbance defined on the probability space $(\mathcal{W}, \mathscr{B}(\mathcal{W}), \Pr_{\boldsymbol{w}})$. We consider three distributions for the disturbance: 1) A Gaussian distribution $\boldsymbol{w}_k \sim \mathcal{N}(0,\Sigma)$, where $\Sigma = 0.01I$; 2) A beta distribution $\boldsymbol{w}_k \sim 0.1\text{Beta}(\alpha,\beta)$, with PDF $f(x\,|\,\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$ where Γ is the Gamma function and shape parameters $\alpha = 2$, $\beta = 0.5$; and 3) An exponential distribution $\boldsymbol{w}_k \sim 0.01\text{Exp}(\alpha)$, with $\alpha = 3$ and PDF $f(x\,|\,\alpha) = \alpha\exp(-\alpha x)$. For the purpose of validation against a known model, the control policy was chosen to be $\pi(x) = 0$ The target set and safe set are defined as $\mathcal{T} = [-0.5, 0.5]^2$ and $\mathcal{K} = [-1,1]^2$.

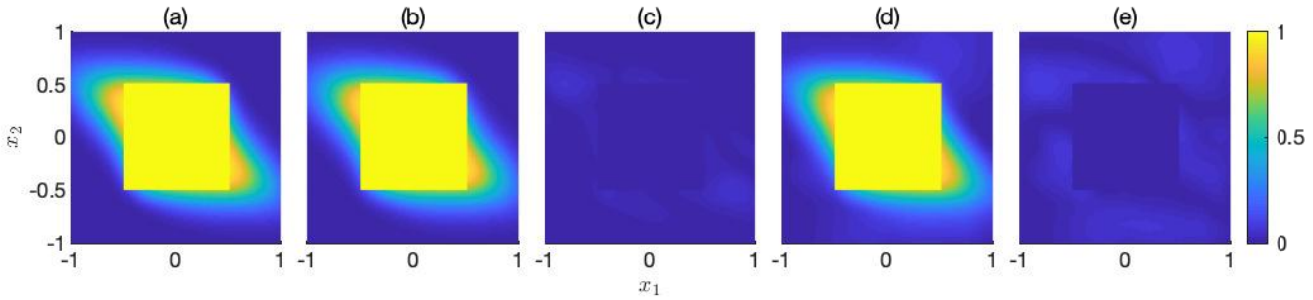

Fig. 1. (a) Dynamic-programming-based solution for a double integrator system with a Gaussian disturbance over the horizon $N = 5$. (b) First-hitting time safety probabilities for a double integrator system computed without RFF. (c) Absolute error between (a) and (b). (d) First-hitting time safety probabilities for a double integrator using Algorithm 1, where $D = 15{,}000$. (e) Absolute error between (a) and (d).

For the 2-D chain of integrators with a Gaussian disturbance, in order to compare against a known "truth" model, we computed the safety probabilities using a dynamic programming solution implemented in [29] with a time horizon of $N = 5$ (Fig. 1(a)). Following [17], we then computed the safety probabilities using β in (13) (without RFF) using a sample $\mathcal{S}$ of size $M = 2{,}500$ (Fig. 1(b)) in order to compare against the quality of the approximation obtained using RFF. The absolute error between the approximation and the dynamic programming solution is shown in Fig. 1(c), and the maximum absolute error was 0.0748. We then generated $D = 15{,}000$ frequency samples from $\Lambda(\omega)$ and computed the safety probabilities using γ in (21) (with RFF) for the same sample $\mathcal{S}$ according to Algorithm 1 (Fig. 1(d)). The absolute error between the approximation computed using RFF and the dynamic programming solution is shown in Fig. 1(e), and the maximum absolute error was 0.0907. From Theorem 2, the approximation error increases as the time horizon increases. Thus, to demonstrate the increase in error over longer time horizons, we recomputed the safety probabilities over a time horizon of $N = 50$. The absolute error of the approximation vs. dynamic programming over a longer time horizon $N = 50$ is 0.0836, while the absolute error using RFF is 0.2223.

We then computed the safety probabilities for the same system with a beta distribution disturbance and an exponential distribution disturbance for a time horizon of $N = 50$. The results are shown in Fig. 2. Because Algorithm 1 is agnostic to the complexities of the disturbance, handling arbitrary disturbances is straightforward.

As expected, Algorithm 1 produced a higher error estimate of the safety probabilities due to the kernel approximation. The quality of the approximation is dependent on M and D, and in some cases, the number of frequency samples D required to approximate the kernel can mean Algorithm 1 does not provide better computational efficiency. However, when $D \ll M$, or when the system is high-dimensional, RFF can significantly reduce the computational burden. By choosing a lower value of D, we exchange numerical accuracy for lower computation times of the algorithm.

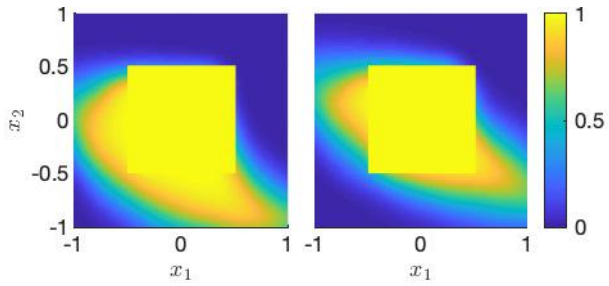

Fig. 2. First-hitting time safety probabilities for a double integrator system with a beta distribution disturbance (left) and an exponential distribution disturbance (right) over the horizon $N = 50$.

B. Planar Quadrotor

We implemented Algorithm 1 on a planar quadrotor system, as well as a million-dimensional repeated planar quadrotor system, comprised of 170,000 dynamically decoupled six-dimensional planar quadrotors. This problem can be interpreted as a simplification of formation control for a large swarm of quadrotors, where we compute the safety probabilities for the entire swarm as the quadrotors are controlled to reach a particular configuration. The nonlinear dynamics of a single quadrotor are given by

$$\begin{aligned} m\ddot{x} &= -(u_1 + u_2)\sin(\theta) \\ m\ddot{y} &= (u_1 + u_2)\cos(\theta) - mg \\ \mathcal{I}\ddot{\theta} &= r(u_1 - u_2) \end{aligned} \tag{27}$$

where x is the lateral position, y is the vertical position, θ is the pitch, and we have the constants inertia $\mathcal{I} = 2$, length $r = 2$, mass $m = 5$, and $g = 9.8$ is the gravitational constant.

For a single quadrotor, the state space is $\mathcal{X} \subset \mathbb{R}^6$, with state vector given by $\boldsymbol{z} = [x, \dot{x}, y, \dot{y}, \theta, \dot{\theta}]^\top$, and the input space is $\mathcal{U} \subset \mathbb{R}^2$, with input vector $u = [u_1, u_2]^\top$. The input is chosen to be a reference tracking controller, computed using a linearization of the system dynamics about a hover point. We discretize the nonlinear dynamics in time using an Euler approximation with sampling time $T = 0.25$, and add an affine disturbance $\boldsymbol{w}$. The disturbance is a Markov process with elements $\boldsymbol{w}_k$ defined on the probability space $(\mathcal{W}, \mathscr{B}(\mathcal{W}), \Pr_{\boldsymbol{w}})$. We consider two distributions for the

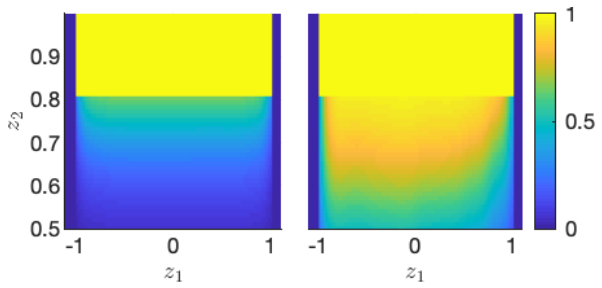

Fig. 3. First-hitting time safety probabilities for a planar quadrotor system with a Gaussian disturbance (left) and with a beta distribution disturbance (right) over the horizon $N = 5$.

TABLE I
COMPUTATION TIME

System	Dim. [n]	Without RFF	With RFF	Dyn. Prog.
Integrator	2	2.30 s	22.94 s	65.78 s
Quadrotor	6	0.62 s	15.24 s	–
Quadrotor	1,020,000	1.23 h	44.59 s	–

disturbance: 1) A Gaussian distribution $\boldsymbol{w}_k \sim \mathcal{N}(0, \Sigma)$, with variance $\Sigma = \text{diag}(1\times10^{-3}, 1\times10^{-5}, 1\times10^{-3}, 1\times10^{-5}, 1\times10^{-3}, 1\times10^{-5})$; and 2) A beta distribution $\boldsymbol{w}_k \sim \text{Beta}(\alpha, \beta)$ with shape parameters $\alpha = 2$, $\beta = 0.5$. The beta disturbance has a non-zero mean, and can be interpreted as wind, such that the dynamics are biased in a particular direction. For a single planar quadrotor, the safe set and target set are defined as $\mathcal{K} = \{\boldsymbol{z} \in \mathbb{R}^6 : |z_1| < 1, 0 \leq z_3 < 0.8\}$, and $\mathcal{T} = \{\boldsymbol{z} \in \mathbb{R}^6 : |z_1| < 1, z_3 \geq 0.8\}$. For the repeated quadrotor system, we define the safe sets and target sets as a series of parallel tubes, such that no quadrotor may enter into the safe set of an adjacent quadrotor. This means the quadrotors must all reach an altitude of 0.8 while remaining within their respective tube.

We first computed the safety probabilities for a single quadrotor in order to demonstrate the capabilities of Algorithm 1 to handle nonlinear dynamics. We generated a sample $\mathcal{S}$ consisting of $M = 1{,}000$ observations of the single quadrotor system with a Gaussian disturbance and took $D = 15{,}000$ frequency samples from $\Lambda(\omega)$. We then computed the safety probabilities using Algorithm 1 over a time horizon of $N = 5$ and then repeated this procedure using the beta distribution disturbance. The results are shown in Fig. 3. As expected, the algorithm was able to compute the safety probabilities due to the fact that Algorithm 1 does not exploit any knowledge of the underlying dynamics.

For the repeated quadrotor system, we first computed the safety probabilities without RFF in order to demonstrate the reduced computational complexity of Algorithm 1 for high-dimensional systems. We generated a sample $\mathcal{S}$ of $M = 1{,}000$ observations drawn i.i.d. from the stochastic kernel of the repeated quadrotor system with a beta distribution disturbance, and computed the safety probabilities (without RFF) over a time horizon of $N = 1$ from a single initial condition, to demonstrate feasibility of the approach. We repeated this procedure 7 times and averaged the computation time over all trials to obtain an average computation time of 1.23 hours. One advantage of using the Matlab performance testing framework is that it provides a 95% confidence of the computation time of the algorithm by running the code multiple times to obtain a mean computation time with a 5% margin of error. This means that the computation times we obtain are accurate to within a small margin of error, and a good indication of the relative performance of the algorithms. We then compared this performance against Algorithm 1 (with RFF) using the same procedure. We generated $D = 15{,}000$ frequency samples from $\Lambda(\omega)$ and computed the safety probabilities using Algorithm 1 over the same time horizon and the same initial condition. Using the same averaging approach, we obtained an average computation time of 44.59 seconds. We obtained comparable results for the Gaussian disturbance case. As shown in Table I, computation time is reduced by two orders of magnitude for the high-dimensional repeated quadrotor system.

VI. CONCLUSIONS & FUTURE WORK

We presented an algorithm based on random Fourier features to compute the stochastic reachability first-hitting time safety probabilities for high-dimensional Markov control processes. This approach is applicable to arbitrary disturbances and is model-free, meaning it does not rely upon a known stochastic kernel. We demonstrated it on a million-dimensional system to showcase the efficiencies of the computation. We plan to extend this to safe controller synthesis using kernel distribution embeddings.

REFERENCES

[1] S. Summers and J. Lygeros, "Verification of discrete time stochastic hybrid systems: A stochastic reach-avoid decision problem," *Automatica*, vol. 46, no. 12, pp. 1951–1961, 2010.

[2] N. Kariotoglou, S. Summers, T. Summers, M. Kamgarpour, and J. Lygeros, "Approximate dynamic programming for stochastic reachability," in *Proc. European Ctrl. Conf.*, 2013, pp. 584–589.

[3] G. Manganini, M. Pirotta, M. Restelli, L. Piroddi, and M. Prandini, "Policy search for the optimal control of Markov decision processes: A novel particle-based iterative scheme," *IEEE Trans. Cybern.*, vol. 46, no. 11, pp. 2643–2655, 2015.

[4] K. Lesser, M. Oishi, and R. S. Erwin, "Stochastic reachability for control of spacecraft relative motion," in *Proc. IEEE Conf. Dec. & Ctrl.*, 2013, pp. 4705–4712.

[5] S. Soudjani, C. Gevaerts, and A. Abate, "FAUST^2: Formal abstractions of uncountable-state stochastic processes," in *Int. Conf. on Tools and Algorithms for the Construction and Anal. of Syst.*, 2015, pp. 272–286.

[6] A. Vinod, V. Sivaramakrishnan, and M. Oishi, "Piecewise-affine approximation-based stochastic optimal control with Gaussian joint chance constraints," in *Proc. Amer. Ctrl. Conf.*, 2019, pp. 2942–2949.

[7] H. Sartipizadeh, A. Vinod, B. Açıkmeşe, and M. Oishi, "Voronoi partition-based scenario reduction for fast sampling-based stochastic reachability computation of LTI systems," *Proc. Amer. Ctrl. Conf.*, pp. 37–44, 2018.

[8] A. Vinod, B. HomChaudhuri, C. Hintz, A. Parikh, S. Buerger, M. Oishi, G. Brunson, S. Ahmad, and R. Fierro, "Multiple pursuer-based intercept via forward stochastic reachability," in *Proc. Amer. Ctrl. Conf.*, 2018, pp. 1559–1566.

[9] A. Vinod, S. Rice, Y. Mao, M. Oishi, and B. Açıkmeşe, "Stochastic motion planning using successive convexification and probabilistic occupancy functions," in *Proc. IEEE Conf. Dec. & Ctrl.*, 2018, pp. 4425–4432.

[10] A. Vinod and M. Oishi, "Scalable underapproximative verification of stochastic LTI systems using convexity and compactness," in *Proc. Hybrid Syst.: Comput. and Ctrl.*, 2018, pp. 1–10.
[11] ——, "Scalable underapproximation for the stochastic reach-avoid problem for high-dimensional LTI systems using Fourier transforms," *IEEE Ctrl. Syst. Letters.*, vol. 1, no. 2, pp. 316–321, 2017.
[12] S. Bak, H.-D. Tran, and T. Johnson, "Numerical verification of affine systems with up to a billion dimensions," in *Proc. Hybrid Syst.: Comput. and Ctrl.*, 2019, pp. 23–32.
[13] S. Bak and P. S. Duggirala, "HyLAA 2.0: A verification tool for linear hybrid automaton models of cyber-physical systems," in *IEEE Real-Time Syst. Symp.*, 2018.
[14] ——, "HyLAA: A tool for computing simulation-equivalent reachability for linear systems," in *Proc. Hybrid Syst.: Comput. and Ctrl.*, 2017, pp. 173–178.
[15] D. Bertsekas and S. Shreve, *Stochastic optimal control: the discrete time case*. Academic Press, 1978.
[16] A. Smola, A. Gretton, L. Song, and B. Schölkopf, "A Hilbert space embedding for distributions," in *Int. Conf. Algorithmic Learn. Theory*, 2007, pp. 13–31.
[17] A. Thorpe and M. Oishi, "Model-free stochastic reachability using kernel distribution embeddings," *IEEE Ctrl. Syst. Letters.*, 2019.
[18] A. Rahimi and B. Recht, "Random features for large-scale kernel machines," in *Adv. in Neural Inf. Process. Syst.*, 2008, pp. 1177–1184.
[19] ——, "Weighted sums of random kitchen sinks: Replacing minimization with randomization in learning," in *Adv. in Neural Inf. Process. Syst.*, 2009, pp. 1313–1320.
[20] A. Berlinet and C. Thomas-Agnan, "Reproducing kernel Hilbert spaces in probability and statistics," 2004.
[21] N. Aronszajn, "Theory of reproducing kernels," *Trans. of the American Mathematical Society*, vol. 68, no. 3, pp. 337–404, 1950.
[22] S. Grünewälder, G. Lever, L. Baldassarre, S. Patterson, A. Gretton, M. Pontil, J. Langford, and J. Pineau, "Conditional mean embeddings as regressors," in *Int. Conf. on Mach. Learn.*, 2012, pp. 1823–1830.
[23] S. Grünewälder, G. Lever, L. Baldassarre, M. Pontil, and A. Gretton, "Modelling transition dynamics in MDPs with RKHS embeddings," in *Int. Conf. on Mach. Learn.*, 2012, pp. 1603–1610.
[24] C. Micchelli and M. Pontil, "On learning vector-valued functions," *Neural Computation*, vol. 17, no. 1, pp. 177–204, 2005.
[25] W. Rudin, *Fourier analysis on groups*. Wiley, 1962, vol. 121967.
[26] Z. Li, J.-F. Ton, D. Oglic, and D. Sejdinovic, "Towards a unified analysis of random Fourier features," in *Int. Conf. on Mach. Learn.*, 2019, pp. 3905–3914.
[27] L. Song, J. Huang, A. Smola, and K. Fukumizu, "Hilbert space embeddings of conditional distributions with applications to dynamical systems," in *Int. Conf. on Mach. Learn.*, 2009, pp. 961–968.
[28] A. Rudi and L. Rosasco, "Generalization properties of learning with random features," in *Adv. in Neural Inf. Process. Syst.*, 2017, pp. 3215–3225.
[29] A. Vinod, J. Gleason, and M. Oishi, "SReachTools: a Matlab stochastic reachability toolbox," in *Proc. Hybrid Syst.: Comput. and Ctrl.*, 2019, pp. 33–38.