

Data-Driven Reachability Analysis with Christoffel Functions

Alex Devonport, Forest Yang, Laurent El Ghaoui, Murat Arcak

Electrical Engineering and Computer Sciences

University of California, Berkeley

{alex.devonport, forestyang, elghaoui, arcak}@berkeley.edu

Abstract—We present an algorithm for data-driven reachability analysis that estimates finite-horizon forward reachable sets for general nonlinear systems using level sets of a certain class of polynomials known as Christoffel functions. The level sets of Christoffel functions are known empirically to provide good approximations to the support of probability distributions: the algorithm uses this property for reachability analysis by solving a probabilistic relaxation of the reachable set computation problem. We also provide a guarantee that the output of the algorithm is an accurate reachable set approximation in a probabilistic sense, provided that a certain sample size is attained. We also investigate three numerical examples to demonstrate the algorithm’s capabilities, such as providing non-convex reachable set approximations and detecting holes in the reachable set.

I. INTRODUCTION

A popular and effective way to guarantee the safety of a system in the face of uncertainty is *reachability analysis*, a set-based method that characterizes all possible evolutions of the system by computing reachable sets. Many algorithms in reachability analysis use detailed system information to compute a sound approximation to the reachable set, that is an approximation guaranteed to completely contain (or be contained in) the reachable set. However, in many important applications, such as complex cyber-physical systems that are only accessible through simulations or experiments, this detailed system information is not available, so these algorithms cannot be applied.

Applications such as these motivate *data-driven* reachability analysis, which studies algorithms to estimate reachable sets using the type of data that can be obtained from experiments and simulations. These algorithms have the advantage of being able to estimate the reachable sets of any system whose behavior can be simulated or measured experimentally, without requiring any additional mathematical information about the system. The main disadvantage of data-driven reachability algorithms is that generally they cannot provide the same type of soundness guarantees as traditional reachability analysis algorithms; however, they can still guarantee accuracy of the estimates in a probabilistic sense with high confidence.

Data-driven reachability is a rapidly growing area of research within reachability analysis. Many recent developments focus either on providing probabilistic guarantees of correctness for data-driven methods that estimate the reachable set directly from data, for instance using results from statistical learning theory [1] or scenario optimization [2], [3], [4], [5], [6], [7]. Others incorporate data-driven elements

into more traditional reachability approaches, for instance estimating entities such as discrepancy functions [8] or differential inclusions [9]. Finally, other developments include incorporating data-driven reachability into verification tools for cyber-physical systems [8], [10].

This paper investigates a data-driven reachability algorithm that directly estimates the reachable set from data using the sublevel sets of an empirical inverse Christoffel function, and provides a probabilistic guarantee of accuracy for the method using statistical learning-theoretic methods. Christoffel functions are a class of polynomials defined with respect to measures on $\mathbb{R}^n$: a single measure defines a family of Christoffel function polynomials. When the measure in question is defined by a probability distribution on $\mathbb{R}^n$ the level sets of Christoffel functions are known empirically to provide tight approximations to the support. This support-approximating quality has motivated the use of Christoffel functions in several statistical applications, such as density estimation [11], [12] and outlier detection [13]. Additionally, the level sets have been shown, using the plug-in approach [14], to converge exactly to the support of the distribution (in the sense of Hausdorff measure) when the degree of the polynomial approaches infinity, and when the true probability distribution is available [12]. When the true probability distribution is *not* known, as is typically the case in data analysis, the Christoffel function can be empirically estimated using a point cloud of independent and identically distributed (iid) samples from the distribution: this *empirical Christoffel function* still provides accurate estimates for the support, and some convergence results in this case are also known [15].

The contribution of this paper is twofold. First, we provide an algorithm which uses the level sets of a Christoffel function to estimate a reachable set using a point cloud of iid samples from the reachable set, which can be obtained through simulations by a Monte Carlo sampling scheme. Second, we provide a guarantee of the probabilistic accuracy of the reachable set estimate produced by the algorithm: provided that a certain (finite) sample size is attained, the level set provided by the algorithm is guaranteed to achieve a user-specified level of probabilistic accuracy with high confidence. Unlike the convergence results of [12], [15], this result holds for finite sample sizes and finite degrees.

Notation

Given vectors $a, b \in \mathbb{R}^n$, a multidimensional interval (“interval” for brevity) is the set $[a, b] = \{x \in \mathbb{R}^n | a \leq x \leq b\}$.

$b\}$, where $\leq$ is the standard partial order $\mathbb{R}^n$. Given a vector x , a subscript x_i denotes the i^{th} element of x . Given an ordered multiset of vectors (a collection of points in $\mathbb{R}^n$ for instance), a superscript $x^{(i)}$ denotes the i^{th} member of the multiset. For $x \in \mathbb{R}^n$, the vector $z_k(x) \in \mathbb{R}^{\binom{n+k}{k}}$ denotes the vector of monomials of degree $\leq k$, including degree zero, evaluated at x : for instance, if $n = 2$ and $k = 2$, then $z_k(x) = [1 \ x_1 \ x_2 \ x_1x_2 \ x_1^2 \ x_2^2]^\top$. The space of polynomials of degree $\leq d$ in n variables is denoted $\mathbb{R}[x]_d^n$; note that elements of z_d , treated as polynomials, form a basis for $\mathbb{R}[x]_d^n$.

II. PRELIMINARIES

A. Probabilistic Reachability Analysis

Consider a dynamical system with a state transition function $\Phi(t_1; t_0, x_0, d)$ that maps an initial state $x(t_0) = x_0 \in \mathbb{R}^n$ at time t_0 to a unique final state at time t_1 , under a disturbance $d : [t_0, t_1] \rightarrow \mathbb{R}^w$. For instance, when the system state dynamics $\dot{x}(t) = f(t, x(t), d(t))$ are known and have unique solutions on the interval $[t_0, t_1]$, then $\Phi(t_1; t_0, x_0, d)$ is just $x(t_1)$, where x is the solution of the state dynamics with initial condition $x(t_0) = x_0$. In addition to representing exogenous disturbances, the disturbance signal d may account for deviations of an input from a nominal control law.

For the problem of forward reachability analysis, we are also given an *initial set* $\mathcal{X}_0 \subset \mathbb{R}^n$, a set $\mathcal{D}$ of allowed disturbances and a time range $[t_0, t_1]$. The *forward reachable set* is then defined as the set of all states to which the system can transition in the time range $[t_0, t_1]$ with initial states in $\mathcal{X}_0$ and disturbances in $\mathcal{D}$, that is the set

$$R_{[t_0, t_1]} = \{\Phi(t_1; t_0, x_0, d) : x_0 \in \mathcal{X}_0, d \in \mathcal{D}\}.$$

To tackle the problem of estimating the forward reachable set by statistical means, we add probabilistic structure to the reachability problem that corresponds to taking random independent samples from the reachable set. Specifically, we take random variables X_0 and D that take values on $\mathcal{X}_0$ and $\mathcal{D}$ respectively. These random variables then induce a random variable $\Phi(t_1; t_0, X_0, D)$ over the forward reachable set, whose probability measure we denote as μ .

Remark 1: The random variables X_0 and D may have a physical significance, if the initial states, inputs, or disturbances are known to behave randomly in the problem at hand. However, they do not need to: they may be considered as *instrumental distributions* whose purpose is to provide a consistent rule for selecting initial states and disturbances at random.

The measure $\mu(A)$ of a set $A \in \mathbb{R}^n$ has an intuitive interpretation: if we take samples x_0 and d of the random variables X_0 and D , then the vector $\Phi(t_1; t_0, x_0, d)$ lies in A with probability $\mu(A)$. Additionally, the smallest set of measure 1 is the reachable set. This interpretation motivates $\mu(A)$ as a measure of *probabilistic accuracy*: if a set $A \subseteq \mathbb{R}^n$ has a greater measure $\mu(A)$ than a set $B \subseteq \mathbb{R}^n$, then A is a more accurate approximation of the reachable set than B , in

the sense that it “misses” less of the probability mass than B does. In the probabilistic version of the forward reachability problem, our goal is to find reachable set approximations $\hat{R}_{[t_0, t_1]}$ such that $\mu(\hat{R}_{[t_0, t_1]})$ is close to 1. Formally, we look to solve the following problem.

Problem 1: Given the state transition function $\Phi(t_1; t_0, x_0, u)$, time range $[t_0, t_1]$, initial set $\mathcal{X}_0$, and disturbance set $\mathcal{D}$, the random variables X_0 and D , and an *accuracy level* $\epsilon \in (0, 1)$, compute a set $\hat{R}_{[t_0, t_1]}$ such that $\mu(\hat{R}_{[t_0, t_1]}) \geq 1 - \epsilon$.

Selecting a set with high measure under μ is not sufficient to ensure a reasonable estimate, since the trivial solution $\hat{R}_{[t_0, t_1]} = \mathbb{R}^n$ satisfies $\mu(\hat{R}_{[t_0, t_1]}) = 1$. To avoid this problem we require some regularization, such as requiring that $\hat{R}_{[t_0, t_1]}$ be compact and penalizing estimates with high volume.

B. Christoffel Functions

Given a finite measure μ on $\mathbb{R}^n$ and a positive integer k , the Christoffel function of order k is defined as the ratio

$$\kappa(x) = \frac{1}{z_k(x)^\top M^{-1} z_k(x)},$$

where M is the matrix of moments

$$M = \int_{\mathbb{R}^n} z_k(x) z_k(x)^\top d\mu(x)$$

and $z_k(x)$ is the vector of monomials of degree $\leq k$. We assume throughout that M is positive definite, ensuring that M^{-1} exists. The Christoffel function has several important application in approximation theory, where its asymptotic properties are used to prove the regularity and consistency of Fourier series of orthogonal polynomials. For our purposes, it is more convenient to use the *inverse Christoffel function*

$$\kappa(x)^{-1} = z_k(x)^\top M^{-1} z_k(x),$$

which is a polynomial of degree $2k$. In Problem 1, and more generally in the problem of estimating a probability distribution from samples, μ is a probability measure which we do not *a priori* know. In this case, we instead use an empirical estimate of μ constructed from a collection of independently and identically distributed (iid) samples $x^{(i)}$, $i = 1, \dots, N$ samples from μ , namely

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^N \delta_{x^{(i)}},$$

where δ_x is the *Dirac measure* satisfying $\int f(y) d\delta_x(y) = f(x)$. The measure $\hat{\mu}$ itself defines a Christoffel function, whose inverse

$$\begin{aligned} C(x) &= \hat{\kappa}^{-1}(x) = z_k(x)^\top \hat{M}^{-1} z_k(x) \\ &= z_k(x)^\top \left(\frac{1}{N} \sum_{i=1}^N z_k(x^{(i)}) z_k(x^{(i)})^\top \right)^{-1} z_k(x), \end{aligned}$$

is called the *empirical inverse Christoffel function*. The matrix $\hat{M}$ is positive definite (and hence $\hat{M}^{-1}$ exists) if $N \geq \binom{n+k}{k}$ and the $x^{(i)}$ do not all belong to the zero set of a single degree k polynomial.

III. CHRISTOFFEL FUNCTION LEVEL SETS AS REACHABLE SET APPROXIMATIONS

The ability of level sets of Christoffel functions to estimate the support of probability distributions motivates Algorithm 1 as a data-driven strategy for solving Problem 1. Specifically, Algorithm 1 computes an empirical inverse Christoffel function $C(x)$ and a level parameter $\alpha \in \mathbb{R}$, and returns the sublevel set $\{x \in \mathbb{R}^n : C(x) \leq \alpha\}$ as a proposed solution to Problem 1.

Algorithm 1: Data-driven reachable set estimation by a sublevel set of an empirical inverse Christoffel function.

Input: Transition function Φ of a system with state dimension n ; random variables X_0 and D defined on $\mathcal{X}_0$ and $\mathcal{D}$ respectively; time range $[t_0, t_1]$; probabilistic guarantee parameters ϵ and δ ; Christoffel function order k .

Output: Set $\hat{R}_{[t_0, t_1]}$ representing an ϵ -accurate reachable set estimate with confidence $1 - \delta$.
Set number of samples

$$N = \left\lceil \frac{5}{\epsilon} \left(\log \frac{4}{\delta} + \binom{n+2k}{n} \log \frac{40}{\epsilon} \right) \right\rceil.$$

forall $i \in \{1, \dots, N\}$ **do**

Take iid samples $x_0^{(i)}$ and $d^{(i)}$ from X_0 and D respectively;
evaluate $x_f^{(i)} = \Phi(t_1; t_0, x_0^{(i)}, d^{(i)})$.

end

Compute the matrix $\hat{M}^{-1}$ and level parameter α , where

$$\hat{M} = \frac{1}{N} \sum_{i=1}^N z_k(x_f^{(i)}) z_k(x_f^{(i)})^\top,$$

$$\alpha = \max_{i=1, \dots, N} z_k(x_f^{(i)})^\top \hat{M}^{-1} z_k(x_f^{(i)}).$$

Record the set

$$\hat{R}_{[t_0, t_1]} = \{x \in \mathbb{R}^n : z_k(x)^\top \hat{M}^{-1} z_k(x) \leq \alpha\}$$

as the reachable set estimate.

Since Algorithm 1 is a randomized algorithm, it is possible that a particular run will produce an invalid solution to Problem 1. However, Theorem 1 guarantees that the probability that this occurs is no greater than δ , a parameter that the user can specify in advance.

Theorem 1: Let C denote the empirical inverse Christoffel function for a point cloud $x^{(1)}, \dots, x^{(N)}$ of iid samples from μ , i.e.

$$C(x) = z_k(x)^\top \left(\frac{1}{N} \sum_{i=1}^N z_k(x^{(i)}) z_k(x^{(i)})^\top \right)^{-1} z_k(x),$$

and let $\alpha = \max_i C(x^{(i)})$. Let μ^N denote the joint probability measure corresponding to N iid samples from μ . If

$$N \geq \frac{5}{\epsilon} \left(\log \frac{4}{\delta} + \binom{n+2k}{n} \log \frac{40}{\epsilon} \right), \quad (1)$$

then

$$\mu^N \left(\{ (x^{(1)}, \dots, x^{(N)}) : \mu(\{x \in \mathbb{R}^n : C(x) \leq \alpha\}) \geq 1 - \epsilon \} \right) \geq 1 - \delta.$$

This means that, with probability $\geq 1 - \delta$, the α -sublevel set of $C(x)$ contains at least $1 - \epsilon$ of the probability mass of μ . The probability $1 - \delta$ is the *confidence* that the solution is valid. For instance, suppose we set $\delta = 10^{-9}$; then Theorem 1 gives us the confidence that there is less than a one in a billion chance that Algorithm 1 will fail to solve Problem 1.

The proof of this result is based on the following two results from statistical learning theory.

Lemma 1 ([16], Theorem 7.2): Let V be a vector space of functions $g : \mathbb{R}^n \rightarrow \mathbb{R}$ with dimension m . Then the class of sets

$$\text{Pos}(V) = \{ \{x | g(x) \geq 0\}, g \in V \}$$

has Vapnik–Chervonenkis (VC) dimension m .

Lemma 2 ([17], Corollary 4): Let $\mathcal{C}$ be a class of sets with VC dimension m . For a set $c \in \mathcal{C}$, let $\hat{\ell}(c) = \frac{1}{N} \sum_{i=1}^N 1\{x^{(i)} \notin c\}$ be the empirical error from a sample of N iid samples from μ , and let $\ell(c) = \mathbb{E}_\mu[1\{X \notin c\}] = 1 - \mu(c)$ be the generalization error. If

$$N \geq \frac{5}{\epsilon} \left(\log \frac{4}{\delta} + m \log \frac{40}{\epsilon} \right),$$

and if $\hat{\ell}(c) = 0$, that is if all of the points $x^{(i)}$, $i = 1, \dots, N$ are contained in the concept c , then $\mu^N(\{(x^{(1)}, \dots, x^{(N)}) : \ell(c) \leq \epsilon\}) \geq 1 - \delta$.

Theorem 1 follows from Lemmas 1 and 2 because the set $c = \{x \in \mathbb{R}^n | C(x) \leq \alpha\}$ belongs to the class $\mathcal{C} = \text{Pos}(\mathbb{R}[x]_d^n)$ and satisfies $\hat{\ell}(c) = 0$, and because the dimension of $\mathbb{R}[x]_d^n$ is $\binom{n+2k}{n}$.

In addition to providing a high-confidence solution to Problem 1, Algorithm 1 also achieves the regularization goals mentioned at the end of Section II-A. In particular, the estimate $\hat{R}_{[t_0, t_1]}$ produced by Algorithm 1 is compact, since it is a sublevel set of the sum-of-squares polynomial $z(x)^\top \hat{M}^{-1} z(x)$. Furthermore, the level parameter α can equivalently be defined as the solution to the optimization problem

$$\arg \min_{\alpha > 0} \alpha$$

subject to $z_k(x^{(i)})^\top \hat{M}^{-1} z_k(x^{(i)}) \leq \alpha, \quad i = 1, \dots, N.$

In this problem, α acts as a penalty term for the volume of the sublevel set, since the volume increases monotonically with increasing α .

Remark 2: In some reachability problems, we are only interested in computing a reachable set for a subset of the state

variables. For example, suppose the state is $(x_1, \dots, x_n) \in \mathbb{R}^n$, and we wish to verify a safety specification involving only the states $x_1, \dots, x_m$, where $m < n$: a reachable set for the states $x_1, \dots, x_m$ would suffice for this problem. In cases like this, Algorithm 1 can be modified to use only the first m elements of the samples $x_f^{(i)}$. The output of the algorithm is then an empirical inverse Christoffel function with domain $\mathbb{R}^m$ whose sublevel set $\hat{R}_{[t_0, t_1]}$ estimates the reachable set for the reduced set of states. In the sequel, we refer to this application of Algorithm 1 as the *reduced-state variant* of Algorithm 1.

IV. EXAMPLES

This section demonstrates Algorithm 1's ability to make accurate estimates of forward reachable sets with three numerical examples. We demonstrate how the parallel nature of the algorithm can be leveraged to improve computation times by running all experiments on two computing platforms: (i) a laptop with 4 2.6 GHz cores; and (ii) an instance of the AWS EC2 computing platform `c5.24xlarge`, a virtual machine with 96 3.6 GHz cores.

A. Chaotic Nonlinear Oscillator

The first example is a reachable set estimation problem for the nonlinear, time-varying system with dynamics

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\alpha y + x - x^3 + \gamma \cos(\omega t), \end{aligned} \quad (2)$$

with states $x, y \in \mathbb{R}$ and parameters $\alpha, \gamma, \omega \in \mathbb{R}$. This system is known as the *Duffing oscillator*, a nonlinear oscillator which exhibits chaotic behavior for certain values of α, γ , and ω , for instance

$$\alpha = 0.05, \quad \gamma = 0.4, \quad \omega = 1.3.$$

The initial is the interval such that $x(0) \in [0.95, 1.05]$, $y(0) \in [-0.05, 0.05]$, and we take X_0 to be the uniform random variable over this interval. The time range is $[t_0, t_1] = [0, 100]$.

We use Algorithm 1 to compute a reachable set for (2) using an order $k = 10$ empirical inverse Christoffel function with accuracy and confidence parameters $\epsilon = 0.05$, $\delta = 10^{-9}$. With these parameters, (1) states that $N = 156,626$ samples are required to ensure that Theorem 1 holds for the reachable set estimate. Total computation times for this example were 39 minutes on the laptop, and 41 seconds on `c5.24xlarge`.

Figure 1 shows the reachable set estimate for the Duffing oscillator system with the problem data given above, and the point cloud of 156,626 samples used to compute the empirical inverse Christoffel function and the level parameter α . The reachable set estimate is neither convex nor simply connected, closely following the boundaries of the cloud of points and excluding an empty region within the cloud of points.

To experimentally verify that the assertion of Proposition 1 holds for the reachable set estimate, we compute an *a posteriori* estimate of the accuracy of the empirical inverse

Christoffel function sublevel set. To do this, we first compute a new set of sample points of size N_{ap} . Denoting by N_{out} the number of new samples that lie outside of the reachable set estimate, we can compute the empirical accuracy of a reachable set approximation as $1 - N_{out}/N_{AP}$. We use $N_{AP} = 46,052$ sample points to make the *a posteriori* estimate. This sample size ensures that a one-sided Chernoff bound holds, which guarantees that empirical accuracy is within 1% of the true with 99.99% confidence. The *a posteriori* empirical accuracy computed with this sample is $1 - (2 \times 10^{-5})$, ensuring that the true accuracy of the reachable set estimate is at least $0.99 - 2 \times 10^{-5}$ with 99.99% confidence. This is well in excess of the 0.95 accuracy guaranteed by Theorem 1.

B. Planar Quadrotor Model

The next example is a reachable set estimation problem for horizontal position and altitude in a nonlinear model of the planar dynamics of a quadrotor used as an example in [18], [19]. The dynamics for this model are

$$\begin{aligned} \ddot{x} &= u_1 K \sin(\theta) \\ \ddot{h} &= -g + u_1 K \cos(\theta) \\ \ddot{\theta} &= -d_0 \theta - d_1 \dot{\theta} + n_0 u_2, \end{aligned}$$

where x and h denote the quadrotor's horizontal position and altitude in meters, respectively, and θ denotes its angular displacement (so that the quadrotor is level with the ground at $\theta = 0$) in radians. The system has 6 states, which we take to be x, h, θ , and their first derivatives. The two system inputs u_1 and u_2 (treated as disturbances for this example) represent the motor thrust and the desired angle, respectively. The parameter values used (following [19]) are $g = 9.81$, $K = 0.89/1.4$, $d_0 = 70$, $d_1 = 17$, and $n_0 = 55$. The set of initial states is the interval such that

$$\begin{aligned} x(0) &\in [-1.7, 1.7], & \dot{x}(0) &\in [-0.8, 0.8], \\ h(0) &\in [0.3, 2.0], & \dot{h}(0) &\in [-1.0, 1.0], \\ \theta(0) &\in [-\pi/12, \pi/12], & \dot{\theta}(0) &\in [-\pi/2, \pi/2], \end{aligned}$$

the set of inputs is the set of constant functions $u_1(t) = u_1$, $u_2(t) = u_2 \forall t \in [t_0, t_1]$, whose values lie in the interval

$$u_1 \in [-1.5 + g/K, 1.5 + g/K], \quad u_2 \in [-\pi/4, \pi/4],$$

and we take X_0 and D to be the uniform random variables defined over these intervals. The time range is $[t_0, t_1] = [0, 5]$. We take probabilistic parameters $\epsilon = 0.05$, $\delta = 10^{-9}$. Since the goal of this example is to estimate a reachable set for the horizontal position and altitude only, we are interested in a reachable set for a subset of the state variables, namely x and h . As mentioned in Remark 2, Algorithm 1 can be used to estimate a reachable set for x and h in two ways: we can either compute a Christoffel function estimate for the reachable set and take the "shadow projection" of the estimate onto x and h , or we could compute a Christoffel function estimate for x and h directly using the reduced-state variant of Algorithm 1 with the (x, h) components of the reachable set data. To compare the relative accuracy

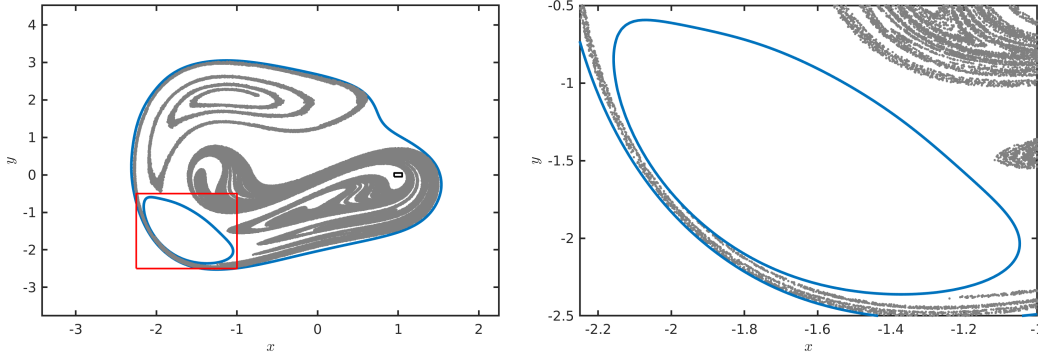


Fig. 1. *Left*: reachable set estimate for the Duffing oscillator system (blue contour), the cloud of 156,626 samples used to compute the empirical inverse Christoffel function (grey points), and the initial set (black box). *Right*: enlarged version of the region in the left plot enclosed by the red box, showing the region excluded from the reachable set.

and computational expense of these methods, we compute a reachable set estimate for (x, h) using both methods.

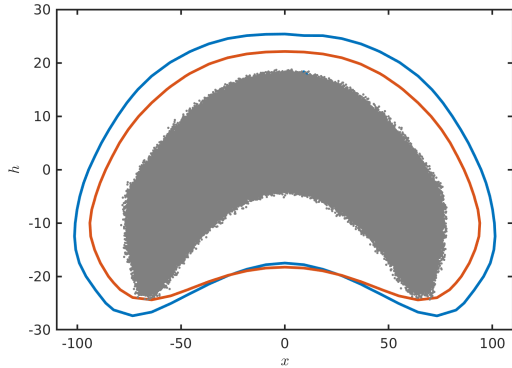


Fig. 2. Reachable set estimates for the horizontal position and altitude of the planar quadrotor model, computed by projecting the output of Algorithm 1 onto (x, h) (blue) and using the modification of Algorithm 1 mentioned in Remark 2, where the algorithm is run using only the (x, h) components of the data (orange).

Figure 2 shows the reachable set estimates computed using both methods using order $k = 4$ inverse empirical Christoffel functions. Both reachable estimates turn out to be similar, though the estimate using the modification of Remark 2 is slightly tighter and significantly less computationally expensive. Running Algorithm 1 with the full state dimension $n = 6$ and order $k = 4$ with the ϵ and δ above requires $N = 2,009,600$ samples: using the reduced-state variant brings the effective state dimension to $n = 2$, and the sample size to $N = 32,292$. The computation times in the full-state case were 77 minutes on the laptop and 2 minutes on c5.24xlarge; in the reduced-state case, computation times were 78 seconds on the laptop and 2 seconds on c5.24xlarge. This shows that Algorithm 1's ability to work on subsets of the state space can speed up computations in cases where only a subset of state variables are of interest.

C. Monotone Traffic Model

The final example is a special case of a continuous-time road traffic analysis problem used as a reachability benchmark in [20], [21], [22]. This problem investigates the density of traffic on a single lane over a time range over four periods of duration T using a discretization of the cell transmission model that divides the road into n equal segments. The spatially discretized model is an n -dimensional dynamical system with states $x_1, \dots, x_n$, where x_i represents the density of traffic in the i^{th} segment. Traffic enters segment through x_1 and flows through each successive segment before leaving through segment n . The state dynamics are

$$\begin{aligned} \dot{x}_1 &= \frac{1}{T} (d - \min(c, vx_1, w(\bar{x} - x_2))) \\ \dot{x}_i &= \frac{1}{T} (\min(c, vx_{i-1}, w(\bar{x} - x_i)) \\ &\quad - \min(c, vx_i, w(\bar{x} - x_{i+1}))), \quad (i = 2, \dots, n-1) \\ \dot{x}_n &= \frac{1}{T} (\min(c, vx_{n-1}, w(\bar{x} - x_n)/\beta) - \min(c, vx_n)), \end{aligned} \quad (3)$$

where v represents the free-flow speed of traffic, c the maximum flow between neighboring segments, $\bar{x}$ the maximum occupancy of a segment, and w the congestion wave speed. The input u represents the influx of traffic into the first node. For the reachable set estimation problem, we use a model with $n = 6$ states, and take $T = 30$, $v = 0.5$, $w = 1/6$, and $\bar{x} = 320$. The initial set is the interval such that $x_i(0) \in [100, 200]$, $i = 1, \dots, n$, the set of disturbances is the set of constant disturbances with values in the range $d \in [40/T, 60/T]$, and X_0 and D are the uniform random variables over these sets. The time range is $[t_0, t_1] = [0, 4T]$.

The system dynamics (3) are *monotone*, or order-preserving, meaning that if two initial conditions $x^{(1)}(0)$, $x^{(2)}(0)$ and disturbances $d^{(1)}, d^{(2)}$ satisfy $x^{(1)}(0) \leq x^{(2)}(0)$ (where $\leq$ is the standard partial order) and $d^{(1)}(t) \leq d^{(2)}(t)$, $t \in [0, T]$, then $x^{(1)}(T) \leq x^{(2)}(T)$. This monotonicity allows for a convenient interval over-approximation of the reachable set. If $\underline{x}, \bar{x}$ are the lower and upper bounds of the interval of initial states, and $\underline{d}, \bar{d}$ are the lower and

upper bounds on the values admitted by the disturbance signal, then $[\Phi(t_1; t_0, \underline{x}, \underline{d}), \Phi(t_1; t_0, \bar{x}, \bar{d})]$ is the smallest interval that contains the entire reachable set. While this over-approximation is easy to compute, and the best possible over-approximation by an interval, it is in general a conservative over-approximation because reachable set may only occupy a small volume of the interval. Since the empirical Inverse Christoffel function method can accurately detect the geometry of the reachable set, we use this method to compare the shape of the reachable set to the best interval over-approximation. In particular, we use the reduced-state variant of Algorithm 1 to compute a reachable set for the traffic densities x_5 and x_6 at the end of the road, using an order $k = 10$ empirical inverse Christoffel function with accuracy and confidence parameters $\epsilon = 0.05$, $\delta = 10^{-9}$. Computation times for this example were 10 minutes on the laptop and 2 minutes on c5.24xlarge.

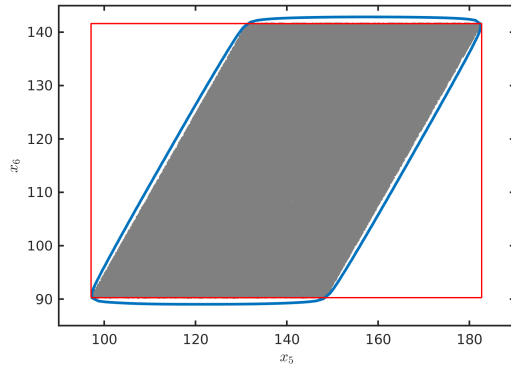


Fig. 3. Reachable set estimate for the monotone traffic model with an order 10 empirical inverse Christoffel function (blue), compared to the tight interval over-approximation (red). The reachable set estimate was computed with Algorithm 1 using samples projected onto states x_5 and x_6 .

Figure 3 compares the reachable set estimate computed with Algorithm 1 to the projection of the tight interval over-approximation computed using the monotonicity property of the traffic system. The figure indicates that the tight interval over-approximation of the reachable set is a somewhat conservative over-approximation, since the reachable set has approximately the shape of a parallelotope whose sides are not axis-aligned.

V. CONCLUSION

Algorithm 1 demonstrates that Christoffel functions, in addition to being useful in data analysis, can also be used as tools to provide principled, data-driven solutions to control-theoretic problems. While Theorem 1 assures that the proposed algorithm is a sound approach to solving reachability problems with data, and the examples of Section IV demonstrate that the algorithm can provide accurate reachable set approximations, we believe it represents only the first step in applying Christoffel functions to data-driven reachability. For instance, the *a posteriori* analysis of Section IV-A suggests the sample bound of Theorem 1 is conservative, and could

be significantly improved by applying some of the special properties of Christoffel functions.

In addition, this paper did not explore how kernel methods can be used alongside Christoffel functions. Although we have defined the Christoffel function using the standard monomial basis vector $z_k(x)$, the Christoffel function is in fact invariant to changes in polynomial coordinates. For instance, $z_k(x)$ could be replaced with the feature vector $\phi_k(x)$ of the polynomial kernel $(1 + x^\top x)^k$, that is the monomial vector $\phi_k(x)$ such that $\phi(x)^\top \phi(x) = (1 + x^\top x)^k$. By an application of the kernel trick, this approach can be extended to kernels with infinite-dimensional feature spaces, as in [13]. However, the statistical learning-theoretic proof in this paper covers only the finite-dimensional case: providing finite-sample statistical guarantees for the infinite-dimensional case is a topic for future research.

ACKNOWLEDGMENTS

This work was supported in part by the grants ONR N00014-18-1-2209, AFOSR FA9550-18-1-0253, NSF ECCS-1906164.

REFERENCES

- [1] A. Devonport and M. Arcak, “Data-driven reachable set computation using adaptive Gaussian process classification and Monte Carlo methods,” in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 2629–2634.
- [2] G. R. Marzeglia, J. Scott, L. Magni, R. D. Braatz, and D. M. Raimondo, “A hybrid stochastic-deterministic approach for active fault diagnosis using scenario optimization,” *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 1102–1107, 2014.
- [3] Y. Yang, J. Zhang, K.-Q. Cai, and M. Prandini, “Multi-aircraft conflict detection and resolution based on probabilistic reach sets,” *IEEE Transactions on Control Systems Technology*, vol. 25, no. 1, pp. 309–316, 2016.
- [4] D. Ioli, A. Falsone, H. Marianne, B. Axel, and M. Prandini, “A smart grid energy management problem for data-driven design with probabilistic reachability guarantees,” in *4th International Workshop on Applied Verification of Continuous and Hybrid Systems*, vol. 48, 2017, pp. 2–19.
- [5] H. Sartipzadeh, A. P. Vinod, B. Açikmeşe, and M. Oishi, “Voronoi partition-based scenario reduction for fast sampling-based stochastic reachability computation of linear systems,” in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 37–44.
- [6] L. Hewing and M. N. Zeilinger, “Scenario-based probabilistic reachable sets for recursively feasible stochastic model predictive control,” *IEEE Control Systems Letters*, vol. 4, no. 2, pp. 450–455, 2019.
- [7] A. Devonport and M. Arcak, “Estimating reachable sets with scenario optimization,” ser. *Proceedings of Machine Learning Research*, A. M. Bayen, A. Jadbabaie, G. Pappas, P. A. Parrilo, B. Recht, C. Tomlin, and M. Zeilinger, Eds., vol. 120. PMLR, 10–11 Jun 2020, pp. 75–84.
- [8] C. Fan, B. Qi, S. Mitra, and M. Viswanathan, “DryVR: data-driven verification and compositional reasoning for automotive systems,” in *International Conference on Computer Aided Verification*. Springer, 2017, pp. 441–461.
- [9] F. Djeumou, A. P. Vinod, E. Goubault, S. Putot, and U. Topcu, “On-the-fly control of unknown smooth systems from limited data,” *arXiv preprint arXiv:2009.12733*, 2020.
- [10] B. Qi, C. Fan, M. Jiang, and S. Mitra, “DryVR 2.0: a tool for verification and controller synthesis of black-box cyber-physical systems,” in *Proceedings of the 21st International Conference on Hybrid Systems: Computation and Control (part of CPS Week)*, 2018, pp. 269–270.
- [11] J. B. Lasserre and E. Pauwels, “The empirical Christoffel function in statistics and machine learning,” *arXiv preprint arXiv:1701.02886*, 2017.
- [12] —, “The empirical Christoffel function with applications in data analysis,” *Advances in Computational Mathematics*, vol. 45, no. 3, pp. 1439–1468, 2019.

- [13] A. Askari, F. Yang, and L. E. Ghaoui, "Kernel-based outlier detection using the inverse Christoffel function," *arXiv preprint arXiv:1806.06775*, 2018.
- [14] A. Cuevas and R. Fraiman, "A plug-in approach to support estimation," *The Annals of Statistics*, vol. 25, no. 6, pp. 2300–2312, 1997.
- [15] E. Pauwels, M. Putinar, and J.-B. Lasserre, "Data analysis from empirical moments and the Christoffel function," *Foundations of Computational Mathematics*, pp. 1–31, 2020.
- [16] R. M. Dudley, "Central limit theorems for empirical measures," *The Annals of Probability*, pp. 899–929, 1978.
- [17] T. Alamo, R. Tempo, and E. F. Camacho, "Randomized strategies for probabilistic solutions of uncertain feasibility and optimization problems," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2545–2559, 2009.
- [18] I. M. Mitchell, J. Budzis, and A. Bolyachevets, "Invariant, viability and discriminating kernel under-approximation via zonotope scaling," in *Proceedings of the 22nd ACM International Conference on Hybrid Systems: Computation and Control*, 2019, pp. 268–269.
- [19] P. Bouffard, "On-board model predictive control of a quadrotor helicopter: Design, implementation, and experiments," 2012. [Online]. Available: <http://www.eecs.berkeley.edu/Pubs/TechRpts/2012/EECS-2012-241.html>
- [20] S. Coogan and M. Arcak, "A benchmark problem in transportation networks," *arXiv preprint arXiv:1803.00367*, 2018.
- [21] P.-J. Meyer, A. Devonport, and M. Arcak, "Tira: toolbox for interval reachability analysis," in *Proceedings of the 22nd ACM International Conference on Hybrid Systems: Computation and Control*. ACM, 2019, pp. 224–229.
- [22] A. Devonport, M. Khaled, M. Arcak, and M. Zamani, "PIRK: scalable interval reachability analysis for high-dimensional nonlinear systems," in *Computer Aided Verification*, S. K. Lahiri and C. Wang, Eds. Cham: Springer International Publishing, 2020, pp. 556–568.