Safe Learning of Regions of Attraction for Uncertain, Nonlinear Systems with Gaussian Processes

Felix Berkenkamp, Riccardo Moriconi, Angela P. Schoellig, and Andreas Krause

Abstract— Control theory can provide useful insights into the properties of controlled, dynamic systems. One important property of nonlinear systems is the region of attraction (ROA), a safe subset of the state space in which a given controller renders an equilibrium point asymptotically stable. The ROA is typically estimated based on a model of the system. However, since models are only an approximation of the real world, the resulting estimated safe region can contain states outside the ROA of the real system. This is not acceptable in safety-critical applications. In this paper, we consider an approach that learns the ROA from experiments on a real system, without ever leaving the true ROA and, thus, without risking safety-critical failures. Based on regularity assumptions on the model errors in terms of a Gaussian process prior, we use an underlying Lyapunov function in order to determine a region in which an equilibrium point is asymptotically stable with high probability. Moreover, we provide an algorithm to actively and safely explore the state space in order to expand the ROA estimate. We demonstrate the effectiveness of this method in simulation.

I. Introduction

Estimating the region of attraction (ROA) of an equilibrium point is an important problem when analyzing nonlinear systems. It specifies the region within which a given control law renders an equilibrium point asymptotically stable. Thus, the system can be operated safely within the ROA. This is of practical importance, since controllers for safety-critical systems are required to guarantee safety over a certain domain of operation, before they can be implemented on the real system. Typically, ROAs are estimated based on a model of the system. However, due to model errors, ROAs of the real system can be drastically different, raising questions about the viability of such purely model-based methods.

In this paper, we address the problem of safely estimating the ROA from experiments on a real system, while always remaining in the real ROA with high probability. As a result, the equilibrium point of the closed-loop system is attractive for any state visited throughout the learning process. This avoids failures and makes the method applicable to safety-critical domains. To expand the estimated safe region over time, we actively select states that improve our estimate of the underlying uncertain model.

Felix Berkenkamp and Andreas Krause are with the Learning & Adaptive Systems Group (www.las.ethz.ch), Department of Computer Science, ETH Zurich, Switzerland. Email: {befelix, krausea}@ethz.ch

Riccardo Moriconi is with the Department of Mechanical Engineering, ETH Zurich, Switzerland. Email: rmoricon@ethz.ch

Angela P. Schoellig is with the Dynamic Systems Lab (www.dynsyslab.org), Institute for Aerospace Studies, University of Toronto, Canada. Email: schoellig@utias.utoronto.ca

Related work: In the literature, ROAs are computed based on system models [1]. One typical approach is to use the level sets of a Lyapunov function in order to quantify the ROA [2]. Given a Lyapunov function, the problem of finding the ROA reduces to a line search for the maximum safe level set of the Lyapunov function. For polynomial systems, Lyapunov functions can be found efficiently by solving a system of linear matrix inequalities (LMIs) [3]. A relaxation to Lyapunov-like functions for ROA computation is given in [4] and a review of numerical methods to compute Lyapunov functions can be found in [5]. The approach in [6] considered quadratic systems and the problem of testing whether a given polytope lies inside the ROA. Beyond Lyapunov functions, the ROA can be found via constraint solving [7] or by optimizing over state trajectories [8]. The ROA can also be approximated using sampling [9].

An area that is of particular relevance to this paper is the computation of ROAs for systems with uncertainties. In [10], quadratic Lyapunov functions for uncertain, linear systems were used, while [11] considered polynomial Lyapunov functions. A more general approach for systems with a polynomial dependence on uncertainties within a polytope is shown in [12]. Similar ideas are used in robust control [13].

All the previous methods for estimating ROAs are based on a model with fixed uncertainties. The method proposed in this paper uses learning in order to reduce the model uncertainty over time. Learning the dynamics of nonlinear systems from data has been considered before. In particular, Gaussian processes (GPs [14]) have been of interest in this area, since they provide uncertainty information about the estimated dynamic model. This can be used to derive high-probability safety guarantees. For example, [15] combines GP regression with robust control theory to provide stability and performance guarantees for a linear controller applied to a nonlinear system, while [16] uses reachability analysis to compute robust invariant sets.

In our method, we use ideas from Bayesian optimization [17], set of optimization methods that aim to find the global optimum of an *a priori* unknown function based on noisy evaluations. Bayesian optimization algorithms that are based on GP models provably converge close to the global optimum [18]. In this setting, [19], [20] consider safety in terms of high-probability constraint satisfaction, with a generalization to multiple constraints in [21].

Our contribution: In this paper, we combine ideas from GP learning, safe Bayesian optimization, and ROA computation based on Lyapunov functions for uncertain systems. Using previous results on model-based ROA computation as a

starting point, we compute ROA estimates for the real system by approximating the model uncertainties with a GP on a discrete domain. We use ideas from safe Bayesian optimization in order to actively learn about the real dynamics from experiments executed in the estimated ROA. As we learn more about the dynamics of the real system, the uncertainty in the estimate decreases and the ROA expands, until we reach a desired accuracy. As a result, we are able to learn an accurate estimate of the ROA of a general, nonlinear system through experiments, without leaving the ROA and, thus, without violating safety requirements.

II. PROBLEM STATEMENT

We consider a nonlinear, continuous-time system,

$$\dot{x}(t) = \underbrace{f(x(t), u(t))}_{\text{a priori model}} + \underbrace{g(x(t), u(t))}_{\text{unknown model}}, \tag{1}$$

where $x(t) \in \mathcal{X} \subseteq \mathbb{R}^q$ is the state at time t within a connected set $\mathcal{X}$ and $u(t) \in \mathcal{U} \subseteq \mathbb{R}^p$ is the control input. The system dynamics consist of a *known* model $f(x, u)$ and an *unknown* model $g(x, u)$. The latter accounts for unknown dynamics and model uncertainties. We assume a control policy $u = \pi(x)$ is given, which has been designed for the prior model, $f(\cdot)$. The resulting, closed-loop dynamics are denoted by $f_\pi(x) := f(x, \pi(x))$ and $g_\pi(x) := g(x, \pi(x))$.

The goal is to estimate the ROA of (1) under the control policy, $\pi(x)$, based on experiments on the real system. We want to actively learn the *a priori* unknown dynamics, g_π, without leaving the ROA. This is, of course, impossible without further assumptions about the model in (1).

We assume that the unknown model, g_π, has low 'complexity', as measured under the norm of a *reproducing kernel Hilbert space* (RKHS, [22]). An RKHS $\mathcal{H}_k(\mathcal{X})$ is a complete subspace of $L_2(\mathcal{X})$, the space of square integrable functions, that includes functions of the form $g_\pi(x) = \sum_i \alpha_i k(x, x_i)$ with $\alpha_i \in \mathbb{R}$ and representer points $x_i \in \mathcal{X}$, where k (and the subscript k) refer to a symmetric, positive-definite kernel function $k(\cdot, \cdot)$. The RKHS has an inner product $\langle \cdot, \cdot \rangle_k$, which obeys the reproducing property: $\langle g_\pi(\cdot), k(x, \cdot) \rangle_k = g_\pi(x)$ for all $g_\pi \in \mathcal{H}_k(\mathcal{X})$. The induced RKHS norm $\|g_\pi\|_k^2 = \langle g_\pi, g_\pi \rangle_k$ measures smoothness of g_π with respect to the kernel $k(\cdot, \cdot)$, [18]. For universal kernels [22], members of $\mathcal{H}_k(\mathcal{X})$ can uniformly approximate any continuous function on any compact subset of $\mathcal{X}$. We make assumptions about the complexity of g_π.

Assumption 1. *The function g_π has bounded RKHS norm with respect to a continuously differentiable, bounded kernel $k(x, x')$; that is, $\|g_\pi\|_k \leq B_g$.*

Moreover, we make the following, standard assumption on the *a priori* model, $f_\pi(x)$:

Assumption 2. *The function f_π is Lipschitz continuous with Lipschitz constant L_f and bounded in $\mathcal{X}$ with $\|f_\pi\|_\infty \leq B_f$.*

In the following sections, we use a Lyapunov function in order to compute the ROA. In practice, one may compute good Lyapunov functions during the experimentation phase

using, for example, the methods in [5]. Here, we consider a fixed, given Lyapunov function. We assume the following:

Assumption 3. *The origin is an equilibrium point of* (1) *with $f_\pi(0) = g_\pi(0) = 0$.*

Assumption 4. *A fixed, two-times continuously differentiable Lyapunov function $V(x)$ is given. Moreover, there exists a constant $c_0 > 0$ such that $\frac{\partial V(x)}{\partial x}(f_\pi(x) + g_\pi(x)) < 0$ for all $x \in \mathcal{S}_0 = \mathcal{V}(c_0)$, where $\mathcal{V}(c) = \{x \in \mathcal{X} \mid V(x) \leq c\}$.*

Assumption 3 implies that the origin is an equilibrium point of (1) with the unknown model $g_\pi(x)$. Without this assumption, there is no hope to prove that the origin is attractive, but showing boundedness would be possible. Assumption 4 implies that we have chosen the policy, $\pi(x)$, and the Lyapunov function, $V(x)$, such that the origin of (1) is locally asymptotically stable. This is, for example, fulfilled if the origin of the *a priori* dynamics $\dot{x} = f_\pi(x)$ is locally asymptotically stable and $\frac{\partial g_\pi(x)}{\partial x}$ is zero at the origin; that is, the prior dynamics dominate close to the origin.

Lastly, we require measurements of $g_\pi(x)$ in order to learn about the unknown dynamics. That is, we must be able to measure x and $\dot{x}$ ($f_\pi(\cdot)$ is known). In practice, we may consider a discrete-time approximation of $\dot{x}$, which only requires measurements of the state, x, rather than its derivative, $\dot{x}$. We discuss practical aspects in Sec. VII.

Assumption 5. *We have access to measurements, $\hat{g}_\pi(x) = \dot{x} - f_\pi(x) + \omega$, which are corrupted by zero-mean, independent and bounded noise, $\|\omega\| \leq \sigma$.*

For ease of notation, we consider a one-dimensional system of the form of (1) in the following. We explain how to generalize the results to multiple dimensions in Sec. VI.

III. GAUSSIAN PROCESSES (GPs)

The function $g_\pi(x)$ in (1) is unknown *a priori*. Given Assumption 1, we use a GP to approximate this unknown function over its domain $\mathcal{X}$, see [18] and Lemma 1 below. The following introduction to GPs is based on [14].

GPs are a nonparametric regression method from machine learning, where the goal is to find an approximation of a nonlinear map $g_\pi: \mathcal{X} \to \mathbb{R}$ from a state x to the function value $g_\pi(x)$. This is accomplished by considering the function values $g_\pi(x)$ to be random variables, so that any finite number of them have a joint Gaussian distribution [14].

The Bayesian, nonparametric regression is based on a prior mean function and a covariance function, $k(x, x')$, which defines the covariance of any two function values, $g_\pi(x)$ and $g_\pi(x')$, $x, x' \in \mathcal{X}$. The latter is also known as the kernel. In this work, the mean is zero, since any prior knowledge about the dynamics is captured by $f_\pi(x)$; that is, $g_\pi(x) \sim \mathcal{GP}(0, k(x, x'))$. In general, the choice of kernel function is problem-dependent and encodes assumptions about the unknown function. A discussion of kernel choices for dynamic systems can be found in [23]. In the following, we use the same kernel as in Assumption 1.

We can obtain the posterior distribution of a function value $g_\pi(x)$ at an arbitrary state $x \in \mathcal{X}$ by conditioning the

GP distribution of g_π on a set of n past measurements, $y_n = (\hat{g}_\pi(x_1), \ldots, \hat{g}_\pi(x_n))$ at states $\mathcal{D}_n = \{x_1, \ldots, x_n\}$, where $\hat{g}_\pi(x) = \dot{x} - f_\pi(x) + \omega$ with $\omega \sim \mathcal{N}(0, \sigma^2)$. The posterior over $g_\pi(x)$ is a GP distribution again, with mean $\mu_n(x)$, covariance $k_n(x, x')$, and variance $\sigma_n(x)$:

$$\mu_n(x) = k_n(x)(K_n + \mathbb{I}_n \sigma^2)^{-1} y_n, \tag{2}$$

$$k_n(x, x') = k(x, x') - k_n(x)(K_n + \mathbb{I}_n \sigma^2)^{-1} k_n^{\mathrm{T}}(x'), \tag{3}$$

$$\sigma_n^2(x) = k_n(x, x), \tag{4}$$

where the vector $k_n(x) = \big(k(x, x_1), \ldots, k(x, x_n)\big)$ contains the covariances between the new input, x, and the states in $\mathcal{D}_n$, $K_n \in \mathbb{R}^{n \times n}$ is the positive-definite kernel matrix with $[K_n]_{(i,j)} = k(x_i, x_j)$, $i, j \in \{1, \ldots, n\}$, and $\mathbb{I}_n \in \mathbb{R}^{n \times n}$ is the identity matrix.

Assumption 1 allows us to model $g_\pi(x)$ as a GP. In particular, we have the following result:

Lemma 1. *Supposed that $\|g_\pi\|_k^2 \leq B_g$ and that the zero-mean noise ω is uniformly bounded by σ. Choose $\beta_n = 2B_g + 300\gamma_n \log^3(n/\delta)$, where γ_n is defined in the paragraph below. Then, for all $n \geq 1$ and $x \in \mathcal{X}$, it holds with probability at least $1 - \delta$, $\delta \in (0, 1)$, that*

$$|g_\pi(x) - \mu_{n-1}(x)| \leq \beta_n^{1/2} \sigma_{n-1}(x). \tag{5}$$

Proof. See [18, Theorem 6]. $\qquad\square$

Lemma 1 allows us to make high-probability statements about the function values of $g_\pi(x)$, even though the GP model makes different assumptions than we do in Assumption 1 (e.g., about the noise, ω). The scalar β_n depends on the information capacity, $\gamma_n = \max_{x_1, \ldots, x_n} I(g_\pi, y_n)$, which is the maximal mutual information that can be obtained about the GP prior from n noisy samples y_n at states $\mathcal{D}_n$. In [18], it was shown that γ_n has a sublinear dependence on n for many commonly used kernels and that it can be efficiently approximated up to a constant. As a result, even though β_n is increasing with n, we are able to learn about the true values of $g_\pi(x)$ over time by making appropriate choices for data samples in $\mathcal{D}_n$, see Sec. V.

IV. LYAPUNOV STABILITY

In this section, we show how the assumptions in Sec. II can be used in order to compute the ROA for the nonlinear system in (1) based on a GP model of g_π from Sec. III. For now, we base our analysis on a general GP model of $g_\pi(x)$ from Sec. III with n arbitrary measurements of g_π. We actively select states for new measurements in Sec. V.

We start by observing the following:

Lemma 2. *The function g_π is Lipschitz continuous with Lipschitz constant L_g and bounded by $\|g_\pi\|_\infty \leq B_g \|k\|_\infty$.*

Proof. Boundedness by [22, Lemma 4.23]. From Assumption 1, $|g_\pi(x) - g_\pi(x')|^2 \leq \|g_\pi\|_k^2 d_k(x, x')$, where $d_k(x, x') = k(x, x) + k(x', x') - 2k(x, x')$ is the kernel metric [22, Lemma 4.28]. Since $k(\cdot, \cdot)$ is continuously differentiable and bounded, $|g_\pi(x) - g_\pi(x')|^2 \leq L_g^2 |x - x'|^2$, where $L_g^2 = 2B_g \|k\|_\infty \|\frac{\partial k}{\partial x}\|_\infty$. $\qquad\square$

Since the closed-loop dynamics in (1) are Lipschitz continuous based on Lemma 2 and Assumption 2, existence and uniqueness of the solutions of (1) are guaranteed.

The goal of this section is to quantify the ROA based on the Lyapunov function in Assumption 4. From Lyapunov stability theory we have the following [2]:

Lemma 3. *The origin of the dynamics in (1) is asymptotically stable within a level set, $\mathcal{V}(c) = \{x \in \mathcal{X} \mid V(x) \leq c\}$ with $c \in \mathbb{R}_{>0}$, if, for all $x \in \mathcal{V}(c)$,*

$$\dot{V}(x) = \frac{\partial V(x)}{\partial x}\big(f_\pi(x) + g_\pi(x)\big) < 0. \tag{6}$$

The goal of this section is to use the GP model of $g_\pi(x)$ from Sec. III in order to find the largest c such that (6) holds true within $\mathcal{V}(c)$ with high probability. The existence of such a value $c_0 > 0$ is ensured by Assumption 4.

In order to quantify the ROA using (6), we need to evaluate $\dot{V}(x)$, which depends on the GP model of $g_\pi(x)$. From Sec. III we know that the GP posterior distribution over g_π is Gaussian. As a consequence, since $\dot{V}(x)$ in (6) is affine in $g_\pi(x)$, $\dot{V}(x)$ itself is a GP. In particular, we have that $\dot{V}(x)$ is normally distributed with mean $\mu_{n,\dot{V}}(x)$ and standard deviation $\sigma_{n,\dot{V}}(x)$ given by:

$$\mu_{n,\dot{V}}(x) = \frac{\partial V(x)}{\partial x}\big(\mu_n(x) + f_\pi(x)\big), \tag{7}$$

$$\sigma_{n,\dot{V}}(x) = \left|\frac{\partial V(x)}{\partial x}\right| \sigma_n(x), \tag{8}$$

where $\mu_n(x)$ and $\sigma_n(x)$ are the GP predictions of the unknown dynamics, $g_\pi(x)$, from (2) and (4). We can use (7) and (8) to directly make predictions about $\dot{V}(x)$ in (6) based on measurements of $g_\pi(x)$.

Since (7) and (8) provide a probabilistic estimate of $\dot{V}(x)$, we cannot expect to make deterministic statements about stability. Instead, we consider the confidence intervals of the GP model of $\dot{V}(x)$. We denote the lower and upper confidence intervals after $(n-1)$ measurements of $g_\pi(x)$ by

$$l_n(x) := \mu_{\dot{V}, n-1}(x) - \beta_n^{1/2} \sigma_{\dot{V}, n-1}, \tag{9}$$

$$u_n(x) := \mu_{\dot{V}, n-1}(x) + \beta_n^{1/2} \sigma_{\dot{V}, n-1}, \tag{10}$$

respectively. The confidence intervals are parameterized by the scalar β_n. In the following, we assume that β_n is chosen according to Lemma 1, which allows us to say that $\dot{V}(x)$ takes values within the interval $[l_n(x), u_n(x)]$ with high probability (at least $1 - \delta$).

Based on these confidence intervals, we can see from (6) that $\dot{V}(x) < 0$ holds within $\mathcal{V}(c)$ for some $c > 0$ with high probability, if $u_n(x) < 0$ for all $x \in \mathcal{V}(c)$. Unfortunately, it is impossible to evaluate the upper bound (10) everywhere in a continuous domain. Nevertheless, it is possible to evaluate predictions of $\dot{V}(x)$ from (7) and (8) at a finite number of points. In the following, we exploit the continuity properties of $\dot{V}(x)$ in order to derive stability properties within a continuous domain based on a finite number of predictions. In particular, we use the following property of $\dot{V}$:

Lemma 4. *The function $\dot{V}(x)$ is Lipschitz continuous with Lipschitz constant L.*

Proof. Based on the Lipschitz continuity of (1) from Assumption 2 and Lemma 2, we expand $|\dot{V}(x) - \dot{V}(x')|$ to

$$\left| \frac{\partial V(x)}{\partial x}(f_\pi(x) + g_\pi(x)) - \frac{\partial V(x')}{\partial x}(f_\pi(x') + g_\pi(x')) \right|$$
$$\leq \left| f_\pi(x) + g_\pi(x) \right| \left| \frac{\partial V(x)}{\partial x} - \frac{\partial V(x')}{\partial x'} \right|$$
$$+ \left| \frac{\partial V(x')}{\partial x'} \right| \left| f_\pi(x) + g_\pi(x) - f_\pi(x') - g_\pi(x') \right|,$$
$$\leq (B_f + B_g \|k\|_\infty) L_{\partial V} |x - x'| + L_V(L_g + L_f)|x - x'|,$$
$$:= L|x - x'|,$$

where $L_V = \|\frac{\partial V(x)}{\partial x}\|_\infty$ and $L_{\partial V} = \|\frac{\partial^2 V(x)}{\partial x^2}\|_\infty$ are the Lipschitz constants of V and its first derivative. These are guaranteed to exist by Assumption 4, since a continuous function obtains a maximum over a bounded domain. $\square$

The continuity of $\dot{V}(x)$ allows us to evaluate predictions of $\dot{V}(x)$ from (7) and (8) only at a finite number of points, without loosing guarantees.

Lemma 5. *Let $\mathcal{X}_\tau \subset \mathcal{X}$ be a discretization of $\mathcal{X}$ with $|x - [x]_\tau| \leq \tau/2$ for all $x \in \mathcal{X}$. Here, $[x]_\tau$ denotes the closest point in $\mathcal{X}_\tau$ to $x \in \mathcal{X}$. Choosing β_n according to Lemma 1, the following holds with probability at least $(1 - \delta)$ for all $x \in \mathcal{X}$ and all $n \geq 1$:*

$$\left| \dot{V}(x) - \mu_{\dot{V},\,n-1}([x]_\tau) \right| \leq \beta_n^{1/2} \sigma_{\dot{V},\,n-1}([x]_\tau) + L\tau. \quad (11)$$

Proof. The result follows from the Lipschitz continuity of $\dot{V}(x)$ in Lemma 4 and is similar to [18, Lemma 5.7]. $\square$

Lemma 5 provides high-probability bounds on $\dot{V}$ on the continuous domain $\mathcal{X}$ using the GP confidence intervals (7) and (8) on the discrete set $\mathcal{X}_\tau$. It achieves this by using the Lipschitz property of $\dot{V}$ to generalize from $\mathcal{X}_\tau$ to $\mathcal{X}$. The discretization accuracy τ trades off the accuracy for the reduced computational cost of computing the confidence intervals for all states in $\mathcal{X}_\tau$. Combining the previous results lets us argue about the stability:

Lemma 6. *With a discretization of $\mathcal{X}$, $\mathcal{X}_\tau$, according to Lemma 5 and with β_n according to Lemma 1, the origin of (1) is asymptotically stable within $\mathcal{V}(c)$ for some $c > 0$ with probability at least $(1 - \delta)$ if, for all $x \in \mathcal{V}(c) \cap \mathcal{X}_\tau$,*

$$\mu_{\dot{V},\,n-1}(x) + \beta_n^{1/2} \sigma_{\dot{V},\,n-1}(x) = u_n(x) < -L\tau. \quad (12)$$

Proof. This is a consequence of Lemmas 1, 3 and 5. $\square$

Given the previous results, after $(n-1)$ evaluations of g_π, it suffices to maximize c such that the condition in Lemma 6 holds for all discretized states in $\mathcal{V}(c)$:

Theorem 1. *Under the assumptions of Lemma 6, pick*

$$c_n = \underset{c \in \mathbb{R}_{>0}}{\operatorname{argmax}} \; c, \quad \text{subject to (12) for all } x \in \mathcal{V}(c) \cap \mathcal{X}_\tau. \quad (13)$$

Algorithm 1: Safe ROA exploration

Inputs: Domain $\mathcal{X}$ and discretization with τ, $\mathcal{X}_\tau$
   GP prior $k(x, x')$
   Initial safe set $\mathcal{S}_0 \subseteq \mathcal{X}$

1 **for** $n = 1, \ldots$ **do**
2 $\quad$ $c_n \leftarrow \underset{c > 0}{\operatorname{argmax}} \; c$, subject to
$\quad\quad\quad u_n(x) < -L\tau$, for all $x \in \mathcal{V}(c) \cap \mathcal{X}_\tau$
3 $\quad$ $\mathcal{S}_n \leftarrow \mathcal{S}_0 \cup \mathcal{V}(c_n)$
4 $\quad$ $x_n \leftarrow \operatorname{argmax}_{x \in \mathcal{S}_n} \sigma_{n-1}(x)$
5 $\quad$ Update GP with measurement of $\hat{g}_\pi(x_n)$

Then, $\mathcal{S}_n = \mathcal{S}_0 \cup \mathcal{V}(c_n)$ is contained within the ROA of (1) for all $n \geq 1$ with probability at least $(1 - \delta)$.

Proof. The set $\mathcal{S}_0$ from Assumption 4 is contained in the ROA deterministically, while $\mathcal{V}(c_n)$ is contained with probability at least $(1 - \delta)$ by Lemma 6. The result follows. $\square$

Theorem 1 provides a way to compute an estimate of the ROA via an efficient binary search to solve the optimization problem (13). The discretization in Lemma 6 allows us to consider high-probability confidence intervals on $\mathcal{X}_\tau$ only, while we generalize using the Lipschitz constant from Lemma 4. A finer discretization (smaller value of τ) decreases the conservativeness of the ROA estimate. This is different from the number of experiments, n, which decreases the uncertainty in the model estimate as we obtain more data. In particular, as more information about g_π becomes available from measurements, the ROA increases beyond the initial, deterministic safe set in Assumption 4.

V. ACTIVE LEARNING

In the previous section, we provided an estimate of the ROA based on a GP model of the unknown dynamics of $g_\pi(x)$ in (1), which was informed by arbitrarily selected measurements/experiments. In this section, we actively expand the ROA estimate by selecting new, safe states within the current ROA estimate at which to obtain measurements.

As we update the model with new data, we can compute the current estimate of the ROA, $\mathcal{S}_n$, using Theorem 1. In order to expand the ROA from the initial estimate $\mathcal{S}_0$, we need to obtain measurements at states that are relevant for learning g_π. Here, we assume that we have access to a control method that drives us to desired states without leaving the ROA. Any method that applies the policy $\pi(x)$ on the boundary of $\mathcal{S}_n$ fulfills this safety requirement [16]. Based on ideas from safe Bayesian optimization [19], we use the uncertainty estimate of the GP. We select

$$x_n = \underset{x \in \mathcal{S}_n}{\operatorname{argmax}} \; \sigma_{n-1}(x) \quad (14)$$

as the next state to evaluate according to (4) within a current estimate of the safe set, $\mathcal{S}_n$. At x_n, the GP model is the most uncertain about the unknown dynamics, $g_\pi(x)$. The idea behind this selection criterion is that we want to learn about the most uncertain state in order to increase the ROA estimate. By decreasing the uncertainty about the unknown
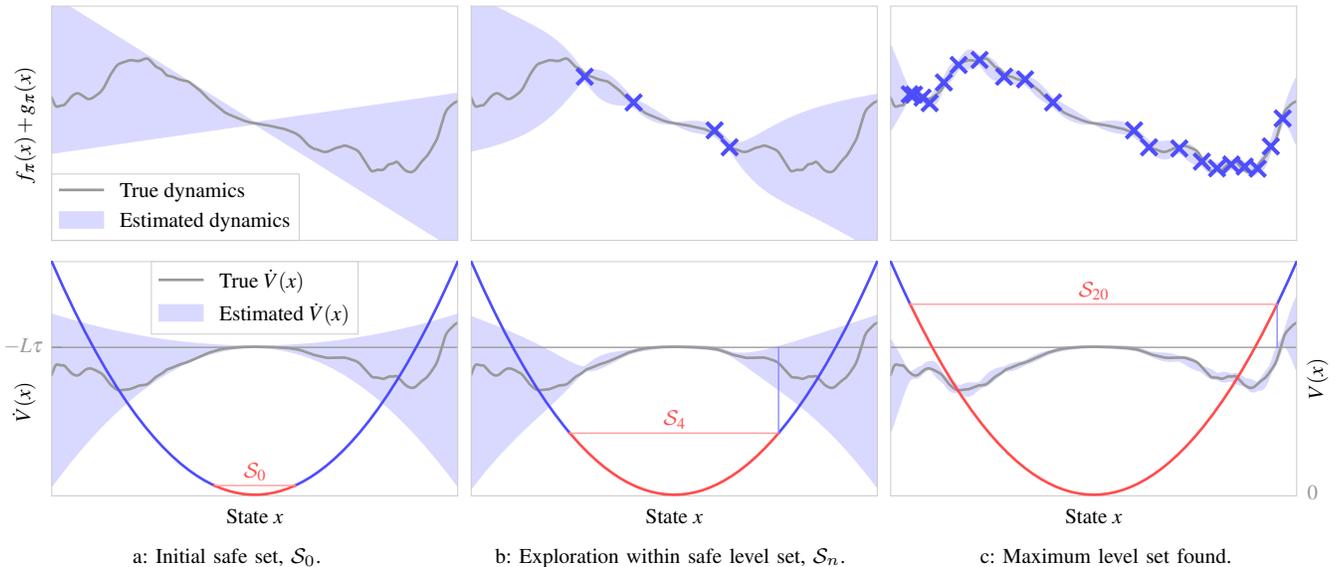
Fig. 1. One-dimensional example of Algorithm 1. Initially, in Fig. 1a, the estimate of the dynamics is uncertain (top row, blue shade is the 2σ confidence interval). As a result, the ROA (red part of the Lyapunov function, bottom plot) consists only of the initial safe set, $\mathcal{S}_0$; larger level sets of the Lyapunov function are unsafe (blue part). The algorithm actively selects new states at which to obtain measurements of the dynamics, which causes the safe set to increase in Fig. 1b. Eventually, this procedure leads to the largest safe level set in Fig. 1c. Only states inside the ROA are evaluated during the learning.

dynamics, $g_\pi(x)$, over time, the safe region expands eventually. We summarize the method in Algorithm 1 and an example is shown in Fig. 1.

It follows from Theorem 1 that every state chosen by (14) lies inside the ROA of (1) with high probability. For a variant of Algorithm 1 it is possible to prove an even stronger result: that the maximum, safe level set can be found to some accuracy (with probability at least $(1 - \delta)$) after a finite number of experiments. The corresponding variant of Algorithm 1 builds up the ROA estimate using the Lipschitz constant. More details are given in the appendix in Theorem 2.

VI. EXTENSION TO MULTIPLE DIMENSIONS

So far, we have considered a one-dimensional system in (1). The preceding analysis directly transfers to multiple dimensions, at the expense of more cumbersome notation. The only non-trivial parts of the extension are GP models with vector predictions and the choice of β_n in Lemma 1. The main observation is that vector-valued functions can be modeled as a single GP over an extended state space, $\mathcal{X} \times \mathcal{I}$, where integer elements in $\mathcal{I}$ index the output dimension of $g_\pi(x)$. At each iteration, we obtain q measurements of $g_\pi(x)$; one for each dimension of the state. As a result, $\beta_{n'} = 2B_g + 300\gamma'_{nq} \log^3(nq/\delta)$ increases at a faster rate of $n' = nq$ compared to β_n in Lemma 1. The information capacity, γ', is measured relative to the combined function over $\mathcal{X} \times \mathcal{I}$. Refer to [21] for more details.

VII. DISCUSSION

Algorithm 1 allows for the safe exploration of the ROA without leaving the ROA for a general, uncertain, nonlinear system. While the discretization in Lemma 5 may be conservative and computing the predictions of the GP model of $\dot{V}$ at the discretization points is computationally intensive, this

needs to be compared to the generality of the statements we have made. In particular, the only assumption we have made about the unknown dynamics, $g_\pi(x)$ in (1), is Assumption 1, which is very general.

In the analysis, we have assumed that a method exists, which drives the system to any state x_n selected by (14), without leaving the ROA. In practice, this must be further restricted by reachability properties, see [16]. These can be included as an additional constraint; that is, if $\mathcal{R}$ is the safely reachable set, we select x_n from $\mathcal{S}_n \cap \mathcal{R}$ in (14).

Algorithm 1 can be made more data-efficient by only evaluating states close to the boundary of the level set, without loosing guarantees. See [19] and [21] for details.

A discrete-time variant of Algorithm 1 that determines safety with $V(f_{\pi,d}(x) + g_{\pi,d}(x)) - V(x) < 0$ only requires measurements of the states, x, rather than derivatives, $\dot{x}$, in Assumption 5. While mapping a GP model of $g_\pi(x)$ through a nonlinear Lyapunov function renders this method not analytically tractable, methods based on approximate uncertainty propagation may work well in practice.

VIII. EXPERIMENTS

In this section. we evaluate Algorithm 1 in a simulated experiment. We only provide a high-level overview of the experiment. For details refer to the documentation and source code at http://github.com/befelix/lyapunov-learning.

We consider an inverted pendulum with angle θ, mass $m = 0.15\,\text{kg}$, length $l = 0.5\,\text{m}$, and friction coefficient $\mu = 0.05\,\text{Nms/rad}$. The dynamics are given by

$$\ddot{\theta}(t) = \frac{mgl\sin\theta(t) - \mu\dot{\theta}(t) + u(t)}{ml^2}, \qquad (15)$$

where $u(t)$ is the torque applied to the pendulum. The torque is limited so that the real system cannot recover from states with $|\theta| > 30\,\text{deg}$. The state is $x = (\theta, \dot{\theta})$. We assume that
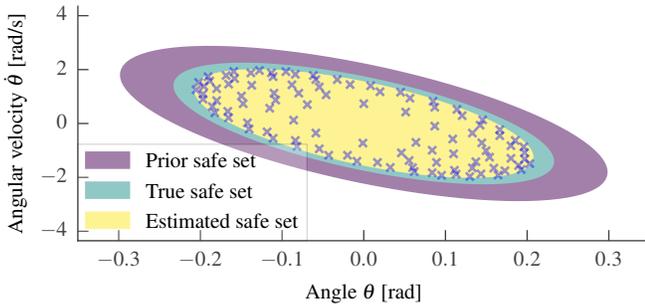
Fig. 2. Prior, true, and estimated ROA level sets after 100 data points (blues crosses). The prior model estimates a safe region (purple) that is larger than the true ROA (green), thus it includes unsafe states. In contrast, Algorithm 1 provides a conservative estimate (yellow), since it considers states with $\dot{V}(x) \le -L\tau$, rather than $\dot{V}(x) \le 0$. The level set could be increased by discretizing the space with a smaller value of τ.

we only know a linear approximation of the dynamics in (15) for the upright equilibrium point, where additionally friction is neglected and the mass is $0.05 \, \text{kg}$ lighter. We use a Linear Quadratic Regulator based on this linear approximation in order to control the origin, $x = 0$, and use the corresponding quadratic Lyapunov function to determine the ROA of (15).

We model the error in (15) as a GP, see Sec. III. In particular, we use the product of a linear and a Matérn kernel for the GP model, which encodes nonlinear functions that are two-times differentiable and take values that are bounded by a linear function from above and below with high probability. Details about kernel choice for dynamic systems can be found in [23] and a one-dimensional sample function of the kernel used here is in the upper plot of Fig. 1a. In practice, the more assumptions are made about the model error, the faster the learning algorithm will converge [23].

In the analysis, we assumed g_π to have bounded RKHS norm in Assumption 1. Here, we model g_π as a sample function of the GP and set $\beta_n^{1/2} = 2$ for all iterations. We use $\tau = 0.002$ and high-probability, local Lipschitz constants encoded by the kernel with Lemma 4. The initial safe set is $\mathcal{S}_0 = \{(\theta, \dot{\theta}) \in \mathbb{R}^2 \mid |\theta| \le 5 \deg, |\dot{\theta}| \le 10 \deg /\text{s}\}$.

The results can be seen in Fig. 2. While the prior model with the wrong mass and friction parameters estimates a safe set that is too large (includes unsafe states), Algorithm 1 provides a conservative estimate. As we gain more data about the dynamics and if we discretize with smaller values of τ, the estimate improves and, in the limit, converges to the true level set. Overall, Algorithm 1 provides a powerful tool to learn the ROA of general nonlinear systems from experiments, without leaving the safe region encoded by the Lyapunov function.

IX. CONCLUSION

We introduced a learning algorithm that, based on an initial, approximate model and a corresponding Lyapunov function, is able to learn about the real ROA through experiments on the real system, without leaving the true ROA with high probability. While some of the assumptions in Sec. II may be restrictive for practical application, the results in this paper should be understood as a theoretical foundation for learning algorithms that learn without risking instability.

REFERENCES

[1] A. Zecevic and D. D. Siljak, *Control of Complex Systems*, ser. Communications and Control Engineering. Springer, 2010.

[2] H. K. Khalil and J. W. Grizzle, *Nonlinear Systems*. Prentice Hall, 1996, vol. 3.

[3] B. Tibken, "Estimation of the domain of attraction for polynomial systems via LMIs," in *Proc. of the IEEE Conference on Decision and Control (CDC)*, vol. 4, 2000, pp. 3860–3864.

[4] S. Ratschan and Z. She, "Providing a basin of attraction to a target region of polynomial systems by computation of Lyapunov-like functions," *SIAM Journal on Control and Optimization*, vol. 48, no. 7, pp. 4377–4394, 2010.

[5] P. Giesl and S. Hafstein, "Review on computational methods for Lyapunov functions," *Discrete and Continuous Dynamical Systems, Series B*, vol. 20, no. 8, pp. 2291–2337, 2015.

[6] F. Amato, C. Cosentino, and A. Merola, "On the region of attraction of nonlinear quadratic systems," *Automatica*, vol. 43, no. 12, pp. 2119–2123, 2007.

[7] H. Burchardt and S. Ratschan, "Estimating the region of attraction of ordinary differential equations by quantified constraint solving," in *Proc. of the WSEAS International Conference on Dynamical Systems and Control*, 2007, pp. 241–246.

[8] D. Henrion and M. Korda, "Convex computation of the region of attraction of polynomial control systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 297–312, 2014.

[9] R. Bobiti and M. Lazar, "A sampling approach to constructing Lyapunov functions for nonlinear continuous-time systems," in *Proc. of the Conference on Decision and Control (CDC)*, 2016.

[10] A. Trofino, "Robust stability and domain of attraction of uncertain nonlinear systems," in *Proc. of the American Control Conference (ACC)*, vol. 5, 2000, pp. 3707–3711.

[11] D. Coutinho, A. Trofino, and M. Fu, "Guaranteed cost control of uncertain nonlinear systems via polynomial Lyapunov functions," *IEEE Transactions on Automatic Control*, vol. 47, no. 9, pp. 1575–1580, 2002.

[12] U. Topcu, A. K. Packard, P. Seiler, and G. Balas, "Robust region-of-attraction estimation," *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 137–142, 2010.

[13] K. Zhou and J. C. Doyle, *Essentials of Robust Control*. Prentice Hall, 1998, vol. 104.

[14] C. E. Rasmussen and C. K. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[15] F. Berkenkamp and A. P. Schoellig, "Safe and robust learning control with Gaussian processes," in *Proc. of the European Control Conference (ECC)*, 2015, pp. 2501–2506.

[16] A. K. Akametalu, S. Kaynama, J. F. Fisac, M. N. Zeilinger, J. H. Gillula, and C. J. Tomlin, "Reachability-based safe learning with Gaussian processes," in *Proc. of the IEEE Conference on Decision and Control (CDC)*, 2014, pp. 1424–1431.

[17] J. Mockus, *Bayesian Approach to Global Optimization*, ser. Mathematics and Its Applications, M. Hazewinkel, Ed. Springer, 1989, vol. 37.

[18] N. Srinivas, A. Krause, S. M. Kakade, and M. Seeger, "Gaussian process optimization in the bandit setting: No regret and experimental design," in *Proc. of the International Conference on Machine Learning (ICML)*, 2010, pp. 1015–1022.

[19] Y. Sui, A. Gotovos, J. W. Burdick, and A. Krause, "Safe exploration for optimization with Gaussian processes," in *Proc. of the International Conference on Machine Learning (ICML)*, 2015, pp. 997–1005.

[20] F. Berkenkamp, A. P. Schoellig, and A. Krause, "Safe controller optimization for quadrotors with Gaussian processes," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 493–496, (to appear), arXiv:1509.01066 [cs.RO].

[21] F. Berkenkamp, A. Krause, and Angela P. Schoellig, "Bayesian optimization with safety constraints: Safe and automatic parameter tuning in robotics." arXiv, 2016, arXiv:1602.04450 [cs.RO].

[22] A. Christmann and I. Steinwart, *Support Vector Machines*, ser. Information Science and Statistics. Springer, 2008.

[23] F. Berkenkamp, A. Krause, and A. P. Schoellig, "Kernel selection for reliable and data-efficient learning of Gaussian process dyanmic models." arXiv, 2016.

A. Full Exploration Proof

We consider a variant of Algorithm 1 that determines the safe set, $\mathcal{S}_n$, starting from $\mathcal{S}_0$ in Assumption 4, using the Lipschitz constant directly. This allows us to prove exploration guarantees.

First, let us consider the best ROA approximation that we could hope to achieve. We have a probabilistic model of $g_\pi(x)$, which means we cannot expect to explore the full ROA. Instead, we consider learning the true ROA up to some error, ϵ. An algorithm with knowledge up to this error, could use the Lipschitz constant of $\dot{V}$ from Lemma 4 in order to learn which discrete states close to $\mathcal{S}_0$ fulfill the requirement of $\dot{V}(x) < -L\tau$ from Lemma 6. We define a general operator on a set that determines these states,

$$R_{\epsilon,\dot{V}}(\mathcal{S}) = \mathcal{S} \cup \{x \in \mathcal{X}_\tau \,|\, \exists x' \in R_l(\mathcal{S}): \tag{16}$$
$$\dot{V}(x) + \epsilon + L\|x - x'\| < -L\tau\},$$

where $R_l(\mathcal{S})$ is an operator that selects the maximum level set of the Lyapunov function contained in $\mathcal{S}$,

$$R_l(\mathcal{S}) = \mathcal{V}\left(\underset{c>0}{\operatorname{argmax}}\ c, \quad \text{subject to } \mathcal{V}(c) \cap \mathcal{X}_\tau \subseteq \mathcal{S}\right). \tag{17}$$

Thus, (16) adds a state $x \in \mathcal{X}_\tau$ to $\mathcal{S}$ if there exists a state x' in the current level set estimate of the ROA that can be used in order to determine that $\dot{V}(x) < -L\tau$ via the Lipschitz constant. If we apply this operator iteratively, in the limit we reach the maximum ROA that can be determined using knowledge up to ϵ. That is, with $R^i_{\epsilon,\dot{V}}(\mathcal{S}) = R_{\epsilon,\dot{V}}(R^{i-1}_{\epsilon,\dot{V}}(\mathcal{S}))$ and $R^1_{\epsilon,\dot{V}}(\mathcal{S}) = R_{\epsilon,\dot{V}}(\mathcal{S})$, the maximum safe set that could be determined is

$$R_l(\overline{R}_{\epsilon,\dot{V}}(\mathcal{S}_0)), \quad \text{where } \overline{R}_{\epsilon,\dot{V}}(\mathcal{S}) = \lim_{i \to \infty} R^i_{\epsilon,\dot{V}}(\mathcal{S}). \tag{18}$$

With the baseline defined, in the following we provide an algorithm that achieves this baseline using the results from Sec. IV. Instead of defining the ROA directly in terms of the GP confidence interval of $\dot{V}$, $\mathcal{Q}_{n,\dot{V}}(x) = [l_n(x), u_n(x)]$ as in Theorem 1, we consider the intersection of these intervals, $\mathcal{C}_n(x) = \mathcal{Q}_{n,\dot{V}}(x) \cap \mathcal{C}_{n-1}(x)$. We initialize these intervals so that points in $\mathcal{S}_0$ are safe; that is, $\mathcal{C}_0(x) = [-\infty, 0)$ if $x \in \mathcal{S}_0$ and $[-\infty, \infty]$ otherwise. As a consequence of Lemma 1, we have that $\dot{V}(x) \in \mathcal{C}_n(x)$ for all $n \geq 1$ with probability at least $(1 - \delta)$. Besides ensuring that points in $\mathcal{S}_0$ are always considered safe, this definition guarantees that the estimated safe set will not decrease over time. We define $l_{n,c}(x) = \min_{x \in \mathcal{X}} \mathcal{C}_n(x)$ and $u_{n,c}(x) = \max_{x \in \mathcal{X}} \mathcal{C}_n(x)$.

Based on these confidence intervals, after $(n - 1)$ measurements we can quantify the following data points as having a value $\dot{V}(x)$ that is sufficiently negative,

$$\mathcal{S}_{n,\dot{V}} = \mathcal{S}_0 \cup \{x \in \mathcal{X} \,|\, \exists x' \in R_l(S_{n-1,\dot{V}}): \tag{19}$$
$$u_{n,c}(x') + L\|x - x'\| < -L\tau\},$$

and estimate the ROA based on this set as

$$\mathcal{S}_n = R_l(\mathcal{S}_{n,\dot{V}}). \tag{20}$$

Intuitively this selection criterion is similar to the one of the baseline in (16), except that (19) uses the GP confidence intervals instead of perfect model knowledge. This allows us to prove that we explore the maximum level set up to ϵ accuracy:

Theorem 2. *Under the assumptions of Theorem 1, choose β_n as in Lemma 1 and let n^* be the smallest positive integer so that*

$$\frac{n^*}{\beta_{n^*}\gamma_{n^*}} \geq \frac{CL^2_{\partial V}(|\overline{R}_{0,\dot{V}}(\mathcal{S}_0)| + 1)}{\epsilon^2}, \tag{21}$$

where $C = 8/\log(1 + \sigma^{-2})$. For any $\epsilon > 0$, and $\delta \in (0, 1)$, under the selection criterion (14) *within $\mathcal{S}_n$ from* (20), *the following holds jointly with probability at least $(1 - \delta)$:*

(i) $\mathcal{S}_n$ is contained in the ROA of (1) *for all $n \geq 1$,*
(ii) $R_l(\overline{R}_{\epsilon,\dot{V}}(\mathcal{S}_0)) \subseteq \mathcal{S}_{n^*} \subseteq R_l(\overline{R}_{0,\dot{V}}(\mathcal{S}_0))$.

Proof. Statement *(i)* follows from Theorem 1. For *(ii)*, we have from [19, Lemma 5] that if the safe set does not increase, then the maximum uncertainty of g_π and $\dot{V}$ within the safe set is bounded by $\epsilon/\|\partial V/\partial x\|_\infty = \epsilon/L_{\partial V}$ and ϵ, respectively, after a finite number of iterations. At that point, either the safe set increases similarly to the baseline, or we have explored the full safe set [19, Lemma 7]. Applying this $\left(|\overline{R}_{0,\dot{V}}(\mathcal{S}_0)| + 1\right)$ times provides the result [19, Cor. 4]. $\square$

That is, after a finite number of evaluations, n^*, we explore at least as much as the baseline up to accuracy ϵ, but not more than we could determine as safe if we knew the function perfectly; that is, the baseline with $\epsilon = 0$.