2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Data-Driven Permissible Safe Control with Barrier Certificates

Rayan Mazouz*[1], John Skovbekk*[1], Frederik Baymler Mathiesen[2], Eric Frew[1], Luca Laurenti[2], and Morteza Lahijanian[1]

***Abstract*— This paper introduces a method of identifying a maximal set of safe strategies from data for stochastic systems with unknown dynamics using barrier certificates. The first step is learning the dynamics of the system via Gaussian Process (GP) regression and obtaining probabilistic errors for this estimate. Then, we develop an algorithm for constructing piecewise stochastic barrier functions to find a maximal permissible strategy set using the learned GP model, which is based on sequentially pruning the worst controls until a maximal set is identified. The permissible strategies are guaranteed to maintain probabilistic safety for the true system. This is especially important for learned systems, because a rich strategy space enables additional data collection and complex behaviors while remaining safe. Case studies on linear and nonlinear systems demonstrate that increasing the size of the dataset for learning grows the permissible strategy set.**

I. Introduction

In an era defined by the increasing integration of Artificial Intelligence (AI) into systems, ensuring the safety of stochastic systems with black-box components has become a major concern. These systems, characterized by uncertain and unpredictable dynamics, are ubiquitous across various domains, from autonomous vehicles [1] to surgical robotics [2]. For such systems, the notion of the *safety invariant set*, which represents regions of the system's state space where non-myopic safety constraints are guaranteed to be satisfied, emerges as a fundamental concept in providing safety guarantees. Particularly, identifying the set of *permissible control strategies*, within these invariant sets holds critical importance, as it provides a comprehensive understanding of the system's operational boundaries and enables complex behaviors while guaranteeing safety (via *shielding*). However, finding such a strategy set poses a major challenge, especially if the system is unknown (due to black-box components). In this work, we aim to provide a data-driven method for identifying a maximal set of permissible strategies that guarantee an unknown system remains inside a safe set.

Our approach is based on stochastic barrier functions (SBFs) [3], [4]. These functions provide a systematic method to bounding the system's behavior within a safe set, even in the presence of uncertainties or disturbances. We also utilize Gaussian process (GP) regression to learn the unknown dynamics, which enable probabilistic bounding of the learning error [5]–[7]. Our key insight in dealing with the computational challenge of identifying permissible strategies is that local treatment of the uncertainty (namely, stochasticity and learning error) is needed, and a particular form of functions that enables such a treatment is piecewice (PW). Hence, we employ the recent results in PW-SBFs [8] to reason about subsets of controls. Particularly, we show a formulation of PW-SBFs that allow admissibility assessment of local control subsets for each compact set of states via linear programming, achieving computational efficiency. Based on these results, we propose an algorithm that, given a set of input-output data on the dynamics, safe and initial sets, and a lower-bound on the safety probability, returns a maximal permissible strategies set by iteratively removing inadmissible control sets until a fixed-point is reached. We show soundness of the algorithm, validate the theoretical guarantees and illustrate that the approach works on unknown systems with linear and nonlinear stochastic dynamics.

In short, this work makes the following contributions:

- a data-driven framework for computation of a maximal set of permissible strategies,
- extension of SBFs to provide safety control invariant sets for GP regressed models on continuous control sets,
- a series of case studies on both linear and nonlinear systems that demonstrate the efficacy of the method and validate the theoretical guarantees.

A. Related Work

Probabilistic safety invariance is an essential property to enforce the safety of stochastic systems [4], [9]–[11]. For Markovian systems, these invariant sets can be found using linear programs (LPs) and mixed integer LPs for finite and infinite horizon probabilistic invariance, respectively [10]. The sound application of these program requires knowledge of the transition kernel, which is not available for unknown systems as in our setting. While a sampling-based procedure for probabilistic invariant sets of deterministic, stable black-box systems is available [9], it does not admit controlled stochastic systems that may be inherently unstable.

GP regression is notable for its flexibility in learning unknown systems and quantifying the uncertainty in safety-critical settings [5], [11]–[13]. In addition to learning and estimating safe and stable regions for unknown systems [5], [12], they have been used to learn state-dependent modelling uncertainties for LTI control systems to subsequently find probabilistic invariant sets under fixed feedback control [11]. However, none of the previous can identify the full set of controls that leads to safety invariance over a horizon.

This work was supported in part by NSF grant 2039062.

*Equal contributions

[1]Authors are with the University of Colorado Boulder, USA {firstname.lastname}@colorado.edu

[2]Authors are with Delft University of Technology, Netherlands {f.b.mathiesen, l.laurenti}@tudelft.nl

Stochastic Barriers Functions (SBFs) can provide bounds on the probability of remaining safe from an initial set [3], [4], even in the case a system is learned from data [14], [15]. Synthesis based on GP regression over states and controls can provide a single strategy with a lower-bound of safety [15]. In addition, SBFs can be used to find invariant sets for uncontrolled systems with bounded disturbances over an infinite horizon [4]. Commonly, SBF synthesis relies on SOS optimization [3], [16] or neural barrier training [17], [18]. While both methods are deemed useful, they each have limitations. A novel and more efficient formulation relies on piecewise (PW) SBFs [8]. In this work, we adapt and extend PW-SBFs to enable computation of permissible controls.

II. Problem Formulation

Consider a stochastic dynamical system with dynamics

$$\begin{aligned} \mathbf{x}_{k+1} &= f(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}, \\ k \in \mathbb{N}, \quad \mathbf{x}_k &\in \mathbb{R}^n, \quad \mathbf{u}_k \in U \subset \mathbb{R}^m, \end{aligned} \tag{1}$$

where U is a compact control set, $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ the vector field, and $\mathbf{w}$ represents noise taking values in $W \subseteq \mathbb{R}^n$. Noise $\mathbf{w}$ is assumed to be an independent and identically distributed (i.i.d.) sub-Gaussian random variable with known probability density function $p_{\mathbf{w}}$. The choice of control $\mathbf{u}_k \in U$ is performed by a stationary (feedback) *control strategy* $\pi : \mathbb{R}^n \to U$. We denote the set of all strategies by Π.

In this paper, we assume that vector field f is *unknown*. Instead, we consider a given dataset $D = \{(x_i, u_i, x'_i)\}_{i=1}^M$, consisting of M input-output i.i.d. samples from System (1), such that $x'_i = f(x_i, u_i) + w_i$, where w_i is a realization of noise $\mathbf{w}$. We aim to learn f using D. To that end, we impose the following assumption, which ensures f is a well-behaved analytical function on a compact set.

Assumption 1 (RKHS Continuity)**.** *Let a compact set $Y \subset \mathbb{R}^n \times U$ and $\kappa : \mathbb{R}^{n+m} \times \mathbb{R}^{n+m} \to \mathbb{R}$ be a given kernel. Define $H_\kappa(Y)$ as the reproducing kernel Hilbert space (RKHS) of functions over Y corresponding to κ with norm $\|\cdot\|_\kappa$ [19]. Then, for each $i \in \{1, \ldots, n\}$, $f^i \in H_\kappa(Y)$ and a constant $C_i > 0$, $\|f^i\|_\kappa \leq C_i$, where f^i is the i-th component of f.*

Assumption 1 is standard in the literature, asserting that f can be learned via GPs with an appropriate kernel. For instance, when κ is a universal kernel, e.g. squared exponential kernel, it includes a class of functions that is dense w.r.t. the continuous functions over a compact set [20], [21].

Given sources of uncertainty, we focus on probabilistic analysis of System (1). We define a probability measure over the trajectories of the system. For a measurable set $X \subset \mathbb{R}^n$, the one-step *transition kernel* under a strategy $\pi \in \Pi$ is

$$T(X \mid x, \pi) := \int_{\mathbb{R}^w} \mathbf{1}_X(f(x, \pi(x)) + w) p_{\mathbf{w}}(w) dw, \tag{2}$$

where $\mathbf{1}_X$ is the indicator function for X such that $\mathbf{1}_X(x) = 1$ if $x \in X$ and $\mathbf{1}_X(x) = 0$ if $x \notin X$ [22]. Then, for a time horizon $N \in \mathbb{N}$, a strategy π, and an initial condition $x_0 \in \mathbb{R}^n$, we consider the probability measure $\Pr$ over the trajectories of System (1) such that for $X_0, X_k \subseteq X$

$$\begin{aligned} &\Pr[\mathbf{x}_0 \in X_0] = \mathbf{1}_{X_0}(x_0), \\ &\Pr[\mathbf{x}_k \in X_k \mid \mathbf{x}_{k-1} = x, \pi] = T(X_k \mid x, \pi). \end{aligned} \tag{3}$$

Definition 1 (Probabilistic Safety)**.** *Given a measurable safe set $X_s \subset \mathbb{R}^n$ and initial set $X_0 \subseteq X_s$, the* probabilistic safety *of System* (1) *under a control strategy π for N time steps is*

$$P_s(X_s, X_0, N, \pi) = \inf_{x_0 \in X_0} \Pr[\mathbf{x}_k \in X_s \ \forall k \leq N \mid \mathbf{x}_0 = x_0, \pi].$$

It is enough to find a single strategy that satisfies a threshold on P_s to claim that a system is safe. However, identifying multiple strategies that guarantee a minimum P_s can unlock more complex behaviors for the system. This is especially important for learning-enabled systems (e.g., the one considered in this work) because under such strategies more data can be collected in a safe manner to improve the learning model. To this end, we define permissible strategies.

Definition 2 (Permissible strategy)**.** *Given a safety threshold p, control strategy $\pi \in \Pi$ is called* permissible *if $P_s(X_s, X_0, N, \pi) \geq p$.*

The goal is to find a maximal set of permissible strategies.

Problem 1. *Given dataset D obtained from System* (1)*, safe set $X_s \subset \mathbb{R}^n$, initial set $X_0 \subseteq X_s$, horizon N, and safety threshold p, find a maximal set of strategies $\Pi_s \subseteq \Pi$ such that every $\pi \in \Pi_s$ is permissible, i.e., $P_s(X_s, X_0, N, \pi) \geq p$.*

Problem 1 poses several challenges. Firstly, it seeks a maximal set of permissible strategies for an unknown System (1), which itself is a stochastic process. Secondly, control set $U \subset \mathbb{R}^m$ is continuous, making admissibility analysis of π difficult due to its range being an uncountable, infinite set. Furthermore, set X_s can be non-convex, possibly leading to an overall non-convex optimization problem.

Approach overview. Our approach to Problem 1 is based on GP regression and SBFs. GP regression allows learning of the dynamics with formal learning-error quantification, and SFBs provide a formal methodology to prove invariance of stochastic systems. To address issues pertaining to continuity of U and non-convexity of X_s, we utilize the recently-developed techniques for *piecewise constant* (PWC) SBFs [8]. We show that based on these functions, an inner-approximation of the set of permissible strategies can be obtained for System (1).

III. Preliminaries

In this section, we provide a brief background on GP regression and PWC-SBFs, which are core to our framework.

A. Gaussian Process Regression

A GP is a collection of random variables X, any finite subset of which are jointly Gaussian [21]. We describe GPs of single-dimension output, noting that multidimensional outputs are similar. GPs are often interpreted as distributions over a function space, which are defined completely by a mean function $\hat{f} : \mathrm{X} \to \mathbb{R}$ and covariance (kernel) function

$\kappa : \mathrm{X} \times \mathrm{X} \to \mathbb{R}$. GP regression is a Bayesian approach to conditioning a prior GP model $(\hat{f}_0, \kappa_0)$ on a dataset $D = (\mathrm{X}, \mathrm{Y})$ to find a posterior model $(\hat{f}_D, \kappa_D)$ defined as

$$\hat{f}_D(x) = \hat{f}_0(x) + \kappa_0(x, \mathrm{X})K^{-1}(\mathrm{Y} - \hat{f}_0(\mathrm{X})),$$
$$\kappa_D(x, x') = \kappa_0(x, x') - \kappa_0(x, \mathrm{X})K^{-1}\kappa_0(\mathrm{X}, x'),$$

where $\kappa_0(\mathrm{X}, x) = [\kappa_0(\mathrm{X}_i, x), \ldots, \kappa_0(\mathrm{X}_M, x)]^T$ (similarly defined vectors of both arguments), $K = \kappa_0(\mathrm{X}, \mathrm{X}) + \sigma^2 I$, and σ^2 is the variance on observations Y. This posterior is a distribution of functions that best describes data D. Identical input points to the kernel yield $\kappa_D^{-1/2}(x, x) = \sigma_D(x)$.

B. Piecewise Constant Stochastic Barrier Functions

Consider a stochastic process $\mathbf{z}_{k+1} = F(\mathbf{z}_k, \mathbf{v})$, where state $\mathbf{z}_k \in \mathbb{R}^{n_z}$, noise $\mathbf{v} \in \mathbb{R}^v$ is a random variable with distribution $p_{\mathbf{v}}$, and vector field $F : \mathbb{R}^{n_z} \times \mathbb{R}^v \to \mathbb{R}^{n_z}$ is almost everywhere continuous. For a measurable compact set $Z \subset \mathbb{R}^{n_z}$, let $T_z(Z \mid z)$ and Pr be the stochastic transition kernel and probability measure for process $\mathbf{z}_k$ defined similarly to (2) and (3), respectively.

Further, consider partition $Z_1, \ldots, Z_{K_z}$ of Z such that F is continuous over each Z_i, and for scalars $b_i \in \mathbb{R}_{\geq 0}$ with $i = \{1, ..., K_z\}$, define PWC function $B : \mathbb{R}^{n_z} \to \mathbb{R}_{\geq 0}$ as

$$B(z) = \begin{cases} b_i & \text{if } z \in Z_i, \\ 1 & \text{otherwise.} \end{cases} \tag{4}$$

The following theorem establishes the conditions under which B is a *PWC-SBF* for process $\mathbf{z}$.

Theorem 1 ([8, Theorem 2]). *PWC function $B(z)$ in Eq. (4) is a SBF for process $\mathbf{z}$ with safe sets $Z \subset \mathbb{R}^{n_z}$, initial set $Z_0 \subseteq Z$, and partition $Z_1, \ldots, Z_{k_z}$ of Z, if $\exists \eta, \beta \geq 0$ s.t.*

$$b_i \leq \eta \qquad \forall i \in \{1, ..., k_z\} : Z_i \cap Z_0 \neq \emptyset, \tag{5a}$$
$$\sum_{j=1}^{K_z} b_j \cdot T_z(Z_j \mid z) + T_z(\mathbb{R}^{n_z} \setminus Z \mid z) \leq b_i + \beta \qquad \forall i \in \{1, ..., k_z\}, \forall z \in Z_i. \tag{5b}$$

Then, for $N \in \mathbb{N}_{\geq 0}$, it follows that $\Pr[\forall k \leq N,\ \mathbf{z}_k \in Z \mid \mathbf{z}_0 \in Z_0] \geq 1 - (\eta + \beta N)$.

The advantage of SBFs is that, by satisfying static conditions, they allow for probabilistic reasoning about stochastic dynamical systems without having to unfold their trajectories.

IV. PWC-SBFs for Invariance via GP Regression

In this section, we introduce our main approach. First, we show that how GP regression can be used to bound transition kernel T. Then, we propose a method to use the kernel bounds with Theorem 1 to identify permissible strategy sets with PWC-SBFs and provide soundness guarantees.

A. Learning Transition Kernel Bounds

a) Learning f: To begin, we learn function f in System (1) using GP regression on the state and control spaces using dataset D. For brevity, let

$$z = (x, u) \quad \text{with} \quad z \in Z = \mathbb{R}^n \times U.$$

Then, we estimate each component of f with a GP map $\hat{f}_D^i : Z \to \mathbb{R}$ for $i = \{1, \ldots, n\}$. Assumption 1 allows for uniform probabilistic error bounds of $\hat{f}_D^i$ [6]. Namely, for scalar $\epsilon^i \geq 0$ for $1 \leq i \leq n$ and compact set $\tilde{Z} \subset Z$,

$$\Pr\left[|f^i(z) - \hat{f}_D^i(z)| \leq \epsilon^i\right] \geq 1 - \delta \qquad \forall z \in \tilde{Z}, \tag{6}$$

where δ satisfies $\epsilon^i = \alpha(\delta)\sigma_D^i(z)$ as defined in [6, Theorem 2]. The term $\alpha(\delta)$ scales the posterior according to the desired probability threshold and the RKHS properties of the function f, including its RKHS norm that can be bounded using its Lipschitz continuity [13]. This probabilistic error bound is the foundation for computing bounds on kernel T.

b) Bounding the Transition Kernel: Directly computing kernel T for every state in a region in Theorem 1 is computationally intractable. Instead, we compute bounds of T using $\hat{f}_D^i$, employing techniques from [7], [13]. Consider a partition $X_1, \ldots, X_K$ of state X_s in K compact sets, and partition $U_1, \ldots, U_L$ of control space U in L compact sets

$$\cup_{i=1}^K X_i = X_\mathrm{s} \quad \text{and} \quad \cup_{l=1}^L U_l = U$$

all of which are non-overlapping. The product of X_i and U_l yields a partition of the state-control space, i.e., $Z_{il} = X_i \times U_l$. Then, the transition kernel from a state-control partition Z_{il} to a compact set $X_j \subset \mathbb{R}^n$ of states can be bounded by finding the extrema of the transition kernel,

$$\underline{p}_{ij}^l \leq \min_{(x,u) \in Z_{il}} T(X_j \mid x, u), \ \ \overline{p}_{ij}^l \geq \max_{(x,u) \in Z_{il}} T(X_j \mid x, u). \tag{7}$$

We partition the domains of uncertainty, i.e., the additive noise W and the learning error $\mathbb{R}_{\geq 0}^n$ in (6) [7]. This method seeks all the uncertainty values to guarantee entering (avoiding) the target set to compute a non-trivial lower (upper) extrema bounds. Let $Q = W \times \mathbb{R}_{\geq 0}^n$ be the product space of the respective domains, and let $\mathcal{C} = \{C_1, \ldots, C_p\}$ be non-overlapping partitions of Q. Also, let the learned image, $Post(Z_{il}, C)$, be all of the points in Z_{il} propagated through $\hat{f}_D$ with all uncertainties in $C \in \mathcal{C}$ accounted for, i.e.,

$$Post(Z_{il}, C) = \{\hat{f}_D(z) + c \mid z \in Z_{il}, c \in C\}.$$

Proposition 1 bounds kernel T for an uncertainty partition $\mathcal{C}$.

Proposition 1 ([7, Theorem 1]). *Let $\mathcal{C}$ be a partition of the product space $Q = W \times \mathbb{R}_{\geq 0}^n$ of System (1), and Z_{il} be a compact subset of Z. Then, the one-step transition kernel to a compact set $X \subset \mathbb{R}^n$ from Z_{il} is bounded by*

$$\min_{x,u \in Z_{il}} T(X \mid x, u) \geq \sum_{C \in \mathcal{C}} \mathbf{1}(Post(Z_{il}, C) \subseteq X) \Pr(C),$$
$$\max_{x,u \in Z_{il}} T(X \mid x, u) \leq \sum_{C \in \mathcal{C}} \mathbf{1}(Post(Z_{il}, C) \cap X = \emptyset) \Pr(C),$$

where $\mathbf{1}(\cdot)$ returns 1 if the argument is true and 0 otherwise.

Remark 1. *As the uncertainty structure of System* (1) *is additive, an optimal partition $\mathcal{C}$ can be found that gives rise to the tightest bounds of T. This requires only three partitions of the uncertainty space in each dimension [7].*

B. Permissible Strategy Set Synthesis

Using PWC-SBFs and the bounds on T, we extend Theorem 1 to find the maximal permissible strategy set Π_s for System (1). Consider the lower and upper bounds on the transition kernel $\underline{p}^l_{ij}$ and $\overline{p}^l_{ij}$ in (7). The set of all feasible values for $T(X_j \mid x, u)$, $\forall x \in X_i, \forall u \in U_l$ is

$$\mathcal{P}^l_i = \Big\{ p^l_i = (p^l_{i1}, \dots, p^l_{iK}, p^l_{i\mathrm{u}}) \in [0,1]^{K+1} \quad s.t. \sum_{j=1}^{K} p^l_{ij} + p^l_{i\mathrm{u}} = 1, \; \underline{p}^l_{ij} \le p^l_{ij} \le \overline{p}^l_{ij} \quad \forall j \in \{1, \dots, K, \mathrm{u}\} \Big\}. \tag{8}$$

The following theorem sets up an optimization problem for synthesizing permissible strategy set Π_s using PWC-SBF.

Theorem 2 (PWC-SBF for Permissible Strategies)**.** *Consider System* (1) *with safe set $X_\mathrm{s} \subset \mathbb{R}^n$, initial set $X_0 \subseteq X_\mathrm{s}$, dataset D, and safety bound p. Given K-partitions of X_s, let the set of PWC functions $\mathcal{B}_K$ be in the form* (4)*, with $B_i(x) = b_i \in \mathbb{R}_{\ge 0}$, for every $i \in \{1, \dots, K\}$. Further, given L-partitions of U, let $\mathcal{P}^l = \prod_{i=1}^{K} \mathcal{P}^l_i$, where $\mathcal{P}^l_i$ is the set of feasible transition kernel in* (8)*, with the bounds computed by Proposition 1, for every $l \in \{1, \dots, L\}$. Finally, let $B^* \in \mathcal{B}_K$ be a solution to the following optimization problem*

$$B^* = \arg \min_{B \in \mathcal{B}_K} \max_{\{\mathcal{P}^l\}_{l=1}^{L}} \eta + N\beta \tag{9}$$

subject to

$$b_i \le \eta \qquad \forall i : X_i \cap X_0 \neq \emptyset, \tag{10a}$$

$$\sum_{j=1}^{K} b_j \cdot p^l_{ij} + p^l_{i\mathrm{u}} \le b_i + \beta^l_i \quad \forall i, \; \forall l, \; \forall p^l_i \in \mathcal{P}^l_i, \tag{10b}$$

$$0 \le \beta^l_i \le \beta, \qquad \forall i, \; \forall l. \tag{10c}$$

Then,

1) *B^* constitutes a stochastic barrier certificate for System* (1) *that guarantees safety probability $P_s(X_s, X_0, N, \pi) \ge 1 - (\eta + \beta N)$, $\forall \pi \in \Pi$;*
2) *if $\beta^l_i \le (1 - \eta - p)/N$, then every choice of $u \in U_l$ is permissible for every $x \in X_i$ of System* (1)*.*

Proof Sketch. The optimization is an exact solution to the problem in Theorem 1. It is straightforward that B^* is a proper barrier certificate, and that for every strategy $\pi \in \Pi$, $P_s(X_s, X_0, N, \pi) \ge 1 - (\eta + \beta N)$, where $\beta = \max \beta^l_i$. Further, it follows that if $\beta^l_i \le (1 - \eta - p)/N$, state-control partition Z_{il} is rendered safe $\forall (x, u) \in X_i \times U_l$. Since the transition kernel bounds of the learned system in (8) are contained in the true system distribution [13, Theorem 1], the probabilistic safety guarantees hold for System (1). □

The above theorem provides a method of identifying permissible strategies, which requires solving a minimax optimization problem. This problem includes bilinear terms, i.e., $b_j \cdot p^l_{ij}$ in Condition (10b). Nevertheless, by introducing dual variables, the minimax problem becomes a simple linear program (LP) as shown in [8], which guarantees efficiency. Furthermore, following the same approach developed in [8], the resulting LP can be efficiently solved by a Counter Example Guided Synthesis (CEGS) approach or a gradient descent method. Our evaluations in Section V use CEGS.

Algorithm 1: PWC-SBF based Control Invariant Sets

Input : Initial set X_0, state-control space $\mathcal{Z}$, time horizon N, feasible transition kernel sets $\tilde{\mathcal{P}} = \{\mathcal{P}^l\}_{l=1}^{L}$, and probability threshold p.
Output: Permissible strategy set Π_s.

1 $\eta, \beta \leftarrow$ BARRIER$(X_0, \mathcal{Z}, N, \tilde{\mathcal{P}})$ ▷ Theorem 2
2 **while** $1 - (\eta + N\beta) < p$ **do**
3 $\quad Z'_{il}, \mathcal{P}^l_i \leftarrow$ Identify region with $\max \beta^l_i$
4 $\quad \mathcal{Z} \leftarrow \mathcal{Z} \setminus Z'_{il}$ ▷ Remove worst region
5 $\quad \tilde{\mathcal{P}} \leftarrow \tilde{\mathcal{P}} \setminus \mathcal{P}^l_i$ ▷ Remove kernel bounds
6 $\quad \eta, \beta \leftarrow$ BARRIER$(X_0, \mathcal{Z}, N, \tilde{\mathcal{P}})$ ▷ Theorem 2
7 $\quad$ **if** *all control sets removed for X_i* **then**
8 $\quad\quad$ **return** False
9 $\quad \Pi_s \leftarrow \{X_i \to U_l \mid \forall (X_i, U_l) \in \mathcal{Z}\}$
10 **return** Π_s

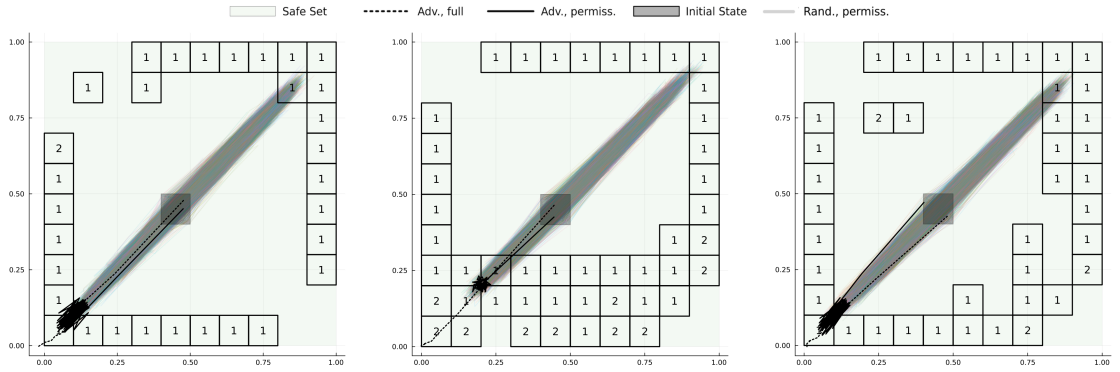
We propose an algorithm that computes a maximal set of permissible strategies $\Pi_s \subseteq \Pi$ w.r.t. to partitions of the safe set and control space. An overview is shown in Alg. 1, taking as input the initial set X_0, time horizon N, set $\mathcal{Z} = \{Z_{il} = X_i \times U_l \mid 1 \le i \le K, \; 1 \le l \le L\}$, the bounds on the transition kernel T, and the desired safety threshold p. On Line 1, the algorithm computes an initial PWC-SBF using Theorem 2. Next, the while loop is instantiated, with a termination condition that probability criteria p must be met.

Using the PWC-SBF, the worst state-control pair Z'_{il} corresponding to $\max \beta^l_i$, along with the probability bounds P^l_i, are identified on Line 3. These are subsequently removed from their respective sets on Lines 4 and 5. Intuitively, this means that the algorithm sequentially prunes control region U_l that is deemed not permissible for a given state partition X_i, corresponding to $\max \beta^l_i$. Then, a new PWC-SBF is synthesized on Line (6) using Theorem 2. A check procedure on Line (7) is incorporated to make sure that for all regions X_i, there exists at least one control region U_l inside updated space $\mathcal{Z}$. If the latter is true, the set of permissible strategies Π_s is updated accordingly on Line (9). Else, the algorithm terminates as `False`. A successful algorithm termination (Line 10) implies that the probability threshold p is met.

Corollary 1. *Alg. 1 terminates in finite time. If it terminates with the return statement on Line 10, then it returns Π_s, which is guaranteed to include only permissible strategies.*

The proof is a direct implication of Theorem 2.

Remark 2. *Alg. 1 can also compute the probabilistic control invariant set. This is established by running the algorithm for every $X_i \in X_\mathrm{s}$, $i \in \{1, \dots, K\}$ as the initial set.*



(a) Known: $P_{\mathrm{s}} \geq 1 - 5.83 \times 10^{-5}$ (b) GP 500 - $P_{\mathrm{s}} \geq 1 - 2.39 \times 10^{-4}$ (c) GP 2000 - $P_{\mathrm{s}} \geq 1 - 1.93 \times 10^{-5}$

Fig. 1: Linear system case studies using the known system and two datasets. Regions with # of removed control regions and simulated trajectories over 100 steps are shown in each plot. Adversarial trajectories show permissible sets maintain safety.

V. Case Studies

We demonstrate the efficacy of Alg. 1 by finding permissible strategy sets for linear and nonlinear systems learned from data. We perform GP regression to estimate the system and then find the permissible sets. Each GP uses zero mean and squared exponential kernel priors. To determine the quality of our solution, we compute the permissible strategy set of the known linear system on the same partitions.

In all experiments, the initial state region of interest is $X_0 = [0.4, 0.5]^2$, the safe set is $X_{\mathrm{s}} = [0.0, 1.0]^2$, and the noise $\mathbf{w}$ is Gaussian zero-mean with variance 0.01^2 and zero covariance in each dimension. Permissible sets are constructed with Alg. 1 using a convergence threshold of $p = 1 - 10^{-4}$. Then, this set is used to generate bounds on the probability of safety for $N = 100$ time step trajectories.

The safety probability bounds are validated by Monte Carlo simulations randomly initialized in X_0 and propagated by random sampling of the controls from the permissible set Π_s. To further validate the guarantees of the approach, adversarial trajectories are generated by choosing control inputs that drive the system towards exiting the safe set X_{s} from both the full control set and the permissible set.

A. Linear System

First, consider the 2D system $\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k + \mathbf{w}_k$, where $A = 0.5I$, $B = (1, 1)^T$, and $u \in [0.0, 0.5]$. From the initial set, this system primarily traverses around the diagonal $x_1 = x_2$ line. The safe set is discretized using a uniform grid with length 0.1, and the control range is partitioned into five intervals. Computing each permissible set takes approximately 70-100 seconds for this system.

Fig. 1 shows the safe and the initial sets, and the final results for the known and learned systems. Every cell indicates the number of removed control sets. In each case, we simulated 1000 trajectories using randomly sampled permissible strategies, resulting in zero violations. The trajectories are shown in gray color. Observe that, while the system can reach towards the edges of the safe set, the permissible strategy set ensures the system remains inside the safe set.

To further demonstrate the efficacy of the permissible strategy set, we compare trajectories that choose the most adversarial actions according to the full control set (shown in dashed black) and the permissible strategy set (shown in solid black). Note that without restricting the strategies, the system can always be driven outside of the safe set.

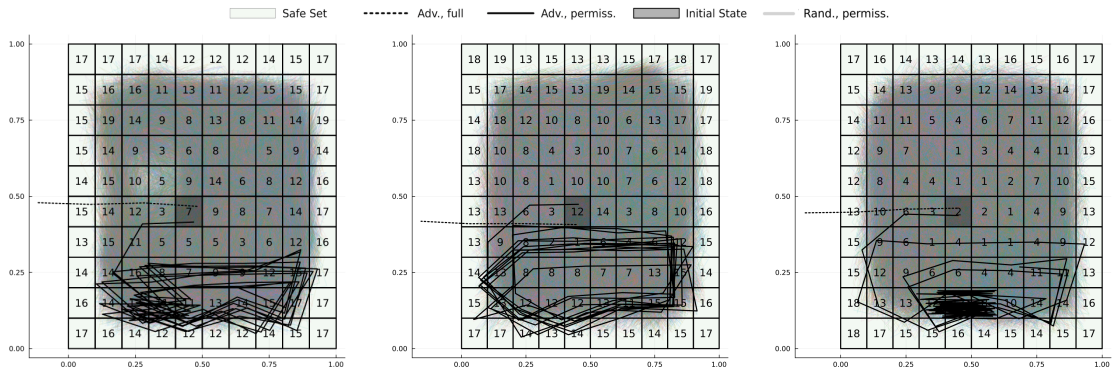
Figs. 1b and 1c show the result using a GPs with 500 and 2000 datapoints, respectively. For the smaller dataset, more control partitions are removed from the permissible set, resulting in trajectories that cover a smaller range of the safe set. The 2000 datapoints model results in an permissible set that more resembles the known system, with fewer removed control partitions. The probabilities of exiting the safe set in 100 time-steps is very small, and the learned system with 2000 datapoints has a tighter upper bound than the known system shown in Fig. 1a. This is likely due to convergence-related hyperparameters in the optimization procedure.

The richness of the permissible set is indicated by the proportion of the control space it contains. The set for a known system encompasses 93.6% of the system's total possible actions, serving as a benchmark. With 500 datapoints, the procedure recognizes 88.8% of the entire action space, while incorporating up to 2000 datapoints enhances this figure to 91.2%. This indicates that with more data, the permissible set progressively approximates that of the known system.

B. Nonlinear System

Consider the nonlinear dynamical system $\mathbf{x}_{k+1} = \mathbf{x}_k + 0.2 \cdot (\cos(\mathbf{u}_k),\ \sin(\mathbf{u}_k))^T + \mathbf{w}_k,\ u \in [-\pi, \pi]$. This is a simplified Dubin's car model, where the heading is set directly, allowing movement in any direction. The state-space partitioning is the same as above. The control range is partitioned into 20 intervals. Computing the permissible set takes ≈1-1.5 hours for each dataset.

Fig. 2 illustrates the comparison of permissible strategy sets obtained for the system using datasets of 500, 1000, and 1500 datapoints. In constructing the permissible sets, the algorithm removed numerous control intervals that would lead the system to approach the safe set boundary. Despite



(a) GP 500 - $P_{\mathrm{s}} \geq 1 - 8.96 \times 10^{-3}$ (b) GP 1000 - $P_{\mathrm{s}} \geq 1 - 9.10 \times 10^{-3}$ (c) GP 1500 - $P_{\mathrm{s}} \geq 1 - 9.95 \times 10^{-3}$

Fig. 2: Nonlinear system case studies using various sizes of datasets. Regions with # of removed control regions and simulated trajectories over 100 steps are shown in each plot. Adversarial trajectories show permissible control sets maintain safety.

the variations in dataset size, each permissible set constrains the probability of exiting the safe set within a similar bound.

The resulting permissible sets for the system learned with 500 and 1000 datapoints maintained 39.5% and 40.1% of all available controls, respectively. Adding 500 datapoints increased the available controls to 49.5%. Out of the 1000 sampled trajectories none exited the safe set, and only the adversarial trajectory using the full control set violated safety. The strict bound on β during permissible set synthesis removed a majority of the controls to ensure safety. Relaxing the condition could add more strategies to the permissible set.

VI. Conclusion

In this work, we introduce a data-driven framework for the computation of a maximal set of permissible strategies for ensuring the safety of unknown stochastic systems. The unknown model with a continuous control space is regressed from data using Gaussian processes. Then, the regressed model is used to synthesize a piecewise stochastic barrier function, which in turn identifies theoretically-sound permissible control sets. The safety guarantees are validated in case studies of linear and nonlinear systems. Future work includes the incorporation of other regression approaches, and applications to online safe exploration and learning.

References

[1] S. Shalev-Shwartz, S. Shammah, and A. Shashua, "On a Formal Model of Safe and Scalable Self-driving Cars," *arXiv preprint arXiv:1708.06374*, 2017.

[2] J. Guiochet, M. Machin, and H. Waeselynck, "Safety-critical Advanced Robots: A survey," *Robotics and Autonomous Systems*, vol. 94, pp. 43–52, 2017.

[3] C. Santoyo, M. Dutreix, and S. Coogan, "A barrier function approach to finite-time stochastic system verification and control," *Automatica*, vol. 125, p. 109439, 2021.

[4] Y. Yu, T. Wu, B. Xia, J. Wang, and B. Xue, "Safe probabilistic invariance verification for stochastic discrete-time dynamical systems," in *2023 62nd IEEE Conference on Decision and Control (CDC)*. IEEE, 2023, pp. 5804–5811.

[5] F. Berkenkamp, M. Turchetta, A. Schoellig, and A. Krause, "Safe model-based reinforcement learning with stability guarantees," *Advances in neural information processing systems*, vol. 30, 2017.

[6] S. R. Chowdhury and A. Gopalan, "On kernelized multi-armed bandits," in *International Conference on Machine Learning*. PMLR, 2017, pp. 844–853.

[7] J. Skovbekk, L. Laurenti, E. Frew, and M. Lahijanian, "Formal abstraction of general stochastic systems via noise partitioning," *IEEE Control Systems Letters*, 2023.

[8] R. Mazouz, F. B. Mathiesen, L. Laurenti, and M. Lahijanian, "Piecewise stochastic barrier functions," *arXiv preprint arXiv:2404.16986*, 2024.

[9] Z. Wang and R. M. Jungers, "Data-driven computation of invariant sets of discrete time-invariant black-box systems," *arXiv preprint arXiv:1907.12075*, 2019.

[10] Y. Gao, K. H. Johansson, and L. Xie, "Computing probabilistic controlled invariant sets," *IEEE Transactions on Automatic Control*, vol. 66, no. 7, pp. 3138–3151, 2020.

[11] P. Griffioen, A. Devonport, and M. Arcak, "Probabilistic invariance for gaussian process state space models," in *Learning for Dynamics and Control Conference*. PMLR, 2023, pp. 458–468.

[12] A. Lederer and S. Hirche, "Local asymptotic stability analysis and region of attraction estimation with gaussian processes," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 1766–1771.

[13] J. Jackson, L. Laurenti, E. Frew, and M. Lahijanian, "Formal verification of unknown dynamical systems via gaussian process regression," *arXiv preprint arXiv:2201.00655*, 2021.

[14] P. Jagtap, S. Soudjani, and M. Zamani, "Formal synthesis of stochastic systems via control barrier certificates," *IEEE Transactions on Automatic Control*, vol. 66, no. 7, pp. 3097–3110, 2020.

[15] R. Wajid, A. U. Awan, and M. Zamani, "Formal synthesis of safety controllers for unknown stochastic control systems using gaussian process learning," in *Learning for Dynamics and Control Conference*. PMLR, 2022, pp. 624–636.

[16] R. Mazouz, K. Muvvala, A. Ratheesh, L. Laurenti, and M. Lahijanian, "Safety Guarantees for Neural Network Dynamic Systems via Stochastic Barrier Functions," *Advances in Neural Information Processing Systems*, 2022.

[17] F. B. Mathiesen, S. C. Calvert, and L. Laurenti, "Safety Certification for Stochastic Systems via Neural Barrier Functions," *IEEE Control Systems Letters*, vol. 7, pp. 973–978, 2022.

[18] C. Dawson, S. Gao, and C. Fan, "Safe control with learned certificates: A survey of neural lyapunov, barrier, and contraction methods for robotics and control," *IEEE Transactions on Robotics*, 2023.

[19] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Information-theoretic regret bounds for gaussian process optimization in the bandit setting," *IEEE transactions on information theory*, vol. 58, no. 5, pp. 3250–3265, 2012.

[20] I. Steinwart, "On the influence of the kernel on the consistency of support vector machines," *Journal of machine learning research*, vol. 2, no. Nov, pp. 67–93, 2001.

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

[22] D. P. Bertsekas and S. Shreve, *Stochastic optimal control: the discrete-time case*. Athena Scientific, 2004.