TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Mathiesen, F. B., Romao, L., Calvert, S. C., Laurenti, L., & Abate, A. (2026). A data-driven approach for safety quantification of non-linear stochastic systems with unknown additive noise distribution. *Automatica*, *193*, Article 113187. https://doi.org/10.1016/j.automatica.2026.113187

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

Automatica 193 (2026) 113187

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

A data-driven approach for safety quantification of non-linear stochastic systems with unknown additive noise distribution ☆

Frederik Baymler Mathiesen [a,*,1], Licio Romao [b,1], Simeon C. Calvert [d], Luca Laurenti [a,c], Alessandro Abate [b]

[a] Delft Center for Systems and Control, TU Delft, Netherlands
[b] Department of Computer Science, University of Oxford, United Kingdom
[c] The Italian Institute of Artificial Intelligence (AI4I), Italy
[d] Department of Transport & Planning, TU Delft, Netherlands

ARTICLE INFO

Article history:
Received 9 October 2024
Received in revised form 5 April 2026
Accepted 1 June 2026

ABSTRACT

In this paper, we present a novel data-driven approach to quantify safety for non-linear, discrete-time stochastic systems with unknown noise distribution. We define safety as the probability that the system remains in a given region of the state space for a given time horizon and, to quantify it, we present an approach based on Stochastic Barrier Functions (SBFs). In particular, we introduce an inner approximation of the stochastic program to design an SBF in terms of a chance-constrained optimisation problem, which allows us to leverage the scenario approach theory to design an SBF from samples of the system with Probably Approximately Correct (PAC) guarantees. Our approach leads to tractable, robust linear programs, which enable us to assert safety for nonlinear models with a confidence that was otherwise deemed infeasible with existing methods. To further mitigate the computational complexity of our approach, we exploit the structure of the system dynamics and rely on spatial data structures to accelerate the construction and solution of the underlying optimisation problem. We show the efficacy and validity of our framework in several benchmarks, showing that our approach can obtain tighter certificates compared to state-of-the-art with a confidence that is several orders of magnitude higher.

© 2026 Published by Elsevier Ltd.

1. Introduction

Safety-critical applications, such as autonomous driving, require provable guarantees of safety, as undesirable behaviours may lead to catastrophic outcomes with long-term economic costs. As a consequence, asserting the safety of complex non-linear noisy systems has been the focus of many recent approaches (Abate, 2017; Cosner et al., 2023; Prajna et al., 2007). However, these approaches generally suffer from exponential complexity with respect to the dimension of the state space. Besides, generally, these approaches require that the distribution of the noise affecting the system is either known and Gaussian or of bounded support (Prajna et al., 2007; Santoyo et al., 2021). Unfortunately, in practice, the noise characteristics of the system are often not known (Abate, 2017; Badings, Romao, Abate, Parker et al., 2023; Salamati et al., 2021). This leads to the main question of this paper: how can we compute formal safety certificates for non-linear stochastic systems with unknown noise distribution?

Following the existing literature (Santoyo et al., 2021; Steinhardt & Tedrake, 2012), we define safety as the probability that the system will remain within a given safe set for a given time horizon. Common approaches to quantify safety for non-linear stochastic systems either rely on abstraction methods (Badings, Romao, Abate, Parker et al., 2023; Cauchi et al., 2019) or on SBFs (Mathiesen, Calvert et al., 2023; Santoyo et al., 2021; Steinhardt & Tedrake, 2012). Abstraction-based methods build a discrete representation (i.e., a variant of a Markov chain) of the underlying stochastic system via the discretisation of its state space (Badings, Romao, Abate, Parker et al., 2023; Cauchi et al., 2019). Then, value iteration is performed on such a representation to verify properties and synthesise controllers, which can be mapped back on the original system by relying on simulation relations between the discrete representation and the underlying system (Girard & Pappas, 2007; Haesaert & Soudjani, 2021). Various

☆ The material in this paper was partially presented at 62nd IEEE Conference on Decision and Control (CDC), Dec. 13–15, Singapore, 2023. This paper was recommended for publication in revised form by Associate Editor Antoine Girard under the direction of Editor Maria Elena Valcher.

* Corresponding author.
E-mail addresses: f.b.mathiesen@tudelft.nl (F.B. Mathiesen), licio@dtu.dk (L. Romao), s.c.calvert@tudelft.nl (S.C. Calvert), l.laurenti@tudelft.nl (L. Laurenti), aabate@cs.ox.ac.uk (A. Abate).

[1] Frederik Baymler Mathiesen and Licio Romao contributed equally to this work.

https://doi.org/10.1016/j.automatica.2026.113187
0005-1098/© 2026 Published by Elsevier Ltd.

approaches have been proposed that combine abstractions and data-driven methods, including distributionally robust methods (Gracia et al., 2023), the scenario approach (Badings, Romao, Abate, Parker et al., 2023), Chernoff bounds (Lavaei et al., 2023), and Gaussian process regression (Jackson et al., 2020). However, a common drawback of these approaches is the need to partition the state space of the original stochastic system and to solve the value iteration on the resulting discrete abstraction, which leads to the well-known state-space explosion problem (Cauchi et al., 2019).

In contrast, SBFs (Kushner, 1967; McAllister & Rawlings, 2020) are Lyapunov-like functions, whose level sets allow one to bound the probability that a dynamical system will remain safe for a given time horizon, without the need to explicitly evolve the system over time (Santoyo et al., 2021; Steinhardt & Tedrake, 2012). By not requiring an analytical solution to the system's governing equations, SBFs represent a promising technique to efficiently quantify the safety of stochastic systems. However, one of the main challenges in SBF design is in finding an SBF that does not lead to overly conservative results (Laurenti & Lahijanian, 2026). In fact, designing an SBF requires the solution to a stochastic optimisation problem whose nature depends on the class of dynamics under study. In literature, synthesis of SBFs is usually performed with convex optimisation, in particular sum-of-squares (SoS) optimisation (Jagtap et al., 2020; Prajna et al., 2007; Santoyo et al., 2021) and Linear Programming (LP) for piece-wise constant SBFs (Mazouz et al., 2026), or with deep learning (Mathiesen, Calvert et al., 2023). However, these papers make restrictive assumptions on the noise distribution and on the system dynamics, e.g. Gaussian noise and linear or polynomial dynamics. More recently, some papers relax the assumption on the noise distribution by means of data-driven approaches (Salamati et al., 2021; Salamati & Zamani, 2023), which in addition to the level of safety also induces a (formally quantified) confidence. In particular, Salamati et al. (2021) uses SAA and Chebyshev's inequality to synthesise an SBF. This alternative approach, however, suffers from a sample complexity that is linear in the inverse of the confidence. Recent work on synthesising SBFs purely from trajectory data (Schön et al., 2024), i.e., without assuming any knowledge of the dynamics, has used Conditional Mean Embeddings, whose bottleneck is a computational complexity of $O(N^3)$ where N is the number of samples.

Our approach departs from previous techniques to ensure safety for non-linear stochastic dynamical systems using SBFs. First, we present an inner approximation to the stochastic program commonly used to design an SBF in terms of a chance-constrained optimisation problem. The feasible set of the latter is contained in the feasible set of the former. Then, by restricting to the class of piece-wise affine stochastic barrier functions[2] and relying on uncertain linear relaxations of non-linear systems (Xu et al., 2020), we show that the resulting chance-constrained problem can be reformulated into a robust LP problem (Boyd & Vandenberghe, 2004). This reformulation allows us to employ the scenario approach theory to devise a data-driven framework to synthesise an SBF,[3] and consequently obtain safety guarantees to the trajectories of the system. The resulting approach is data-efficient, as it only requires a limited amount of samples from the noise distribution that is logarithmic in the negative inverse of the confidence, which is in contrast with Salamati et al. (2021) where, as previously mentioned, the sample complexity is instead proportional in the inverse of the confidence. Furthermore, our approach is applicable to piece-wise uncertain dynamics and is scalable due to its LP representation. We also introduce an a priori sample discarding procedure and spatial indexing for efficient model construction to improve scalability of the proposed framework. Our experiments show competitive performance on various systems, including a model of a vehicle in windy conditions and various Neural Network Dynamical Models (NNDMs). Our numerical analysis illustrates how our approach substantially outperforms state-of-the-art comparable methods in terms of both the tightness of bounds and the amount of data required. Overall, our main contributions are:

- We develop a data-driven technique for the design of barrier functions that relies on a chance-constrained inner approximation of stochastic programs.
- We use the scenario approach theory and Linear Bound Propagation techniques to synthesise an SBF for non-linear dynamical systems.
- We present an efficient computational architecture to construct the resulting optimisation problem using spatial indexing methods, such as R-trees, for faster set intersection computations.
- We benchmark the proposed framework, showing its advantages with respect to approaches in the literature, including instances of Neural Network Dynamical Models (NNDMs).

A conference version of this paper appeared in Mathiesen, Romao et al. (2023). This paper significantly expands the preliminary results in Mathiesen, Romao et al. (2023), which focused on developing data-driven SBF synthesis techniques only for PWA dynamics. In this paper, we deal with general non-linear dynamics, which necessitates non-trivial extensions of the techniques in Mathiesen, Romao et al. (2023) to handle uncertain PWA models. Furthermore, here we show how to leverage spatial indexing and sample discarding to enhance scalability of our framework. We also newly include detailed proofs, as well as a substantially extended empirical analysis, including experiments on NNDMs.

The structure of this paper is as follows. The problem statement is described in Section 2. Section 3 covers preliminary results from convex analysis, scenario optimisation, and piece-wise affine relaxations of non-linear functions. Section 4 defines piece-wise linear SBFs and how they can be used to certify safety for non-linear stochastic systems. Section 5 details the main theoretical results, including a reformulation of the stochastic program to synthesise an SBF in to an inner chance-constrained approximation. In Section 6, we apply scenario approach theory towards a data-driven synthesis of a PWA SBF. In Section 7 we show how to compute the necessary polyhedral over-approximations and two methods to improve the scalability, namely the convex hull over sample set and spatial indexing. Empirical studies are reported in Section 8.

2. Problem formulation

Notation Let $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, and $\mathbb{N}$ represent, respectively, the set of real, non-negative real, and natural numbers. We denote by $\{x_1, \ldots, x_m\}$ the elements in a finite set and the convex hull of a finite set as $\text{conv}(x_1, \ldots, x_m)$. Let Ω be an abstract space, $\mathcal{F}$ be a σ-algebra defined on this set, and $\mathbb{P}$ be a probability measure; we denote by $(\Omega, \mathcal{F}, \mathbb{P})$ the associated complete probability (or uncertainty) space. A random variable η with values in $\mathbb{R}^n$ is a measurable function $\eta : \Omega \to \mathbb{R}^n$, with $\mathbb{R}^n$ equipped with its standard Borel σ-algebra. A realisation of η is denoted by $\eta(\omega)$, for some $\omega \in \Omega$. For a given set Ω, we denote by $\mathcal{P}(\Omega)$ the set of probability measures defined over Ω. The supremum norm $\|\cdot\|_\infty$ on a function $f : \mathbb{R}^n \to \mathbb{R}$ is defined as $\|f\|_\infty = \sup_{x \in \mathbb{R}^n} |f(x)|$.

[2] It is known that with sufficiently many pieces a piece-wise affine function can formally approximate any continuous function arbitrarily well.

[3] Salamati et al. (2021) use the scenario approach, in addition to SAA, to sample a robust convex program for synthesising SBFs. This application of scenario approach theory, and the underlying problem, is fundamentally different to our proposal where the scenario approach is used to obtain a solution to a stochastic program.

2.1. Problem formulation

We consider the following discrete-time, stochastic non-linear system

$$x(k+1) = f(x(k)) + \eta, \quad x(0) = \bar{x}, \tag{1}$$

where $k \in \mathbb{N}$ denotes time, $\bar{x} \in \mathbb{R}^n$ is the initial condition, and $f : \mathbb{R}^n \to \mathbb{R}^n$ is a Lipschitz continuous function. η is a random variable defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and, at each time instance, an independent and identically distributed (i.i.d.) realisation is drawn and added to the nominal dynamics represented by the function f. We further assume that the probability distribution of η is absolutely continuous with respect to the Lebesgue measure of $\mathbb{R}^n$ for a well-defined probability density function. Throughout this paper, we assume that the probability distribution of η is unknown.

Remark 1. The assumption that f is known exactly can be relaxed. In fact, as we will detail in Section 5.1, our approach only requires f to be represented as a piece-wise uncertain affine function, thus, allowing for, e.g. parametric uncertainty, in f. □

Because of the i.i.d. assumption on η, we can write the dynamics of System (1) using its kernel representation. A *stochastic kernel* is a measurable map from $\mathbb{R}^n$ onto the space of probability measures $\mathcal{P}(\mathbb{R}^n)$, $T : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ (Bertsekas & Shreve, 2004). In particular, the stochastic kernel associated with System (1) is given by

$$T(X \mid x) = \int_{\Omega} \mathbf{1}_X(f(x) + \eta(\omega))\mathbb{P}(d\omega), \tag{2}$$

where $X \subset \mathbb{R}^n$ is a Borel set of $\mathbb{R}^n$, and $\mathbf{1}_X$ is the indicator function, i.e., $\mathbf{1}_X(x) = 1$ if $x \in X$, and 0 otherwise. In other words, for a fixed $x \in \mathbb{R}^n$ representing the current state, the stochastic kernel returns the probability distribution of the state in the next time step. For $K \in \mathbb{N}$ and $\bar{x} \in \mathbb{R}^n$, we denote by $(\mathbb{R}^n)^K = \mathbb{R}^n \times \cdots \times \mathbb{R}^n$ the K-fold Cartesian product of $\mathbb{R}^n$. Then, the stochastic kernel (2) induces a unique measure on $(\mathbb{R}^n)^K$ given by the unique extension (due to Kolmogorov's extension theorem (Tao, 2013, Theorem 2.4.3)) of the measure

$$\mathbb{P}^{\bar{x}}(X_1, \ldots, X_K) = \int_{X_1} \left(\prod_{k=2}^{K} \int_{X_k} T(d\xi_k | \xi_{k-1}) \right) T(d\xi_1 \mid \bar{x}),$$

which represents the probability of a trajectory starting at $\bar{x}$, under the dynamics of System (1), being in the set X_k, $k = 1, \ldots, K$ at the various time steps. We now have all the ingredients to define the notion of probabilistic safety that is crucial to our approach.

Definition 1 (*Probabilistic Safety Santoyo et al., 2021*)**.** Let K be a non-negative integer and $X_s \subseteq X$ be a compact set representing the safe set. Then, for a given initial state $\bar{x}$, we define probabilistic safety as $\zeta(X_s, K \mid \bar{x}) = \mathbb{P}^{\bar{x}}(X_s, \ldots, X_s)$.

Our main objective in this paper is to obtain a uniform lower bound on the probabilistic safety of System (1) for all initial conditions in a given set X_0.

Problem 1. Let $D = (\omega_1, \ldots, \omega_N)$ be i.i.d. samples from η. Then, for a given compact set $X_0 \subseteq X$, a safe set $X_s \subseteq X$, and a time horizon $K \in \mathbb{N}$, find $\rho \in (0, 1]$ such that, with high confidence, $\zeta(X_s, K) = \inf_{\bar{x} \in X_0} \zeta(X_s, K \mid \bar{x}) \geq \rho$.

Note that, as the noise is additive, the assumption of having i.i.d. noise samples in Problem 1 is equivalent to having i.i.d. full measurements of the state. Exact computation of probabilistic safety for System (1) is generally infeasible, even when η is restricted to being a Gaussian random variable (Cauchi et al., 2019). Consequently, it is obvious that the more general setting considered in this paper, where the noise distribution is arbitrary and possibly unknown, requires approximations. We should also stress that, while in Problem 1 we focus on safety, the techniques developed in this paper to obtain its solution can also be applied to perform verification of more complex temporal properties. In fact, safety is the dual to, and can be reformulated as, reachability (Baier & Katoen, 2008),[4] and complex temporal specifications such as LTLf (De Giacomo & Vardi, 2013) can be reduced to reachability by representing the formula as an automaton and checking reachability on the product system between the automaton and the system (Haesaert & Soudjani, 2021; Jagtap et al., 2020).

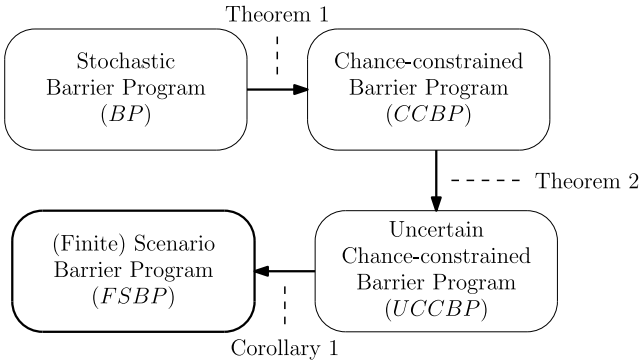


Fig. 1. Summary of the proposed framework for safety verification of nonlinear systems. Theorem 1 (Section 5) guarantees that the associated stochastic program to find an SBF can be solved as a chance-constrained program via constraint tightening. Non-linearities can be dealt with by means of a PWA abstraction as in Theorem 2 (Section 5.1). The scenario approach is applied to use available data for high-confidence certificates; strong duality is essential to relax the problem to Linear Programming in Corollary 1 (in Section 6).

Approach We leverage the availability of data D to synthesise a Stochastic Barrier Function (SBF) that, with high confidence, provides a bound on probabilistic safety. Our approach is summarised in Fig. 1 and is based on a novel reformulation of the optimisation problem to find an SBF into a chance-constrained problem (detailed in Section 5). To deal with the non-linearity of a system, we develop in Section 5.1 a formal (i.e. with a quantified error) approximation of System (1) as an uncertain piece-wise affine (PWA) system. Such an approximation is then used in Section 6 to reformulate the chance-constrained problem into a robust linear program, which can be efficiently solved using duality.

3. Preliminaries

Our approach requires some results from convex optimisation and scenario optimisation, which we summarise below for convenience.

3.1. Convex analysis and linear programming

Given a matrix $H \in \mathbb{R}^{p \times n}$ and a vector $h \in \mathbb{R}^p$, a polyhedral set in *half-space representation* is denoted by $P = \{x \in \mathbb{R}^n : Hx \leq h\}$. The vertex representation of P is $P = \text{conv}(x_1, \ldots, x_m)$ where

[4] The safety probability is one minus the probability of reaching the unsafe set.

vert$(P) = \{x_1, \ldots, x_m\}$ are the vertices of P. The following class of robust, or semi-infinite, LPs is crucial to our developments:

$$\begin{aligned} \min_{z\in\mathbb{R}^d} \quad & c^\top z \\ \text{s.t.} \quad & (Az+a)^\top x \le Bz+b, \quad \text{for all } x \in P, \end{aligned} \tag{3}$$

where $c \in \mathbb{R}^d$, $A \in \mathbb{R}^{n\times d}$, $a \in \mathbb{R}^n$, $B \in \mathbb{R}^{1\times d}$, $b \in \mathbb{R}$, and $P \subset \mathbb{R}^n$ is a polyhedral set. By relying on standard duality arguments, we can recast the semi-infinite optimisation problem (3) as a regular LP as shown in the following proposition.

Proposition 2 (*Boyd and Vandenberghe (2004, Exercise 5.17)*)**.** *Denote by $\mathcal{Z} = \{z \in \mathbb{R}^d : (Az+a)^\top x \le Bz+b, \text{ for all } x \in P\}$ the feasible set of (3). Define the optimisation problem*

$$\begin{aligned} \min_{z\in\mathbb{R}^d,\lambda\ge 0} \quad & c^\top z \\ \text{s.t.} \quad & h^\top \lambda \le Bz+b \\ & H^\top \lambda = Az+a, \end{aligned} \tag{4}$$

whose (projected) feasible set is given by $\mathcal{Z}' = \{z : \exists \lambda \in \mathbb{R}^p_{\ge 0},\ h^\top\lambda \le Bz+b,\ H^\top\lambda = (Az+a)\}$. Then, it holds that $\mathcal{Z} = \mathcal{Z}'$.

3.2. Scenario optimisation

The scenario approach theory allows one to certify the probability of constraint violation associated with chance-constrained optimisation problems (Campi & Garatti, 2008). Let $\eta : \Omega \to \mathbb{R}^q$ be a random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $D = (\omega_1, \ldots, \omega_N)$ be i.i.d. samples from $\mathbb{P}$, which live in the space $(\Omega^N, \otimes_N\mathcal{F}, \mathbb{P}^N)$, where Ω^N is the N-fold Cartesian product of Ω, and $\otimes_N\mathcal{F}$ is the product sigma algebra generated by the sigma algebra $\mathcal{F}$, and $\mathbb{P}^N$ represents the induced measure on Ω^N. Then, consider the scenario program

$$\begin{aligned} \min_{z\in\mathbb{R}^d} \quad & c^\top z \\ \text{s.t.} \quad & g(z,\eta(\omega)) \le 0, \quad \text{for all } \omega \in D, \end{aligned} \tag{5}$$

where $g(z,\eta) : \mathbb{R}^d \times \mathbb{R}^q$ is convex in z for each value of η and measurable in η for each value of z.

Assumption 1. We assume almost surely with respect to the measure $\mathbb{P}^N$ that: (i) the feasible set $\mathcal{Z} = \{z : g(z,\eta(\omega)) \le 0, \forall\omega \in D\}$ has non-empty interior, and (ii) the optimal solution of program (5) exists and is unique.

The feasible set of every program considered in this paper has by construction a non-empty interior and uniqueness can always be enforced with a tie-break rule. We denote the unique solution of (5) by $z^\star(D)$, which is a random variable on the space Ω^N. A key result within the scenario approach theory establishes an upper bound on the tail distribution of the constraint violation probability associated with $z^\star(D)$.

Proposition 3 (*Campi & Garatti, 2008*)**.** *Consider the scenario program (5) and suppose Assumption 1 holds. Then, for any $\epsilon \in (0,1)$, we have that*

$$\mathbb{P}^N\{D \in \Omega^N : V(z^\star(D)) > \epsilon\} \le \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i},$$

where $V(z) = \mathbb{P}\{\omega \in \Omega : g(z,\eta(\omega)) > 0\}$ is the violation probability of $z \in \mathbb{R}^d$.

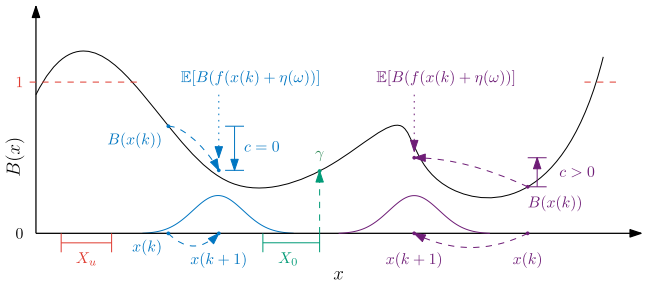


Fig. 2. The figure is borrowed from Mathiesen, Calvert et al. (2023). An SBF $B(x)$ is a non-negative function that is greater than 1 in an unsafe region X_u, which is the complement of the safe set X_s. The variable γ is an upper bound for $B(x)$ over an initial region X_0. The upper bound for the expected increase in $B(x)$ after one step of (1) over the safe set X_s is denoted c. Proposition 5 shows that $\zeta(X_s, K) \ge \max\{1-(\gamma+cK), 0\}$.

3.3. Uncertain piece-wise affine relaxations

Uncertain PWA relaxations are key to our method. This type of relaxation allows one to treat complex non-linear functions as uncertain PWA functions, simplifying analysis and optimisation. An uncertain PWA relaxation for a function f is a collection of local linear relaxations, that is, $\underline{A}_i x + \underline{b}_i \le f(x) \le \overline{A}_i x + \overline{b}_i$ for all x in a convex region Q_i, given a partition $\mathcal{Q} = \{Q_1, \ldots, Q_\ell\}$. We formalise the existence of an uncertain PWA relaxation in what follows.

Proposition 4. *Let $\mathcal{Q} = \{Q_1, \ldots, Q_\ell\}$ be a given convex partition of a compact set $X \subset \mathbb{R}^n$. Then, for any function $f : \mathbb{R}^n \to \mathbb{R}^n$ that is locally Lipschitz on X, there exists an uncertain PWA function $\hat{f}(x,\alpha) = \max\{\hat{f}_1(x,\alpha), \ldots, \hat{f}_\ell(x,\alpha)\}$, where $\alpha \in [0,1]$ and $\hat{f}_i(x,\alpha) = A_i(\alpha)x + b_i(\alpha)$ if $x \in Q_i$ with $A_i(\alpha) = \alpha\underline{A}_i + (1-\alpha)\overline{A}_i$ and $b_i(\alpha) = \alpha\underline{b}_i + (1-\alpha)\overline{b}_i$, and zero otherwise, such that it holds that $f(x) \in \{\hat{f}(x,\alpha) : \alpha \in [0,1]\}$ for all $x \in X$.*

A detailed proof can be found in Appendix A.2. Note that there are various possible approaches to construct matrices $\underline{A}_i, \overline{A}_i$ and vectors $\underline{b}_i, \overline{b}_i$ for each region Q_i including first-order Taylor models (Joldes, 2011) and based on local Lipschitz constants (Clarke, 1981), which is used in the proof of Proposition 4. In this paper, we use a state-of-the-art technique, called Linear Bound Propagation (LBP) (Xu et al., 2020). The core idea of LBP is to recursively propagate linear bounds backward through the computation graph representing the function f.

4. Stochastic barrier functions

In this paper, we use Stochastic Barrier Functions to provide safety guarantees for System (1).

Definition 2 (*Stochastic Barrier Function*)**.** Given a safe set X_s and a set of initial conditions $\bar{X}$, a measurable function $B : \mathbb{R}^n \to \mathbb{R}_{\ge 0}$ is called a Stochastic Barrier Function (SBF) for System(1) if there exist non-negative constants γ, c satisfying

$$B(x) \le \gamma, \qquad \text{for all } x \in X_0, \tag{6}$$

$$B(x) \ge 1, \qquad \text{for all } x \in \mathbb{R}^n \setminus X_s, \tag{7}$$

$$\mathbb{E}\{B(f(x)+\eta(\omega))\} \le B(x) + c, \qquad \text{for all } x \in X_s. \tag{8}$$

A pictorial representation of an SBF is presented in Fig. 2. Inequality (8) requires that for all $x \in X_s$, the expected value of the barrier function at the next step cannot increase more than c. As shown in Proposition 5 below, we can leverage results from martingale theory to obtain a lower bound on $\zeta(X_s, K)$.

Proposition 5 (*Kushner, 1967, Chapter 3, Theorem 3, Steinhardt & Tedrake, 2012, Section 2.2*)**.** *For a safe $X_s \subset \mathbb{R}^n$ and a set of initial conditions $X_0 \subset X_s$, let the function $B : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ be an SBF for System (1), and let the positive integer K be a time horizon. Then, it holds that $\zeta(X_s, K) \geq \max\{1 - (\gamma + cK), 0\}$.*

The safety guarantees provided by Proposition 5 can be broken down into the initial set level and the bound on the expected increase. The intuition is that together these two terms can be used to bound the expected value of B after K steps, starting from any state in the initial set. Since the unsafe set is captured by the 1-superlevel set of B, the bound on the expected value of B can subsequently be used to bound the probability of entering the unsafe set within K time steps, or equivalently, bound the number of steps to reach the unsafe set. Importantly, the consequence of this formulation, and its derivation via bounding the expected value of B, is that the guarantee on the safety probability decreases linearly with an increasing horizon, namely with the scale factor c. To obtain an optimal barrier function for a fixed horizon, one can formulate the search as the following infinite-dimensional stochastic optimisation problem

$$\min_{B \in \mathcal{M}, c, \gamma} \quad \gamma + cK \quad \text{s.t.} \quad (6),\ (7),\ (8), \tag{BP}$$

where $\mathcal{M}$ represents the set of non-negative measurable functions in $\mathbb{R}^n$. Two challenges emerge when solving the barrier program (BP) to solve Problem 1: (i) $\mathcal{M}$ is an infinite-dimensional space, and (ii) constraint (8) involves computing an expectation over an unknown probability distribution. To address the first challenge, we restrict the search for barrier functions to the class of piece-wise affine functions given by

$$\mathcal{M}' = \{B : \mathbb{R}^n \to \mathbb{R}_{\geq 0} : B(x, \theta) = \max\{B_1(x, \theta), \dots, B_\ell(x, \theta)\}\}, \tag{9}$$

with $\theta \in \mathbb{R}^{\ell(n+1)}$ representing the (finite) set of parameters $(u_i, v_i) \in \mathbb{R}^{(n+1)}$ that define the barrier function, i.e., $B_i(x, \theta) = u_i^\top x + v_i$ if $x \in Q_i$ and zero otherwise where the polyhedral sets $Q_i = \{x \in \mathbb{R}^n : H_i x \leq h_i\}$, for all $i \in \{1, \dots, \ell\}$, constitute a partition of $\mathbb{R}^n$. While various classes of barrier functions have been considered in the literature (Prajna et al., 2007; Santoyo et al., 2021), in this work, we focus on PWA barrier functions, as they are expressive enough to approximate any non-linear function arbitrarily well with a sufficient number of parameters and, as we will show in Section 6, this choice leads to an LP program to synthesise an SBF, thus guaranteeing efficiency. For the second challenge, in the next section we develop a new approach to create a chance-constrained approximation of (BP) using a constraint tightening technique.

5. An inner chance-constrained approximation of the barrier program (BP)

Solving (BP) is challenging, even in the case that the probability distribution underpinning (8) is known. In this section, we show how to relax (BP) using a reformulation in terms of a chance-constrained optimisation problem whose feasible set is contained in the set of feasible solutions of (BP). Such ideas have never been used in this context. Hence, we depart completely from the existing approaches, which either rely on approximating Constraint (8) with the empirical distribution, make strong assumptions about the noise distribution to analytically compute the expectation and recast (8) as a convex constraint, or rely on convex over-approximations of the expectation to (conservatively) verify (8). Furthermore, due to the inner approximation of (BP) in terms of a chance-constrained problem, our approach opens the road to use the tools of scenario optimisation discussed in Section 3.2 to obtain strong sample complexity guarantees on the safety probability of System (1). To this end, fix any $B \in \mathcal{M}$ and $x \in \mathbb{R}^n$, and let $E \in \mathcal{F}$ be a measurable set. Then, observe that

$$\begin{aligned}\mathbb{E}\{B(f(x) + \eta(\omega))\} = &\int_E B(f(x) + \eta(\omega))\mathbb{P}(d\omega) \\ &+ \int_{E^c} B(f(x) + \eta(\omega))\mathbb{P}(d\omega).\end{aligned} \tag{10}$$

Under the assumption that $B(x) \leq M$ for any $x \in \mathbb{R}^n$, which can always be enforced for SBFs, the second term on the right-hand side of (10) can be bounded by $M\mathbb{P}\{E^c\}$. Our main idea then is to choose a particular E^c that allows us to control the right-hand side of (10), and conquer enough of $\mathcal{F}$ to ensure the desired confidence. This is formalised in the next lemma.

Lemma 1. *For $B \in \mathcal{M}$ with $\|B\|_\infty = M$, let c and η be as in (BP). Define the set*

$$E = \{\omega \in \Omega : B(f(x) + \eta(\omega)) + \xi \leq B(x) + c, \ \text{for all } x \in X_s\}, \tag{11}$$

for some constant $\xi \geq 0$. If there exists an $\epsilon \in (0, 1)$ such that $\xi \geq M\frac{\epsilon}{1-\epsilon}$ and $\mathbb{P}\{E^c\} \leq \epsilon$, then we have $\mathbb{E}\{B(f(x)+\eta(\omega))\} \leq B(x)+c$ for all $x \in X_s$.

Proof. The proof of our result is based on (10). We first remark that E is a measurable set (see Appendix A.3). Next, fix any $x \in X_s$ and define the set

$$E_x = \{\omega \in \Omega : B(f(x) + \eta(\omega)) + \xi \leq B(x) + c\}. \tag{12}$$

Notice that, by definition, we have $E \subset E_x$. Furthermore, the set E_x is also measurable due to the assumptions on B, f and η, and the fact that measurability is closed under composition. Using (10), we then obtain

$$\mathbb{E}\{B(f(x) + \eta(\omega))\} \leq (B(x) + c - \xi)\mathbb{P}\{E_x\} + M\mathbb{P}\{E_x^c\}, \tag{13}$$

where the first term on the right-hand side of (13) follows from the definition of E_x in (12), and the second by the uniform boundedness condition on B. It remains to show that $(B(x)+c-\xi)\mathbb{P}\{E_x\}+M\mathbb{P}\{E_x^c\} \leq B(x) + c$. To this end, it suffices to show that

$$-\xi\mathbb{P}\{E_x\} + M\mathbb{P}\{E_x^c\} \leq 0 \tag{14}$$

holds, since $B(x) + c$ is non-negative and $\mathbb{P}\{E_x\} \leq 1$ by definition. Substituting the two inequalities $\mathbb{P}\{E_x\} \geq \mathbb{P}\{E\} \geq 1 - \epsilon$ and $\mathbb{P}\{E_x^c\} \leq \mathbb{P}\{E^c\} \leq \epsilon$ into the left-hand side of (14) we obtain $-\xi\mathbb{P}\{E_x\} + M\mathbb{P}\{E_x^c\} \leq -\xi(1 - \epsilon) + M\epsilon$, whose right-hand side is less than or equal to zero due to the fact that $\xi \geq M\frac{\epsilon}{1-\epsilon}$. □

Lemma 1 is enabled by the chance-constraint tightening variable ξ through the condition that $\xi \geq M\frac{\epsilon}{1-\epsilon}$. A reader may question how to choose ξ and ϵ. Choosing ϵ is a trade-off between assigning less probability mass to the uniform upper bound, which is desirable as the uniform upper bound represents a worst-case for barrier value at the next step, is and the amount of data required when applying the scenario approach theory. In our experiments (see Section 8), we employ $\epsilon = 0.005$. Once ϵ is chosen, the optimal choice of ξ to minimise c is $\xi = M\frac{\epsilon}{1-\epsilon}$, which follows from the fact that ξ is on the left-hand side of the inner inequality of E (and E_x) with c on the right-hand side and $M\frac{\epsilon}{1-\epsilon}$ is the smallest allowed value of ξ. An immediate consequence of Lemma 1 is the fact that we can obtain an inner approximation of the optimisation problem (BP) in terms of a chance-constrained optimisation problem.

Theorem 1. *Consider the dynamical system given in System (1). Then, we have that for all $\epsilon \in (0, 1)$ and positive integers K and $M \geq 1$ such that $\xi \geq M\frac{\epsilon}{1-\epsilon}$, the feasible set of*

$$\begin{aligned} \min_{B\in\mathcal{M},c,\gamma} \quad & \gamma + cK \\ s.t. \quad & (6), \quad (7), \quad \gamma \geq 0, \quad c \geq 0, \\ & B(x) \leq M, \ \forall x \in \mathbb{R}^n \\ & \mathbb{P}\{\omega \in \Omega : B(f(x) + \eta(\omega)) + \xi \\ & \quad \leq B(x) + c, \ \forall x \in X_s\} \geq 1 - \epsilon, \end{aligned} \tag{CCBP}$$

is contained in the feasible set of (BP).

Proof. Constraints (6) and (7) are shared between (BP) and the chance-constrained barrier program (CCBP). Thus, it is sufficient to show that the chance-constraint of (CCBP) implies that constraint (8) is satisfied. Rewriting the chance-constraint as $\mathbb{P}\{E\} \geq 1 - \epsilon$ where E is defined as in Lemma 1, it holds that $\mathbb{P}\{E^c\} \leq \epsilon$. From this, the assumption $\xi \geq M\frac{\epsilon}{1-\epsilon}$, and the constraint that B is uniformly bounded by M, the conditions of Lemma 1 are satisfied. Thus, it holds that if the chance constraint of (CCBP) is satisfied, then constraint (8) is satisfied. □

Unfortunately, the computational burden of solving (CCBP) is not negligible, mainly due to the quantification over all $x \in X_s$ and to the non-linear function f. In Section 5.1, we will show how the framework developed in Section 3.3 can alleviate this burden.

5.1. Over-approximating general non-linear systems with uncertain PWA dynamics

In Section 3.3, we showed that any locally Lipschitz function can be over-approximated by uncertain PWA functions, and by extension non-linear systems of the form of System (1) can be over-approximated by uncertain PWA systems. More formally, an uncertain PWA over-approximation of System (1) is described as follows.

$$x(k+1) \in F(x(k)) + \eta(k), \tag{15}$$

where $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a set-valued mapping defined as $F(x(k)) = \left\{\hat{f}(x(k), \alpha) : \alpha \in [0, 1]\right\}$, with $\hat{f}$ being an uncertain PWA function with the uncertain parameter α and such that for any $x \in \mathbb{R}^n$ it holds that $f(x) \in F(x)$. Such a definition of a uncertain PWA over-approximation of System (1) guarantees that between System (1) and PWA System (15) there is a *behavioural inclusion* relation, that is, for any $x \in \mathbb{R}^m$ there exists $\alpha \in [0, 1]$ such that $\hat{f}(x, \alpha) = f(x)$. Intuitively, we would like to synthesise an SBF for System (15) and rely on the behavioural relation described above to ensure that the synthesised SBF also is an SBF for System (1). This is what we show in the following Theorem.

Theorem 2. *Consider the dynamical system given in (1) and assume access to an uncertain PWA over-approximation (15) of the system. Then, we have that for all $\epsilon \in (0, 1)$ and positive integer K, if there exist $M \geq 1$ and $\xi \geq M\frac{\epsilon}{1-\epsilon}$, then the feasible set of*

$$\begin{aligned} \min_{B\in\mathcal{M},c,\gamma} \quad & \gamma + cK \\ s.t. \quad & (6), \quad (7), \quad \gamma \geq 0, \quad c \geq 0, \\ & B(x) \leq M, \ \forall x \in \mathbb{R}^n \\ & \mathbb{P}\{\omega \in \Omega : B(y + \eta(\omega)) + \xi \\ & \quad \leq B(x) + c, \ \forall x \in X_s, \ \forall y \in F(x)\} \geq 1 - \epsilon, \end{aligned} \tag{UCCBP}$$

is contained in the feasible set of (BP).

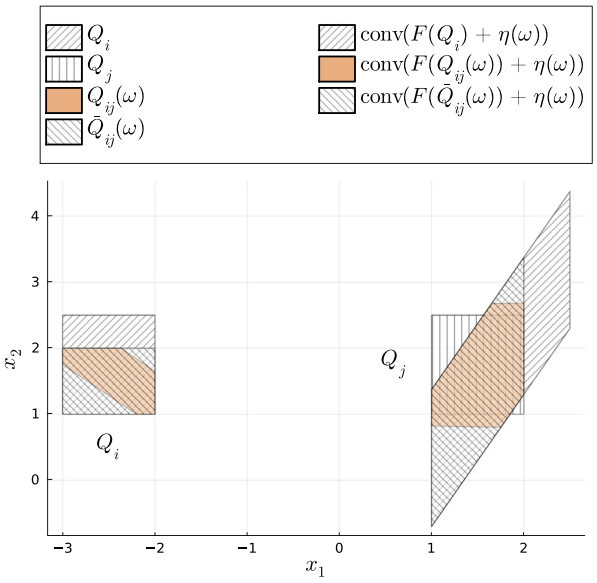


Fig. 3. Given regions Q_i, Q_j and a realisation of the noise ω, the set $Q_{ij}(\omega)$ represents the subset of $x \in Q_i$ such that $\hat{f}(x, \alpha) + \eta(\omega) \in Q_j$ for some $\alpha \in [0, 1]$. That is, $Q_{ij}(\omega)$ is the subset of Q_i that can reach Q_j given the realisation of the noise ω. Unfortunately, $Q_{ij}(\omega)$ is not easily computable (see Fig. 4) and thus in Section 7.1, we will compute an over-approximation $\bar{Q}_{ij}(\omega)$.

Proof. Constraints (6) and (7) are imposed directly in the uncertain chance-constrained barrier program (UCCBP). Since $f(x) \in F(x)$ implies that there exists an $\alpha \in [0, 1]$ such that $\hat{f}(x, \alpha) = f(x)$, it holds that $B(f(x) + \eta(\omega)) \leq \sup_{y\in F(x)} B(y + \eta(\omega))$. Therefore, the chance constraint of (UCCBP) implies the chance constraint of (CCBP). By Theorem 1 and transitivity of the subset relation, the feasible set of (UCCBP) is contained in the feasible set of (BP). □

6. Data-driven approximation of (UCCBP) for the class of uncertain dynamical systems

To reformulate (UCCBP) as a robust LP problem, we need to introduce some mathematical notation. Let $\mathcal{Q} = \{Q_1, \ldots, Q_\ell\}$ be a partition of the state space associated with a barrier function, as described in Section 4. Let us also denote four collections of indices by

$$\begin{aligned} & I = \{1, \ldots, \ell\}, \\ & I_{X_0} = \{i \in I : Q_i \cap X_0 \neq \emptyset\}, \\ & I_{X_s} = \{i \in I : Q_i \cap X_s \neq \emptyset\}, \\ & I_{X_u} = \{i \in I : Q_i \cap (\mathbb{R}^n \setminus X_s) \neq \emptyset\}, \end{aligned} \tag{16}$$

which represent the collection of all indices, and the elements of $\mathcal{Q}$ with non-empty intersections with the initial state, the safe set, and unsafe set, respectively. Finally, for each pair $(i, j) \in I_{X_s} \times I$, we denote the set

$$Q_{ij}(\omega) = \left\{x \in Q_i : \exists y \in F(x), y + \eta(\omega) \in Q_j\right\}, \tag{17}$$

representing the subset of Q_i with i belonging to I_{X_s} and that is mapped to Q_j under a given realisation of the noise. A pictorial example of $Q_{ij}(\omega)$ can be found in Fig. 3. Leveraging the results of Theorem 2 and Proposition 3 and using the notation we have introduced so far, we obtain the following intermediate result.

The goal of Lemma 2 below is two fold: (i) to impose piecewise constraints on the barrier synthesis and (ii) apply scenario approach theory to obtain a tractable solution in the face of an unknown noise distribution.

Lemma 2. *Let $D = \{\omega_1, \dots, \omega_N\}$ be a collection of N independent samples from $\mathbb{P}$. Fix $\epsilon \in (0,1)$, $M \geq 1$, and $\xi \geq M\frac{\epsilon}{1-\epsilon}$, and consider the optimisation program*

$$
\begin{aligned}
\min_{z} \quad & \gamma + cK \\
s.t. \quad & \gamma \geq 0, \quad c \geq 0, \\
& B_i(x,\theta) \in [0, M], \quad \forall x \in Q_i,\ i \in I, \\
& B_i(x,\theta) \leq \gamma, \quad \forall x \in Q_i \cap X_0,\ i \in I_{X_0}, \\
& B_i(x,\theta) \geq 1, \quad \forall x \in Q_i,\ i \in I_{X_u}, \\
& B_j(y+\eta(\omega),\theta) + \xi \leq B_i(x,\theta) + c, \\
& \forall (\omega, i, j) \in D \times I_{X_s} \times I,\ y \in F(x),\ x \in Q_{ij}(\omega)
\end{aligned}
\tag{SBP}
$$

where $z = (\gamma, c, \theta)$ is the collection of decision variables in (UCCBP) and $d = |z| = 2 + \ell(n+1)$ is the number of decision variables. Then, with probability at least $1-\beta$, where β is the right-hand side of the inequality in Proposition 3, the optimal solution of the scenario barrier program (SBP) satisfies the constraints of Definition 2.

The relevance of Lemma 2 towards the general framework proposed in the paper is: (1) it establishes a sufficient condition to enforce the constraints of (UCCBP) by restricting the attention to each element of the partition $\mathcal{Q}$ individually; (2) it allows us to use the duality results of Section 3.1 to obtain a tractable reformulation of (SBP) into a robust LP. In the rest of this section, we discuss how to obtain this finite, yet equivalent, representation of (SBP). The reformulation of the first three semi-infinite constraints of (SBP) follows standard dualisation arguments, i.e. using Proposition 2, thus we defer their explicit reasoning to the proof of Corollary 1. The remaining constraint $B(y+\eta(\omega),\theta) + \xi \leq B(x,\theta) + c$ for all $\omega \in D$, $x \in X_s$, $y \in F(x)$, requires more care, yet Lemma 2 also enables a computationally tractable reformulation of this. To this end, consider $(i,j) \in I_{X_s} \times I$ and $\omega \in D$. A challenge in reformulating this constraint as a linear constraint is that the set $Q_{ij}(\omega)$ is not a polyhedron or even convex (see Fig. 4). For now, we assume access to a polyhedral over-approximation $Q_{ij}(\omega) \subset \overline{Q}_{ij}(\omega) = \{x : H_{ij\omega}x \leq h_{ij\omega}\}$. Then, if we impose the constraint for all $x \in \overline{Q}_{ij}(\omega)$, then it trivially follows that it also holds for all $x \in Q_{ij}(\omega)$. How to compute $\overline{Q}_{ij}(\omega)$ will be discussed in Section 7.1.

Proposition 6. *Let $(i,j) \in I_{X_s} \times I$ and $\omega \in D$ be given. Assume that there exists two affine functions such that $\underline{A}_i x + \underline{b}_i \leq F(x) \leq \overline{A}_i x + \overline{b}_i$ for all $x \in Q_i$. Then, it holds that $B_j(y+\eta(\omega),\theta) + \xi \leq B_i(x,\theta) + c$ for all $x \in \overline{Q}_{ij}(\omega)$, $y \in F(x)$ iff the following constraints hold*

$$
\begin{aligned}
& h_{ij\omega}^\top \underline{\lambda}_{ij\omega} \leq v_i - v_j - u_j^\top(\underline{b}_i + \eta(\omega)) + c - \xi, \\
& H_{ij\omega}^\top \underline{\lambda}_{ij\omega} = \underline{A}_i^\top u_j - u_i, \\
& h_{ij\omega}^\top \overline{\lambda}_{ij\omega} \leq v_i - v_j - u_j^\top(\overline{b}_i + \eta(\omega)) + c - \xi, \\
& H_{ij\omega}^\top \overline{\lambda}_{ij\omega} = \overline{A}_i^\top u_j - u_i,
\end{aligned}
$$

where $\underline{\lambda}_{ij\omega}$, $\overline{\lambda}_{ij\omega}$ are a non-negative dual variables.

Collecting together all finite sets of constraints, a finite representation of the semi-infinite program (SBP) is given as

$$
\begin{aligned}
\min_{z} \quad & \gamma + cK \\
\text{s.t.} \quad & \gamma \geq 0,\ c \geq 0, \\
& \text{(Non-negativity)} \\
& h_i^\top \underline{\nu}_i \leq v_i,\ H_i^\top \underline{\nu}_i = -u_i, \ \text{for all } i \in I, \\
& \text{(Uniform upper bound)} \\
& h_i^\top \overline{\nu}_i \leq M - v_i,\ H_i^\top \overline{\nu}_i = u_i, \ \text{for all } i \in I, \\
& \text{(Initial set)} \\
& h_{i0}^\top \mu_i^0 \leq \gamma - v_i,\ H_{i0}^\top \mu_i^0 = u_i, \ \text{for all } i \in I_{X_0}, \\
& \text{(Unsafe set)} \\
& h_i^\top \mu_i^u \leq v_i - 1,\ H_i^\top \mu_i^u = -u_i, \ \text{for all } i \in I_{X_u}, \\
& \text{(One-step constraints)} \\
& h_{ij\omega}^\top \underline{\lambda}_{ij\omega} \leq v_i - v_j - u_j^\top(\underline{b}_i + \eta(\omega)) + c - \xi, \\
& H_{ij\omega}^\top \underline{\lambda}_{ij\omega} = \underline{A}_i^\top u_j - u_i, \\
& h_{ij\omega}^\top \overline{\lambda}_{ij\omega} \leq v_i - v_j - u_j^\top(\overline{b}_i + \eta(\omega)) + c - \xi, \\
& H_{ij\omega}^\top \overline{\lambda}_{ij\omega} = \overline{A}_i^\top u_j - u_i, \ \text{for all } \omega \in D, \\
& \qquad \text{for all } (i,j) \in I_{X_s} \times I,
\end{aligned}
\tag{FSBP}
$$

where $\underline{\nu}_i$, $\overline{\nu}_i$, μ_i^0, μ_i^u, $\underline{\lambda}_{ij\omega}$, $\overline{\lambda}_{ij\omega}$ are non-negative dual variables. (H_{i0}, h_{i0}) denotes the half-space representation of $Q_i \cap X_0$.

In Corollary 1, we combine the results we introduced and show how probabilistic safety can be computed via the scenario approach using samples of $\eta(\omega)$.

Corollary 1. *Assume that $D = \{\omega_1, \dots, \omega_N\}$ is a collection of N independent samples from the noise distribution $\mathbb{P}$. Fix $\epsilon \in (0,1)$, $M \geq 1$, and $\xi \geq M\frac{\epsilon}{1-\epsilon}$, and let β be defined as Proposition 3 where $d = 2 + \ell(n+1)$ is the number of (non-dual) decision variables in (FSBP). Consider the optimal (primal) solution $z^\star(D) = (c^\star, \gamma^\star, \theta^\star)$ of (FSBP). Then, with confidence $1-\beta$, it holds that*

$$\zeta(X_s, K) \geq \max\{1 - (\gamma^\star + c^\star K), 0\}.$$

Proof. It is enough to show an equivalence between (SBP) and (FSBP). To do that, in what follows, we show that each semi-infinite constraint of (SBP) is equivalent to a set of constraints of (FSBP). We start by showing the equivalence for the initial set constraint (Condition $B_i(x,\theta) \leq \gamma, \forall x \in Q_i \cap X_0$, $i \in I_{X_0}$ in (SBP)) to the initial set constraints in (FSBP). For this, let I_{X_0} be the index set for regions intersecting with X_0 and consider its corresponding partition $Q_{i0} = Q_i \cap X_0 = \{x : H_{i0}x \leq h_{i0}\}$, $i \in I_{X_0}$. Then, the condition $B(x,\theta) \leq \gamma$ for all $x \in Q_{i0}$ can be written equivalently as $h_{i0}^\top \mu_i^0 \leq \gamma - v_i$ and $H_i^\top \mu_i^0 = u_i$ where μ_i^0 is a vector-valued non-negative decision variable. This follows from the fact that within the region Q_{i0} the function $B(x,\theta) = B_i(x,\theta)$ is affine and, consequently, by relying on Proposition 2, we can rewrite this robust constraint as two regular linear constraints, with the dual variable μ_i^0. The equivalence for the non-negativity, uniform upper bound, and unsafe set robust constraints (see Definition 2) to regular linear constraints follow a similar argument. Hence, for brevity, we omit the reformulation of these. The dualisation of the remaining constraint follows directly from Proposition 6. Thus, the proof is concluded by noting that the cost to minimise is the same in both (SBP) and (FSBP). □

Remark 2. Corollary 1 can easily be modified to the setting where the noise samples are not known exactly, but they are only known to lie in specified sets. This can be done by simply adding a quantifier for each of these sets in (FSBP). □

Thus, by solving (FSBP), we can certify probabilistic safety with high confidence. Note that naïvely trying to solve (FSBP) can soon become intractable on contemporary hardware, due to both

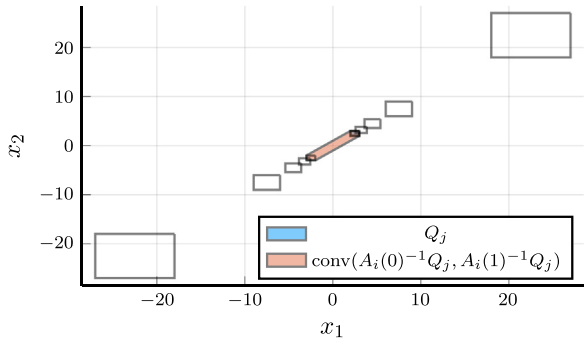


Fig. 4. A pictorial example that $Q_{ij}(\omega)$ is not necessarily convex. In this example, we have $\underline{A}_i = -I$, $\overline{A}_i = I$, and $\underline{b}_i = \overline{b}_i = 0$, which is a valid uncertain affine relaxation of the trivial function $f(x) = 0$ in the non-negative orthant. We consider the region $Q_j = [2, 3]^2$ and plot $A_i(\alpha)^{-1}Q_j$ for 10 uniformly spread values of $\alpha \in [0, 1]$. Note that in this case $\overline{A}_i = I$, hence $Q_j = A_i(1)^{-1}Q_j$ and Q_j (blue) is contained in the convex hull (pink). $Q_{ij}(\omega)$ can take on complex shapes and no method exists for exactly computing $Q_{ij}(\omega)$. Therefore, we compute a sound over-approximation $\overline{Q}_{ij}(\omega) \supset Q_{ij}(\omega)$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

memory requirements and computational time. In particular, the cardinality of the Cartesian product of I_{X_s}, I, and D can already be prohibitively large for relatively small systems. In the next section, we will discuss algorithmic strategies that make (FSBP) computationally tractable.

7. Algorithms for program construction

In this section, we discuss three aspects that allows one to solve (SBP) efficiently. Namely, in Section 7.1, we discuss how to compute a polyhedral over-approximation of $Q_{ij}(\omega)$. In Section 7.2, we introduce an a-priori sample discarding procedure for Problem (FSBP), which guarantees the same optimal solution, but allowing one to consider less samples in the optimisation problem. Finally, in Section 7.3, we employ spatial indexing methods to efficiently find the set of triplets with a non-empty $Q_{ij}(\omega)$. More specifically, we will rely on the fact that, often, for many triplets (i, j, ω), the set $Q_{ij}(\omega)$ is empty, thus the martingale constraint is trivially satisfied. That is, $Q_{ij}(\omega)$ is empty if $\mathrm{im}_i(Q_i, \omega) \cap Q_j = \emptyset$ where the image for region Q_i is defined as $\mathrm{im}(Q_i, \omega) = \{y + \eta(\omega) : x \in Q_i, y \in F(x)\}$.

7.1. Over-approximation of $Q_{ij}(\omega)$

As illustrated in Fig. 3, computing a polyhedral over-approximation of $Q_{ij}(\omega)$ is challenging, as $Q_{ij}(\omega)$ is (possibly) non-convex due to the uncertain[5] affine transformation F. Furthermore, it is not sufficient to compute the convex hull for the vertices of the uncertainty variable α, that is,

$$\mathrm{conv}\left(\left\{x \in Q_i : \hat{f}_i(x, 0) + \eta(\omega) \in Q_j\right\} \cup \left\{x \in Q_i : \hat{f}_i(x, 1) + \eta(\omega) \in Q_j\right\}\right), \tag{18}$$

as shown in Fig. 3. However, we note that by definition $Q_{ij}(\omega) \subseteq Q_i$, that is, Q_i is a, generally conservative, polyhedral over-approximation of $Q_{ij}(\omega)$. Hence, our approach to find a polyhedral over-approximation of $Q_{ij}(\omega)$, denoted $\overline{Q}_{ij}(\omega)$, is to start from Q_i and then iteratively removing subsets of $Q_i \setminus Q_{ij}(\omega)$. To accomplish this, we rely on repeated bisection. To simplify the presentation, in what follows, we assume that Q_i is a hyperrectangle. Note, however, that the procedure generalises to compact polyhedra in half-space representation.

Fig. 5 shows an example of the bisection algorithm for two regions Q_i, Q_j and a given sample ω. The bisection is repeated twice along each axis, once to increase the lower bound, and once to decrease the upper bound. Note that $Q_{ij}(\omega)$, although depicted in Fig. 5, is unknown and possibly non-convex, and further we cannot readily check if $Q_{ij}(\omega) \subset \overline{Q}_{ij}(\omega)$. However, recall that $Q_{ij}(\omega)$ is the subset of region Q_i that under a realisation of the noise ω reaches region Q_j in one time step. As a result, we can instead check if $\mathrm{conv}(\overline{Q}_{ij}(\omega) + \eta(\omega))$ intersects with Q_j. By applying this algorithm, we compute an over-approximation $\overline{Q}_{ij}(\omega)$ of $Q_{ij}(\omega)$. Indeed, as $\overline{Q}_{ij}(\omega)$ is an over-approximation, using $\overline{Q}_{ij}(\omega)$ for the constraints in Proposition 6 yields a sound, although slightly conservative, solution. The computational complexity of the bisection algorithm, for a fixed number of bisection steps t per halfplane and assuming that Q_i is hyperrectangular, $O(2n\ell t)$.

7.2. Convex hull over the sample set

To reduce the number of constraints of (FSBP), observe that the constraints of the problem are affine in $\eta(\omega)$. This implies that the active constraints, also known as support constraints, will always belong to the vertices of the convex hull over $\eta(D) = \{\eta(\omega_1), \ldots, \eta(\omega_N)\}$, enabling a great reduction in the number of constraints in the program. That is, we can impose the constraint only on the following set of samples:

$$\bar{D} = \{\omega \in D : \eta(\omega) \in \mathrm{vert}(\mathrm{conv}(\eta(D)))\}. \tag{19}$$

In the experiments conducted (see Section 8), we find that generally, in practice, the cardinality of $\bar{D}$ is orders of magnitude lower than the cardinality of D. Thus, this can greatly improve the efficiency of our approach.

Remark 3. The method presented in this subsection was discovered independently, but is similar to the approach in Sartipizadeh and Acikmese (2018) with the exception that our method requires an exact convex hull rather than an approximate convex hull, which is important for its sound use in (SBP). Note that our approach for sample reduction works for any scenario program that is affine in the random variable $\eta(\omega)$. □

7.3. Spatial indexing for intersection search

Constructing (FSBP) efficiently is also a non-trivial problem due to memory limits. In fact, a naïve approach to construct the problem is to iterate over all triplets (i, j, ω) in $I_{X_s} \times I \times D$, check if $Q_{ij}(\omega) \neq \emptyset$, and add a set of constraints if the test is positive. This approach is only tractable for smaller problems as it has time complexity $\mathcal{O}(|\mathcal{Q}|^3)$. To reduce the complexity, we can exploit methods from database theory; namely spatial indexing, which is the structuring and querying of spatially distributed data, such as maps, with improved computational complexity (Guttman, 1984).

To apply spatial indexing to our setting, we must first establish the data and query. It holds that $Q_{ij}(\omega) \neq \emptyset$ only if $\mathrm{im}_i(Q_i, \omega) \cap Q_j \neq \emptyset$. Hence, if we search for regions Q_j that intersect with the image $\mathrm{im}_i(Q_i, \omega)$, we find all pairs (i, j) such that $Q_{ij}(\omega) \neq \emptyset$. For spatial indexing, we focus on R-trees as it is a well-studied and widely available method (Guttman, 1984). The idea is to structure the data in a tree structure and at each node store a Minimum Bounding Rectangle (MBR) for the nodes below. Then, querying the tree for the intersection with another region proceeds recursively down the tree, where it is only necessary to search down a branch if the MBR of the branch and the query

[5] If F is a deterministic affine transformation and Q_j is a polyhedron, then $Q_{ij}(\omega)$ is also a polyhedron and analytical methods for computation exist.

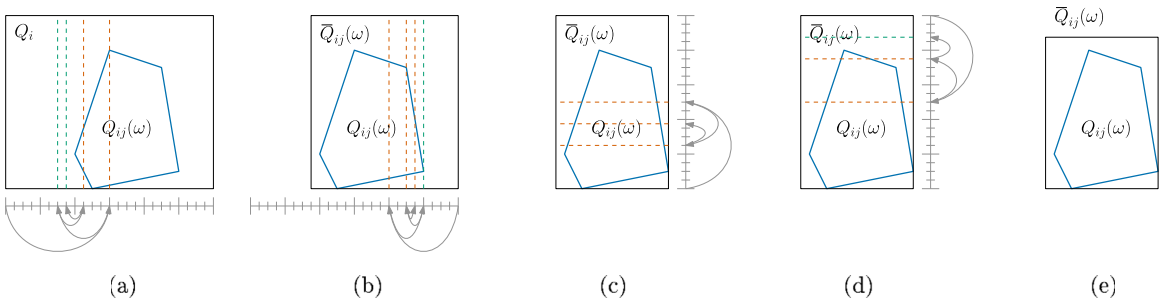


Fig. 5. An example of the bisection algorithm to compute the over-approximation $\overline{Q}_{ij}(\omega)$ of $Q_{ij}(\omega)$. The set $Q_{ij}(\omega)$ is unknown and possibly non-polyhedral, but Q_i is a sound over-approximation. By bisection from either side (first the lower then the upper bound) along each axis we obtain a smaller over-approximation. We start by bisecting for x_1 ((a) and (b)) followed by x_2 ((c) and (d)). This results in the over-approximation $\overline{Q}_{ij}(\omega)$ in (e).

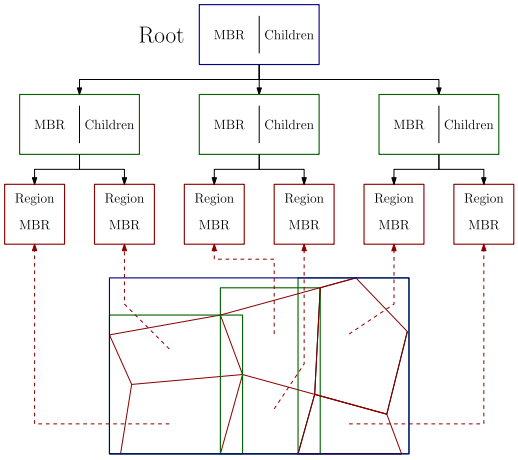


Fig. 6. An example of an R-tree applied to a partitioned state space to allow efficient search for regions intersecting with $\text{im}_i(Q_i, \omega)$. The rectangles are the Minimum Bounding Rectangle for each node in the tree.

intersect, which is an inexpensive operation by the separating hyperplane theorem (Boyd & Vandenberghe, 2004). Fig. 6 shows an example of an R-tree for a partitioned state space. This method improves complexity by efficiently searching for relevant triplets (i, j, ω).

In summary, to use the framework to compute data-driven safety certificates for non-linear systems: let the nominal dynamics f, initial and safe set X_0, X_s, a horizon K, and a dataset of samples D be given. Then start by abstracting the non-linear dynamics f to uncertain PWA dynamics $\hat{f}$ using LBP techniques. Compute the vertices $\bar{D}$ of the convex hull of D and discard all interior points. For each region Q_i, find, using an R-tree, regions that intersect with the image of the dynamics $\text{im}_i(Q_i, \omega)$ and add constraints accordingly. Solve the LP problem (FSBP), then the solution $z^\star(\bar{D}) = (c^\star, \gamma^\star, \theta^\star)$ is a safety certificate $\zeta(X_s, K) \geq 1 - (\gamma^\star + c^\star K)$ with confidence $1 - \beta$ where β is defined as in Proposition 3.

8. Experiments

To investigate the efficacy of our approach, we implemented our framework in Julia[6] and performed an empirical analysis on various benchmarks. The experiments have been conducted on a computer with an Intel i7-1365U CPU and 16 GB RAM, running Linux 6.8.0-52-generic. We start by describing the benchmarks followed by the results. For a comparison with state-of-the-art, we consider the Sample Average Approximation (SAA) method proposed in Salamati et al. (2021) and Salamati and Zamani (2023).[7] As this SBF synthesis method has been developed specifically for linear or polynomial systems, namely through SoS optimisation, in order to provide a general baseline, in the case of non-polynomial and/or uncertain systems, we combine it with the method proposed in Mazouz et al. (2022), to find a valid SBF in the general case. We also compare against (Gracia et al., 2024), which is a framework for data-driven abstraction.

8.1. Benchmarks

First, for all benchmarks, with our method, we use a uniform upper bound on the barrier of $M = 2$. The simplest system considered is an uncertain 1D linear system $x(k+1) = x(k) + b(\alpha) + \eta$ where $b(\alpha) = -0.05 + 0.1\alpha$, i.e. the uncertainty is in $b(\alpha)$ with uncertainty variable $\alpha \in [0, 1]$. Starting from a set $X_0 = \{|x| \leq 0.5\}$, the goal is to stay within a larger set $X_s = \{|x| \leq 2.5\}$. The distribution of the noise is a zero-mean normal distribution with standard deviation 0.01.

We also consider a 2D system from Badings, Romao, Abate and Jansen (2023) representing the longitudinal dynamics of a drone. The coordinates x_1, x_2 are the position and velocity, respectively, and the system has the following dynamics $x(k+1) = \begin{pmatrix} 1 & \tau \\ 0 & 1 - 0.1\tau \end{pmatrix} x(k) + \eta$, where η has a zero-mean normal distribution with diagonal covariance of [0.01 0.01]. The variable τ represents the discretisation step, which is set to $\tau = 1.0$. As with the 1D linear system, we certify safety for a horizon $K = 10$.

The third benchmark represents a model of a vehicle travelling down a straight road when it experiences an (uncertain) gust of wind. The coordinates x_1, x_2 represent, respectively, the longitudinal and lateral position of the vehicle. Similar to the drone system, τ represents the discretisation step. The goal is certify the probability of staying on the road $X_s = \{|x_1| \leq 2.5\}$ for a horizon $K = 10$, when the system evolves according to the following dynamics $x(k+1) = \begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix} x(k) + \begin{pmatrix} \frac{50}{3.6} \cdot \tau \\ \frac{1}{2} a_{lat} \cdot \tau^2 \end{pmatrix} + \eta$, where $\tau = 1$, and $a_{lat} = 0$ for regions where $x_1 \leq 80$ or $x_1 \geq 120$, and $a_{lat} \in [0.0913, 0.364]$ for regions where $x_1 \leq 80$ or $x_1 \geq 120$. η has a zero-mean normal distribution with diagonal covariance of [0.01 0.01].

A 3D benchmark with linear dynamics is the room temperature verification from Salamati et al. (2024). The dynamics are

[6] Code is available at https://github.com/DAI-Lab-HERALD/scenario-barrier under GNU GPLv3 licence.

[7] We compare only with the SAA and not the scenario approach for robust constraints such that the confidence level for both methods pertain to the probability that the c-martingale constraint in Condition (8) is satisfied.

$x(k+1) = Ax(k) + b + \eta$ where

$$A = \begin{pmatrix} 1-\tau(\alpha+\alpha_e) & \tau\alpha & 0 \\ \tau\alpha & 1-\tau(2\alpha+\alpha_e) & \tau\alpha \\ 0 & \tau\alpha & 1-\tau(\alpha+\alpha_e) \end{pmatrix},$$

$b = \begin{pmatrix} \tau\alpha_e T_e & \tau\alpha_e T_e & \tau\alpha_e T_e \end{pmatrix}^\top$ with $\tau = 5$, $\alpha = 0.0062$, $\alpha_e = 0.008$, $T_e = 10$. η has a zero-mean normal distribution with diagonal covariance of [0.0001 0.0001 0.0001]. The initial set is $X_0 = [17, 18]^3$, the safe set is $X_s = [17, 29]^3$, and the time horizon is $K = 3$.

While the previous models were linear, we also considered non-linear models. In particular, we consider Neural Network Dynamical Models (NNDMs) with 1 and 2 hidden layers of 64 neurons each modelling a pendulum taken from Mazouz et al. (2022). Second is the 3D Dubin's car benchmark from Mathiesen, Calvert et al. (2023) for a time horizon $K = 10$. Dubin's car is a unicycle model where the state is (x, y, ϕ) with ϕ being the heading of the vehicle. We consider a grid-based partitioning of 10 segments along each axis, i.e. 1000 regions. The noise is only applied to the last dimension and has a normal distribution with mean of $60 \cdot \frac{\pi}{180} \approx 1.053$ and standard deviation 0.1. Finally, we consider the lane keeping benchmarking from Salamati et al. (2024) where the dynamics are $x(k+1) = x(k) + \tau v \begin{pmatrix} \cos(x(k)_3 + b) \\ \sin(x(k)_3 + b) \\ \sin(b)/l_r \end{pmatrix} + \eta$, with $b = \frac{l_r \tan^{-1}(\delta_f)}{l_r + l_f}$, $\delta_f = 5$, $l_r = l_f = 1.384$, $v = 5$, and $\tau = 0.1$. The noise η has a zero-mean normal distribution with diagonal covariance of [0.0001 0.0001 0.000001]. The initial set is $X_0 = [1, 2] \times [-0.5, 0.5] \times [-0.005, 0.005]$, the safe set is $X_s = [1, 10] \times [-6, 6] \times [-0.05, 0.05]$, and the time horizon is $K = 3$.

8.2. Results

Table 1 reports the results across all benchmarks. Both the safety probability and the computation time are reported averaged over 100 trials, and for all cases, the number of samples is selected to ensure a confidence $1 - 10^{-9}$. We observe that, depending on the system, the method can certify safety to $> 99\%$ with high confidence, e.g. 99.5% certified safety for the NNDM model of a pendulum with 2 layers and 64 neurons. This certification requires relatively few regions of 10–30 segments per axis. Comparing the NNDM pendulum model with 2 and 3 layers (1 and 2 hidden layers, respectively), the complexity of the nominal dynamics impacts both computation time and certifiable safety, e.g. 99.5% safety probability in 45.0s vs. 97.6% safety probability in 78.5s for 480 regions, 2 and 3 layers respectively. This behaviour can be explained by LBP computing wider uncertain affine transformations lead to more non-empty $Q_{ij}(\omega)$. Remarkably, the linearity of the underlying system has little impact on the certifiable safety. This is observed in that both the 1D linear and drone systems exhibit uncertain linear behaviour, yet the 1D linear system is certifiable to 50.8% safety while the drone is certifiable to 99.5%. Furthermore, for the largest systems considered, Dubin's car and lane keeping, which both include trigonometric functions, safety is certified to 99.9%.

To compare against state-of-the-art, we report in Table 2 the certified safety by using our method, SAA (Salamati et al., 2021), and data-driven abstractions (Gracia et al., 2024), averaged over 100 trials. We observe that for our method, certifying for higher confidence (10^{-2} to 10^{-4}) has a small impact in the computation time (1.2% faster for the 1D linear system and 0.5% slower for the pendulum model) and achieves similar levels of certified safety. In contrast, for SAA, going from 10^{-2} to 10^{-3} increases the computation time between 1.3x and 1.9x and become

Table 1
Certified safety and computation time using our approach. Results are reported as the mean over 100 iterations for each case study. n is the dimensionality of the system and ℓ is the number of regions. $\zeta(X_s, K)$ is the certified level of safety for $\beta = 10^{-9}$ where $1 - \beta$ is the level of confidence.

System	n	ℓ	$\zeta(X_s, K)$	Time (s)
Linear	1	27	0.511	0.359
Drone	2	37	0.995	44.7
Vehicle	2	18	0.606	1.89
		42	0.710	4.28
		54	0.828	5.24
		150	0.995	13.6
Room temperature	3	2197	0.998	356
Pendulum (NNDM)		120	0.346	11.4
- 2 layers	2	240	0.752	28.3
- 64 neurons		480	0.995	71.4
Pendulum (NNDM)		120	0.265	28.33
- 3 layers	2	240	0.376	68.9
- 64 neurons		480	0.980	130
Dubin's car	3	1000	0.999	321
Lane keeping	3	2520	1.000	33.6

Table 2
Comparison of our method against Sample Average Approximation (SAA) combined with the method presented in Mazouz et al. (2022), to synthesise SBFs in a data-driven fashion, and against (Gracia et al., 2024) for data-driven abstractions. We do not compare with (Gracia et al., 2024) on the pendulum model, since the Matlab implementation does not support uncertain PWA dynamics. Results are reported as the mean over 100 iterations for each case study. $1 - \beta$ is the confidence in the certificate and $\zeta(X_s, K)$ is the certified level of safety. OOM means the certification procedure exceeded available memory. The number of samples required for our method and Gracia et al. (2024) increases from $\approx 3.1 \cdot 10^6$ to $\approx 3.4 \cdot 10^6$ with β for the drone and pendulum benchmarks and from $\approx 4.3 \cdot 10^6$ to $\approx 4.6 \cdot 10^6$ for the room temperature benchmark, while for SAA, the number grows from $8 \cdot 10^6$ to $8 \cdot 10^{13}$ as $\beta \to 10^{-9}$ for all benchmarks.

System	Method	β	$\zeta(X_s, K)$	Time (s)
Drone	Ours	10^{-2}	0.995	0.403
		10^{-3}	0.995	0.361
		10^{-4}	0.995	0.392
		10^{-9}	0.995	0.398
	SAA	10^{-2}	0.995	0.595
		10^{-3}	0.995	0.798
		10^{-4}	–	OOM
		10^{-9}	–	OOM
	Data-driven abstr.	10^{-2}	0.989	6.84
		10^{-3}	0.988	6.39
		10^{-4}	0.986	6.63
		10^{-9}	0.981	6.70
Room temp.	Ours	10^{-2}	0.995	445
		10^{-3}	0.995	446
		10^{-4}	0.995	437
		10^{-9}	0.995	432
	SAA	10^{-2}	–	OOM
		10^{-3}	–	OOM
		10^{-4}	–	OOM
		10^{-9}	–	OOM
	Data-driven abstr.	10^{-2}	0.997	884
		10^{-3}	0.997	833
		10^{-4}	0.996	886
		10^{-9}	0.995	901
Pendulum (NNDM) - 2 layers - 64 neurons - ReLU act. - 480 regions	Ours	10^{-2}	0.995	38.6
		10^{-3}	0.995	38.7
		10^{-4}	0.995	38.7
		10^{-9}	0.995	38.8
	SAA	10^{-2}	0.903	14.4
		10^{-3}	0.903	27.8
		10^{-4}	–	OOM
		10^{-9}	–	OOM

infeasible already for $\beta = 10^{-4}$. The achieved level of certified safety is also marginally better with our proposed method (e.g. 0.995 vs. 0.903 for the pendulum model). When comparing to data-driven abstractions, the certified probability of safety is comparable or marginally better. However, the computation is considerably faster; 17x for the drone benchmark and 2x for the room temperature benchmark.

9. Conclusion

We have presented a data-driven method to synthesise piece-wise affine Stochastic Barrier Function (SBF) based on a novel inner-approximation, which relies on the scenario theory. Our approach employs Linear Bound Propagation and stochastic approximations to guarantee that the search for a barrier reduces to solving a Linear Programming problem, which can be solved efficiently using the convex hull over the noise samples and spatial indexing for faster searching. As with any method, ours comes with limitations. In fact, in our approach we assume that the noise is additive and that we have i.i.d. full measurements of the state of the system. In particular, we rely on the additivity of the noise to achieve efficient algorithms in practice. How to extend our approach to non-additive noise represents an important open question. Another key limitation of the proposed framework is the reliance on linear relaxations, which as a consequence if the system is highly nonlinear, requires fine partitioning, limiting scalability. Possible approaches to scale to higher dimensions is to rely on compositional methods, such as Nejati et al. (2022), which are orthogonal to our approach and can in principle be combined, or on refinement procedures (Dutreix & Coogan, 2021; Esmaeil et al., 2013). A direction for future research to improve scalability is to consider more complex piece-wise templates for parameterising an SBF, e.g. piece-wise quadratic barrier candidates. Using more complex templates may require fewer pieces and thus improve synthesis performance. For more complex templates, one challenge is that strong duality, on which we have relied heavily, does not necessarily hold. Another direction of future work is to treat the safe control synthesis problem within the proposed framework, where a challenge is that the optimisation problem easily becomes bilinear.

Acknowledgments

This research has been supported by EPSRC grant EP/Y028 872/1, Mathematical Foundations of Intelligence: An "Erlangen Programme" for AI.

Appendix A. Proof and auxiliary results

A.1. Proof of Proposition 2

Proof. Let the sets $\mathcal{Z}$ and

$$\mathcal{Z} = \{z : (Az+a)^\top x \le Bz+b, \text{ for all } x \in P\}, \tag{A.1}$$

$$\mathcal{Z}'' = \{(z,\lambda) : h^\top\lambda \le Bz+b,\ H^\top\lambda = Az+a,\ \lambda \ge 0\}, \tag{A.2}$$

be the feasible sets of (3) and (4), respectively. We need to show that $\mathcal{Z} = \text{proj}_z(\mathcal{Z}'') = \mathcal{Z}'$, where $\mathcal{Z}' = \{z : \exists\lambda \in \mathbb{R}^p_{\ge 0},\ h^\top\lambda \le Bz+b,\ H^\top\lambda = (Az+a)\}$ is the projection of the set $\mathcal{Z}''$ onto its first d coordinates. At the core of this result is the strong duality (Boyd & Vandenberghe, 2004, Section 5.2.1) between the linear programs

$$\max_x \ (Az+a)^\top x \quad \text{s.t.} \quad Hx \le h, \qquad \min_\lambda \ \lambda^\top h \quad \text{s.t.} \quad H^\top\lambda = (Az+a),\ \lambda \ge 0,$$

which states that for any element $x_\star$ in the optimal set of the maximisation problem there exist a $\lambda_\star$ in the optimal set of the minimisation problem such that $(Az+a)^\top x_\star = \lambda_\star^\top h$ and $\lambda_\star^\top(H^\top x_\star - h) = 0$. Vice-versa, for all $\lambda_\star$ in the optimal set of the minimisation problem there exists a $x_\star$ in the optimal set of the maximisation problem such that similar conclusions hold. Pick any element $\bar z \in \mathcal{Z}$. Let $\bar x$ be such that $(A\bar z+a)^\top\bar x = \sup_{x\in P} a(\bar z)^\top x$. Hence, by strong duality, there exists a $\bar\lambda \ge 0$ with $H^\top\bar\lambda = (A\bar z+a)$ such that $\bar\lambda^\top h = (A\bar z+a)^\top\bar x \le B\bar z+b$. In other words, there is a $\bar\lambda$ such that $(\bar z,\bar\lambda) \in \mathcal{Z}''$, which implies $\mathcal{Z} \subseteq \mathcal{Z}'$. For the other direction, consider a tuple $(\bar z,\bar\lambda) \in \mathcal{Z}''$. For this given $\bar z$, pick $\lambda_\star(\bar z) \in \arg\min_{H^\top\lambda = A\bar z+a,\ \lambda\ge 0} \lambda^\top h$, we notice that we have that $\bar z \in \mathcal{Z}'$ and $\sup_{x\in P}(A\bar z+a)^\top x = \lambda_\star(\bar z)^\top h \le \bar\lambda^\top h \le B\bar z+b$. The right-most inequality follows from feasibility of $(\bar z,\bar\lambda)$, the middle inequality by our choice of $\lambda_\star(\bar z)$, the left-most equality by strong duality of linear programming. Then we conclude that $\bar z \in \mathcal{Z}$, thus concluding the proof. □

A.2. Proof of Proposition 4

Proof. Fix $j \in \{0,\dots,n\}$ and a convex region P_i. Then by the mean value theorem for generalised gradients (Clarke, 1981), there exists two hyperplanes $(\underline{A}_i x + \underline{b}_i)_j$, $(\overline{A}_i x + \overline{b}_i)_j$ such that $(\underline{A}_i x + \underline{b}_i)_j \le f(x)_j \le (\overline{A}_i x + \overline{b}_i)_j$ for all $x \in Q_i$. Combining all dimensions, it holds that $\underline{A}_i x + \underline{b}_i \le f(x) \le (\overline{A}_i x + \overline{b}_i)$ for all $x \in Q_i$ for each region Q_i. As a result, by the definition of $\hat f$ it holds that $f(x) \in \{\hat f(x,\alpha) : \alpha \in [0,1]\}$ for all $x \in X$ concluding the proof. □

A.3. Measurability issue of Theorem 1

We need to show that set

$$E = \{\omega \in \Omega : B(f(x)+\eta(\omega)) + \xi \le B(x) + c, \text{ for all } x \in X_s\}, \tag{A.3}$$

is Borel measurable. First of all, notice that due to the fact that $\{Q_1,\dots,Q_\ell\}$ is a finite partition of X_s, we have that $E = \bigcap_{Q_i\in X_s}\{\omega \in \Omega : \sup_{x\in Q_i} B(f(x)+\eta(\omega)) - B_i(x) \le c-\xi\}$. As finite intersections of measurable sets are still measurable and a measurable function maps measurable sets into measurable sets, it is enough to show that each set $E_i = \{\eta \in \mathbb{R}^n : \sup_{x\in Q_i} B(f(x)+\eta) - B_i(x) \le c-\xi\}$ is measurable. In order to do that note that when restricted to Q_i, B_i is a linear function, while by construction $B(x)$ is upper semi-continuous. Furthermore, as composition of an upper semi-continuous function with a continuous one is still upper semi-continuous, we have that both $B(f(x)+\eta)$ and $B(f(x)+\eta)-B_i(x)$ are upper semi-continuous functions. Consequently, by Proposition 7.32 in Bertsekas and Shreve (2004) we have that $g_i(\eta) = \sup_{x\in Q_i} B(f(x)+\eta) - B_i(x)$ is upper semi-continuous. As $g_i(\eta)$ is upper semi-continuous, hence Borel measurable, we have that set E_i is Borel measurable, thus concluding the proof.

A.4. Proof for Proposition 6

Proof. Start by fixing $\alpha \in [0,1]$, the pair $(i,j) \in I_{X_s} \times I$, and the noise sample $\omega \in D$. Then the constraint $B(\hat f(x,\alpha)+\eta(\omega),\theta)+\xi \le B(x,\theta)+c$ for all $x \in X_s$ can be rewritten as $B_j(A_i(\alpha)x + b_i(\alpha) + \eta(\omega),\theta)+\xi \le B_i(x,\theta)+c$, for all $x \in \overline{Q}_{ij}(\omega)$. Using Proposition 2, we rewrite this constraint into a dual formulation so that we can solve it in a lifted space. Thus, they become

$$h_{ij\omega}^\top\lambda_{ij\omega} \le v_i - v_j - u_j^\top(b_i(\alpha)+\eta(\omega)) + c - \xi,$$
$$H_{ij\omega}^\top\lambda_{ij\omega} = A_i(\alpha)^\top u_j - u_i,$$

where $\lambda_{ij\omega}$ is a non-negative dual variable. Since both constraints are affine in α, they will hold for all $\alpha \in [0,1]$ if and only if they hold for the vertices, that is, $\alpha \in \{0,1\}$. We conclude the proof by observing that for each noise sample $\omega \in D$, the union of $\overline{Q}_{ij}(\omega)$ for all $(i,j) \in I_{X_s} \times I$ is a superset of X_s. □

Appendix B. Algorithm for computing over-approximation of $Q_{ij}(\omega)$

Our approach to computing an over-approximation of $Q_{ij}(\omega)$ is detailed in Alg. 1, where we start with a hyperrectangle $\overline{Q}_{ij}(\omega) = Q_i = \{x \in \mathbb{R}^m : l \leq x \leq u\}$, with $\leq$ interpreted element-wise. The idea to reduce the size of $\overline{Q}_{ij}(\omega)$ is to increase l (Line 5–14) and decrease u (Line 16–25) while maintaining the over-approximation of $Q_{ij}(\omega)$. Treating the axes sequentially (Line 3), denoting the current axis by k, we use bisection first to find the largest l_k and then the smallest u_k such that $\overline{Q}_{ij}(\omega) \cap \{x : l_k \leq x_k \leq u_k\}$ is an over-approximation of $Q_{ij}(\omega)$. Then we can replace $\overline{Q}_{ij}(\omega)$ with $\overline{Q}_{ij}(\omega) \cap \{x : l_k \leq x_k \leq u_k\}$ for the next axis (Line 27). To perform the bisection for increasing l_k, we compute the midpoint c_k^l between l_k and u_k (Line 7), denoting them l_k^l and u_k^l respectively, and test if $Q_{ij}(\omega) \subset \overline{Q}_{ij}(\omega) \cap \{x : c_k^l \leq x_k\}$. If true, then we may let $l_k^l = c_k^l$ (Line 10), and if not, let $u_k^l = c_k^l$ (Line 12). This procedure repeats for a fixed number of iterations and is performed mutatis mutandis to decrease u_k. The question remains how to check if $Q_{ij}(\omega) \subset \overline{Q}_{ij}(\omega) \cap \{x : c_k^l \leq x_k\}$, since $Q_{ij}(\omega)$ is unknown. To this end, recall that $Q_{ij}(\omega)$ is the subset of region Q_i that under a realisation of the noise ω reaches region Q_j in one time step. Thus, if the image of the other subregion $\text{im}(\overline{Q}_{ij}(\omega) \cap \{x : x_k \leq c_k^l\}, \omega)$ under the realisation of the noise ω does not intersect Q_j (Line 9), then it holds that $Q_{ij}(\omega) \subset \overline{Q}_{ij}(\omega) \cap \{x : c_k^l \leq x_k\}$.

Algorithm 1 Bisection-based algorithm for computing a subset $\overline{Q}_{ij}(\omega)$ of region Q_i as an over-approximation of $Q_{ij}(\omega)$.

1: **function** PolyPreimage(Q_i, Q_j, $\hat{f}$, ω, t)
2: $\quad\overline{Q}_{ij}(\omega) \leftarrow Q_i$
3: $\quad$**for** $k \leftarrow 1$ to m **do** ▷ For each axis
4:
5: $\quad\quad l_k^l, u_k^l \leftarrow l_k, u_k$ ▷ Increase lower bound
6: $\quad\quad$**for** $s \leftarrow 1$ to t **do**
7: $\quad\quad\quad c_k^l \leftarrow \frac{l_k^l + u_k^l}{2}$
8: $\quad\quad\quad \overline{Q}_{ij}(\omega)' \leftarrow \overline{Q}_{ij}(\omega) \cap \{x : x_k \leq c_k^l\}$
9: $\quad\quad\quad$**if** $\text{im}_i(\overline{Q}_{ij}(\omega)', \omega) \cap Q_j = \emptyset$ **then**
10: $\quad\quad\quad\quad l_k^l \leftarrow c_k^l$
11: $\quad\quad\quad$**else**
12: $\quad\quad\quad\quad u_k^l \leftarrow c_k^l$
13: $\quad\quad\quad$**end if**
14: $\quad\quad$**end for**
15:
16: $\quad\quad l_k^u, u_k^u \leftarrow l_k, u_k$ ▷ Decrease upper bound
17: $\quad\quad$**for** $s \leftarrow 1$ to t **do**
18: $\quad\quad\quad c_k^u \leftarrow \frac{l_k^u + u_k^u}{2}$
19: $\quad\quad\quad \overline{Q}_{ij}(\omega)' \leftarrow \overline{Q}_{ij}(\omega) \cap \{x : x_k \geq c_k^u\}$
20: $\quad\quad\quad$**if** $\text{im}_i(\overline{Q}_{ij}(\omega)', \omega) \cap Q_j = \emptyset$ **then**
21: $\quad\quad\quad\quad u_k^u \leftarrow c_k^u$
22: $\quad\quad\quad$**else**
23: $\quad\quad\quad\quad l_k^u \leftarrow c_k^u$
24: $\quad\quad\quad$**end if**
25: $\quad\quad$**end for**
26:
27: $\quad\quad\overline{Q}_{ij}(\omega) \leftarrow \overline{Q}_{ij}(\omega) \cap \{x : l_k^l \leq x_k \leq u_k^u\}$
28: $\quad$**end for**
29: $\quad$**return** $\overline{Q}_{ij}(\omega)$ ▷ It holds that $Q_{ij}(\omega) \subset \overline{Q}_{ij}(\omega)$
30: **end function**

Appendix C. Benchmark partition details

All methods considered in Section 8, i.e., our method, (Gracia et al., 2024; Salamati et al., 2021), requires a partitioning of the safe set. For our method and Salamati et al. (2021), this corresponds to the piece-wise template of the barrier, while for Gracia et al. (2024), the partitioning is used to build to the finite-state model. We detail in this appendix the partitioning employed for each benchmark. The same partitioning is used for all methods.

Linear. The barrier template considered has 25 regions of equal width, between -2.5 and 2.5. That is, each region is 0.2 wide.

Drone. The safe set for the drone benchmark is partitioned uniformly along each axis into 24 and 20 segments, respectively.

Vehicle. The safe set is partitioned in an axis-aligned grid. Therefore, the regions are described by the discretisation along each dimension:

$$x_1 : \{(-\infty, 0], [0, 10], [10, 80], [80, 120], [120, 150], [150, \infty)\}$$
$$x_2 : \{(-\infty, -2.5], [-2.5, 0.5], [-0.5, 0.5], [0.5, 2.5], [2.5, \infty)\}$$

This discretisation yields 30 regions, which we refine by splitting longitudinally, laterally, or both in equal-width slices. To split longitudinally, we split each of $[0, 10]$, $[10, 80]$, $[80, 120]$, $[120, 150]$ into 5 narrower slices, and to split laterally, we split $[-2.5, 0.5]$, $[-0.5, 0.5]$, $[0.5, 2.5]$ in 20, 10, and 20 narrower slices respectively. The number of regions by splitting longitudinally is 110, laterally is 312, and 1144 for splitting in both directions.

Room temperature. The safe set for the room temperature benchmark is a uniform grid partition of 13 segments along each axis.

NNDM Pendulum. The safe set of the NNDM pendulum benchmark is the set $[-12°, 12°] \times [-1, 1]$. For all partitions, the second axis, corresponding to the angular velocity, is partitioned into 10 segments. Then, the angular position is partitioned into 12, 24, and 48 segments, for 120, 240, and 480 total number of regions, respectively.

Dubin's car. The safe set for the Dubin's car benchmark is partitioned uniformly into 10 segments per axis.

Lane keeping. The safe set for the lane keeping benchmark is partitioned uniformly along each axis into 18, 14, and 10 segments, respectively.

References

Abate, A. (2017). Formal verification of complex systems: Model-based and data-driven methods. In *MEMOCODE*. http://dx.doi.org/10.1145/3127041.3131362.

Badings, T., Romao, L., Abate, A., & Jansen, N. (2023). Probabilities are not enough: Formal controller synthesis for stochastic dynamical models with epistemic uncertainty. In *AAAI*.

Badings, T., Romao, L., Abate, A., Parker, D., Poonawala, H. A., Stoelinga, M., & Jansen, N. (2023). Robust control for dynamical systems with non-gaussian noise via formal abstractions. *JAIR*.

Baier, C., & Katoen, J. P. (2008). *Principles of model checking*. MIT Press Books.

Bertsekas, D. P., & Shreve, S. E. (2004). *Stochastic optimal control: the discrete-time case*. Athena Scientific.

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambride University Press.

Campi, M., & Garatti, S. (2008). The exact feasibility of randomized solutions of uncertain convex programs. *SIAM Journal on Optimization*.

Cauchi, N., Laurenti, L., Lahijanian, M., Abate, A., Kwiatkowska, M., & Cardelli, L. (2019). Efficiency through uncertainty: Scalable formal synthesis for stochastic hybrid systems. In *HSCC*.

Clarke, F. H. (1981). Generalized gradients of Lipschitz functionals. *Advances in Mathematics*, http://dx.doi.org/10.1016/0001-8708(81)90032-3.

Cosner, R. K., Culbertson, P., Taylor, A. J., & Ames, A. D. (2023). Robust safety under stochastic uncertainty with discrete-time control barrier functions. arXiv:2302.07469.

De Giacomo, G., & Vardi, M. Y. (2013). Linear temporal logic and linear dynamic logic on finite traces. In *IJCAI*.

Dutreix, M., & Coogan, S. (2021). Specification-guided verification and abstraction refinement of mixed monotone stochastic systems. *IEEE Transactions on Automatic Control*, http://dx.doi.org/10.1109/TAC.2020.3014142.

Esmaeil, S., Soudjani, Z., & Abate, A. (2013). Adaptive and sequential gridding procedures for the abstraction and the verification of stochastic processes. *SIAM Journal on Applied Dynamical Systems*, *2*, 921–956.
Girard, A., & Pappas, G. (2007). Approximation metrics for discrete and continuous systems. *IEEE TAC*, http://dx.doi.org/10.1109/TAC.2007.895849.
Gracia, I., Boskos, D., Laurenti, L., & Mazo, M. (2023). Distributionally robust strategy synthesis for switched stochastic systems. *HSCC*.
Gracia, I., Laurenti, L., Mazo, M., Jr., Abate, A., & Lahijanian, M. (2024). Temporal logic control for nonlinear stochastic systems under unknown disturbances. arXiv preprint arXiv:2412.11343.
Guttman, A. (1984). R-trees: A dynamic index structure for spatial searching. *SIGMOD Record*, http://dx.doi.org/10.1145/971697.602266.
Haesaert, S., & Soudjani, S. (2021). Robust dynamic programming for temporal logic control of stochastic systems. *IEEE TAC*, http://dx.doi.org/10.1109/TAC.2020.3010490.
Jackson, J., Laurenti, L., Frew, E., & Lahijanian, M. (2020). Safety verification of unknown dynamical systems via Gaussian process regression. *CDC*, http://dx.doi.org/10.1109/CDC42340.2020.9303814.
Jagtap, P., Soudjani, S., & Zamani, M. (2020). Formal synthesis of stochastic systems via control barrier certificates. *IEEE TAC*.
Joldes, M. M. (2011). *Rigorous polynomial approximations and applications* (Theses), Ecole normale supérieure de lyon - ENS LYON.
Kushner, H. J. (1967). *Stochastic stability and control*: *Technical report*, Brown Univ Providence RI.
Laurenti, L., & Lahijanian, M. (2026). A unifying perspective for safety of stochastic systems: From barrier functions to finite abstractions. *IEEE Transactions on Automatic Control*, *71*(2), 769–779. http://dx.doi.org/10.1109/TAC.2025.3597569.
Lavaei, A., Soudjani, S., Frazzoli, E., & Zamani, M. (2023). Constructing MDP abstractions using data with formal guarantees. *IEEE Control Systems Letters*.
Mathiesen, F. B., Calvert, S. C., & Laurenti, L. (2023). Safety certification for stochastic systems via neural barrier functions. *IEEE Control Systems Letters*.
Mathiesen, F. B., Romao, L., Calvert, S. C., Abate, A., & Laurenti, L. (2023). Inner approximations of stochastic programs for data-driven stochastic barrier function design. In *CDC*.
Mazouz, R., Mathiesen, F. B., Laurenti, L., & Lahijanian, M. (2026). Piecewise stochastic barrier functions. *Automatica*, *189*, Article 113020. http://dx.doi.org/10.1016/j.automatica.2026.113020.
Mazouz, R., Muvvala, K., Babu, A. R., Laurenti, L., & Lahijanian, M. (2022). Safety guarantees for neural network dynamic systems via stochastic barrier functions. In *NeurIPS*.
McAllister, R. D., & Rawlings, J. B. (2020). *Stochastic Lyapunov functions and asymptotic stability in probability*: *Technical report*, Texas - Wisconsin - California Control Consortium.
Nejati, A., Soudjani, S., & Zamani, M. (2022). Compositional construction of control barrier functions for continuous-time stochastic hybrid systems. *Automatica*, *145*, Article 110513. http://dx.doi.org/10.1016/j.automatica.2022.110513.
Prajna, S., Jadbabaie, A., & Pappas, G. (2007). A framework for worst-case and stochastic safety verification using barrier certificates. *IEEE TAC*.
Salamati, A., Lavaei, A., Soudjani, S., & Zamani, M. (2021). Data-driven safety verification of stochastic systems via barrier certificates. *IFAC-PapersOnLine*.
Salamati, A., Lavaei, A., Soudjani, S., & Zamani, M. (2024). Data-driven verification and synthesis of stochastic systems via barrier certificates. *Automatica*, *159*, Article 111323. http://dx.doi.org/10.1016/j.automatica.2023.111323.
Salamati, A., & Zamani, M. (2023). Safety verification of stochastic systems: A repetitive scenario approach. *IEEE Control Systems Letters*.
Santoyo, C., Dutreix, M., & Coogan, S. (2021). A barrier function approach to finite-time stochastic system verification and control. *Automatica*.
Sartipizadeh, H., & Acikmese, B. (2018). Approximate convex hull based sample truncation for scenario approach to chance constrained trajectory optimization. In *ACC*. http://dx.doi.org/10.23919/acc.2018.8430936.
Schön, O., Zhong, Z., & Soudjani, S. (2024). Data-driven distributionally robust safety verification using barrier certificates and conditional mean embeddings. arXiv:2403.10497.
Steinhardt, J., & Tedrake, R. (2012). Finite-time regional verification of stochastic non-linear systems. *The International Journal of Robotics Research*.
Tao, T. (2013). *An introduction to measure theory*. American Mathematical Society.
Xu, K., Shi, Z., Zhang, H., Wang, Y., Chang, K.-W., Huang, M., Kailkhura, B., Lin, X., & Hsieh, C.-J. (2020). Automatic perturbation analysis for scalable certified robustness and beyond. In *NeurIPS*.



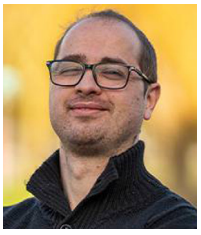
Frederik Baymler Mathiesen obtained his B.Sc. (2019) and M.Sc. (2021) degrees in Computer Science at Aalborg University, Denmark. He is currently pursuing a Ph.D. at Delft Center for Systems and Control, Delft University of Technology. The primary focus of his research concerns formal methods and machine learning for stochastic dynamical systems.



Licio Romao is an Assistant Professor in the Department of Wind and Energy Systems at the Technical University of Denmark (DTU). Previously, he held postdoctoral positions in the Computer Science Department at the University of Oxford and in the Department of Aeronautics and Astronautics at Stanford University. He obtained his Ph.D. from the University of Oxford and was awarded the IET Control & Automation Doctoral Dissertation Prize. Other awards include the 2022 Distinguished Paper Award (AAAI) and the 2024 Hybrid Systems TC Outstanding Student Paper Prize, both as a co-author. His research interests lie in applying mathematical engineering methods to reason about stochastic uncertainty in sequential decision-making and feedback control problems.



Simeon C. Calvert Dr. Simeon Calvert is associate professor of Smart & Automated Driving in the department of Transport & Planning at the TU Delft. He is director and founder of the Automated Driving & Simulation (ADaS) research lab and co-director of the CityAI-lab for research on urban behaviour using AI. He is also a board member of the Centre for Meaningful Human Control over AI, the interdisciplinary research program for responsible autonomous technology. His research is focussed on the impacts of technology on road traffic through experimentation, conceptualisation and simulation.



Luca Laurenti is a tenure track assistant professor at the Delft Center for Systems and Control at TU Delft and head of the Research on Robust and Intelligent Autonomous Systems Lab at AI4I. He received his Ph.D. from the Department of Computer Science, University of Oxford (UK), where he was a member of the Trinity College. Luca has a background in stochastic systems, control theory, formal methods, and artificial intelligence. His research work focuses on developing data-driven systems that are provably robust to interactions with a dynamic and uncertain world.



Alessandro Abate is Professor of Verification and Control in the Department of Computer Science at the University of Oxford. Earlier, he did research at Stanford University and at SRI International, and was Assistant Professor at TU Delft. He received a Ph.D./M.Sc. degrees from UC Berkeley and the University of Padua. His research spans logic, probability, control and decision theory.