Data-driven and Model-based Verification: a Bayesian Identification Approach

Sofie Haesaert [a], Paul M.J. Van den Hof [a], Alessandro Abate [b]

[a]*Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands*

[b]*Department of Computer Science, University of Oxford, Oxford, United Kingdom*

Abstract

This work develops a measurement-driven and model-based formal verification approach, applicable to systems with partly unknown dynamics. We provide a principled method, grounded on reachability analysis and on Bayesian inference, to compute the confidence that a physical system driven by external inputs and accessed under noisy measurements, verifies a temporal logic property. A case study is discussed, where we investigate the bounded- and unbounded-time safety of a partly unknown linear time invariant system.

Key words: Temporal logic properties, Bayesian inference, Linear time-invariant models, Model-based verification, Data-driven validation, Statistical model checking,

1 Introduction

The design of complex, high-tech, safety-critical systems such as autonomous vehicles, intelligent robots, and cyber-physical infrastructures, demands guarantees on their correct and reliable behaviour. Correct functioning and reliability over models of systems can be attained by the use of formal methods. Within the computer sciences, the formal verification of software and hardware has successfully led to industrially relevant and impactful applications [13]. Carrying the promise of a decrease in design faults and implementation errors and of correct-by-design synthesis, the use of formal methods, such as model checking [13], has become a standard in the avionics, automotive, and railway industries [34]. Life sciences [6,14] and general engineering applications [5,11] have also recently pursued the extension of these successful techniques from the computer science: this has required a shift from finite-state to physical and cyber-physical models that are of practical use in nowadays science and technology [23,32].

The strength of formal techniques, such as model checking, is bound to the fundamental requirement of having access to a given model, obtained from the knowledge

of the behaviour of the underlying system of interest. In practice, for most physical systems the dynamical behaviour is known only in part: this holds in particular with biological systems [1] or with classes of engineered systems where, as a consequence, the use of uncertain control models built from data is a common practice [22].

Only limited work within the formal methods community deals with the verification of models with partly unknown dynamics. Classical results [4,19] consider the verification problem for non-stochastic models described by differential equations and with bounded parametric uncertainty. Similarly, but for continuous time probabilistic models, [9,10] explore the parameter space with the objective of model verification (respectively statistical or probabilistic). Whenever full state measurements of the system are available, Statistical Model Checking (SMC) [31,24] replaces model(-based) checking procedures with empirical testing of formalised properties. SMC is limited to fully observable stochastic systems with little or no non-determinism, and may require the gathering a large set of measurements. Extensions towards the inclusion of non-determinism have been studied in [18,25], with preliminary steps towards Markov decision processes. Related to SMC techniques, but bound to finite state models, [12,27,30] assume that the system is encompassed by a finite-state Markov chain and efficiently use data to learn the corresponding model and to verify it. Similarly, [3,8] employ machine learning tech-

niques to infer finite-state Markov models from data over specific logical formulae.

An alternative approach, allowing both partly unknown dynamics over uncountable (continuous) variables and noisy output measurements, is the usage of a Bayesian framework relating the confidence in a formal property to the uncertainty of a model built from data. When applied on nonlinearly parameterised linear time invariant (LTI) models this approach introduces huge computational problems, which as proposed in [16], can only be mitigated by statistical methods. Instead, to obtain reliable and numerical solutions, we propose the use of linearly parameterised model sets defined through orthonormal basis functions to represent these partially unknown systems. This is a broadly used framework in system identification [21,22]: it allows for the incorporation of prior knowledge, while maintaining the benefits (computational aspects) of linear parameterisations. Practically, it has been widely used for the modelling of physical systems, such as the thermal dynamics of buildings [35]. In contrast, in this paper we pursue a promising new numerical approach: instead of employing directly a nonlinearly parameterised model, we embed it in a linearly parameterised one via a series expansion of orthonormal basis functions.

In this contribution we further analyse and extend the related results in [17], obtained for a time-bounded subset of temporal logic properties, to unbounded-time temporal logic properties, and analyse their robustness.

2 General Framework and Problem Statement

In this section, we provide a novel methodology to verify whether a system $\mathbf{S}$ satisfies a specification ψ, formulated in a suitable temporal logic, by integrating the partial knowledge of the system dynamics with data obtained from a measurement set-up around the system.

Let us further clarify this framework. Let us denote with $\mathbf{S}$ a physical system, or equivalently the associated dynamical behaviour. A signal input $u(t) \in \mathbb{U}, t \in \mathbb{N}$, captures how the environment acts on the system. Similarly, an output signal $y_0(t) \in \mathbb{Y}$ indicates how the system interacts with the environment, or alternatively how the system can be measured. Note that the input and output signals are assumed to take values over continuous domains. The system dynamics can be described via mathematical models, which express the behavioural relation between its inputs and outputs. The knowledge of the behaviour of the system is often limited or uncertain, making it impossible to analyse its behaviour via that of a "true" model. In this case, a-priori available knowledge allows to construct a model set $\mathcal{G}$ with elements $\mathbf{M} \in \mathcal{G}$: this model class supports the structured uncertainty as a distribution over a parameterisation $\theta \in \Theta$, $\mathcal{G} = \{\mathbf{M}(\theta)|\theta \in \Theta\}$. The unknown "true" model $\mathbf{M}(\theta^0)$

representing $\mathbf{S}$, is assumed to be an element of $\mathcal{G}$, namely $\theta^0 \in \Theta$: as an example, model sets $\mathcal{G}$ obtained through first principles adhere to this classical assumption.

Samples can be drawn from the underlying physical system via a measurement set-up, as depicted in Figure 1. An experiment consists of a finite number (N_s) of input-output samples drawn from the system, and is denoted by $Z^{N_s} = \{u(t)_{ex}, \tilde{y}(t)_{ex}\}_{t=1}^{N_s}$, where $u(t)_{ex} \in \mathbb{U}$ is the input for the experiment and $\tilde{y}(t)_{ex}$ is a (possibly noisy) measurement of $y_0(t)_{ex}$. In general, the measurement noise can enter non-additively and be a realisation of a stationary stochastic process.[1] We assume that at the beginning of the measurement procedure (say at $t = 0$), the initial condition of the system, encompassed by the initial state $\mathbf{x}(0)_{ex}$ of models in $\mathbf{M}$, is either known, or, when not known, has a structured uncertainty distribution based on the knowledge of past inputs and/or outputs. As reasonable, we implicitly consider only well-defined problems, such that for any model representing the system, given a signal input $u(t)_{ex}$ and an (uncertainty distribution for) $\mathbf{x}(0)_{ex}$, the probability density distribution of the measured signal can be fully characterised.
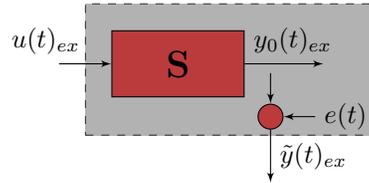

Fig. 1. System and measurement setup. In the measurement setup (grey box) the measured output $\tilde{y}(t)_{ex}$ includes the system output $y_0(t)_{ex}$ and the measurement noise $e(t)$. Data collected from experiments comprises the input $u(t)_{ex}$ and the measured output $\tilde{y}(t)_{ex}$ signals.

The end objective is to analyse the behaviour of system $\mathbf{S}$. We consider properties encoded as specifications ψ and expressed in a temporal logic of choice (to be detailed shortly). Let us remark that the behaviour of $\mathbf{S}$ to be analysed is bound to a set of operating conditions that are pertinent to the verification problem and that will be indexed with ver: this comprises the set of possible input signals $u(t)_{ver}$ (e.g., a white or coloured noise signal, or a non-deterministic signal $u(t)_{ver} \in \mathbb{U}_{ver} \subseteq \mathbb{U}$), and of the set of initial states $\mathbf{x}(0)_{ver} \in \mathbb{X}_{ver}$ for the mathematical models $\mathbf{M}$ reflecting past inputs and/or outputs of the system. The system satisfies a property if the "true" model representing it satisfies it, namely $\mathbf{S} \vDash \psi$ if and only if $\mathbf{M}(\theta^0) \vDash \psi$.

---

[1] Both the operating conditions of the experiment, that is the input signal $u(t)_{ex}$ and the initial state $\mathbf{x}(0)_{ex}$, and the measurements have been indexed with ex to distinguish them from the operating conditions of interest for verification, to be discussed shortly.

In this work we consider the satisfaction of a property $\mathbf{M}(\theta) \vDash \psi$ as a *binary-valued mapping* from the parameter space Θ. More generally, when in addition to the measurements of the system also its transitions are disturbed by stochastic noise, then property satisfaction is a mapping from the parameter space Θ to the interval $[0, 1]$, and quantifies the probability that the model $\mathbf{M}(\theta)$ satisfies the property. This mapping generalises the definition of the satisfaction function introduced in [9], and is now stated as follows.

Definition 1 (Satisfaction Function) *Let $\mathcal{G}$ be a set of models $\mathbf{M}$ that is indexed by a parameter $\theta \in \Theta$, and let ψ be a formula in a suitable temporal logic. The satisfaction function $f_\psi : \Theta \to [0, 1]$ associated with ψ is*

$$f_\psi(\theta) = \mathbf{P}\left(\mathbf{M}(\theta) \vDash \psi\right). \tag{1}$$

Let us assume that the satisfaction function f_ψ is measurable and entails a decidable verification problem (e.g., a model checking procedure) for all $\theta \in \Theta$.

Problem 1 *For a partly unknown physical system $\mathbf{S}$, under prior knowledge on the system given as a parameterised model class $\mathcal{G}$ supporting an uncertainty distribution over the parameterisation, gather possibly noisy data drawn from the measurement setup and verify properties on $\mathbf{S}$ expressed in a temporal logic of choice, with a formal quantification of the confidence of the assertion.*

2.1 A Bayesian Framework for Data-driven Modelling and Verification

Consider Problem 1. Denote loosely with $\mathbf{P}(\cdot)$ and $p(\cdot)$ respectively a probability measure and a probability density function, both defined over a continuous domain. We employ Bayesian probability calculus [26] to express the confidence in a property as a measure of the uncertainty distribution defined the set $\mathcal{G}$. By adopting the Bayesian framework, uncertainty distributions are handled as probability distributions of random variables. Therefore the confidence in a property is computed as a probability measure $\mathbf{P}(\cdot)$ via the densities $p(\cdot)$ over the uncertain variables.

Proposition 1 (Bayesian Confidence) *Given a specification ψ and a data set Z^{N_s}, the confidence that $\mathbf{S} \vDash \psi$ can be quantified via inference as*

$$\mathbf{P}\left(\mathbf{S} \vDash \psi \mid Z^{N_s}\right) = \int_\Theta f_\psi(\theta) p\left(\theta | Z^{N_s}\right) d\theta. \tag{2}$$

where f_ψ is the satisfaction function given in (1). The a-posteriori uncertainty distribution $p\left(\theta | Z^{N_s}\right)$, given the data set Z^{N_s}, is based on parametric inference over θ as

$$p\left(\theta | Z^{N_s}\right) = \frac{p\left(Z^{N_s} | \theta\right) p(\theta)}{\int_\Theta p(Z^{N_s} | \theta) p(\theta) d\theta}, \tag{3}$$

which presumes an uncertainty distribution $p(\theta)$ over the parameter set Θ, representing the prior knowledge.

The statement can be formally derived based on standard Bayesian calculus, as in [26]. We have chosen to employ a Bayesian framework, as per (3), since it allows to reason explicitly over the uncertain knowledge on the system and to work with the data acquired from the measurement setup. This leads to the efficient incorporation of the available knowledge and to its combination with the data acquisition procedure, in order to compute the confidence on the validity of a given specification over the underlying system. As a special instance, this result can be employed for Bayesian hypothesis testing [36]. As long as the mapping f_ψ is measurable, the models in the model set (and hence the system represented by it) can be characterised by either probabilistic or non-probabilistic dynamics.

Remark 2 *In statistical model checking [24,31], the objective is to replace the computationally tolling verification of a system over bounded-time properties by the empirical (statistical) testing of the relevant specifications over finite executions drawn from the system. In contrast, our problem statement tackles the problem of efficiently incorporating data with prior knowledge, for the formal (deductive) verification of the behaviour of a system with partly unknown dynamics – as such our overall verification approach is, as claimed, both data-driven and model-based. Moreover, by separating the operational conditions in an experiment from those of importance for the verification procedure, the system can be verified over non-deterministic inputs, encompassing as such both controller and disturbance inputs, or modelling errors.*

2.2 Computational Approaches

The Bayesian approach is widely applicable to different types of properties and models, however its computational complexity might in practice limit its implementation. In the literature the satisfaction function is related to the exploration of a parameter set over the validity of a formal property $f_\psi(\theta)$, and has been studied for autonomous models in continuous time in [4,15,19]. Analytical solutions to the parametric inference equation (3) can be found if the prior is a conjugate distribution. For linear dynamical systems, closed-form solutions are given inter alia in [28]. In general (2)-(3) in Proposition 1 lack analytical solutions, and the assessment of the satisfaction function (1) may be computationally intensive. Statistical methods such as the one proposed in [16] on a similar Bayesian approach lead to involved computations and introduce additional uncertainty from Monte Carlo techniques.

On the contrary, in the next section, we propose a novel computational approach over discrete-time linear time-

invariant systems. By exploiting linear parameterisations analytical solutions of both the parametric inference and the satisfaction function are characterised for properties expressed within a fragment of a temporal logic.

3 LTL Verification of LTI systems

Consider a system $\mathbf{S}$ that can be represented by a class of finite-dimensional dynamical models that evolve in discrete-time, and are linear, time-invariant (LTI), and not probabilistic. These models depend on input and output signals ranging over $\mathbb{R}^m$ and $\mathbb{R}^p$, respectively, and on variables $\mathbf{x_S}(t)$ taking values in an Euclidean space, $\mathbf{x_S}(t) \in \mathbb{X} \subseteq \mathbb{R}^n$, where n, the state dimension, is the model order. The behaviour of such a system is encompassed by state-space models $(A_\mathbf{S}, B_\mathbf{S}, C_\mathbf{S}, D_\mathbf{S})$ as

$$\mathbf{S} : \begin{cases} \mathbf{x_S}(t+1) = A_\mathbf{S}\mathbf{x_S}(t) + B_\mathbf{S}u(t), \\ y_0(t) \quad\;\; = C_\mathbf{S}\mathbf{x_S}(t) + D_\mathbf{S}u(t), \end{cases} \quad (4)$$

where matrices $A_\mathbf{S}, B_\mathbf{S}, C_\mathbf{S}, D_\mathbf{S}$ are of appropriate dimensions. Let us remark that LTI systems represent the most common modelling framework in control theory, a key framework leading towards generalisations to more complicated (e.g., nonlinear) dynamical models. The experimental measurement setup, as depicted in Figure 1, consists of the signals $u(t)_{ex}$ and $\tilde{y}(t)_{ex} = y_0(t)_{ex} + e(t)$, representing the inputs and the measured outputs, respectively, and where $e(t)$ is an additive zero-mean, white, Gaussian-distributed measurement noise with covariance Σ_e that is uncorrelated from the inputs. N_s samples are collected within a data set $Z^{N_s} = \{u(t)_{ex}, \tilde{y}(t)_{ex}\}_{t=1}^{N_s}$.

System properties are expressed, over a finite set of atomic propositions $p_i \in AP$, $i = 1, \dots, |AP|$, in Linear-time Temporal Logic [2]. LTL formulae are built recursively via the syntax $\psi ::= \text{true} \mid p \mid \neg\psi \mid \psi \wedge \psi \mid \psi \vee \psi \mid \bigcirc\psi \mid \psi \,\mathsf{U}\, \psi$. Let $\pi = \pi(0), \pi(1), \pi(2), \dots \in \Sigma^{\mathbb{N}^+}$ be a string composed of letters from the alphabet $\Sigma = 2^{AP}$, and let $\pi_t = \pi(t), \pi(t+1), \pi(t+2), \dots$ be a subsequence of π, then the satisfaction relation between π and ψ is denoted as $\pi \vDash \psi$ (or equivalently $\pi_0 \vDash \psi$). The semantics for the satisfaction are defined recursively over π_t and the LTL syntax as

$$
\begin{aligned}
\text{(true)} \;\; \pi_t &\vDash \text{true} & &\Leftrightarrow \text{true} \\
\text{(atomic prop.)} \;\; \pi_t &\vDash p & &\Leftrightarrow p \in \pi(t) \\
\text{(negation)} \;\; \pi_t &\vDash \neg\psi & &\Leftrightarrow \pi_t \nvDash \psi \\
\text{(conjunction)} \;\; \pi_t &\vDash \psi_1 \wedge \psi_2 & &\Leftrightarrow \pi_t \vDash \psi_1 \text{ and } \pi_t \vDash \psi_2 \\
\text{(disjunction)} \;\; \pi_t &\vDash \psi_1 \vee \psi_2 & &\Leftrightarrow \pi_t \vDash \psi_1 \text{ or } \pi_t \vDash \psi_2 \\
\text{(next)} \;\; \pi_t &\vDash \bigcirc\psi & &\Leftrightarrow \pi_{t+1} \vDash \psi \\
\text{(until)} \;\; \pi_t &\vDash \psi_1 \,\mathsf{U}\, \psi_2 & &\Leftrightarrow \exists i \in \mathbb{N} : \pi_{t+i} \vDash \psi_2, \\
& & &\quad \text{and } \forall j \in \mathbb{N} : \\
& & &\quad 0 \le j < i, \pi_{t+j} \vDash \psi_1
\end{aligned}
$$

Denote the k-bounded and unbounded invariance operator as $\square^k \psi = \bigwedge_{i=0}^{k} \bigcirc^i \psi$ and $\square\psi = \neg(\text{true} \,\mathsf{U}\, \neg\psi)$, respectively.

Of interest are formal properties encoded on the input-output behaviour of the system, and over a time horizon $t \ge 0$. The output $y_0(t)_{ver} \in \mathbb{Y}$ is labeled by a map $L : \mathbb{Y} \to \Sigma$, which assigns letters α in the alphabet Σ via half spaces on the output, as

$$L(y_0(t)_{ver}) = \alpha \in \Sigma \;\Leftrightarrow\; \bigwedge_{p_i \in \alpha} A_{p_i} y_0(t)_{ver} \le b_{p_i}, \quad (5)$$

for given $A_{p_i} \in \mathbb{R}^{1 \times p}$, $b_{p_i} \in \mathbb{R}$ that is, sets of atomic propositions are associated to polyhedra over $\mathbb{Y} \subset \mathbb{R}^p$. Let us underline that properties are defined over the behaviour $y_0(t)_{ver}$ of the system, and not over the noisy measurements $\tilde{y}(t)_{ex}$ of the system in the measurement setup. Additionally, for the verification problem the input signal is modelled as a bounded signal $u(t) \in \mathbb{U}_{ver}$, and represents possible external non-determinism of the environment acting on the system.

3.1 Model Set Selection

As a first step we need to embed the a-priori available knowledge on the underlying system within a parameterised model set, under a prior distribution. The use of linearly parameterised model sets defined through orthonormal basis functions to represent partially unknown systems is a broadly used framework in system identification: it allows for the incorporation of prior knowledge, while maintaining the benefits (computational aspects) of linear parameterisations. Practically, it has been widely used for the modelling of physical systems, such as the thermal dynamics of buildings [35,29]. Note that although the goal of parameter exploration in formal verification has recently attracted quite some attention [4,15,19], there are as of yet no general scalable results for the computation of the satisfaction function for nonlinearly-parameterised discrete-time LTI models. Whilst in general linear time-invariant models with uncertain parameters do not map onto a linearly-parameterised model set, we argue that a linearly-parameterised model set can encompass a relevant class of models. For instance, any asymptotically stable LTI model can be represented uniquely by its (infinite) impulse response [20], and the coefficients of the impulse response define a linear parameterisation for this model. Further, the coefficients of the impulse response converge to zero, so that a truncated set of impulse coefficients can provide a good approximate model set with a finite-dimensional, linear parameterisation. This is only one possible instance of modelling by a finite set of orthonormal basis functions [21, Chapters 4 and 7],[33], which can be selected to optimally incorporate prior knowledge: we conclude that, as an alternative to the use of a nonlinearly parameterised set of models, structural information (even when inexact) can be used to select a

set of orthonormal basis functions, whose finite truncation defines a finite-dimensional linearly-parameterised model set indexed over the coefficients of the basis functions. Thus, in the following we consider a linearly parameterised model set $\mathcal{G}$ that encapsulates system $\mathbf{S}$, and specifically $\mathcal{G} = \{(A, B, C(\theta), D(\theta)), \theta \in \Theta\}$.

A system, or equivalently the mathematical model that represents it, satisfies a property if all the words generated by the model satisfy that property. Since properties are encoded over the external (input-output) behaviour of the system $\mathbf{S}$, which is the behaviour of $\mathbf{M}(\theta^0), \theta^0 \in \Theta$, we can equivalently assert that any property ψ is verified by the system, $\mathbf{S} \vDash \psi$, if and only if it is verified by the unknown model representing the system, namely $\mathbf{M}(\theta^0) \vDash \psi$. Introduce Θ_ψ to be the feasible set of parameters, such that for every parameter in that set the property ψ holds, i.e., $\forall \theta \in \Theta_\psi : \mathbf{M}(\theta) \vDash \psi$. As such Θ_ψ is characterised as the level set of the satisfaction function f_ψ, $\Theta_\psi = \{\theta \in \Theta : f_\psi(\theta) = 1\}$.

3.2 Safety Verification of Bounded-time Properties

Models $\mathbf{M}$ in the class $\mathcal{G}$ have the following representation $(A, B, C(\theta), 0)$:

$$\mathbf{M}(\theta): \quad \begin{cases} \mathbf{x}(t+1) = A\mathbf{x}(t) + Bu(t), \\ \hat{y}(t, \theta) = C(\theta)\mathbf{x}(t), \end{cases} \tag{6}$$

and are parameterised by $\theta \in \Theta \subset \mathbb{R}^{pn} : \theta = \text{vec}(C)$ with a prior probability distribution $p(\theta)$. In addition to this strictly proper model class we will also allow for proper model $(A, B, C(\theta), D(\theta))$ where both the C and the D-matrices are parameterised and the parameterisation is $\theta = \text{vec}([C\ D])$). For a given initial condition $\mathbf{x}(0)$ and input sequence, the output of the "true" model $\hat{y}(t, \theta^0)$ is equal to the system output $y_0(t)$.

Given a measurement set-up as in Figure 1 with unknown parameter θ^0. Then $u(t)_{ex}$ and $\tilde{y}(t)_{ex}$ represent the input and the measured output, respectively, and $e(t)$ is an additive zero-mean, white, Gaussian-distributed measurement noise with covariance Σ_e that is uncorrelated from the input. Furthermore $u(t)$ is assumed to be uncorrelated with the noise $e(t)$. From this set-up N_s samples are collected in a data set $Z^{N_s} = \{u(t)_{ex}, \tilde{y}(t)_{ex}\}_{t=1}^{N_s}$.

Therefore given the operating conditions of the experiment set-up the measured signal $\tilde{y}(t)_{ex}$ can be fully characterised: its probability density, conditional on the parameters θ, is

$$p\left(Z^{N_s}|\theta\right) = \prod_{t=1}^{N_s} p\left(\tilde{y}(t)_{ex}|\theta\right)$$

$$= \frac{1}{\sqrt{|\Sigma_e|^{N_s}(2\pi)^{pN_s}}} \exp\Bigg[$$
$$- \frac{1}{2}\sum_{t=1}^{N_s}(\hat{y}(t,\theta) - \tilde{y}(t)_{ex})^T \Sigma_e^{-1}(\hat{y}(t,\theta) - \tilde{y}(t)_{ex})\Bigg]$$

and can be directly used in Proposition 1. This conditional density $p\left(Z^{N_s}|\theta\right)$ depends implicitly on the given initial state $\mathbf{x}(0)_{ex}$ and, for the case of a given uncertainty distribution for $\mathbf{x}(0)_{ex}$, $p\left(Z^{N_s}|\theta\right)$ should be marginalised as a latent variable [28]. The a-posteriori uncertainty distribution is obtained as the analytical solution of the parametric inference in (3) [28].

Recall now that for a given specification ψ, we seek to determine a feasible set of parameters Θ_ψ, such that the corresponding models admit property ψ, namely $\mathbf{M}(\theta) \vDash \psi$, $\forall \theta \in \Theta_\psi$. Since models $\mathbf{M}(\theta)$ have a linearly-parameterised state space realisation as per (6), it follows that when the set of initial states and inputs $\mathbb{X}_{ver}$ and $\mathbb{U}_{ver}$ are bounded polyhedra, the verification of a class of safety properties expressed by formulae with labels as in (5) leads to a set of feasible parameters Θ_ψ that is a polyhedron, which can be easily computed. More precisely, the following theorem can be derived.

Theorem 3 ([17]) *Given a bounded polyhedral set (or equivalently a polytope) of initial states $\mathbf{x}(0) \in \mathbb{X}_{ver}$ and of inputs $u(t) \in \mathbb{U}_{ver}$ for $t \geq 0$, and considering a labelling map as in (5), then the feasible set Θ_ψ of the parameterised model set (6) results in a polyhedron for properties ψ composed of the LTL fragment $\psi ::= \alpha | \bigcirc \psi | \psi_1 \wedge \psi_2$, with $\alpha \in \Sigma$.*

Proof[of Theorem 3] Let $\otimes$ denote the Kronecker product. Consider the input set $\mathbb{U}_{ver}$ to be the convex hull of U, i.e. $\text{conv}(U) = \mathbb{U}_{ver}$. Similarly let the set of initial states be $\text{conv}(X_{ver}) = \mathbb{X}_{ver}$. Let the model set be given as $\mathbf{M}(\theta) = (A, B, C(\theta), D)$. We will temporarily assume that D is set equal to zero. Afterwards we will show how to work with a parameterised D. Note that the syntax fragment $\psi ::= \alpha | \bigcirc \psi | \psi_1 \wedge \psi_2$ with $\alpha \in \Sigma = 2^{AP}$ is equivalent to $\psi ::= p | \bigcirc \psi | \psi_1 \wedge \psi_2$ with $p \in AP$.

1. We claim that for every specification ψ composed from the syntax fragment $\psi ::= p | \bigcirc \psi | \psi_1 \wedge \psi_2$ and $\theta \in \Theta$, the words generated by a model $\mathbf{M}(\theta) = (A, B, C(\theta), 0)$ with state $\mathbf{x}(t)$ satisfy the specification ψ, denoted $< \mathbf{M}(\theta), \mathbf{x}(t) > \vDash \psi$, if and only if

$$\left(\left(I_{n_\psi} \otimes \mathbf{x}(t) \right)^T N_\psi + K_\psi \right) \theta \leq B_\psi. \tag{7}$$

The matrices $N_\psi \in \mathbb{R}^{nn_\psi \times np}$, $K_\psi \in \mathbb{R}^{n_\psi \times np}$, $B_\psi \in \mathbb{R}^{n_\psi}$ in the above satisfaction relation have dimensions that are functions of the parametrisation and of the property dependent "dimension" n_ψ, and are obtained inductively

over the syntax of the specification.

For any *atomic propositions* the model starting from state $\mathbf{x}(t)$ satisfies a property p_i, i.e., $< \mathbf{M}(\theta), \mathbf{x}(t) > \vDash p_i \Leftrightarrow A_{p_i} y \leq b_{p_i}$, with $A_{p_i} \in \mathbb{R}^{1 \times p}$ and $b_{p_i} \in \mathbb{R}$ we construct the matrices N_{p_i}, K_{p_i} and B_{p_i} as follows. Consider $y(t)$ for a given $x(t)$ then

$$A_{p_i} y(t) = A_{p_i} C(\theta) \mathbf{x}(t) = \mathbf{x}(t)^T (I_n \otimes A_{p_i}) \theta.$$

This yields $N_{p_i} = (I_n \otimes A_{p_i}) \in \mathbb{R}^{n \times np}$, $K_{p_i} = O_{1 \times np} \in \mathbb{R}^{1 \times np}$, and $B_{p_i} = b_{p_i} \in \mathbb{R}^{1 \times 1}$.

The *next* operation $\bigcirc \psi_1$ with matrices $(N_{\psi_1}, K_{\psi_1}, D_{\psi_1}, b_{\psi_1})$ yields matrices

$$N_{\bigcirc \psi_1} = \mathbf{1}_{|U|} \otimes \left(I_{n_{\psi_1}} \otimes A^T\right) N_{\psi_1},$$
$$K_{\bigcirc \psi_1} = \mathcal{U} \left(I_{n_{\psi_1}} \otimes B\right)^T N_{\psi_1} + \mathbf{1}_{|U|} \otimes K_{\psi_1},$$
$$B_{\bigcirc \psi_1} = \mathbf{1}_{|U|} \otimes B_{\psi_1},$$

where the i-th set of n_{ψ_1} rows of $\mathcal{U} \in \mathbb{R}^{|U| n_{\psi_1} \times m}$ is defined as

$$\left(I_{n_{\psi_1}} \otimes u_i^T\right) \text{ with } u_i \in U$$

and where $n_{\bigcirc \psi_1} = |U| n_{\psi_1}$. This can be derived as

$$< \mathbf{M}(\theta), \mathbf{x}(t) > \vDash \bigcirc \psi \Leftrightarrow \forall u(t) \in \mathbb{U}_{ver} :$$
$$\left(\left(I_{n_{\psi_1}} \otimes \mathbf{x}(t+1)\right)^T N_{\psi_1} + K_{\psi_1}\right) \theta \leq B_{\psi_1},$$
$$\Leftrightarrow \forall u(t) \in \mathbb{U}_{ver} :$$
$$\left(\left(I_{n_{\psi_1}} \otimes A\mathbf{x}(t)\right)^T N_{\psi_1}\right.$$
$$\left. + \left(I_{n_{\psi_1}} \otimes Bu(t)\right)^T N_{\psi_1} + K_{\psi_1}\right) \theta \leq B_{\psi_1}.$$

Since the above is an affine function in $u(t)$, the image of every $u(t) \in \mathrm{conv}(U) = \mathbb{U}_{ver}$ can be expressed as a convex combination of the values at the vertices $u_i \in U$, c.f. [6]. Then an equivalent expression is

$$\Leftrightarrow \forall u_i \in U : \left(\left(I_{n_{\psi_1}} \otimes A\mathbf{x}(t)\right)^T N_{\psi_1}\right.$$
$$\left. + \left(I_{n_{\psi_1}} \otimes u_i\right)^T \left(I_{n_{\psi_1}} \otimes B\right)^T N_{\psi_1} + K_{\psi_1}\right) \theta \leq B_{\psi_1}$$

which can be rewritten as

$$\Leftrightarrow \left(\mathbf{1}_{|U|} \otimes \left(I_{n_{\psi_1}} \otimes A\mathbf{x}(t)\right)^T N_{\psi_1} + \mathcal{U} \left(I_{n_{\psi_1}} \otimes B\right)^T N_{\psi_1}\right.$$
$$\left. + \mathbf{1}_{|U|} \otimes K_{\psi_1}\right) \theta \leq \mathbf{1}_{|U|} \otimes B_{\psi_1}.$$

Having obtained $K_{\bigcirc \psi}$, $D_{\bigcirc \psi}$, and $b_{\bigcirc \psi}$, now rewrite the first term to obtain $N_{\bigcirc \psi}$:

$$\mathbf{1}_{|U|} \otimes \left(I_{n_{\psi_1}} \otimes \mathbf{x}^T(t)\right) \left(I_{n_{\psi_1}} \otimes A^T\right) N_{\psi_1}$$
$$= \left(I_{|U|} \mathbf{1}_{|U|}\right) \otimes \left(I_{n_{\psi_1}} \otimes \mathbf{x}^T(t)\right) \left(I_{n_{\psi_1}} \otimes A^T\right) N_{\psi_1}$$
$$= \left(I_{|U| n_{\psi_1}} \otimes \mathbf{x}^T(t)\right) \left(\mathbf{1}_{|U|} \otimes \left(I_{n_{\psi_1}} \otimes A^T\right) N_{\psi_1}\right).$$

The *and* operation $\psi_1 \wedge \psi_2$ for $(N_{\psi_1}, K_{\psi_1}, D_{\psi_1}, b_{\psi_1})$ and $(N_{\psi_2}, K_{\psi_2}, D_{\psi_2}, b_{\psi_2})$ with $n_{\psi_1 \wedge \psi_2} = (n_{\psi_1} + n_{\psi_2})$ gives

$$N_{\psi_1 \wedge \psi_2} = \begin{bmatrix} N_{\psi_1} \\ N_{\psi_2} \end{bmatrix}, K_{\psi_1 \wedge \psi_2} = \begin{bmatrix} K_{\psi_1} \\ K_{\psi_2} \end{bmatrix}, \; B_{\psi_1 \wedge \psi_2} = \begin{bmatrix} B_{\psi_1} \\ B_{\psi_2} \end{bmatrix}.$$

This can be derived from

$$< \mathbf{M}(\theta), \mathbf{x}(t) > \vDash \psi_1 \wedge \psi_2$$
$$\Leftrightarrow \bigwedge_{i \in \{1,2\}} \left(\left(I_{n_{\psi_i}} \otimes \mathbf{x}(t)\right)^T N_{\psi_i} + K_{\psi_i}\right) \theta \leq B_{\psi_i}$$
$$\Leftrightarrow \left(\left(I_{n_{\psi_1 \wedge \psi_2}} \otimes \mathbf{x}(t)\right)^T \begin{bmatrix} N_{\psi_1} \\ N_{\psi_2} \end{bmatrix} + \begin{bmatrix} K_{\psi_1} \\ K_{\psi_2} \end{bmatrix}\right) \theta \leq \begin{bmatrix} B_{\psi_1} \\ B_{\psi_2} \end{bmatrix}.$$

2. The matrix-valued function

$$\left(\left(I_{n_\psi} \otimes \mathbf{x}(0)\right)^T N_\psi + K_\psi\right) \theta$$

is affine in $\mathbf{x}^T(0)$ (for a fixed θ), therefore its value at the initial condition $\mathbf{x}(0) \in \mathbb{X}_{ver}$ is a convex combination of the function values at the vertices X_{ver} of $\mathbb{X}_{ver}$. Thus the satisfaction relation $< \mathbf{M}(\theta), \mathbf{x}(0) > \vDash \psi$ represented by the multi-affine inequality holds uniformly over $\mathbf{x}(0) \in \mathbb{X}_{ver}$ if and only if it holds for the vertices of $\mathbb{X}_{ver}$. This gives a set of affine inequalities in θ, thus the feasible set Θ_ψ is a polyhedron and is given as

$$\left\{\theta \in \Theta : \bigwedge_{\mathbf{x}_i \in X_{ver}} \left(\left(I_{n_\psi} \otimes \mathbf{x}_i\right)^T N_\psi + K_\psi\right) \theta \leq B_\psi\right\}.$$

The set Θ_ψ is a polyhedron, since it is formed by a finite set of half spaces.

3. To prove Theorem 3 we need to extend the results to models with parameterised D. The dynamics of model (A, B, C, D) with both C and D fully parameterised can be reformulated as

$$\begin{bmatrix} \mathbf{x}(t+1) \\ u(t+1) \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ u(t) \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} u(t+1)$$
$$y(t) = \begin{bmatrix} C & D \end{bmatrix} \mathbf{x}(t).$$

Using the new matrices $(\tilde{A}, \tilde{B}, \tilde{C}(\theta), 0)$ the obtained results still hold. For part **2.** set of vertices X_{ver} needs to be extended with the vertices of U as $X_{ver} \times U$. $\quad \square$

In the computation of the feasible set, the faces of the polyhedron Θ_ψ are shown to be a function of the ver-

tices [2] of the bounded set of initial states $\mathbb{X}_{ver}$ and of the set of inputs $\mathbb{U}_{ver}$, and are also expected to grow in number as a function of the time horizon of the property. The result in Theorem 3 is valid for any finite composition of the LTL fragment $\psi ::= \alpha | \bigcirc \psi | \psi_1 \wedge \psi_2$, as such it only holds for finite horizon properties. Properties defined over the infinite horizon will be the objective of Section 3.4.

3.3 Case Study: Bounded-Time Safety Verification

Consider a system **S** and verify whether the output $y_0(t)_{ver}$ remains within the interval $\mathcal{I} = \left[-0.5, \ 0.5 \right]$, labeled as ι, for the next 5 time steps, under $u(t)_{ver} \in \mathbb{U}_{ver} = [-0.2, \ 0.2]$ and $\mathbf{x}(0)_{ver} \in \{0_2\} = \mathbb{X}_{ver}$. Introduce accordingly the alphabet $\Sigma = \{\iota, \tau\}$ and the labelling map $L : L(y) = \iota, \forall y \in \mathcal{I}, L(y) = \tau, \forall y \in \mathbb{Y} \setminus \mathcal{I}$. Now check whether the following LTL property holds: $\mathbf{S} \vDash \bigwedge_{i=1}^{5} (\bigcirc)^i \iota$. We assume that system **S** can be represented as an element of a model set $\mathcal{G}$ with transfer functions characterised by second-order Laguerre-basis ones [20] (a special case of orthonormal basis functions), which translates to the following parameterised state-space representation:

$$\mathbf{x}(t+1) = \begin{bmatrix} a & 0 \\ 1-a^2 & a \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} \sqrt{1-a^2} \\ (-a)\sqrt{1-a^2} \end{bmatrix} u(t), \quad (8)$$

$$\hat{y}(t,\theta) = \theta^T \mathbf{x}(t) .$$

The parameter set is chosen as $\theta \in \Theta = [-10, 10]^2$, whereas the coefficient a is chosen to be equal to 0.4. We select, as prior available knowledge on the system, a uniform distribution $p(\theta)$ on the model class, and pick a known variance $\sigma_e^2 = 0.5$ for the white additive noise on the measurement. The set of feasible parameters $\Theta_\psi \subset \Theta$ is represented in Figure 2 and is computed according to Theorem 3. Based on the prior available knowledge, the confidence associated to $\theta_0 \in \Theta_\psi$ amounts to 0.0165 [3]. In comparison to this value, after doing an experiment on the system with "true parameter" $\theta_0 = [1 \ 0]^T$ (Figure 2) and with input signal $u(t)_{ex}$, a realisation of a white noise with a uniform distribution over $[-0.2, 0.2]$, and measuring $\tilde{y}(t)_{ex}$ for 200 consecutive time instances the uncertainty distribution is refined as $p\left(\theta | Z^{N_s}\right)$. The resulting confidence (2) in the property is increased to 0.779.

Along this line of experiments, we have repeated the test 100 times, for several instances of the parameter θ^0 characterising the underlying system **S**. In all instances, after obtaining 200 measurements the a-posteriori confidence

Table 1
Mean (μ) and variance (σ^2) of the confidence obtained from 100 experiments with 200 measurements each.

θ^0	μ	σ^2	θ^0	μ	σ^2
$\begin{bmatrix} -1 & -1 \end{bmatrix}^T$	0.348	0.073	$\begin{bmatrix} 1 & -1 \end{bmatrix}^T$	0.491	0.085
$\begin{bmatrix} -1 & 0 \end{bmatrix}^T$	0.705	0.060	$\begin{bmatrix} 1 & 0 \end{bmatrix}^T$	0.730	0.056
$\begin{bmatrix} -1 & 1 \end{bmatrix}^T$	0.492	0.086	$\begin{bmatrix} 1 & 1 \end{bmatrix}^T$	0.339	0.065

represents the confidence in the safety of the system, as displayed in Table 1 via mean and variance terms.
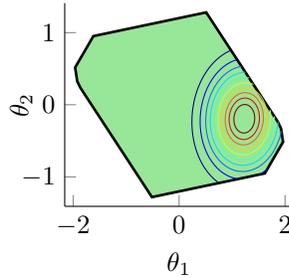
Fig. 2: Feasible set of parameters in Θ, and contour lines of the quantity $p\left(\theta | Z^{N_s}\right)$, obtained for $\theta^0 = [1 \ 0]^T$.

3.4 Verifying Unbounded-Time Properties Using Invariant Sets

In this section we extend the approach unfolded in Section 3.2, to hold on the LTL fragment $\psi ::= \alpha | \bigcirc \psi | \psi_1 \wedge \psi_2$ with additionally the *unbounded* invariance (safety) operator. Recall the form of the k-bounded and of the unbounded invariance operators, namely $\square^k \psi = \bigwedge_{i=0}^{k} \bigcirc^i \psi$ and $\square \psi = \neg(\texttt{true} \ \mathsf{U} \ \neg \psi)$ respectively. The extension from a k-bounded operator, covered by the result in Theorem 3, to the unbounded invariance one, is based on the concept of robust positive invariance [7, Def. 4.3], recalled next.

Definition 2 *For the system* $\mathbf{x}(t+1) = A\mathbf{x}(t) + Bu(t)$, *the set* $\mathcal{S} \subseteq \mathbb{X}$ *is said to be robustly positively invariant if, for all* $\mathbf{x}(0) \in \mathcal{S}$ *and* $u(t) \in \mathbb{U}$, *the condition* $\mathbf{x}(t) \in \mathcal{S}$ *holds for all* $t \geq 0$.

Recall that the feasible set Θ_ψ is defined as the set of parameters for which property ψ holds, namely $\forall \theta \in \Theta_\psi :$ $\mathbf{M}(\theta) \vDash \psi$. The satisfaction relation $\mathbf{M}(\theta) \vDash \psi$ depends implicitly on the set of initial states $\mathbf{x}(0) \in \mathbb{X}_{ver}$ and on the set of inputs $\mathbb{U}_{ver}$. Let us extend the definition of the feasible set to explicitly account for its dependence on the set of initial conditions: given a bounded and convex set $\mathcal{S} \subset \mathbb{X}$, let $\Theta_\psi(\mathcal{S})$ be defined as the set of parameters in Θ for which the parameterised models $\mathbf{M}(\theta)$ initialised with $\mathbf{x}(0) \in \mathcal{S}$ satisfy ψ over input signals $u(t) \in \mathbb{U}_{ver} \ t \geq 0$. Hence the feasible set Θ_ψ can be written as a function of the set of initial states $\mathbb{X}_{ver}$, that is $\Theta_\psi(\mathbb{X}_{ver})$. Thus the extended map $\Theta_\psi(\cdot)$ takes

---

[2] A polytope can be written as the convex hull of a finite set of *vertices*.

[3] This is obtained by numerical computation of (2) with probability distribution $p(\theta)$. ntegrals are solved via the numerical integration tool in `Matlab`.

subsets of the state space into subsets of the parameter space. Note that if $\mathcal{S}$ is a robustly positively invariant set that includes the set of initial states $\mathbb{X}_{ver} \subseteq \mathcal{S}$, then for all $\theta \in \Theta_\psi(\mathcal{S})$ the models $\mathbf{M}(\theta)$ satisfy ψ over all infinite-time model traces $\mathbf{x}(t)$: this allows to state that $\mathbf{M}(\theta) \vDash \Box\psi$. We can show that the following holds.

Lemma 4 *The function $\Theta_\psi(\cdot) : 2^\mathbb{X} \to 2^\Theta$, for specifications obtained as $\psi ::= \alpha \mid \bigcirc\psi \mid \psi_1 \wedge \psi_2$, is monotonically decreasing: that is if $\mathcal{S}_1 \subseteq \mathcal{S}_2$, then $\Theta_\psi(\mathcal{S}_2) \subseteq \Theta_\psi(\mathcal{S}_1)$.*

Proof We leverage the notation used in the proof of Theorem 1. Provided that the parameterised model is given as $(A, B, C(\theta), 0)$, we show that any $\theta \in \Theta_\psi(\mathcal{S}_2)$ is also an element of $\theta \in \Theta_\psi(\mathcal{S}_1)$. Suppose $\mathcal{S}_2$ has a finite number of vertices $\mathbf{x}_i \in \mathcal{V}(\mathcal{S}_2)$, then for any $\theta \in \Theta_\psi(\mathcal{S}_2)$:

$$\bigwedge\nolimits_{\mathbf{x}_i \in \mathcal{V}(\mathcal{S}_2)} \left((I_{n_\psi} \otimes \mathbf{x}_i)^T N_\psi + K_\psi \right) \theta \leq B_\psi$$

and for every $\mathbf{x} \in \mathcal{S}_2$

$$\left((I_{n_\psi} \otimes \mathbf{x})^T N_\psi + K_\psi \right) \theta \leq B_\psi.$$

Since the vertices $\mathbf{x}_j \in \mathcal{V}(\mathcal{S}_1)$ are also elements of $\mathcal{S}_2$, then

$$\bigwedge\nolimits_{\mathbf{x}_j \in \mathcal{V}(\mathcal{S}_1)} \left((I_{n_\psi} \otimes \mathbf{x}_j)^T N_\psi + K_\psi \right) \theta \leq B_\psi$$

and $\theta \in \Theta_\psi(\mathcal{S}_1)$. This reasoning can be trivially extended to include parameterised D matrices. Increasing the number of vertices of $\mathcal{S}_1$ and $\mathcal{S}_2$, does not change the result, hence the same holds if $\mathcal{S}_1$ and $\mathcal{S}_2$ are convex sets. $\quad\square$

Based on the result in Lemma 4, we conclude that the maximal feasible set $\Theta_{\Box\psi}$ is obtained as a mapping from the minimal robustly positively invariant set $\mathcal{S}$ that includes $\mathbb{X}_{ver}$: $\Theta_{\Box\psi} = \Theta_\psi(\mathcal{S})$. This leads next to consider under which conditions such minimal robustly positively invariant set $\mathcal{S}$ can be exactly computed or approximated.

Feasible set for invariance properties with $\mathbb{X}_{ver} = \{0_n\}$

For $\mathbb{X}_{ver} = \{0_n\}$, assuming a bounded interval $\mathbb{U}_{ver}$ with the origin in its interior, and under some basic assumptions on the dynamics (to be shortly discussed), the minimal robustly positively invariant set can be shown to be a bounded and convex set that includes the origin [7]. Maintaining the condition of $\mathbb{U}_{ver}$ being bounded and having the origin in its interior, we first consider the case that $\mathbb{X}_{ver} = \{0_n\}$ and characterise $\mathcal{S}$ via tools available from set theory in systems and control; thereafter we look at extensions to more general sets of initial states $\mathbb{X}_{ver}$.

Assume that $\mathbb{U}_{ver}$ includes the origin, and denote the forward reachability mappings initialised with $\mathcal{R}^{(0)} := \{0_n\} \subset \mathbb{X}$ as

$$\mathcal{R}^{(i)} := \mathrm{Post}(\mathcal{R}^{(i-1)}), \tag{9}$$

with set operation $\mathrm{Post}(X) := \{\mathbf{x}' = A\mathbf{x} + Bu, \mathbf{x} \in X, u \in \mathbb{U}\}$. Denote the limit reachable set as $\mathcal{R}^\infty = \lim_{i\to\infty} \mathcal{R}^{(i)}$. From literature we recall that properties of these i-step reachable sets, as given in [7] include the following: for a reachable pair (A, B) and an asymptotically stable matrix A, the ∞-reachable set $\mathcal{R}^\infty$ is bounded and convex [7, Proposition 6.9]. The k-step reachable set converges to the ∞-reachable set via (9), since it is monotonically increasing $\mathcal{R}^{(i)} \subseteq \mathcal{R}^{(i+1)}$. Moreover, $\mathcal{R}^\infty$ is the minimal robustly positively invariant set for the system, so that any positively invariant set includes $\mathcal{R}^\infty$ [7, Proposition 6.13]. Thus, starting from $\mathbf{x}(0) = 0_n$, all $\mathbf{x}(t) \in \mathcal{R}^\infty$, and furthermore of interest to this work we conclude that $\Theta_{\Box^k\psi} = \Theta_\psi(\mathcal{R}^{(k)})$ and $\Theta_{\Box\psi} = \Theta_\psi(\mathcal{R}^\infty)$.

Feasible set for invariance properties under polytopic sets of initial states

More generally, if $\mathbb{X}_{ver} \subseteq \mathcal{R}^\infty$ and ceteris paribus, then $\mathcal{R}^\infty$ is the minimal robustly positively invariant set that includes $\mathbb{X}_{ver}$, and $\Theta_\psi(\mathcal{R}^\infty) = \Theta_{\Box\psi}$. For finite iterations the reachable sets $\mathcal{R}^{(i)}$ are polytopes, and if $\mathcal{R}^{(i)} = \mathcal{R}^{(i+1)}$, then $\mathcal{R}^{(i)} = \mathcal{R}^\infty$. Though the iterations can stop in finite time, in general the number of iterations to obtain $\mathcal{R}^\infty$ can be infinite. Whilst the minimal robustly positively invariant set is not necessarily closed or a polytope, there exist methods to approximate $\mathcal{R}^\infty$ as detailed in [7]. For instance, for stable systems, $\mathcal{R}^{(k)}$ is shown to converge to $\mathcal{R}^\infty$, in the sense that for all $\epsilon > 0$ there exists $\bar{k}$ such that for $k \geq \bar{k}$, $\mathcal{R}^{(k)} \subseteq \mathcal{R}^\infty \subseteq (1 + \epsilon)\mathcal{R}^{(k)}$ [7, Proposition 6.9].

Recall that the maximal feasible set $\Theta_{\Box\psi}$ is obtained as a mapping from the minimal robustly positively invariant set $\mathcal{S}$ including $\mathbb{X}_{ver}$, that is $\Theta_{\Box\psi} = \Theta_\psi(\mathcal{S})$. Let us extend the study to the case where the conditions $\mathbb{X}_{ver} = \{0_n\}$ or its extension $\mathbb{X}_{ver} \subseteq \mathcal{R}^\infty$ do not apply, while the condition on the bounded set $\mathbb{U}_{ver}$ is maintained, that is $0 \in \mathbb{U}_{ver}$. Consider the more general case where the set of initial states is a polytope but not necessarily a subset of $\mathcal{R}^\infty$. Denote the union of the forward reachability mappings initialised with $\mathcal{R}_{\mathbb{X}_{ver}}^{(0)} := \mathbb{X}_{ver} \subseteq \mathbb{X}$ as

$$\mathcal{R}_{\mathbb{X}_{ver}}^{(i)} := \mathcal{R}_{\mathbb{X}_{ver}}^{(i-1)} \cup \mathrm{Post}(\mathcal{R}_{\mathbb{X}_{ver}}^{(i-1)}). \tag{10}$$

This set is also known in the literature as the *reach tube*. The corresponding set for infinite time is denoted as $\mathcal{R}_{\mathbb{X}_{ver}}^\infty = \lim_{i\to\infty} \mathcal{R}_{\mathbb{X}_{ver}}^{(i)}$. Notice that if $\mathbb{X}_{ver} \subseteq \mathcal{R}^\infty$, then $\mathcal{R}^\infty = \mathcal{R}_{\mathbb{X}_{ver}}^\infty$. The iteration is monotonically increasing $\mathcal{R}_{\mathbb{X}_{ver}}^{(i)} \subseteq \mathcal{R}_{\mathbb{X}_{ver}}^{(i+1)}$, and whenever $\mathcal{R}_{\mathbb{X}_{ver}}^{(i)} = \mathcal{R}_{\mathbb{X}_{ver}}^{(i+1)}$ it stops

after a finite number of iterations with $\mathcal{R}_{\mathbb{X}_{ver}}^{\infty} = \mathcal{R}_{\mathbb{X}_{ver}}^{(i)}$. Of course, also in this more general case, the number of iterations can be unbounded, however the convergence properties of $\mathcal{R}^{(i)}$ extend seamlessly to the case of sets $\mathcal{R}_{\mathbb{X}_{ver}}^{(i)}$. Since $\mathcal{R}_{\mathbb{X}_{ver}}^{(i)}$ is a union of polytopes, it is not guaranteed to be a convex set. Still, it can be shown via the proof of Theorem 3 that the computation of the feasible set $\Theta_{\psi}(\mathcal{S})$ boils down to that of $\Theta_{\psi}(\text{conv}(\mathcal{S}))$.

Remark 5 *Let us illustrate the convergence property for sets $\mathcal{R}_{\mathbb{X}_{ver}}^{(i)}$ as follows. For every vertex $\mathbf{x}^i(0) \in \mathbb{X}_{ver}$, select a decomposition $\mathbf{x}_r^i + \mathbf{x}_s^i$ with $\mathbf{x}_r^i \in \mathcal{R}^{\infty}$, which minimises $\|\mathbf{x}_s^i\|$ for a chosen vector norm $\|\cdot\|$. Since every element $\mathbf{x}(0) \in \mathbb{X}_{ver}$ is a convex combination of the vertices $\mathbf{x}^i(0)$, it follows that for all $\mathbf{x}(0) \in \mathbb{X}_{ver}$:*

$$\mathbf{x}(0) = \sum_i a_i \mathbf{x}^i(0) = \sum_i a_i \mathbf{x}_r^i(0) + \sum_i a_i \mathbf{x}_s^i(0)$$
$$\in \text{conv}(\mathbf{x}_r^i(0)) + \text{conv}(\mathbf{x}_s^i(0)) \subseteq \mathcal{R}^{\infty} + \bar{\mathbb{X}}_{ver},$$

with $\sum_i a_i = 1$ for $a_i \geq 0$ and where $\bar{\mathbb{X}}_{ver} = \text{conv}(\mathbf{x}_s^i(0))$. We obtain that $\mathbb{X}_{ver} \subseteq \mathcal{R}^{\infty} + \bar{\mathbb{X}}_{ver}$, and that the minimal positively invariant set $\mathcal{R}_{\mathbb{X}_{ver}}^{\infty}$ can be bounded by $\mathcal{R}^{\infty} + \lim_{k \to \infty} \bigcup_{i=0}^{k} A^i \bar{\mathbb{X}}_{ver}$. Under condition of asymptotic stability on A, necessary for $\mathcal{R}^{\infty}$ to be a bounded and convex polytope, $A^i \bar{\mathbb{X}}_{ver}$ will converge to $\{0_n\}$. Thus, the iteration $\mathcal{R}_{\mathbb{X}_{ver}}^{(k)}$ is monotonically increasing and bounded, hence it converges. If $\bar{\mathbb{X}}_{ver}$ includes the origin in its interior then there exists a finite iteration such that $\bigcup_{i=0}^{k} A^i \bar{\mathbb{X}}_{ver} = \bigcup_{i=0}^{k+1} A^i \bar{\mathbb{X}}_{ver}$. Moreover, for any reachable pair (A, B) and asymptotically stable A, the closure of the minimal robustly positively invariant set $\mathcal{R}_{\mathbb{X}_{ver}}^{\infty}$ includes the origin.

Robust approximations of the feasible set via $\Theta_{\psi}(\cdot)$

In order to exploit convergence in the computation of the feasible set for invariance properties, we need to bound the error incurred with the use of approximations of the sets $\mathcal{R}_{\mathbb{X}_{ver}}^{\infty}$ or $\mathcal{R}^{\infty}$. Let $\mathcal{B}$ denote a unit ball centred at the origin and let the Hausdorff distance between sets $\mathcal{R}_1$ and $\mathcal{R}_2$ be defined as

$$\delta_H(\mathcal{R}_1, \mathcal{R}_2) = \inf\{\epsilon \geq 0 | \mathcal{R}_1 \subseteq \mathcal{R}_2 + \epsilon\mathcal{B}, \mathcal{R}_2 \subseteq \mathcal{R}_1 + \epsilon\mathcal{B}\}.$$

We can show that the following holds.

Lemma 6 *Consider a polytope $\mathcal{R}$, and a property ψ comprised of $\psi ::= \alpha | \bigcirc \psi | \psi_1 \wedge \psi_2$, with $\alpha \in \Sigma$, for which $\Theta_{\psi}(\mathcal{R})$ is a non-empty polytope with vertices v_i and the origin in its interior. Let A be bounded as $\|A\|_2 \leq 1$. Then for any $\epsilon_x \geq 0$,*

$$\Theta_{\psi}(\mathcal{R} + \epsilon_x\mathcal{B}) \subseteq \Theta_{\psi}(\mathcal{R}) \subseteq \Theta_{\psi}(\mathcal{R} + \epsilon_x\mathcal{B}) + \epsilon_{\theta}\mathcal{B} \quad (11)$$

if $\epsilon_{\theta} \geq \dfrac{\epsilon_x \epsilon_p \max_i(\|v_i\|)^2}{1 + \epsilon_x \epsilon_p \max_i(\|v_i\|)}$, for $\epsilon_p := \max_{p \in AP} \dfrac{\|A_p\|_2}{|b_p|}$.

Proof 1. $\Theta_{\psi}(\mathcal{R} + \epsilon_x\mathcal{B}) \subseteq \Theta_{\psi}(\mathcal{R})$
Based on the definition of this set (c.f. the proof of Theorem 3), the set operation $\Theta_{\psi}(\cdot)$ is monotonically decreasing. Therefore $\Theta_{\psi}(\mathcal{R} + \epsilon_x\mathcal{B}) \subseteq \Theta_{\psi}(\mathcal{R})$ holds.

2. $\Theta_{\psi}(\mathcal{R}) \subseteq \Theta_{\psi}(\mathcal{R} + \epsilon_x\mathcal{B}) + \epsilon_{\theta}\mathcal{B}$
Consider the case where the model is $(A, B, C(\theta), 0)$. To prove (11), we first find a ϵ_{θ} as a function of ϵ_x such that

$$\Theta_{\psi}(\mathcal{R}) \subseteq \Theta_{\psi}(\mathcal{R} + \epsilon_x\mathcal{B}) + \epsilon_{\theta}\mathcal{B}. \quad (12)$$

Let v_i be the vertices of the polytope $v_i \in \mathcal{V}(\Theta_{\psi}(\mathcal{R}))$, then (12) holds if and only if $v_i \in \Theta_{\psi}(\mathcal{R} + \epsilon_x\mathcal{B}) + \epsilon_{\theta}\mathcal{B}$. Equivalently, this means that there exists a $r_{\theta} \in \epsilon_{\theta}\mathcal{B}$ such that $v_i - r_{\theta} \in \Theta_{\psi}(\mathcal{R} + \epsilon_x\mathcal{B})$. This is equivalent to demanding that for every $\mathbf{x}_j^T \in \mathcal{V}(\mathcal{R})$, $v_i \in \mathcal{V}(\Theta_{\psi}(\mathcal{R}))$ and $r_x \in \epsilon_x\mathcal{B}$, there exists a vector $r_{\theta} \in \epsilon_{\theta}\mathcal{B}$:

$$\left((I_{n_{\psi}} \otimes (\mathbf{x}_j^T + r_x^T))N_{\psi} + K_{\psi}\right)(v_i - r_{\theta}) \leq B_{\psi}$$
$$\Leftrightarrow \quad \left((I_{n_{\psi}} \otimes \mathbf{x}_j^T)N_{\psi} + K_{\psi}\right)(v_i - r_{\theta})$$
$$+ \left((I_{n_{\psi}} \otimes r_x^T)N_{\psi}\right)(v_i - r_{\theta}) \leq B_{\psi}.$$

Take $(v_i - r_{\theta}) = (1 - \alpha_i)v_i$ with $\alpha_i \in [0, 1)$, then

$$\left((I_{n_{\psi}} \otimes \mathbf{x}_j^T)N_{\psi} + K_{\psi}\right)(1 - \alpha_i)v_i$$
$$+ \left((I_{n_{\psi}} \otimes r_x^T)N_{\psi}\right)(1 - \alpha_i)v_i \leq B_{\psi}$$
$$\Leftrightarrow \quad (1 - \alpha_i)(I_{n_{\psi}} \otimes r_x^T)N_{\psi}v_i \leq \alpha_i B_{\psi}. \quad (13)$$

Separate the matrix N_{ψ} and B_{ψ} into its block matrices $N_{\psi}^j = [N_{\psi}]_{\{1+(j-1)n:nj\} \times \{1:n\}}$ and $B^j = [B_{\psi}^j]_j$ such that inequality (13) is equivalent to the set of inequalities

$$(1 - \alpha_i)r_x^T N_{\psi}^j v_i' \leq \alpha_i b^j, \quad \text{for } j = 1, \ldots, n_{\psi} \quad (14)$$
$$\Leftrightarrow \quad r_x^T N_{\psi}^j v_i' \leq \frac{\alpha_i}{(1 - \alpha_i)} b^j \quad . \quad (15)$$

Given that $0 \in \Theta_{\psi}(\mathcal{R})$, it follows that $b_j \geq 0$ for $j = 1, \ldots, n_{\psi}$

$$\max_j \left(r_x^T N_{\psi}^j v_i'\right)(b^j)^{-1} \leq \frac{\alpha_i}{(1 - \alpha_i)} \quad .$$

The term on the left can be upper bounded based on the Cauchy-Schwarz inequality

$$\max_j \left(r_x^T N_{\psi}^j v_i'\right)(b^j)^{-1} \leq \max_j \|(N_{\psi}^j)^T r_x\|_2 \|v_i'\|_2 (b^j)^{-1}$$
$$\leq \max_j \|(N_{\psi}^j)^T\|_2 \|r_x\|_2 \|v_i'\|_2 (b^j)^{-1} \text{ and } \|r_x\|_2 \leq \epsilon_x$$
$$\leq \epsilon_x \epsilon_p \|v_i'\|_2.$$

The last inequality follows from the introduction of the precision of the labelling, denoted as ϵ_p, and defined as

$$\epsilon_p = \max_{p \in AP} \frac{\|A_p\|_2}{|b_p|}. \qquad (16)$$

Remember that $\|L \otimes K\|_2 = \|L\|_2 \|K\|_2$. Then based on Theorem 3 and on the condition $\|A\|_2 \leq 1$, it can be shown that

$$\max_j \|(N_\psi^j)^T\|_2 |b^j|^{-1} \leq \max_{p \in AP} \frac{\|A_p\|_2}{|b_p|}.$$

Note that $\frac{\alpha_i}{(1-\alpha_i)}$ monotonically increases with α_i for $\alpha_i \in [0,1)$. Therefore a bound on α_i can be found as

$$\alpha_i = (\epsilon_x \epsilon_p \|v_i\|)/(1 + \epsilon_x \epsilon_p \|v_i\|) \text{ for } j = 1, \ldots, n_\psi. \quad (17)$$

It follows that (12) holds if

$$\epsilon_\theta = \max(\|v_i\|_2) \frac{\epsilon_x \epsilon_p \max(\|v_i\|_2)}{1 + \epsilon_x \epsilon_p \max(\|v_i\|_2)}. \qquad (18)$$

For the case that the model is parameterised in both S and D, i.e., $(A, B, C(\theta), D(\theta))$ the derivation is a bit more cumbersome but can be repeated with no change to the end result. $\qquad \square$

Let us briefly discuss the conditions under which Lemma 6 is applicable. The condition that $\Theta_\psi(\mathcal{R})$ is not empty is raised to avoid the trivial case where $\Theta_\psi(\mathcal{R}) = \emptyset$ (11) holds for all ϵ_θ. The condition that $\Theta_\psi(\mathcal{R})$ is a polytope and hence bounded is necessary to obtain a bounded Hausdorff distance. This distance quantifies the difference between two sets, and is a necessary step to bound the approximation error. The requirement that $\Theta_\psi(\mathcal{R})$ includes the origin is a sufficient condition and relates to well-posedness for bounded input sets including the origin. When considering invariance properties defined for $0 \in \mathbb{U}_{ver}$ and for any polytope $\mathbb{X}_{ver}$, the requirement that $0_n \in \Theta_\psi(\cdot)$ is necessary for $\Theta_{\square\psi}$ to be non-empty: this can be intuitively illustrated by noting that under an assumption of asymptotic stability for A, for any θ and for $u(\cdot) = 0$ the output $\hat{y}(t, \theta)$ of the model in (6) converges to 0. Hence for a property to be satisfied under these conditions it should at least hold for the zero output, which is equivalent to demanding that it holds for $\theta = 0_n$. For any atomic proposition $p_i \in AP$ (see Equation (5)) it can be shown that there is an invertible mapping between the row vectors, proportional to the normals of the faces of the polyhedral set $\Theta_{p_i}(\mathbf{x}(0))$, and the initial state $\mathbf{x}(0)$. Therefore, if $\mathcal{R}^{(k)}$ has the origin in its interior, then $\Theta_{p_i}(\mathcal{R}^{(k)})$ has to be bounded, and as a consequence so has any feasible set comprising this atomic proposition. This holds for $k \geq n$ if (A, B) is a reachable pair and if $\mathbb{U}_{ver}$ has 0 in its interior. Under the same conditions there exists a k such that $\mathcal{R}^{(k)}_{\mathbb{X}_{ver}}$ has

0_n in its interior. The generalisation to the case dealing with an Hausdorff distance of the feasible set for invariance properties with a set of inputs $0 \notin \mathbb{U}_{ver}$ is outside of the scope of this work.

Convergence properties

We can employ Lemma 6 to bound the Hausdorff distance between $\Theta_\psi(\mathcal{R}^{(k)}_{\mathbb{X}_{ver}})$ and $\Theta_{\square\psi}$. If $\mathbb{X}_{ver} = \{0_n\}$ and the spectral radius of A is strictly less than 1 (that is $\rho(A) < 1$), then the Hausdorff distance can be bounded as

$$\delta_H(\mathcal{R}^{(k)}, \mathcal{R}^\infty) \leq \epsilon(k) := \|A^k\|_2 \max_{u \in \mathbb{U}}(|u|)c_1, \qquad (19)$$

with c_1 a bound on $\sum_{i=0}^\infty \|A^i B\|$, which is the peak-to-peak performance of the dynamical system formed by (A, B). In case that $\mathbb{X}_{ver} \not\subseteq \mathcal{R}^\infty$ then the forward reachable iteration can be rewritten as

$$\mathcal{R}^{(k)}_{\mathbb{X}_{ver}} = \Big(\bigcup_{i=0}^k A^i \mathbb{X}_{ver} \Big) + \mathcal{R}^{(k)}.$$

The Hausdorff norm can be bounded as

$$\delta_H(\mathcal{R}^{(k)}_{\mathbb{X}_{ver}}, \mathcal{R}^\infty_{\mathbb{X}_{ver}}) \leq \epsilon(k) + \|A^{k+1}\|_2 \delta_H(\mathbb{X}_{ver}, \{0_n\}).$$

Note that for $\rho(A) < 1$ the norm $\|A^k\|_2 \to 0$ for $k \to \infty$. In case the conditions of Lemma 6 on $\mathcal{R}^{(k)}_{\mathbb{X}_{ver}} \subseteq \mathbb{X}$ and $\Theta_\psi(\mathcal{R}^{(k)}_{\mathbb{X}_{ver}})$ hold, the Hausdorff distance $\delta_H(\Theta_{\square^k \psi}, \Theta_{\square\psi})$ can be bounded by

$$\|A^k\|_2 \max_i(\|v_i\|)^2 \epsilon_p \big(\max_{u \in \mathbb{U}}(|u|)c_1 + \|A\|\delta_H(\mathbb{X}_{ver}, \{0_n\}) \big). \tag{20}$$

Use in the verification of unbounded-time properties

Based on the convergence properties of the feasible set, the asymptotic behaviour of the confidence computed in Proposition 1 can be stated.

Corollary 7 (Convergence) *Under the conditions of Lemma 6, for a Gaussian distribution $p(\theta) \sim \mathcal{N}(\mu_\theta, R_\theta)$ with a covariance $R_\theta \succ 0$, $\mathbf{P}(\theta \in \Theta_{\square^k \psi}) \to \mathbf{P}(\theta \in \Theta_{\square\psi})$ for $k \to \infty$.*

Proof[of Corollary 7] For a strictly positive R_θ, the Gaussian density distribution takes finite values over the parameter space, therefore the convergence of a monotonically-decreasing polytope over the parameter space induces the convergence of the associated probability measure. $\qquad \square$

Theorem 3 can now be generalised to include unbounded-time invariance properties as follows.

Theorem 8 *Consider a polytopic set of initial states $x(0) \in \mathbb{X}_{ver}$, inputs $u(t) \in \mathbb{U}_{ver}$ for $t \geq 0$, and a labelling map as in (5). Let $\hat{\mathcal{R}}^\infty_{\mathbb{X}_{ver}}$ be a polytopic superset of the minimal robustly positively invariant set that includes $\mathbb{X}_{ver}$, denoted as $\mathcal{R}^\infty_{\mathbb{X}_{ver}}$; then the feasible set admits a polyhedral subset $\hat{\Theta}_\psi \subset \Theta_\psi$ for every specification ψ expressed within the LTL fragment $\psi := \alpha | \bigcirc \psi | \psi_1 \wedge \psi_2 | \Box \psi$, and if $\hat{\mathcal{R}}^\infty_{\mathbb{X}_{ver}} = \mathcal{R}^\infty_{\mathbb{X}_{ver}}$ then $\hat{\Theta}_\psi = \Theta_\psi$.*

Proof Every property $\phi ::= p | \bigcirc \psi | \psi_1 \wedge \psi_2 | \Box \psi$ with $p \in AP$ can be rewritten as $\Box \psi_1 \wedge \psi_2$ where ψ_1 and ψ_2 have syntax $\psi ::= p | \bigcirc \psi | \psi_1 \wedge \psi_2$.

For the set of initial states $\mathbb{X}_{ver}$, a property ψ is invariant

$$\langle \mathbf{M}(\theta), \mathbf{x}(0) \rangle \vDash \Box \psi, \ \forall \mathbf{x}(0) \in \mathbb{X}_{ver}$$

if and only if $\forall x \in \mathcal{R}^\infty_{\mathbb{X}_{ver}} : \langle \mathbf{M}(\theta), x \rangle \vDash \psi$. Let $\hat{\mathcal{R}}^\infty_{\mathbb{X}_{ver}}$ be a polytopic superset of $\mathcal{R}^\infty_{\mathbb{X}_{ver}}$ with a finite set of vertices $v_\mathcal{R} \in V_\mathcal{R}$, then the subset approximation of the feasible set $\Theta_{\Box \psi}$ follows as $\Theta_{\Box \psi} \supseteq \hat{\Theta}_{\Box \psi} =$

$$\left\{ \theta \in \Theta : \bigwedge_{v_\mathcal{R} \in V_\mathcal{R}} \left((I_{n_\psi} \otimes v_\mathcal{R}^T) N_\psi + K_\psi \right) \theta \leq b_\psi \right\}$$

where $\hat{\Theta}_{\Box \psi} \subseteq \Theta_{\Box \psi}$. Note that if $\hat{\mathcal{R}}^\infty_{\mathbb{X}_{ver}} = \mathcal{R}^\infty_{\mathbb{X}_{ver}}$ then $\hat{\Theta}_{\Box \psi} = \Theta_{\Box \psi}$. The feasible set of $\Box \psi_1 \wedge \psi_2$ is equal to $\Theta_{\Box \psi_1 \wedge \psi_2} = \Theta_{\Box \psi_1} \cap \Theta_{\psi_2}$. And $\Theta_{\Box \psi_1 \wedge \psi_2}$ can be upper and lower bounded as $\hat{\Theta}_{\Box \psi_1} \cap \Theta_{\psi_2} \subseteq \Theta_{\Box \psi_1 \wedge \psi_2} \subseteq \Theta_{\Box^k \psi_1} \cap \Theta_{\psi_2}$ with $k \in \mathbb{N}$. This proves Theorem 8 for the case where the model is $(A, B, C(\theta), 0)$. The additional parameterisation of D does not change the reasoning. $\square$

The extension beyond the LTL fragment discussed above may lead to feasible sets that are in general not convex and are therefore beyond the scope of this work.
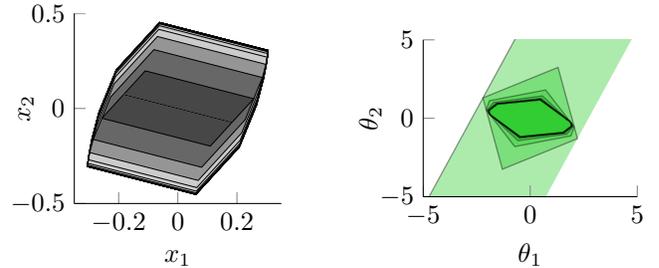
3.5 Case Study (cont.): Unbounded-Time Safety Verification

We study convergence properties for the safety specification ι considered in the case study in Section 3.3 maintaining the same operating conditions as before for the safety verification and the experiment. In Figure 3a the forward reachability sets $\mathcal{R}^{(k)}$ with $k = 1, \ldots, 20$ are obtained for the model dynamics in (8). Figure 4 (upper plot) displays bounds $\epsilon(k)$ on the Hausdorff distances $\delta_H(\mathcal{R}^{(k)}, \mathcal{R}^\infty)$ computed with (19): starting from a slanted line segment for $\mathcal{R}^{(1)}$ as in Figure 3a, it can be observed that the forward reachable sets $\mathcal{R}^{(k)}$ converge rapidly, as confirmed with the error bound displayed in Figure 4 (upper plot).

Based on $\mathcal{R}^{(k)}$, the feasible set for the k-bounded invariance $\Box^k \iota$ can be computed as $\Theta_{\Box^k \iota} = \Theta_\iota (\mathcal{R}^{(k)})$. The feasible sets $\Theta_{\Box^k \iota}$ with $k = 1, \ldots, 20$ are plotted in Figure 3b. Observe that the feasible set $\Theta_{\Box^1 \iota}$ is not bounded, but for $k \geq 2$ the feasible sets are bounded and, as expected, decrease in size with time. In Figure 4 (middle plot) bounds on the Hausdorff distances $\delta_H(\Theta_{\Box \iota}, \Theta_{\Box^k \iota})$ are given for $k = 2, \ldots, 20$ (no finite bound is computed for the index $k = 1$, since for that instance the feasible set is not bounded). Let us conclude this case study looking at confidence quantification, as a function of the time horizon. Figure 4 (lower plot) represents the confidence over the property $\mathbf{P}\left(\theta \in \Theta_{\Box^k \iota} \mid Z^{N_s} \right)$, for indices $k = 1, \ldots, 20$. Unlike the case discussed in Section 3.3, which focused on looking at statistics of the confidence via mean and variance drawn over multiple experiments, we zoom in on asymptotic properties by considering a data set Z^{N_s} comprising a single trace made up of 200 measurements, simulated under the same conditions as in Section 3.3, and with $\theta_0 = [1 \ 0]^T$. From the resulting probability density distribution $p\left(\theta \mid Z^{N_s} \right)$, it may be observed that the confidence converges rapidly to a nonzero value.

3.6 Discussion on the Generalisation of the Results

The discussed approach based on polytopes allows for analytical expressions of the feasible set, however the implementation may not scale to models with very large dimension: in particular, the number of half-planes characterising the feasible set may increase with the time bound of the LTL formula ψ (that is, with the repeated application of the $\bigcirc$ operator), and with the cardinality of the atomic propositions in the alphabet Σ. Still, note that these computations are essentially quite similar to known reachability computations, therefore the method is extendable well beyond the 2-dimensional case study, especially when applying sophisticated reachabil-

(a) The first 20 iterations of the forward reachable set $\mathcal{R}^{(k)}$, $k = 1, \ldots, 20$ for the case study. The reachable sets grow in size from dark grey ($k = 1$) to light grey ($k = 20$), so that $\mathcal{R}^{(k-1)} \subseteq \mathcal{R}^{(k)}$.

(b) The feasible sets for the k-bounded invariance property $\Box^k \iota$, with $k = 1, \ldots, 20$, obtained for the case study.

Fig. 3. Reachable and feasible sets for unbounded-time verification problem.

$$\epsilon(k) \geq \delta_H(\mathcal{R}^{(k)}, \mathcal{R}^{(\infty)})$$

$$\epsilon_\theta(k) \geq \delta_H(\Theta_{\square\iota}, \Theta_{\square^k\iota})$$

$$\mathbf{P}\left(\theta \in \Theta_{\square^k\iota} \mid Z^{N_s}\right)$$
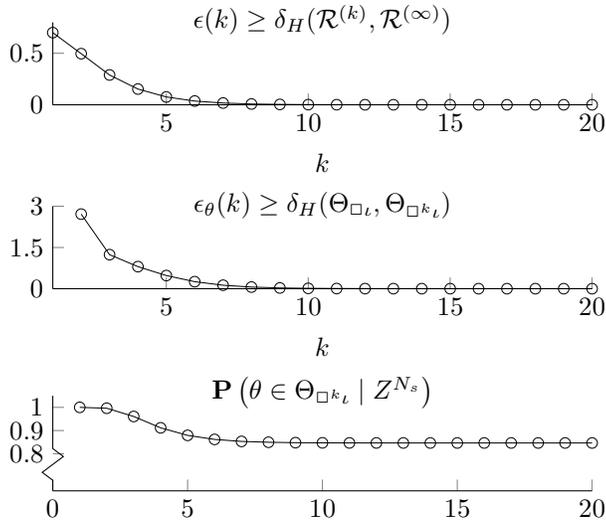
Fig. 4. (Upper plot) Error bound on the approximation level of the k-th forward reachable sets, which is such that $\mathcal{R}^{(\infty)} \subseteq \mathcal{R}^{(k)} + \epsilon(k)$ for $k = 1, \ldots, 20$. (Middle plot) The Hausdorff distance $\epsilon_\theta(k)$ between $\Theta_{\square^k\psi}$ and $\Theta_{\square\psi}$ with $k = 2, \ldots, 20$, obtained for the case study.(Lower plot) Confidence that $\mathbf{S} \vDash \square^k\iota$ for $k = 1, \ldots, 20$ for the case in Section 3.3, with a new experiment consisting of 200 samples collected as Z^{N_s}.

ity analysis tools in the literature. Therefore the discussed limitations related to the current implementation of the approach, ought to be dealt with in the future by the use of tailored and less naïve computational approaches.

In the discussion of model selection, we hinted at possible generalisation beyond linearly-parameterised model sets. Future extension will deal with hybrid models, since when systems are not linear, their (local) behaviour is often well approximated with piecewise-linear dynamical models.

This paper has discussed the formal verification of physical systems with partly unknown dynamics, by introducing a Bayesian framework allowing for the efficient incorporation of measurement data and prior information within a verification procedure based on safety analysis. This formal approach has allowed for the computation of the confidence level over the validity of a property of interest on the unknown system. The method has been applied to the verification of LTI models of systems over bounded and unbounded safety properties, and its computational overhead has been discussed at length.

Looking forward, current work targets the extension of the applicability of tractable solutions to model-based and data-driven verification over complex physical systems. We are presently working to extensions of the considered set of logic formulae of interest, and plan to employ experiment design to optimise the input-output

signal interaction for efficient data usage over general classes of models, as initially attempted in [17]. Additionally, the design of control policies that optimise properties of interest over partly unknown systems is topic of current work.

References

[1] A. Abate, R. C. Hillen, and S. A. Wahl. Piecewise affine approximation of fluxes and enzyme kinetics from in-vivo ^{13}C labeling experiments. *International Journal of Robust and Nonlinear Control*, pages 1120–1139, 2012. Special Issue on System Identification for Biological Systems.

[2] C. Baier and J.-P. Katoen. Principles of model checking. *MIT Press*, 2008.

[3] E. Bartocci, L. Bortolussi, and G. Sanguinetti. Learning temporal logical properties discriminating ECG models of cardiac arrhythmias. *CoRR*, abs/1312.7523, 2013.

[4] G. Batt, C. Belta, and R. Weiss. Model checking genetic regulatory networks with parameter uncertainty. In *HSCC*, pages 61–75. Springer, 2007.

[5] C. Belta, A. Bicchi, M. Egerstedt, E. Frazzoli, E. Klavins, and G. J. Pappas. Symbolic planning and control of robot motion [grand challenges of robotics]. *Robotics Automation Magazine, IEEE*, pages 61–70, Mar. 2007.

[6] C. Belta, L. C. G. J. M. Habets, and V. Kumar. Control of multi-affine systems on rectangles with applications to hybrid biomolecular networks. In *Conf.on CDC*, pages 534–539, 2002.

[7] F. Blanchini and S. Miani. *Set-Theoretic Methods in Control*. Birkhäuser Basel, 1st edition, 2007.

[8] L. Bortolussi and G. Sanguinetti. Learning and designing stochastic processes from logical constraints. In *QEST*, pages 89–105. Springer, 2013.

[9] L. Bortolussi and G. Sanguinetti. Smoothed model checking for uncertain continuous time Markov chains. *CoRR*, abs/1402.1450, 2014.

[10] L. Brim, M. Češka, S. Dražan, and D. Šafránek. Exploring parameter space of stochastic biochemical systems using quantitative model checking. In N. Sharygina and H. Veith, editors, *CAV*, volume 8044 of *LNCS*, pages 1–17. Springer, 2013.

[11] J. W. Burdick, N. du Toit, A. Howard, C. Looman, J. Ma, R. M. Murray, and T. Wongpiromsarn. Sensing, navigation and reasoning technologies for the DARPA urban challenge. Technical report, DTIC Document, 2007.

[12] Y. Chen and T. D. Nielsen. Active learning of Markov decision processes for system verification. In *Conf. on Machine Learning and Applications*, pages 289–294, 2012.

[13] E. M. Clarke. The birth of model checking. In *25 Years of Model Checking*, pages 1–26, 2008.

[14] D. Del Vecchio and E. D. Sontag. Engineering principles in bio-molecular systems: From retroactivity to modularity. *European Journal of Control*, pages 389 – 397, 2009.

[15] G. Frehse, S. K. Jha, and B. H. Krogh. A counterexample-guided approach to parameter synthesis for linear hybrid automata. In *HSCC*, pages 187–200. Springer Berlin Heidelberg, 2008.

[16] B. M. Gyori, D. Paulin, and S. K. Palaniappan. Probabilistic verification of partially observable dynamical systems. *CoRR*, abs/1411.0976, 2014.

[17] S. Haesaert, P. M. J. Van den Hof, and A. Abate. Data-driven property verification of grey-box systems by Bayesian experiment design. In *American Control Conference*, pages 1800–1805, 2015.

[18] D. Henriques, J. G. Martins, P. Zuliani, A. Platzer, and E. M. Clarke. Statistical model checking for Markov decision processes. In *QEST*, pages 84–93, 2012.

[19] T. Henzinger and H. Wong-Toi. Using hytech to synthesize control parameters for a steam boiler. In *Formal Methods for Industrial Applications*, pages 265–282. Springer Berlin Heidelberg, 1996.

[20] P. S. C. Heuberger, P. M. J. Van den Hof, and O. H. Bosgra. A generalized orthonormal basis for linear dynamical systems. *Automatic Control, IEEE Transactions on*, 40(3):451–465, 1995.

[21] P. S. C. Heuberger, P. M. J. Van den Hof, and B. Wahlberg. *Modelling and identification with rational orthogonal basis functions*. Springer London, 2005.

[22] H. Hjalmarsson. From experiment design to closed-loop control. *Automatica*, pages 393–438, 2005.

[23] E. A. Lee. Cyber physical systems: Design challenges. In *Proc. of Object Oriented Real-Time Distributed Computing*, pages 363–369. IEEE Computer Society, 2008.

[24] A. Legay, B. Delahaye, and S. Bensalem. Statistical model checking: An overview. In H. Barringer, Y. Falcone, B. Finkbeiner, K. Havelund, I. Lee, G. Pace, G. Roşu, O. Sokolsky, and N. Tillmann, editors, *Runtime Verification*, volume 6418 of *LNCS*, pages 122–135. Springer Berlin Heidelberg, 2010.

[25] A. Legay and S. Sedwards. Lightweight Monte Carlo algorithm for Markov decision processes. *CoRR*, abs/1310.3609, 2013.

[26] D. V. Lindley. The philosophy of statistics. *Journal of the Royal Statistical Society: Series D (The Statistician)*, pages 293–337, 2000.

[27] H. Mao and M. Jaeger. Learning and model-checking networks of I/O automata. In *Proc. of Asian Conference on Machine Learning*, 2012.

[28] V. Peterka. Bayesian Approach to System Identification. *Trends Prog. Syst. Identif.*, 1981.

[29] B. C. Reginato, R. Z. Freire, G. H. D. C. Oliveira, N. Mendes, and O. Abadie, Marc. Predicting the temperature profile of indoor buildings by using orthonormal basis functions. In *Conf. on Building Performance Simulation Association*, United Kingdom, 2009.

[30] K. Sen, M. Viswanathan, and G. Agha. Learning continuous time Markov chains from sample executions. In *QEST*, pages 146–155, 2004.

[31] K. Sen, M. Viswanathan, and G. Agha. Statistical model checking of black-box probabilistic systems. In R. Alur and D. Peled, editors, *CAV*, volume 3114 of *LNCS*, pages 202–215. Springer, 2004.

[32] P. Tabuada. *Verification and Control of Hybrid Systems: a Symbolic Approach*. Springer, 2009.

[33] P. M. J. Van den Hof, P. S. C. Heuberger, and J. Bokor. System identification with generalized orthonormal basis functions. *Automatica*, pages 1821–1834, 1995.

[34] M. Y. Vardi. From philosophical to industrial logics. In *Proc. of the Indian Conference on Logic and Its Applications*, pages 89–115, Berlin, Heidelberg, 2009. Springer-Verlag.

[35] G. S. Virk and D. L. Loveday. Model-based control for HVAC applications. In *Conf. on Control Applications*, pages 1861–1866. IEEE, 1994.

[36] P. Zuliani, A. Platzer, and E. M. Clarke. Bayesian statistical model checking with application to Stateflow/Simulink verification. *Formal Methods in System Design*, pages 338–367, 2013.

Derivation of the Bounds in Section 3.4

1. Hausdorff distance of forward reachable mappings. We only sketch the method to bound the Hausdorff distance, whereas a more formal derivation can be found in the literature on robustly positively invariant sets [7].

The k-step forward reachable set equals

$$\mathcal{R}^{(k)} := \bigcup_{i=1}^{k} \left\{ \sum_{j=1}^{i} A^{j-1} B u(i-j), \;\; \text{for } u(j) \in \mathbb{U}_{ver} \right\}.$$

For $0 \in \mathbb{U}_{ver}$, the minimal invariant set $\mathcal{R}^{\infty}$ can be written as

$$\mathcal{R}^{(\infty)} := \left\{ \sum_{j=0}^{i-1} A^j B u(j) + A^i \sum_{k=0}^{\infty} A^k B u(k), \; \text{for } u(\cdot) \in \mathbb{U}_{ver} \right\}. \tag{.1}$$

If the spectral radius of a A is strictly smaller than 1, $\rho(A) < 1$, then

$$\mathcal{R}^{(\infty)} \subseteq \mathcal{R}^{(k)} + \epsilon(k) \mathcal{B}, \tag{.2}$$

with

$$A^k \sum_{i=0}^{\infty} A^i B u(k) \subseteq \epsilon(k) \mathcal{B}, \; \text{for } u(\cdot) \in \mathbb{U}_{ver}.$$

Note that $\epsilon(k)$ is bounded for $\rho(A) < 1$. For a matrix A without defective eigenvalues, i.e. where the eigenvectors form a complete basis, this L_1 norm can be easily bounded using the spectral radius of A, by selecting

$$\epsilon(k) = \frac{|\rho(A)|^k}{1 - |\rho(A)|} \|B\|_2 \max_{u \in \mathbb{U}_{ver}} (|u|) \geq \|A^k\|_2 \sum_{i=0}^{\infty} \|A^i B\|_2 |u(k)|.$$

In case that the matrix A is defective, we opt to bound the L_1-norm by exploiting absolute sum of the L_2 induced norm for A^i $i \to \infty$: $\sum_{i=0}^{\infty} \|A^i\|_2$. Note that $\|A^i\|_2$ converges to 0 for $i \to \infty$ since $\rho(A) < 1$, therefore there exists a finite l such that $\|A^l\|_2 < 1$ and we can upper bound the absolute sum as

$$\sum_{i=0}^{\infty} \|A^i\|_2 \leq \left(\sum_{i_1=0}^{l-1} \|A^{i_1}\|_2 \right) \left(\sum_{i_2=0}^{\infty} \|A^l\|_2^{i_2} \right)$$

$$= \left(\sum_{i_1=0}^{l-1} \|A^{i_1}\|_2 \right) \frac{1}{1 - \|A^l\|_2}.$$

Thus in general, the Hausdorff distance can be bounded as

$$\delta_H(\mathcal{R}^{(k)}, \mathcal{R}^{(\infty)}) \le \epsilon(k) = \|A^k\|_2 \max_{u \in \mathbb{U}_{ver}} (|u|) c_1,$$

with $c_1 = \dfrac{\left(\sum_{i_1=0}^{l} \|A^{i_1}\|_2 \right)}{1 - \|A^l\|_2} \|B\|_2$ for l such that $\|A^l\|_2 < 1$. Note that c_1 can be replaced by any bound on the L_1 norm of the dynamical system formed by (A, B).

In case that $\mathbb{X}_{ver} \not\subseteq \mathcal{R}^{\infty}$ then the forward reachable iteration can be rewritten as

$$\mathcal{R}_{\mathbb{X}_{ver}}^{(k)} = \left(\bigcup_{i=0}^{k} A^i \mathbb{X}_{ver} \right) + \mathcal{R}^{(k)},$$

for which we know that

$$\mathcal{R}_{\mathbb{X}_{ver}}^{(\infty)} \subseteq \mathcal{R}_{\mathbb{X}_{ver}}^{(k)} + \epsilon(k) + \|A\|^{k+1} \delta_H(\mathbb{X}_{ver}, \{0\}).$$

Thus the Hausdorff norm is upper bounded as
$\delta_H(\mathcal{R}_{\mathbb{X}_{ver}}^{(k)}, \mathcal{R}_{\mathbb{X}_{ver}}^{(\infty)}) \le \epsilon(k) + \|A^{k+1}\| \delta_H(\mathbb{X}_{ver}, \{0\})$.

2. Hausdorff distance on feasible sets. Suppose that the conditions in Lemma 6 hold for $\mathcal{R}_{\mathbb{X}_{ver}}^{(k)}$, then we can compute a value for ϵ_θ such that $\Theta_\psi(\mathcal{R}_{\mathbb{X}_{ver}}^{(k)}) \subseteq \Theta_\psi(\mathcal{R}_{\mathbb{X}_{ver}}^{(k)} + \epsilon_x \mathcal{B}) + \epsilon_\theta \mathcal{B}$, where ϵ_x is a bound on the Hausdorff distance $\delta_H(\mathcal{R}_{\mathbb{X}_{ver}}^{(k)}, \mathcal{R}_{\mathbb{X}_{ver}}^{(\infty)})$.

The set operation $\Theta_\psi(\cdot)$ is monotonically decreasing, therefore $\Theta_\psi(\mathcal{R}_{\mathbb{X}_{ver}}^{(k)} + \epsilon(k)\mathcal{B}) \subseteq \Theta_{\square\psi} = \Theta_\psi\left(\mathcal{R}_{\mathbb{X}_{ver}}^{\infty}\right) \subseteq \Theta_\psi\left(\mathcal{R}_{\mathbb{X}_{ver}}^{(k)}\right) = \Theta_{\square^k\psi}$, and $\Theta_{\square^k\psi} \subseteq \Theta_\psi(\mathcal{R}_{\mathbb{X}_{ver}}^{(k)} + \epsilon(k)\mathcal{B}) + \epsilon_\theta\mathcal{B} \subseteq \Theta_{\square\psi} + \epsilon_\theta\mathcal{B}$, and

$$\Theta_{\square\psi} \subseteq \Theta_{\square^k\psi} \subseteq \Theta_{\square\psi} + \epsilon_\theta\mathcal{B}.$$

Based on Lemma 6, with $\epsilon_p = \max_{p_i} \frac{|A_{p_i}|}{|b_{p_i}|}$, we obtain

$$\epsilon_\theta = \frac{\epsilon_x \epsilon_p \max_i(\|v_i\|)^2}{1 + \epsilon_x \epsilon_p \max_i(\|v_i\|)} \le \epsilon_x \epsilon_p \max_i(\|v_i\|)^2.$$

Note that since $\|A^k\|_2$ converges to 0 for $k \to \infty$ for $\rho(A) < 1$, and since $\max_i(\|v_i\|)^2$ is not increasing, the error ϵ_θ also converges to 0.