Verifying Stochastic Hybrid Systems with Temporal Logic Specifications via Mori-Zwanzig Model Reduction

Yu Wang, Nima Roohi, Matthew West, Mahesh Viswanathan, and Geir E. Dullerud

Abstract—**We present a scalable methodology to verify stochastic hybrid systems. Using the Mori-Zwanzig reduction method, we construct a finite state Markov chain reduction of a given system. We prove that this reduced Markov chain is approximately equivalent to the original system in a distributional sense. Approximate equivalence of the stochastic hybrid system and its Markov chain reduction means that analyzing the Markov chain with respect to a suitably strengthened property, allows us to conclude whether or not the stochastic hybrid system meets its temporal specification. We present the first statistical model checking algorithms to verify finite state Markov chains against correctness properties expressed in linear inequality linear temporal logic (iLTL), and metric interval temporal logic (MITL).**

I. Introduction

Stochastic hybrid systems, modeling discrete, continuous, and stochastic behavior, arise in many real-world applications ranging from automobiles [1], smart grids [2] and biology [3, 4, 5, 6, 7]. In these contexts, it is often useful to determine if the models meet their time-dependent design goals. However, the verification problem is computationally very challenging — even for systems with very simple dynamics that exhibit no stochasticity, and for the most basic class of safety properties, namely invariants, the problem of determining if a system meets its safety goals is undecidable [8]. The difficulty of the verification problem largely arises from the fact that the state space of such systems has uncountably many states.

The computational challenge posed by the verification problem is often addressed by constructing a simpler finite state model of the system, and then analyzing the finite state model. The finite state model is typically an *abstraction* or a conservative over-approximation of the original system, i.e., every behavior of the system is exhibited by the finite state model, but the finite state model may have additional behaviors that are not system behaviors. This approach has been used to verify [9, 10, 11] and design controllers [12, 13, 14, 15] for non-stochastic systems, and to verify [4, 5, 7, 16, 17] and design controllers [18] for stochastic hybrid systems. For such abstractions, if the finite state model is safe then so is the original system. However, if the finite state model is unsafe, then not much can be concluded about the safety of the original system because the finite state model is an over-approximation.

In this paper, we present a scalable approach to the verification of a class of specifications defined by iLTL or MITL [19, 20] for stochastic hybrid system that relies on constructing a finite state approximation that is "equivalent" to the original system. These specifications reason over the evolution of the probability distribtuions of the systems, and can express a wide class of safety properties.

To verify these specifications, we construct an approximate bi-simulation between stochastic hybrid systems and finite state Markov chains using the Mori-Zwanzig model reduction method [21, 22]. The advantage of bi-simulation is that analyzing the finite state model not only allows us to conclude the safety of the hybrid stochastic system, but also its non-safety. In order to explain the relationship between the Markov chain we construct and the stochastic hybrid system, it is useful to recall that there are two broad approaches to defining the semantics of a stochastic process. One approach is to view a stochastic system as defining a measure space on the collection of executions; by execution here we mean a sequence of states that the system may possibly go through. The other approach is to view the stochastic system as defining a transformation on distributions; in such a view, the behavior of the stochastic model is captured by a sequence of distributions, starting from some initial distribution. For the first semantics (of measures on executions), it has been shown that approxiamte abstractions can be build between finite state Markov chains and certain classes of stochastic hybrid systems [23, 24, 25]. However, it has been observed that constructing an approximate "equivalence" between Markov chains and infinite-state systems is very challenging in general [7].

In this paper, we in contrast show that the Mori-Zwanzig reduction method constructs a finite state Markov chain that is approximately equivalent to a stochastic hybrid system with respect to the second semantics. That is, we show that the distribution on states of the Markov chain at any time, is close to the distribution at the same time defined by the stochastic hybrid system (Theorem 2), even though there might be no (approxiamte) probabilistic path-to-path correspondence between the path space of the stochastic hybrid system and that of the Markov chain, as it is required under the first semantics.

Similar to [24, 25], the Mori-Zwanzig reduction is per-

Yu Wang and Geir E. Dullerud are with Coordinate Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA. Email: yuwang8@illinois.edu, dullerud@illinois.edu

Matthew West is with the Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA. Email: mwest@illinois.edu

Nima Roohi and Mahesh Viswanathan are with the Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA. Email: roohi2@cis.upenn.edu, vmahesh@illinois.edu

This work is supported by NSF CPS Grant 1329991 and AFOSR Grant FA9550-15-1-0059.

formed via partitioning the state space, although the metric for "equivalence" is different. The approximate equivalence by Mori-Zwanzig reduction can be seen to be similar in spirit to the results first established for non-stochastic, stable, hybrid systems [12, 26, 27], and later extended to stochastic dynamical systems [28, 29]. When compared to [28, 29], we consider a more general class of stochastic hybrid systems that have multiple modes and jumps with guards and resets. Second, our reduced system is a Markov chain, whereas in [28, 29] the stochastic system is approximated by a finite state, non-stochastic model. In addition, our notion of distance between the stochastic hybrid system and the reduced system is slightly different.

Having proved that our reduced Markov model is approximately equivalent to the original stochastic hybrid system, we can exploit this to verify stochastic hybrid systems. Approximate equivalence ensures that analyzing the reduced model with respect to a suitably strengthened property, allows us to determine whether the initial stochastic hybrid system meets or violates its requirements. Therefore, a scalable verification approach can be obtained by developing algorithms to verify finite state Markov chains. Since the reduced system, even though finite state, is likely to have a large number of states, we use a statistical approach to verification [30] as opposed to a symbolic one.

In statistical model checking, the model being verified is simulated multiple times, and the drawn simulations are analyzed to see if they constitute a statistical evidence for the correctness of the model. Statistical model checking algorithms have been developed for logics that reason about measures of executions [6, 30, 31, 32]. However, since our reduced Markov chain is only close to the stochastic hybrid system in a distributional sense, we cannot leverage these algorithms. Instead, we develop new statistical model checking algorithms for temporal logics (over discrete and continuous time) that reason about sequences of distributions.

The scalability of our approach depends critically on the way the partition-based Mori-Zwazig model reduction is performed, as it involves numerical integrations on the partitions. For stochastic hybrid systems with nonlinear but polynomial dynamics for the continuous part, the curse of dimensionality for direct numerical integration can be avoided as explicit symbolic solutions for the numerical solution exists. This is demonstrated in Section VI by a case study. Also, Monte-Carlo integrations can be adopted for more general dynamics with considerations on extra statistical errors. Finally, we note that using this approach, we were the first to successfully verify [33] a highly non-linear model including lookup tables of a powertrain control system that was proposed as a challenging problem for verification tools by Toyota engineers [1].

This paper is a generalization and unification of our previous papers [34, 35, 36]. The rest of the article is organized as follows. In Section II we introduce the general setup of the problem, including the definition of continuous-time stochastic hybrid systems and the syntax and semantics of metric interval temporal logic. In Section III we use the Mori-Zwanzig method to reduce the hybrid system to a Markov chain and prove that the temporal logic formulas on the hybrid system can be verified by checking slightly stronger formulas on the Markov chain. In Section IV we develop a statistical model-checking algorithm for actually carrying out the verification. In Section V we consider discrete-time stochastic hybrid systems and derive similar model reduction and model-checking results in this setting. Finally, we conclude in Section VII.

II. Problem Formulation

In the rest of the paper, we denote the set of natural, rational, non-negative rational, real, positive real, and non-negative real numbers by $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{Q}_{\geq 0}$, $\mathbb{R}$, $\mathbb{R}_{>0}$ and $\mathbb{R}_{\geq 0}$ respectively. We denote the essential supremum by ess sup. For $n \in \mathbb{N}$, let $[n] = \{1, 2, \dots, n\}$. For any set $\mathbb{S}$, let $S^{\boldsymbol{\omega}}$ be the set of infinite sequences in $\mathbb{S}$. For $s \in S^{\boldsymbol{\omega}}$, let s_i be the i^{th} element in the sequence. For a finite set A, we denote the cardinality by $|A|$ and its power set by 2^A. The empty set is denoted by $\emptyset$. For $X \subseteq \mathbb{R}^d$, we denote the boundary of X by ∂X. The symbols $\mathbb{P}$ and $\mathbb{E}$ are used for the probability and the expected value, respectively.

A. Stochastic Hybrid System: Continuous-time

In this work, we follow the formal definitions of continuous-time stochastic hybrid systems in [37, 38, 39, 40] as shown in Fig. 1. However, we focus on a Fokker-Planck formulation and interpretation of the model. We denote the continuous and discrete states by $x \in \mathbb{R}^d$ and $q \in \mathcal{Q}$ respectively, where $\mathcal{Q} = \{q_1, \dots, q_m\}$ is a finite set. We call the combination (q, x) the state of the system, and the product set $\mathbb{X} = \mathcal{Q} \times \mathbb{R}^d$ the state space.

For each $q \in \mathcal{Q}$, the state of the system flows in the $\mathbb{A}_q$ and jumps forcedly on hitting the boundary $\mathbb{A}_q$. We assume that each $\mathbb{A}_q$ is open and bounded, and the boundaries $\partial \mathbb{A}_q$ are second-order continuously differentiable. On the flow set, the state x of the system evolves by a stochastic differential equation

$$\mathrm{d}\mathbf{x} = f(\mathbf{q}, \mathbf{x})\mathrm{d}t + g(\mathbf{q}, \mathbf{x})\mathrm{d}B_t, \tag{1}$$

where $\mathbf{q}$ and $\mathbf{x}$ are random processes describing the stochastic evolution of the discrete and continuous states, and B_t is the standard n-dimensional Brownian motion. The vector-valued function f specifies the drift of the state, and the matrix-valued function g describes the intensity of the diffusion [41, 42]. In (1), we assume that $f(q, \cdot)$ and $g(q, \cdot)$ are locally Lipschitz continuous. Meanwhile, the system jumps spontaneously by a non-negative integrable rate function $r(q, x)$. The probability distribution of the jumping target is given by a non-negative integrable target distribution $h(q', x', q, x)$. When the state of the system falls onto the jump set $\mathbb{B}$, the system is forced to jump. The probability distribution of the jumping target is given by a non-negative integrable target distribution $h'(q', x', q, x)$. The two target distributions h and h' defined on two disjoint sets $\mathbb{A}$ and $\mathbb{B}$ are combined into one target transition h defined on the state space $\mathbb{X}$ of the system and satisfying

$$\sum_{q \in \mathcal{Q}} \int_{\mathbb{R}^d} h(q', x', q, x)\mathrm{d}x' = 1. \tag{2}$$

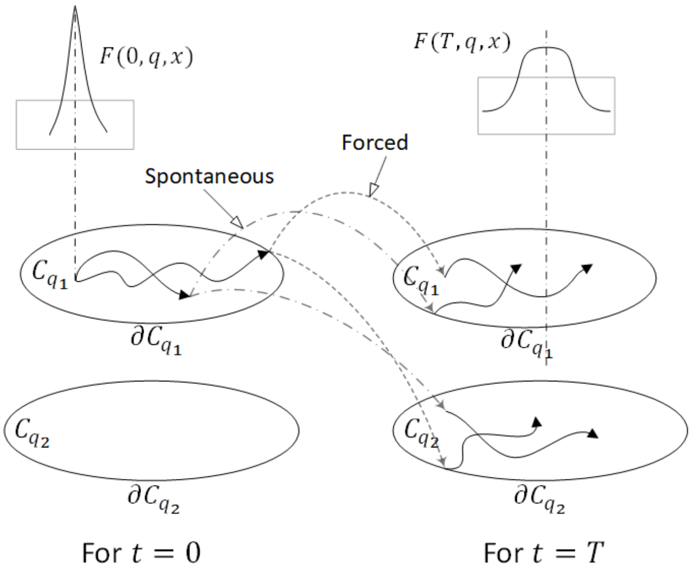


Fig. 1: A continuous-time stochastic hybrid system with two discrete states at time 0 and T.

We are mainly concerned with the dynamics in the flow set $\mathbb{B}$, as the system path immediately jumps back once hitting the jump set $\mathbb{A}$. The probability distribution $F(t,q,x)$ of the state of the system in the flow set $\mathbb{B}$ is determined by the Fokker-Planck equation, which can be derived in the same way as that for jump-diffusion processes [43],

$$\begin{aligned}
&\frac{\partial F(t,q,x)}{\partial t} = L(F(t,q,x)) \\
&= \underbrace{-\sum_{a=1}^{d} \frac{\partial}{\partial x_a}(f_a(q,x)F(t,q,x))}_{\text{drift}} \\
&\quad \underbrace{+\sum_{a=1}^{d}\sum_{b=1}^{d} \frac{\partial^2}{\partial x_a \partial x_b} \sum_{c=1}^{d} \frac{g_{ac}(q,x)g_{cb}(q,x)F(t,q,x)}{2}}_{\text{diffusion}} \\
&\quad \underbrace{-r(q,x)F(t,q,x)}_{\text{jump-out}} \\
&\quad \underbrace{+\sum_{q\in\mathcal{Q}} \int_{\mathbb{R}^d} h(q,x,q',x')r(q',x')F(t,q',x')\mathrm{d}x}_{\text{jump-in}},
\end{aligned} \tag{3}$$

where L is the Fokker-Planck operator for the system. We write symbolically that $F(t,q,x) = e^{tL}F(0,q,x)$. In (3), the four terms on the right hand side describe "drift", "diffusion", "jump-out" and "jump-in", respectively, in which the "jump-in" term includes sources from both spontaneous and forced jumps.

On the other hand, a Fokker-Planck equation with proper boundary conditions that give unique solution defines a stochastic differential equation with jump and diffusion [41, 42]. Therefore, we make the following assumption.

Assumption 1. *In the rest of the paper, we assume that the stochastic hybrid system given in this section is well defined in the sense that it gives a Fokker-Planck equation with a unique solution.*

An invariant distribution of the continuous-time stochastic hybrid system $F_{\text{inv}}(q,x)$ is defined by

$$L(F_{\text{inv}}(q,x)) = 0. \tag{4}$$

In this work, when handling temporal logic specifications of infinite time horizon, we assume that $F(t,q,x)$ converges to the invariant distribution function $F_{\text{inv}}(q,x)$ to ensure that the truth value of the specifications will not change after finite time.

The state of the system is only partially observable. Here, we are interested in observables of the system given by

$$y(t) = \mathbb{E}[y(q(t),x(t))] = \sum_{q\in\mathcal{Q}} \int_{\mathbb{R}^d} \gamma(q,x)F(t,q,x)\mathrm{d}x, \tag{5}$$

where $\gamma(q,x)$ is a weight function on $\mathbb{X}$, which is integrable in x for each $q \in \mathcal{Q}$.

Running Example: Step 1. *Throughout the paper, we use the following example to illustrate the theorems. Consider a continuous-time stochastic hybrid system with two discrete states on $\mathbb{X} = \{1\} \times [0,1] \cup \{2\} \times [2,4]$. It reflects at $x = 0$ and $x = 4$, jumps uniformly to $[2,4]$ when hitting $x = 1$, and jumps uniformly to $[2,4]$ when hitting $x = 2$. It can jump spontaneously at any $x \in \mathbb{X}$ with rate 1 with $r_{\mathbb{A}} = \mathbf{U}_{\mathbb{X}}/3$. In each location, the state of system is governed by the stochastic differential equation*

$$\mathrm{d}\mathbf{x} = \mathrm{d}t + \mathrm{d}B_t, \tag{6}$$

The probability distribution $F(t,q,x)$ of the state evolves by the Fokker-Planck equation

$$\frac{\partial F(t,q,x)}{\partial t} = \frac{\partial F(t,q,x)}{\partial x} + \frac{1}{2}\frac{\partial^2 F(t,q,x)}{\partial x^2} \tag{7}$$

with the boundary conditions

$$\begin{aligned}
&\frac{\partial}{\partial x}F(t,q,0)=0,\\
&\frac{\partial}{\partial x}F(t,q,1)=\frac{1}{2}\int_{[2,4]}\frac{\partial}{\partial t}F(t,q,x)\mathrm{d}x,\\
&\frac{\partial}{\partial x}F(t,q,2)=\int_{[1,2]}\frac{\partial}{\partial t}F(t,q,x)\mathrm{d}x,\\
&\frac{\partial}{\partial x}F(t,q,4)=0.
\end{aligned}\tag{8}$$

Initially, the state of the system is uniformly distributed on $[0,1/2]$.

In the first part of this work, we are interested in verifying temporal properties of continuous-time stochastic hybrid systems. These properties will be specified in the following way: the atomic propositions are inequalities $y\sim c$ ($c\in\mathbb{Q}$, $\sim\in\{<,\leq,\geq,>\}$) on the observables of the system; and they are concatenated by the syntax of Metric Interval Temporal Logic (MITL) [20]. This type of logic is also referred to as Signal Temporal Logic (STL) [44, 45, 46] in the literature. The syntax of MITL is given in Definition 1.

Definition 1 (MITL Syntax)**.** *Let* $\mathbb{I}_{\geq 0}$ *be the set of intervals on* $\mathbb{R}$ *that are both non-negative and non-singleton. An MITL formula is defined using the following BNF form:*

$$\varphi=\bot\mid\top\mid p\mid\varphi\wedge\varphi\mid\varphi\vee\varphi\mid\varphi\mathcal{U}_I\varphi\mid\varphi\mathcal{R}_I\varphi,$$

where $p\in\mathtt{AP}$ *and* $I\in\mathbb{I}_{\geq 0}$.

We note that this syntax does not contain negation ($\neg$), since $\{<,\leq,\geq,>\}$ is closed under negation. For a standard MITL formula, negation on non-atomic formulas can always be pushed inside as part of the atomic propositions. For example, $\neg(y>0)$ is equivalent to $y\leq 0$, $\neg(\varphi_1\vee\varphi_2)$ is equivalent to $(\neg\varphi_1)\wedge(\neg\varphi_2)$, and $\neg(\varphi\mathcal{U}_I\psi)$ is equivalent to $(\neg\varphi)\mathcal{R}_I(\neg\psi)$.

Given a set of atomic propositions, the continuous-time stochastic hybrid system induces a signal $\sigma:\mathbb{R}_{\geq 0}\to 2^{\mathtt{AP}}$, in which $\sigma(t)$ is the set of atomic propositions that hold on the system at time t. The semantics of MITL are defined with respect to the function $\sigma(t)$ as follows.

Definition 2 (MITL Semantics)**.** *Let* φ *be an MITL formula and* f *be a signal* $f:\mathbb{R}_{\geq 0}\to 2^{\mathtt{AP}}$. *The satisfaction relation* $\models$ *between* f *and* φ *is defined according to the following inductive rules:*

$$\begin{array}{lll}
\sigma\models\bot & \text{iff} & \textit{always false}\\
\sigma\models\top & \text{iff} & \textit{always true}\\
\sigma\models y\sim c & \text{iff} & (y\sim c)\in\sigma(0)\\
\sigma\models\varphi\wedge\psi & \text{iff} & (\sigma\models\varphi)\wedge(\sigma\models\psi)\\
\sigma\models\varphi\vee\psi & \text{iff} & (\sigma\models\varphi)\vee(\sigma\models\psi)\\
\sigma\models\varphi\mathcal{U}_I\psi & \text{iff} & \exists t\in I,(\sigma^t\models\psi)\\
 & & \wedge\forall t'\in(0,t),\sigma^{t'}\models\varphi\\
\sigma\models\varphi\mathcal{R}_I\psi & \text{iff} & \forall t\in I,(\sigma^t\models\psi)\quad \textit{or}\\
 & & \exists t\in\mathbb{R}_{>0},(\sigma^t\models\varphi)\\
 & & \quad\wedge\forall t'\in[0,t_1]\cap I,\sigma^{t'}\models\psi
\end{array}$$

where σ^r *is a signal that maps* t *to* $\sigma(t+r)$. *We define* $[\![\varphi]\!]$ *to be the set of signals that satisfy* φ.

Satisfiability and model checking problems for MITL with *abstract* atomic propositions are known to be EXPSPACE-complete [20]. The corresponding decision procedure has a close connection with timed automata.

Definition 3 (Timed Automata [47])**.** *Timed automaton* A *is a tuple* $(\mathtt{Q},\mathtt{X},\Sigma,\mathtt{L},\mathtt{I},\mathtt{E},\mathtt{Q}^{\mathsf{init}},\mathtt{Q}^{\mathtt{final}})$ *where*

- $\mathtt{Q}$ *is a finite non-empty set of* locations.
- $\mathtt{X}$ *is a finite set of* clocks.
- Σ *is a finite* alphabet.
- $\mathtt{L}:\mathtt{Q}\to\Sigma$ *maps each location to the* label *of that location.*
- $\mathtt{I}:\mathtt{Q}\to(\mathtt{X}\to\mathbb{I}_{\geq 0})$ *maps each location to its* invariant *which is the set of possible values of variables in that location, where* $\mathbb{I}_{\geq 0}$ *is the set of intervals on* $\mathbb{R}_{\geq 0}$.
- $\mathtt{E}\subseteq\mathtt{Q}\times\mathtt{Q}\times 2^{\mathtt{X}}$ *is a finite set of* edges *of the form* $e=(s,d,j)$, *where* $s=\mathtt{S}e$ *is* source *of the edge;* $d=\mathtt{D}e$ *is* destination *of the edge; and* $j=\mathtt{J}e$ *is the set of clocks that are* reset *by the edge.*
- $\mathtt{Q}^{\mathsf{init}}\subseteq\mathtt{Q}$ *is the set of* initial locations.
- $\mathtt{Q}^{\mathtt{final}}\subseteq\mathtt{Q}$ *is the set of* final locations.

A *run* of the timed automaton A is a sequence of tuples $(\rho,\tau,\eta)\in\mathtt{Q}\times\mathbb{I}_{\geq 0}\times\mathtt{E}$ with the following conditions holds: (i) $\rho_0\in\mathtt{Q}^{\mathsf{init}}$, *i.e.*, ρ starts from an initial location $\mathtt{Q}^{\mathsf{init}}$; (ii) $(\mathtt{S}\eta_n=\rho_n)\wedge(\mathtt{D}\eta_n=\rho_{n+1})$, *i.e.*, the source and destination of edges η_n are ρ_n and ρ_{n+1}; (iii) $\tau_0,\tau_1,\ldots$ is an ordered and disjoint partition of the time horizon $\mathbb{R}_{\geq 0}$; and (iv) $\forall t\in\tau_n, x\in\mathtt{X}$, we have $\varrho_n(x)+t-\underline{\tau}_n\in\mathtt{I}(\varrho_n,x)$, where $\varrho_{n+1}(x)$ is defined inductively by

$$\varrho_{n+1}(x)=\begin{cases}0, & \text{if } x\in\mathtt{J}\eta_n\\ \varrho_n(x)+\overline{\tau}_n-\underline{\tau}_n, & \text{otherwise}\end{cases}$$

i.e., the clock times must satisfy the invariant of the current location. Here, $\underline{\tau}$ and $\overline{\tau}$ are the lower and upper bound of the interval.

A run satisfying the condition $\mathtt{inf}(\rho)\cap\mathtt{Q}^{\mathtt{final}}\neq\emptyset$, *i.e.*, some location from $\mathtt{Q}^{\mathtt{final}}$ has been visited infinitely many times by ρ, is called an *accepting run* of A. Note that every run of A induces a function f of type $\mathbb{R}_{\geq 0}\to\Sigma$ that maps t to $\mathtt{L}(\rho_n)$, where n is uniquely determined by the condition $t\in\tau_n$. We define the *language* of A, denoted by $\mathtt{Lang}(A)$, to be the set of all functions that are induced by accepting runs of A. The language of Timed Automata is closely related to MITL as follows.

Lemma 1 (MITL to Timed Automata [20])**.** *For any MITL formula* φ, *a timed automaton* A_φ *can be constructed such that* $\mathtt{Lang}(A_\varphi)=[\![\varphi]\!]$, i.e., *the set of functions that satisfy* φ *is exactly those that are induced by accepting runs of* A_φ.

Generally, to check if a signal f satisfies an MITL formula φ, the standard software implementation is to first construct the Timed Automaton A_φ by Lemma 1, and then check if f is included in $\mathtt{Lang}(A_\varphi)$.

Running Example: Step 2. *Following Step 1, we would like* *to check the following MITL formula*

$$\varphi_1 = \mathsf{T}\mathcal{U}\left(y_2(t) > \frac{1}{4}\right), \tag{9}$$

$$\varphi_2 = \left(y_1(t) > \frac{1}{2}\right)\mathcal{U}\left(y_2(t) > \frac{1}{4}\right), \tag{10}$$

where

$$y_1(t) = \sum_{q\in\mathcal{Q}} \int_{\mathbb{R}^d} I_{[0,1]} F(t,q,x)\mathrm{d}x, \tag{11}$$

$$y_2(t) = \sum_{q\in\mathcal{Q}} \int_{\mathbb{R}^d} I_{[2,4]} F(t,q,x)\mathrm{d}x. \tag{12}$$

Finally, the statistical verification algorithms proposed in this work are built upon the following fundamental result: two arbitrary probability distributions y and y' on finite states can be efficiently distinguished with arbitrarily high confidence via sampling. We formulate this result in the following theorem.

III. Model Reduction of Continuous-time Hybrid Systems

A. Reducing the Dynamics

To implement the Mori-Zwanzig model reduction method [21] for continuous-time stochastic systems, we partition the continuous state space into finitely many partitions $\mathbb{S} = \{s_1, \dots, s_n\}$, and treat each of them as a discrete state, similar to [24, 25]. We assume that for each s_i, there exists $q \in \mathcal{Q}$ such that $s_i \subseteq \{q\} \times \mathbb{A}_q$, and denote its measure by $\mu(s_i)$. Let $m(\mathbb{X})$ and $m(\mathbb{S})$ be set of probability distribution functions on $\mathbb{X}$ and $\mathbb{S}$, respectively. Then we can define a projection $P : m(\mathbb{X}) \to m(\mathbb{S})$ and an injection $R : m(\mathbb{S}) \to m(\mathbb{X})$ between $m(\mathbb{X})$ and $m(\mathbb{S})$ by

$$p_j = (PF(q,x))_j = \int_{s_j} F(q,x)\mathrm{d}x, \tag{13}$$

where p_j is the jth element of p, and

$$Rp = \sum_{j=1}^{n} p_j \mathbf{U}_{s_j}, \tag{14}$$

where $\mathbf{U}_{s_j}$ is the uniform distribution on s_j:

$$\mathbf{U}_{s_j}(x) = \begin{cases} \frac{1}{\mu(s_j)}, & \text{if } x \in s_j \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

Here the projection P and the injection R are defined for probability distributions. But they extend naturally to L_1 functions on $\mathbb{X}$ and $\mathbb{S}$ respectively. The projection P is the left inverse of the injection R but not *vice versa*, namely $PR = I$ but $RP \neq I$.

This projection P and injection R can reduce the Fokker-Planck operator to a transition rate matrix on $\mathbb{S}$, and hence reduce the continuous-time stochastic hybrid system into a continuous-time Markov chain.

Theorem 1. *Let $\mathbb{S} = \{s_1, s_2, \dots, s_n\}$ be a partition of the continuous state space $\mathbb{X}$ and P, R be the corresponding projection and injection defined in* (13)-(15). *The Fokker-Planck operator given in* (3) *reduces to the transition rate matrix A of a continuous-time Markov chain on $\mathbb{S}$ by*

$$A = PLR \tag{16}$$

where the transition rate from state s_i to s_j at time t is given by

$$\begin{aligned} A_{ij} = &\int_{\partial s_i \cap \partial s_j} f(q,x)\mathrm{d}x \\ &+ \frac{1}{\mu(s_i)} \int_{s_i} r(q,x)\mathbf{I}_{h(q,x)\in s_j}\mathrm{d}x \end{aligned} \tag{17}$$

for $a, b = 1, \dots, n$, where $\mathbf{I}_{h(q,x)\in s_j} = 1$ when $h(q,x) \in s_j$, and 0 otherwise.

Roughly speaking, the transition rate between two partitions in the same location is the flux of $f(q,x)$ across the boundary and the transition rate between two different locations is the flux of $r(q,x)$.

B. Reducing MITL Formulas

The observables on the continuous-time stochastic hybrid system reduce to the corresponding continuous-time Markov chain using the projection P. Let y be an observable on the continuous-time stochastic hybrid system with weight function $\gamma(q,x)$. To facilitate further discussion, we assume that $\gamma(q,x)$ is invariant under the projection P, i.e. $\gamma(q,x) = P\gamma(q,x)$. We define a corresponding observable y' on the continuous-time Markov chain that derives from the model reduction procedure by

$$\begin{aligned} y'(0) &= \sum_{q\in\mathcal{Q}} \int_{\mathbb{R}^d} \gamma(q,x) PF(0,q,x)\mathrm{d}x \\ &= \sum_{i=1}^{n} \left(\int_{s_i} \gamma(q,x)\mathrm{d}x\right)\left(\int_{s_i} F(0,q,x)\mathrm{d}x\right) \\ &= \sum_{i=1}^{n} r_i p(i) = y'(0). \end{aligned} \tag{18}$$

From now on, we will always denote the corresponding observable on the CTMC by y' for any observable y on the continuous-time stochastic hybrid system.

C. Reduction Error Estimation

For a given observable y with weight function $\gamma(q,x)$, the error of the projection P with respect to the observable y is defined by the maximal possible difference between y and y',

$$\Delta_y = \Big| \sum_{q\in\mathcal{Q}} \int_{\mathbb{R}^d} \gamma(q,x)(F(0,q,x) - RPF(0,q,x))\mathrm{d}x \Big|. \tag{19}$$

Remark 1. *When refining the partition of $\mathbb{X}$, $RP \to I$ in the weak operator topology [48]; that is, any distribution function $F(q,x)$ on the state space, $|\sum_{q\in\mathcal{Q}} \int_{\mathbb{R}^d} \gamma(q,x)(F(q,x) - RPF(q,x))| \to 0$ holds for any measurable weight function $\gamma(q,x)$. Accordingly for* (19), *$\Delta_y \to 0$ for any given y.*

$$\begin{tikzcd}
F(0,q,x) \arrow[r, "e^{Lt}"'] \arrow[d, dashed, "P"'] & F(t,q,x) \\
p(0) \arrow[r, "e^{At}"'] & p(t) \arrow[u, dashed, "R"']
\end{tikzcd}$$

Fig. 2: Diagram for reduction error.

By the definition of Δ_y, we know that, at the initial time, the atomic propositions on the continuous-time stochastic hybrid system and the CTMC have the relations

$$y(0) > c \Longrightarrow y'(0) > c - \Delta_y, \tag{20}$$

$$y(0) < c \Longrightarrow y'(0) < c + \Delta_y, \tag{21}$$

and similarly,

$$y'(0) > c + \Delta_y \Longrightarrow y(0) > c, \tag{22}$$

$$y'(0) < c - \Delta_y \Longrightarrow y(0) < c. \tag{23}$$

To derive the relations of the observables between the continuous-time stochastic hybrid system and the CTMC at any time, we define the reduction error of the observable y at time t due to the model reduction process by

$$\begin{aligned} \Theta_y(t) &= |y(t) - y'(t)| \\ &= \Big| \sum_{q \in Q} \int_{\mathbb{R}^d} \gamma(q,x)(e^{Lt} - Re^{At}P)F(0,q,x)\mathrm{d}x \Big|, \end{aligned} \tag{24}$$

where $F(0,q,x)$ is an initial distribution of the continuous-time stochastic hybrid system and $y'(t)$ is the corresponding observable of $y(t)$ on the CTMC. This reduction error is illustrated in Fig. 2. Note that the diagram is not commutative; actually the difference between going along the two paths is related to the reduction error.

In general, the reduction error $\Theta(t)$ may not be bounded as $t \to \infty$. To find a sufficient condition for boundedness, we define the reduction error of the Fokker-Planck operator L by

$$\delta(t,q,x) = (L - RPL)e^{tRPL}F(0,q,x). \tag{25}$$

Accordingly, we define the integration of $\delta(t,q,x)$ with respect to the weight function $\gamma(q,x)$ by

$$\Lambda_y = \sup_{t \geq 0} \Big| \sum_{q \in Q} \int_{\mathbb{R}^d} \gamma(q,x)(L - RPL)e^{tRPL}F(0,q,x)\mathrm{d}x \Big|, \tag{26}$$

which captures the maximal change of the time derivative of observable y.

A sufficient condition to find a uniform bound over time is that the reduction error of the Fokker-Planck operator $\delta(f(q,x))$ converges exponentially in time for any $f(q,x) \in m(\mathbb{X})$.

Definition 4. *For $\alpha > 0$, $\beta \geq 1$ and a given observable y, the continuous-time stochastic hybrid system is α-contractive with respect to y, if for any initial distribution function $F(0,q,x)$ on the state space, we have*

$$\begin{aligned} &\Big| \sum_{q \in Q} \int_{\mathbb{R}^d} \gamma(q,x)e^{tL}\delta(t,q,x)\mathrm{d}x \Big| \\ &\qquad \leq \beta e^{-\alpha t} \Big| \sum_{q \in Q} \int_{\mathbb{R}^d} \gamma(q,x)\delta(t,q,x)\mathrm{d}x \Big|. \end{aligned} \tag{27}$$

where $\delta(t,q,x)$ is given by (25).

This contractivity condition is to ensure that the model reduction error is bounded over all time, which is required for approximately keeping the truth value of temporal logic specifications of infinite time horizon. Although the condition seems restrictive, it is valid for a relatively wide range of systems including asymptotically stable systems. It is a commonly-used sufficient condition to guarantee the existence and uniqueness of an invariant measure for general dynamical systems, and the contractivity factor α is usually derived case-by-case. Using Definition 4, we obtain the following theorem.

Theorem 2. *If the continuous-time stochastic hybrid system is α-contractive, then for any $t \geq 0$, the reduction error $\Theta_y(t)$ for an observable y satisfies*

$$\Theta_y(t) \leq \frac{\beta \Lambda_y}{\alpha} + \Delta_y. \tag{28}$$

Proof. By Dyson's formula, we can decompose the exponential of L by

$$e^{tL} = e^{tRPL} + \int_{[0,t]} e^{(t-\tau)L}(L - RPL)e^{\tau RPL}\mathrm{d}\tau. \tag{29}$$

This formula, sometimes referred to as Duhamel's principle [21], can be verified by taking time derivatives on both sides. Substituting (29) into (24) gives

$$\begin{aligned} \Theta_y(t) \leq& \Big| \sum_{q \in Q} \int_{\mathbb{R}^d} \gamma(q,x)(e^{tRPL} - Re^{tA}P)F(0,q,x)\mathrm{d}x \Big| \\ &+ \Big| \sum_{q \in Q} \int_{\mathbb{R}^d \times [0,t]} \gamma(q,x)e^{(t-\tau)L}(L - RPL)e^{\tau RPL}F(0,q,x)\mathrm{d}\tau\mathrm{d}x \Big| \end{aligned} \tag{30}$$

Since the projection P and the injection R preserve the L_1 norm, RPL is also a Fokker-Planck operator. Noting $Re^{tA}PF(0,q,x) = e^{tRPL}PF(0,q,x)$, by (19), we see that the first term on the right hand side of (30) is less than Δ_y.

For the second term on the right hand side of (30), by (26)-(27) we have

$$\begin{aligned} \Theta_y(t) &\leq \Delta_y + \Big| \sum_{q \in Q} \int_{\mathbb{R}^d} \int_{[0,t]} \gamma(q,x)e^{(t-\tau)L}\delta(\tau,q,x)\mathrm{d}\tau\mathrm{d}x \Big| \\ &\leq \Delta_y + \Big| \sum_{q \in Q} \int_{\mathbb{R}^d} \int_{[0,t]} \beta e^{-\alpha(t-\tau)}\gamma(q,x)\delta(\tau,q,x)\mathrm{d}\tau\mathrm{d}x \Big| \\ &\leq \frac{\beta \Lambda_y}{\alpha} + \Delta_y. \end{aligned} \tag{31}$$

□

Theorem 2 implies the following relations between the atomic propositions on the continuous-time stochastic hybrid system and the CTMC.

Theorem 3. *If the continuous-time stochastic hybrid system is α-contractive, then we have*

$$y(t) > c \Longrightarrow y'(t) > c - \Big(\frac{\beta \Lambda_y}{\alpha} + \Delta_y\Big), \tag{32}$$

$$y(t) < c \Longrightarrow y'(t) < c + \Big(\frac{\beta \Lambda_y}{\alpha} + \Delta_y\Big), \tag{33}$$

and similarly,

$$y'(t) > c + \Big(\frac{\beta \Lambda_y}{\alpha} + \Delta_y\Big) \Longrightarrow y(t) > c, \tag{34}$$

$$y'(t) < c - \Big(\frac{\beta \Lambda_y}{\alpha} + \Delta_y\Big) \Longrightarrow y(t) < c. \tag{35}$$

The above theorem gives the following result.

Theorem 4. *Given a MITL formula φ on the continuous-time stochastic hybrid system that is α-contractive, it can be strengthened to ψ by replacing the atomic propositions according to* (34)-(35). *If ψ is true on the corresponding CTMC, then φ is true on the continuous-time stochastic hybrid system.*

Running Example: Step 3. *First, we note that the invariant distribution of this process is $F_{\mathrm{inv}} = \mathbf{U}_{\mathbb{X}}/3$. Following Step 2, we partition $\mathbb{X}$ into intervals of length $1/N$. By the above model reduction procedure it reduces to a CTMC with transition rate matrix M given by*

$$M_{ij} = \frac{\delta_{ij}}{4} + \frac{1}{4N} \tag{36}$$

where $i \in [3N]$ and $j \in [3N]$. The invariant distribution F_{inv} remains unchanged, and the MITL formula to check is

$$\varphi_1' = \mathsf{T}\,\mathcal{U}\left(y_2'(t) > \frac{1}{4} + \Theta_y(t)\right) \tag{37}$$

$$\varphi_2' = \left(y_1'(t) > \frac{1}{2} + \Theta_y(t)\right)\mathcal{U}\left(y_2'(t) > \frac{1}{4} + \Theta_y(t)\right) \tag{38}$$

where $\Theta_y(t)$ is the model reduction error and

$$y_1'(t) = \sum_{i=1}^{N} p_i(t), \tag{39}$$

$$y_2'(t) = \sum_{i=2N+1}^{3N} p_i(t). \tag{40}$$

When $N = 30$, we have $\Theta_y(t) \leq 0.02$ from (13) *and* (26).

IV. Statistical Model Checking of MITL

Given a CTMC C and a MITL formula φ with atomic propositions $\mathtt{AP}_\varphi$, we can construct by sampling a timed automaton $T_{C,\mathtt{AP}_\varphi}$ whose reachable locations at time t are labeled by the atomic propositions in φ that are true on C. By $[\![C, \mathtt{AP}_\varphi]\!]$ we denote the singleton set containing the unique signal induced by C and φ. For simplicity, we focus on constructing $T_{C,\{P\}}$ for an atomic formula $P : y = \sum_{i=1}^{n} r_i p_i > c$, denoted by a pair (r, c). Let $f(t)$ be the set of atomic formulas that p satisfies at time t, i.e., $(p, c) \in f(t)$ iff $p(t) > c$. Also, let $T_{C,\{P\}}(t)$ be the set of reachable locations of $T_{C,\{P\}}$ at time t.

Lemma 2. *[49] For any $\alpha, \delta > 0$, and two discrete distributions p and p', there is a test $\mathcal{A}(p, p', \alpha, \delta)$ which runs in time $O\big(n^{2/3}(2\delta)^{-8/3}\log(n/\alpha)\big)$ such that if $\|p - p'\| \leq \max\Big(\frac{\delta^{4/3}}{2^{14/3}\sqrt[3]{n}}, \frac{\delta}{4\sqrt{n}}\Big)$ then the test accepts with probability at least $1 - \alpha$, and if $\|p - p'\| > \delta$ then the test rejects with probability at least $1 - \alpha$.*

We assume an estimation p^* of the invariant distribution p^{inv} are given such that $\forall (p, c) \in \mathtt{AP}, |p^{\mathrm{inv}} - c| > \delta'$ and $|p^{\mathrm{inv}} - p^*| < \frac{\delta'}{3}$ for some $\delta' > 0$. Since $p(t)$ converges to p^{inv} due to contractivity, we know $|p(t) - c| > \delta'$, for large enough t. Using Lemma 2, we can find time T such that for $t > T$, $|p(t) - p^*| < \frac{\delta'}{3}$, namely $|p(t) - p^{\mathrm{inv}}| < \frac{2\delta'}{3}$. Therefore, if $p^{\mathrm{inv}} - c > \delta$ then $p^* - c > 0$. Similarly, if $p^{\mathrm{inv}} - c < -\delta$ then $p^* - c < 0$. Note that exactly one of $p^{\mathrm{inv}} - c > \delta$ and $p^{\mathrm{inv}} - c < -\delta$ is true. Furthermore, $p^* - c > 0$ and $p^* - c < 0$ cannot be both true. Therefore, $p^* - c > 0$ implies $p^{\mathrm{inv}} - c > \delta$, $p^* - c < 0$ implies $p^{\mathrm{inv}} - c < -\delta$, and $p^* - c$ is never zero.

ALGORITHM 1: Truncating time horizon

```
Data: CTMC $(T, y_0)$, estimation of invariant distribution $y^*$,
      MITL formula $\varphi$, parameters $\alpha$, $\gamma$, $\delta$, and $\delta'$
Function DurationOfSimulation
    $t \leftarrow 0$
    $\eta \leftarrow \max_{(p,c)\in\mathtt{AP}} \|p\|_1$
    while $\mathcal{A}\Big(y_t, y^*, \frac{1}{2}\min\{\alpha, \gamma\}, \frac{\delta'}{3\eta}\Big) = \mathtt{failed}$ do
        $t \leftarrow t + 1$;
    end
    return $t$
```

For any $\delta_1 > 0$, let $\Delta = \frac{\delta_1}{3\max\{|\dot{p}_i(t)||t\in[0,T]\}}$. Then, for any $t \in [0, T]$ and $t' \in [t - \Delta, t + \Delta] \cap [0, T]$, we have

1. if $p_i(t) - c > \frac{\delta_1}{3}$ then $p_i(t') - c > 0$,
2. if $p_i(t) - c < -\frac{\delta_1}{3}$ then $p_i(t') - c < 0$,
3. if $|p_i(t) - c| \leq \frac{2\delta_1}{3}$ then $|p_i(t') - c| \leq \delta_1$.

We partition $[0, T)$ into at least $\big\lfloor \frac{T}{2\Delta} \big\rfloor + 1$ intervals, each of size smaller than 2Δ. Let $[t_1, t_2]$ be one of these intervals. We then run $\mathcal{A}$, for $t = \frac{1}{2}(t_1 + t_2)$ to derive

$$\mathtt{res}_1 = \mathcal{A}^{\delta_1/3}\Big(p_i(t), c + \frac{\delta_1}{3}, \alpha', \gamma'\Big),$$

$$\mathtt{res}_2 = \mathcal{A}^{\delta_1/3}\Big(p_i(t), c - \frac{\delta_1}{3}, \alpha', \gamma'\Big),$$

where $\mathcal{A}$ statistically check If $\mathtt{res}_1 = \mathtt{yes}$ then $\forall t' \in [t_1, t_2), (p_i(t') > c)$ holds with bounded error α'. Therefore, we set $T_{C,\{P\}}(t) = \{P\}$. If $\mathtt{res}_2 = \mathtt{no}$ then for any time $t' \in [t_1, t_2]$, we know $p_i(t') < c$ holds with bounded error α'. Therefore, we set $T_{C,\{P\}}(t) = \{\emptyset\}$. Otherwise, for any time t' in the interval, $|p_i(t') - c| \leq \delta_1$ with bounded error $\max(\alpha', \gamma')$. In this case, we set 1) $T_{C,\{P\}}(t) = \{q, q'\}$, 2) $\mathtt{L}(q) = \{P\}$ and $\mathtt{L}(q') = \emptyset$, 3) entry to q or q', and 4) switches between q and q' for arbitrary number of times, while their common invariant permits. The result of the above procedure $\mathtt{res} = \mathcal{A}^{\delta_1,\delta_2}(C, y_0, \varphi, \alpha, \beta)$ satisfies

$$\mathbb{P}[\mathtt{res} = \mathtt{no} \;\; | \; C \models \varphi] \leq \alpha \tag{41}$$

$$\mathbb{P}[\mathtt{res} = \mathtt{yes} \; | \; C \not\models \varphi] \leq \alpha \tag{42}$$

ALGORITHM 2: Constructing the signal for atomic proposition P

```
$h \leftarrow \max\{|\dot{y}_i(t)| \mid t \in [0,T]\}$, $\Delta \leftarrow \frac{\delta_1}{3h}$, $n \leftarrow |\mathtt{AP}|\lceil \frac{T}{\Delta} \rceil$,
  $T_{C,\{P\}} \leftarrow$ an empty automaton, $\mathtt{X} \leftarrow \{t\}$, $q_{\text{last}} \leftarrow \bot$
forall $i \leftarrow 0$ to $\lfloor \frac{T}{\Delta} \rfloor$ do
    $\alpha' \leftarrow \min(\frac{\alpha}{4n}, \frac{\beta}{2n})$, $\beta' \leftarrow \frac{\beta}{n}$
    $\mathtt{res}_1 \leftarrow Algorithm_1^{\delta_1/3}\big(p_i\big((i+\frac{1}{2})\Delta\big), c + \frac{\delta_1}{3}, \alpha', \beta'\big)$
    $\mathtt{res}_2 \leftarrow Algorithm_1^{\delta_1/3}\big(p_i\big((i+\frac{1}{2})\Delta\big), c - \frac{\delta_1}{3}, \alpha', \beta'\big)$
    add a new location $q$ to $\mathtt{Q}$
    if $\mathtt{res}_1 = \mathtt{yes}$ then
        $\mathtt{L}(q) \leftarrow \{P\}$
    else if $\mathtt{res}_2 = \mathtt{no}$ then
        $\mathtt{L}(q) \leftarrow \emptyset$
    else
        $\mathtt{L}(q) \leftarrow \mathtt{unknown}$
    $\mathtt{I}(q) \leftarrow 2i\Delta \leq t < 2(i+1)\Delta$
    if $q_{\text{last}} \neq \bot$ then
        $\mathtt{E} \leftarrow \mathtt{E} \cup \{(q_{\text{last}}, q, \emptyset)\}$
    else
        $\mathtt{Q}^{\mathsf{init}} \leftarrow \{q\}$
    $q_{\text{last}} = q$
end
add a new location $q$ to $\mathtt{Q}$
$\mathtt{I}(q) \leftarrow \mathtt{true}$, $\mathtt{Q}^{\mathtt{final}} \leftarrow \{q\}$
$\mathtt{E} \leftarrow \mathtt{E} \cup \{(q_{\text{last}}, q, \emptyset), (q, q, \emptyset)\}$
if $y^{\text{inv}} > c$ then
    $\mathtt{L}(q) \leftarrow \{P\}$
else
    $\mathtt{L}(q) \leftarrow \emptyset$
$T_{C,\{P\}} \leftarrow$ replace any unknown location in Q with $q$ and $q'$
  labeled $\{P\}$ and $\emptyset$. Duplicate edges from/to $q$ and $q'$
  accordingly
Add $(q, q', \emptyset)$ and $(q', q, \emptyset)$ to E for every split locations in the
  previous step.
return $T_{C,\{P\}}$
```

As for the `unknown` output, let $B^{\delta_1}(y)$ be the δ_1-ball centered at y in the L_∞ norm. The algorithm guarantees that

$$\mathbb{P}[\mathtt{res} = \mathtt{unknown}] \leq \alpha + \beta \tag{43}$$

for all $y' \in B^{\delta_1}(y)$.

Definition 5. *For any $\epsilon > 0$ let $y + B_\epsilon$ be the set of observables achieved by slightly perturbing y. Let C_ϵ be any object with observables in the set $y + B_\epsilon$. We say satisfaction relation of CTMC C and MITL formula φ is* ϵ-robust*, if one of the following is true: 1) For all y' induced by C_ϵ we have $y' \models \varphi$, or 2) For all y' induced by C_ϵ we have $y' \not\models \varphi$. We say satisfaction relation is* robust*, if it is ϵ-robust for some $\epsilon > 0$.*

By definition 5, for any CTMC C and MITL formula φ, if C is robust on φ, iteratively reducing δ_1 in our algorithm guarantees that it will eventually return an answer which is not `unknown` while satisfying conditions (41) and (42).

Running Example: Step 4. *Following Step 3, we run Algorithm 1 on the CTMC and derive that both φ_1' and φ_2' are true. This implies that the formulas φ_1 and φ_2 given Step 2 are true on the system given in Step 1.*

V. Discrete Hybrid Systems

Discretizing the time of the continuous-time stochastic hybrid system gives a discrete-time stochastic hybrid system with the initial distribution $F(0, q, x)$ and transition function $T(q', x', q, x)$ satisfying

$$\sum_{q \in \mathcal{Q}} \int_{\mathbb{R}^d} T(q', x', q, x) \mathrm{d}x' = 1, \tag{44}$$

for any $(q, x) \in \mathbb{X}$. The observable y is defined in the same way as in the continuous-time case.

For discrete-time signals, we use the inequality linear temporal logic (iLTL) [19] whose atomic propositions $\mathtt{AP}$are inequalities as given in Algorithm 1. Again, the negation operator $\neg$ is dropped by pushing it inside and using completeness of $\{<, \leq, \geq, >\}$.

Definition 6 (iLTL Syntax)**.** *The syntax of iLTL formulas is defined using the BNF rule:*

$$\varphi = \bot \mid \top \mid p \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \mathcal{X}\varphi \mid \varphi\mathcal{U}\varphi \mid \varphi\mathcal{R}\varphi,$$

where $p \in \mathtt{AP}$ is an atomic proposition,

For a set of atomic propositions, the discrete-time stochastic hybrid system induces a signal $f : \mathbb{N} \to 2^{\mathtt{AP}}$, where $f(t)$ is the set of atomic propositions holding at time t. According, we define the semantics of iLTL on the system by Definition 7.

Definition 7 (iLTL Semantics)**.** *Let φ be an iLTL formula and f be a discrete-time signal. The satisfaction relation $\models$ between f and φ is inductively defined according to the rules:*

$$\begin{array}{lll}
f \models \bot & \text{iff} & \textit{always false} \\
f \models \top & \text{iff} & \textit{always true} \\
f \models y \sim c & \text{iff} & (y \sim c) \in f(0) \\
f \models \varphi \vee \psi & \text{iff} & (f \models \varphi) \vee (f \models \psi) \\
f \models \varphi \wedge \psi & \text{iff} & (f \models \varphi) \vee (f \models \psi) \\
f \models \mathcal{X}\varphi & \text{iff} & f_{[1,\infty)} \models \varphi \\
f \models \varphi\mathcal{U}\psi & \text{iff} & \exists i \in \mathbb{N}, f^i \models \psi \wedge \forall j \in [i], f^j \models \varphi \\
f \models \varphi\mathcal{R}\psi & \text{iff} & \forall i \in \mathbb{N}, f^i \models \psi \quad \textit{or} \\
 & & \exists i \in \mathbb{N}, f^i \models \varphi \\
 & & \quad \wedge \forall j \in [i+1], f^j \models \psi.
\end{array}$$

We define $[\![\varphi]\!]$ to be the set of signals that satisfy φ.

Verifying the signals can be done by transforming them to Büchi automata [50], which can be viewed as the discrete-time version of Timed Automata in Definition 3.

Definition 8. *A Büchi automaton B is a tuple $\big(\mathtt{S}, \Sigma, \Gamma, \mathtt{S}^{\mathsf{init}}, \mathtt{F}\big)$ where*

- $\mathtt{S}$ *is a finite non-empty set of* states*,*
- Σ *is a finite* alphabet*,*
- $\Gamma \subseteq \mathtt{S} \times \Sigma \times \mathtt{S}$ *is a* transition relation*,*
- $\mathtt{S}^{\mathsf{init}} \subseteq \mathtt{S}$ *is a set of* initial *states,*
- $\mathtt{F} \subseteq \mathtt{S}$ *is a set of* final *states.*

We write $s_1 \xrightarrow{a} s_2$ instead of $(s_1, a, s_2) \in \Gamma$.

The Büchi automaton B takes an infinite sequence $w \in \Sigma^{\boldsymbol{\omega}}$ as an input and accepts it, iff there exists an infinite sequence of states $\rho \in \mathtt{S}^{\boldsymbol{\omega}}$ such that 1) $\rho_0 \in \mathtt{S}^{\mathsf{init}}$, 2) $\forall n \in \mathbb{N}, \rho_n \xrightarrow{w_n} \rho_{n+1}$, and 3) $\mathtt{inf}(\rho) \cap \mathtt{F} \neq \emptyset$, where $\mathtt{inf}(\rho)$ is the set of states that appear infinitely often in $\{\rho_n\}_{n=1}^{\infty}$. An infinite sequence of states is called a *run* of B if it satisfies 1)-2), and an *accepting run* if it satisfies 1)-3). We define *language* of B, denoted by

$$\begin{tikzcd}
F(t,q,x) \arrow[r, "T"'] & F(t+1,q,x) \arrow[d, "P"] \\
p(t) \arrow[u, "R"] \arrow[r, "T_r"'] & p(t+1)
\end{tikzcd}$$

Fig. 3: Diagram for single-step reduction

$\mathtt{Lang}(B)$, to be the set of all infinite sequences in $\Sigma^{\boldsymbol{\omega}}$ that are accepted by B.

Similar to the relation between MITL and Timed Automata (Lemma 1), we introduce following result on the conversion between LTL and Büchi Automata.

Lemma 3 (LTL to Büchi Automata [50, 51, 52]). *For any LTL formula φ, a Büchi automaton B_φ can be constructed such that $\mathtt{Lang}(B_\varphi) = [\![\varphi]\!]$,* i.e., *the set of infinite words that satisfy φ is exactly those that are accepted by B_φ.*

A. Model reduction

The model reduction procedure for discrete-time stochastic hybrid system is similar to that of continuous-time ones as discussed in Sections III-A to III-C, following the three steps of (i) reducing the dynamics by partitioning the state space, (ii) reducing the temporal logic specifications accordingly, and (iii) estimating the model reduction error.

1) Reducing the Dynamics: For a discrete-time stochastic hybrid system, we can reduce it to a finite-state Markov chain by the set-oriented methods [53]. Similar to Section III, let $S = \{s_1, s_2, \ldots, s_n\}$ be a partition of the continuous state space $\mathbb{X}$, and P, R be the corresponding projection and injection operators as given by (13)-(15). As shown in Fig. 3 and Theorem 5, they induce a projection from the Markov kernel $T : m(\mathbb{X}) \to m(\mathbb{X})$ to a Markov kernel $T_r : m(S) \to m(S)$ by

$$T_r = PTR. \tag{45}$$

For multiple steps, the diagram for projection is shown by the non-commutative diagram in Fig. 4.

Theorem 5. *Let $S = \{s_1, \ldots, s_n\}$ be a measurable partition of the state space $\mathbb{X}$. Then the discrete-time stochastic hybrid system reduces to a CTMC (T_r, p_0) by*

$$\begin{aligned} p_0(i) &= \int_{s_i} F(0,q,x)\mathrm{d}x, \\ T_r(i,j) &= \int_{s_i}\int_{s_j} T(q',x',q,x)\mathrm{d}x'\mathrm{d}x. \end{aligned} \tag{46}$$

2) Reduced iLTL: An observable on the discrete stochastic hybrid system can be reduced approximately to an observable on the discrete-time Markov chain by (18). At a time t, discrepancy between $y(t)$ and $y'(t)$ is given by (4).

Lemma 4. *For any $F(q,x) \in m(\mathbb{X})$ and projection operator P, we have*

$$\begin{aligned} y(0) > b + \delta_P(F)\|F\|_\infty &\Longrightarrow y'(0) > b, \\ y'(0) > b + \delta_P(F)\|F\|_\infty &\Longrightarrow y(0) > b, \\ y(0) < b - \delta_P(F)\|F\|_\infty &\Longrightarrow y'(0) < b, \\ y'(0) < b - \delta_P(F)\|F\|_\infty &\Longrightarrow y(0) < b, \end{aligned}$$

where

$$\delta_P(F) = \|F(0,q,x) - PF(0,q,x)\|_{TV}, \tag{47}$$

is the error of projection operator P in total variance, where $\|\cdot\|_{TV}$ is the total variation distance.

Therefore, we can reduce iLTL formulas associated with the discrete-time stochastic hybrid system to iLTL formulas associated with Markov process (T_r, p_0) by replacing the integration with the summation. We call the new temporal logic *reduced iLTL*. This is exactly the form of iLTL proposed in [19].

3) Reduction Error Estimation: First, we note that the projection operator P is contractive.

Lemma 5. *Let $\mathbb{S} = \{s_1, \ldots, s_n\}$ be a measurable partition of $\mathbb{X}$ and P be the projection operator associated with $\mathbb{S}$. For any $F(q,x), F'(q,x) \in m(\mathbb{X})$,*

$$\|PF(q,x) - PF'(q,x)\|_{TV} \le \|F(q,x) - F'(q,x)\|_{TV}. \tag{48}$$

As shown in the non-commutative diagram in Fig. 4, the discrepancy for t steps is

$$\begin{aligned} \Delta_t &= \|PT^{(t)}F(0,q,x) - T_r^{(t)}PF(0,q,x)\|_{\mathrm{TV}} \\ &= \|PT^{(t)}F(0,q,x) - P(TRP)^{(t)}F(0,q,x)\|_{\mathrm{TV}}. \end{aligned} \tag{49}$$

The error bound of t-step projection is given by the following theorem.

Theorem 6. *Given a discrete-time stochastic hybrid system and a projection operator P, the t-step ($t \ge 1$) error of projection*

$$\Delta_t \le \sum_{i=0}^{t-1} \delta_P((TRP)^{(i)}F(0,q,x)), \tag{50}$$

where δ_P is given in (47).

Proof. For $t = 1$, we have,

$$\begin{aligned} \Delta_1 &= \|PTF(0,q,x) - P(TRP)F(0,q,x)\|_{\mathrm{TV}} \\ &\le \|TF(0,q,x) - TRPF(0,q,x)\|_{\mathrm{TV}} \\ &\le \|F(0,q,x) - RPF(0,q,x)\|_{\mathrm{TV}} \\ &= \delta_P(F(0,q,x)). \end{aligned} \tag{51}$$

$$\begin{tikzcd}
F(0,q,x) \arrow[r, "T"'] & F(1,q,x) \arrow[r, dotted] & F(t-1,q,x) \arrow[r, "T"'] & F(t,q,x) \arrow[d, "P"] \\
p(0) \arrow[u, "R"] \arrow[r, dashed, "T_r"'] & p(1) \arrow[r, dotted] & p(t-1) \arrow[r, dashed, "T_r"'] & p(t)
\end{tikzcd}$$

Fig. 4: Diagram for multiple-step reduction

For $t > 1$, with F denoting $F(0, q, x)$, we have

$$\begin{aligned}\Delta_t =&\|PT^{(t)}F - P(TRP)^{(t)}F\|_{\mathrm{TV}}\\ \leq&\|T^{(t)}F - (TRP)^{(t)}F\|_{\mathrm{TV}}\\ \leq&\|T^{(t)}F - T^{(t-1)}(TRP)F\|_{\mathrm{TV}}\\ &+ \|T^{(t-1)}(TRP)F - T^{(t-2)}(TRP)^{(2)}F\|_{\mathrm{TV}}\\ &+ \ldots + \|T(TRP)^{(t-1)}F - (TRP)^{(t)}F\|_{\mathrm{TV}}\\ \leq&\sum_{i=0}^{t-1}\delta_P((TRP)^{(i)}F).\end{aligned} \tag{52}$$

□

When T is strictly contractive, we can derive a tighter error bound.

Theorem 7. *Given a discrete-time stochastic hybrid system, a projection operator P and the corresponding injection R, if the Markov kernel T is strictly contractive by factor $\alpha \in (0, 1)$, then the t-step ($t \geq 1$) error of projection*

$$\Delta_t \leq \frac{\delta_P}{1-\alpha}, \tag{53}$$

where

$$\delta_P = \sup_{i\in\mathbb{N}} \delta_P((TRP)^{(i)}F(0, q, x)). \tag{54}$$

Proof. For $t = 1$, clearly $\Delta_t = \delta_P$. For $t \geq 2$, by (52) and with F denoting $F(0, q, x)$, we have

$$\begin{aligned}\Delta_t \leq&\|T^{(t)}F - T^{(t-1)}(TRP)F\|_{\mathrm{TV}}\\ &+ \|T^{(t-1)}(TRP)F - T^{(t-2)}(TRP)^{(2)}F\|_{\mathrm{TV}}\\ &+ \ldots + \|T(TRP)^{(t-1)}F - (TRP)^{(t)}F\|_{\mathrm{TV}}\\ \leq&(1 + \alpha + \ldots + \alpha^t)\delta_P\\ \leq&\frac{\delta_P}{1-\alpha}.\end{aligned} \tag{55}$$

□

By combining Lemma 4 and Theorem 7, we can derive the following theorem on the relationship between linear inequalities on the original Markov process and linear inequalities on the reduced Markov process.

Theorem 8. *Given a measurable partition $\mathbb{S} = \{s_1, \ldots, s_n\}$ and the corresponding projection operator P, a discrete-time stochastic hybrid system and its reduction (T_r, p_0) satisfies the equations:*

$$y > b + \frac{\delta_P\|F\|_\infty}{1-\alpha} \Longrightarrow y' > b, \tag{56}$$

$$y' > b + \frac{\delta_P\|F\|_\infty}{1-\alpha} \Longrightarrow y > b, \tag{57}$$

$$y < b - \frac{\delta_P\|F\|_\infty}{1-\alpha} \Longrightarrow y' < b, \tag{58}$$

$$y' < b - \frac{\delta_P\|F\|_\infty}{1-\alpha} \Longrightarrow y < b, \tag{59}$$

where δ_p is given by (54) *respectively.*

B. Statistical Model Checking of iLTL

Similar to Section IV, we introduce the statistical model checking procedure of iLTL specifications on the reduced systems. Again, we denote the atomic proposition $y = \sum_{i=1}^n r_i p_i = r \cdot p$ by a pair (r, b). For an iLTL formula φ and a discrete-time Markov chain generating a sequence of distributions $w = p_0 p_1 p_2 \ldots$, define $u = u_0 u_1 u_2 \ldots$ where $u_t = \{(r, b) \in \mathtt{AP}_\varphi \mid r \cdot p_t > b\}$ is the set of atomic propositions that are true at time t. Then $w \models \varphi$, iff $u \models \varphi$. Using Lemma 3, we also know φ can be transformed into a Büchi automaton B_φ such that $[\![\varphi]\!] = \mathtt{Lang}(B_\varphi)$, meaning B_φ accepts exactly those sequences that satisfy φ. This suggests the following algorithm to check if a Markov chain satisfies an iLTL formula φ. Let w be the (unique) sequence of distributions generated by that Markov chain.

1) Construct the sequence $u \in (2^{\mathtt{AP}})^{\boldsymbol{\omega}}$ of atomic propositions that are true at each step of w.
2) Check $u \in \mathtt{Lang}(B_\varphi)$ and return the result as the algorithm output. We know it will be the right answer since $(T_r, p_0) \models \varphi$ iff $w \models \varphi$ iff $u \models \varphi$ iff $u \in \mathtt{Lang}(B_\varphi)$.

In what follows we outline how the above two steps can be accomplished.

a) Constructing the labels for distributions: To construct the set of labels u_t corresponding to the distribution $w_t = p_t$, the simplest algorithm would compute $p_t = T_r^{(t)} p_0$ first and then for every atomic proposition $(r, b) \in \mathtt{AP}_\varphi$, check whether or not $r \cdot p_t > b$ is true. However, this would be expensive for Markov chains with a large number of states. Instead, we compute these labels statistically. First, we can draw samples according to distribution p_t by simulating the Markov chain for t steps. Assuming the elements of each vector r in atomic propositions are from $\{0, 1\}$. In this case, p_t satisfies (r, b) iff the probability of drawing a state s (according to p_t) such that $r_s = 1$ is strictly greater than b. This can be statistically checked by drawing samples from p_t and using either Chernoff bounds, or the Sequential Probability Ratio Test [54] (see [32] and [31]). Such a statistical test usually takes as parameters an indifference parameter $\delta_2 > 0$, and error bounds $\alpha_2, \gamma_2 > 0$. The output of this test is `yes`, `no`, or `unknown` with conditions:

$$\begin{aligned}
\mathbb{P}[\texttt{res} = \texttt{no} \quad &| \; r \cdot p_t > b &] \leq \alpha_2, \qquad (60a)\\
\mathbb{P}[\texttt{res} = \texttt{yes} \quad &| \; r \cdot p_t \not> b &] \leq \alpha_2, \qquad (60b)\\
\mathbb{P}[\texttt{res} = \texttt{unknown} \; &| \; |r \cdot p_t - b| > \delta_2] &\leq \gamma_2. \qquad (60c)
\end{aligned}$$

The parameters $\delta_2, \alpha_2, \gamma_2$ can be made arbitrarily small, though that will increase the number of samples needed. The general case when elements of r can be arbitrary real numbers requires one to estimate the mean of a random variable that is not necessarily Bernoulli. In such a situation, the Sequential Probability Ratio Test cannot be used, but we can use a technique due to Chow and Robbins [55].

b) Running B_φ on the labels: The sequence of labels $u = u_0 u_1 \ldots$ can be constructed statistically, symbol by symbol. However, u is an infinite sequence, and in order to run B_φ on u, u needs to *ultimately periodic*, *i.e.*, there must be finite sequences u_1 and u_2 such that $u = u_1 u_2^{\omega}$. We assume that the mapping defined by the reduced model is contracting and hence the sequence w converges to the invariant distribution p^{inv}. We also assume that p^{inv} is known for some bounded uncertainty. This assumption is readily verified for large classes of important physical models, such as those with energy balance laws that are dissipative in aggregate. For such models, general arguments can be used to derive the prior condition, even in the presence of strong nonlinearity, discontinuous dynamics, or other complexities. Define $\eta = \max_{(r,b)\in\texttt{AP}} \|r\|_1$ to be the size of largest vector in atomic propositions and *wlog.* assume it is positive. We assume $\delta' > 0$ and p^* (an estimate of p^{inv}) are given such that $\forall (r, b) \in \texttt{AP}, |r \cdot p^{\text{inv}} - b| > \delta'$ and $\|p^{\text{inv}} - p^*\|_1 < \frac{\delta'}{3\eta}$ are both true. Note that since w converges to p^{inv} and for any atomic proposition (r, b) we have $|r \cdot p^{\text{inv}} - b| > \delta'$, for large enough t, p_t and p^{inv} will satisfy exactly the same propositions and so u is ultimately periodic. Also, note that if we know p^{inv} precisely then we can easily check if such δ' exists, and if it does, we can easily find the supremum of all such δ'.

Using $\mathcal{A}$, we can outline the overall procedure as follows.

1) Using $\mathcal{A}$, with parameters $\frac{1}{2}\min\{\alpha, \gamma\}$ and $\frac{\delta'}{3\eta}$ find m such that p_m is within distance $\frac{\delta'}{3\eta}$ of p^* and hence within distance $\frac{2\delta'}{3\eta}$ of p^{inv} (see function `NumberOfSamplingSteps` in Algorithm 3). Let $(r, b) \in \texttt{AP}_\varphi$ be an arbitrary atomic proposition, and let p be any distribution for which $\|p - p^{\text{inv}}\| \leq \frac{2\delta'}{3\eta}$. Clearly this includes p^* which is known exactly, and any distribution $p_{m'}$ for $m' \geq m$.

$$|r \cdot p - r \cdot p^{\text{inv}}| \leq \|r\|_1 \times \frac{2\delta'}{3\eta} \leq \frac{2\delta'}{3} \qquad (61)$$

From (61), if $r \cdot p^{\text{inv}} - b > \delta'$ then $r \cdot p - b > \delta' - \frac{2}{3}\delta' > 0$, and if $r \cdot p^{\text{inv}} - b < -\delta'$ then $r \cdot p - b < -\delta' + \frac{2}{3}\delta' < 0$. Therefore, assuming we found a right m, truth values of atomic propositions do not change in $p_m p_{m+1} \ldots$.

2) For each $t < m$ and atomic proposition $(b, r) \in \texttt{AP}$, use $\mathcal{A}$ with parameters $\delta_2 = \delta$, $\alpha_2 = \frac{\alpha}{2m|\texttt{AP}|}$, and $\gamma_2 = \frac{\gamma}{2m|\texttt{AP}|}$, to determine if $r \cdot p_t > b$ (see function `LabelFiniteNumberOfSteps` in Algorithm 3).

3) For every atomic proposition $(b, r) \in \texttt{AP}$, verify if $r \cdot p^* > b$ is true or not (see function `AddLabelsOfInvariantDistribution` in Algorithm 3). Note that there will be no error at this step, since we know p^* precisely. Also, after this step, asg is defined for all $t \in \mathbb{N}$ and $q \in \texttt{AP}$. Whenever $asg(t, q) = \texttt{unknown}$, we consider both possibilities. With this consideration, function asg induces a set of infinite sequences on alphabet $2^{\texttt{AP}}$. Since after m, asg is a constant function, this set can be easily represented by a Büchi automaton $[\![asg]\!]$.

4) Verify if $\texttt{Lang}(B_\varphi)$ and its complement has any intersection with $\texttt{Lang}([\![asg]\!])$. If both intersections are non-empty, it means different choices for `unknown` slots (*i.e.*, `yes` or `no`) can lead to both satisfying φ and not satisfying it. Therefore, the algorithm returns `unknown`. Otherwise, if $\texttt{Lang}(B_\varphi)$ has no intersection with $\texttt{Lang}([\![asg]\!])$, we know that no matter what we choose for `unknown` slots, the result infinite sequence can never satisfy φ, hence the `no` answer. The only remaining case is when $\texttt{Lang}(B_{\neg\varphi})$

ALGORITHM 3: Model checking Markov chains against iLTL formulas

```
Data: Markov chain $(T, p_0)$, estimation of invariant distribution
      $p^*$, iLTL formula $\varphi$, parameters $\alpha$, $\gamma$, $\delta$, and $\delta'$
Result: yes, no, or unknown
Function NumberOfSamplingSteps()
    $t \leftarrow 0$;
    $\eta \leftarrow \max_{(r,b)\in\texttt{AP}} \|r\|_1$;
    while $\mathcal{A}\left(p_t, p^*, \frac{1}{2}\min\{\alpha, \gamma\}, \frac{\delta'}{3\eta}\right) = \textit{failed}$ do
        $t \leftarrow t + 1$;
    end
    return $t$
Function LabelFiniteNumberOfSteps($m \in \mathbb{N}$)
    forall $t \in \{0, 1, \ldots, m-1\}$, $(r, b) \in \texttt{AP}$ do
        $asg(t, (r, b)) \leftarrow \mathcal{A}_2^\delta(p_t, r, b, \frac{\alpha}{2m|\texttt{AP}|}, \frac{\gamma}{2m|\texttt{AP}|})$
    end
    return $asg$
Function AddLabelsOfInvariantDistribution($m \in$
  $\mathbb{N}, asg \in \mathbb{N} \times \texttt{AP} \to \{\texttt{yes}, \texttt{no}, \texttt{unknown}\}$)
    forall $t \in \{m, m+1, \ldots\}$, $(r, b) \in \texttt{AP}$ do
        if $r \cdot p^* > b$ then
            $asg(t, (r, b)) \leftarrow$ yes
        else
            $asg(t, (r, b)) \leftarrow$ no
        end
    end
    return $asg$
Function ModelCheck
    $m \leftarrow$ NumberOfSamplingSteps();
    $asg \leftarrow$ LabelFiniteNumberOfSteps($m$);
    $asg \leftarrow$ AddLabelsOfInvariantDistribution($m, asg$);
    $[\![asg]\!] \leftarrow$ the Büchi automaton that accepts exactly the set
      of infinite paths induced by $asg$;
    if $\texttt{Lang}(B_\varphi) \cap \texttt{Lang}([\![asg]\!]) \neq \emptyset$
      $\wedge\, \texttt{Lang}(B_{\neg\varphi}) \cap \texttt{Lang}([\![asg]\!]) \neq \emptyset$ then
        return unknown
    else if $\texttt{Lang}(B_\varphi) \cap asg = \emptyset$ then
        return no
    return yes
```

has no intersection with $\texttt{Lang}([\![asg]\!])$, and using a similar argument, we know the right answer is $\texttt{yes}$. Note that, $\texttt{Lang}(B_{\neg\varphi})$ is exactly complement of the set $\texttt{Lang}(B_{\varphi})$, a fact which is used multiple times.

Our algorithm $\mathcal{A}$ outlined above provides the following guarantees:

$$\mathbb{P}[\mathcal{A}((T_r, p_0), \varphi, \alpha, \gamma) = \texttt{no} | (T_r, p_0) \models \varphi] \leq \alpha \tag{62a}$$

$$\mathbb{P}[\mathcal{A}((T_r, p_0), \varphi, \alpha, \gamma) = \texttt{yes} | (T_r, p_0) \not\models \varphi] \leq \alpha \tag{62b}$$

$$\begin{aligned}\mathbb{P}\,[\mathcal{A}((T_r, p_0), \varphi, \alpha, \gamma) &= \texttt{unknown} | \\ \nexists (b, r) \in \texttt{AP}, &|r \cdot p_t - b| \leq \delta] \leq \gamma\end{aligned} \tag{62c}$$

The first two inequalities state that probability of having false positive or negative is at most α. The last inequality states that if in all steps that have distributions far enough from the invariant distribution, the actual probability of no atomic proposition (r, b) in φ is too close to b then the probability of returning $\texttt{unknown}$ is at most γ.

The error analysis of our algorithm can be carried out as follows. When the algorithm returns $\texttt{no}$ ($\texttt{yes}$) while the correct answer is $\texttt{yes}$ ($\texttt{no}$), it means that the algorithm made at least one mistake. The probability of finding wrong m is at most $\frac{\alpha}{2}$. Assuming m is computed correctly, the probability of having a step t and an atomic formula $q \in \texttt{AP}$ such that truth value of q at step t is computed incorrectly is at most $(m|\texttt{AP}|)\frac{\alpha}{2m|AP|} = \frac{\alpha}{2}$ (here $\texttt{unknown}$ is considered a correct answer, because it did not effect the output of the algorithm). Therefore, the total error is at most $\alpha = \frac{\alpha}{2} + \frac{\alpha}{2}$.

Similarly, if the algorithm returns $\texttt{unknown}$ while for any step that is far enough from the invariant distribution, we know the actual probability of no atomic proposition is too close to the threshold of that proposition, it means either the algorithm found m incorrectly, or it found $\texttt{unknown}$ for at least one step and one atomic proposition incorrectly. But we know that the probability of making each of these mistakes is at most $\frac{\gamma}{2}$. Thus the probability of incorrectly returning $\texttt{unknown}$ is at most γ.

$\mathcal{A}$ takes δ as one of its parameters. The problem with δ is that one may not know in advance the correct value for δ. Large values cause the algorithm to return $\texttt{unknown}$, and small values make the algorithm slow. In order to solve this problem one can start with a a large value for δ and decrease it when the algorithm returns $\texttt{unknown}$ for that δ.

VI. Case Study

We implemented the proposed model reduction and statistical verification algorithm on high-dimensional stochastic hybrid systems with polynomial dynamics for the continuous states to demonstrate the scalability. In this section we present our experimental results. Consider a piecewise linear jump system under nonlinear perturbation with the continuous state $x(t) \in \mathbb{R}^n$ and the discrete state $q(t) \in [m]$ with $m \in \mathbb{N}$. The continuous dynamics is,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = (A_{q(t)} + c_{q(t)}\|x(t)\|_\infty)x(t) \tag{63}$$

where $A_i \in \mathbb{R}^{n\times n}$ is Hurwitz and $c_i > 0$ for $i \in [m]$. The discrete state jumps spontaneously with rate λ_1 from j to $j-1$ for $j = 2, \ldots, m$ and with rate λ_2 from j to $j+1$ for $j = 1, \ldots, m-1$. Initially, the continuous state is distributed uniformly on the hypercube $C = \{x(0) \in \mathbb{R}^n \mid \|x(t)\|_\infty \leq K\}$; and the discrete state $q(0)$ uniformly on $[m]$.

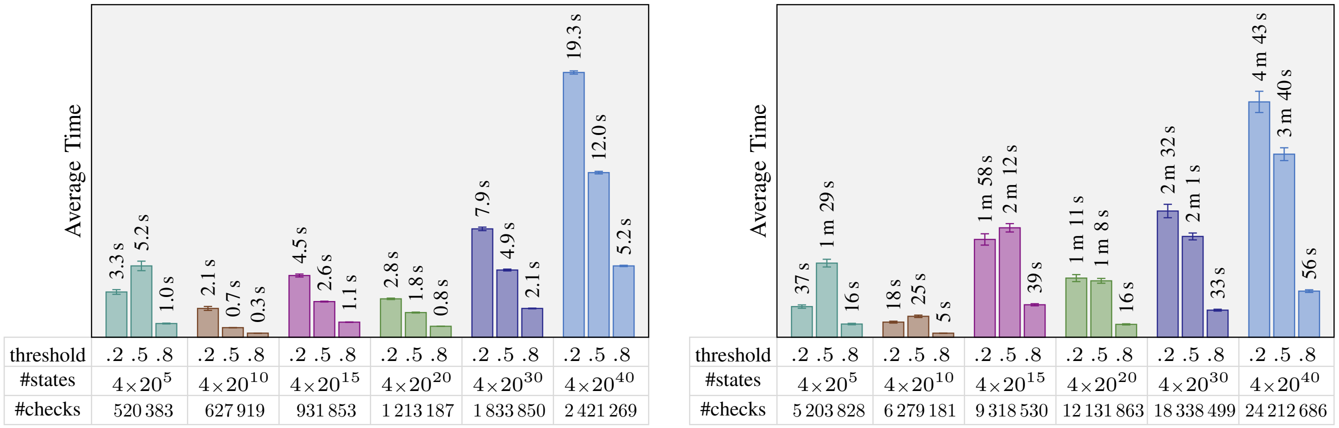
Assume that the elements of the dynamical matrices A_i are non-positive, so $x(t) \in C$ for all $t \in \mathbb{R}$. Therefore, we can partition the state space into $(2\eta)^n \times m$, each of length $1/\eta$. The hypercubes are indexed by $(i_1, \ldots, i_n, j)$ with $|i_k| \in \{-\eta, \ldots, -1, 1, \ldots, \eta\}$, $j \in [m]$, and $k \in [n]$. The transition probability rates are zero except

$$\begin{aligned}&\lambda((i_1, \ldots, i_n, j) \to (i_1, \ldots, i_n, j-1)) = \lambda_1\\ &\lambda((i_1, \ldots, i_n, j) \to (i_1, \ldots, i_n, j+1)) = \lambda_2\\ &\lambda((i_1, \ldots, i_k+1, \ldots, i_n, j) \to (i_1, \ldots, i_k, \ldots, i_n, j)) =\\ &\quad \int_S \frac{(A_j x)_k}{\eta^2}\mathrm{d}x_1 \ldots \mathrm{d}x_{k-1}\mathrm{d}x_{k+1} \ldots \mathrm{d}x_n + c_j K \max_k \frac{|i_k|}{\eta^3}.\end{aligned} \tag{64}$$

The desired property is $\top\mathcal{U}_{[0,T]}\big(w(F(t,q,x))>p\big)$, where T is a time bound (could be ∞), p is a probability threshold, and $w(\cdot)$ is the indicator function on a non-convex predicate stating exactly two elements of the continuous state are more than $\lceil K/2 \rceil$ away from the origin (formally, the predicate holds for a continuous state x iff $|\{i \in [n] \mid |x_i| \geq \lceil K/2 \rceil\}| = 2$). It asserts that before time T, a probability distribution will be reached such that the probability of a state x in that distribution satisfying the aforementioned predicate is larger than p.

We ran Algorithm 2 on multiple instances of this problem. In all of our experiments, $\lambda_1 = 0.03$, $\lambda_2 = 0.02$, $K = 1$, $\eta = 10$, and $\alpha = \beta = \delta_1 = 0.1$. We also fixed number of discrete states (m) to be 4. However, dimension of continuous state are chosen from the set $\{5, 10, 15, 20, 30, 40\}$. Our setting results in CTMCs with very high number of states. For example, our smallest example has 1.28×10^7 many states, and our largest example has more than 4.39×10^{52} many states. In all of our experiments, $c_1 = 0.1$, $c_2 = 0.2$, $c_3 = 0.3$, and $c_4 = 0.4$. Each instances of our problem requires 4 square matrices as well. We generated random Hurwitz matrices of appropriate dimension for them. Finally, we used maximum eigenvalue of the random matrices as maximum rate of changes ($\max\{\dot{y}_i(t) \mid t \in [0, T]\}$) in our algorithm.

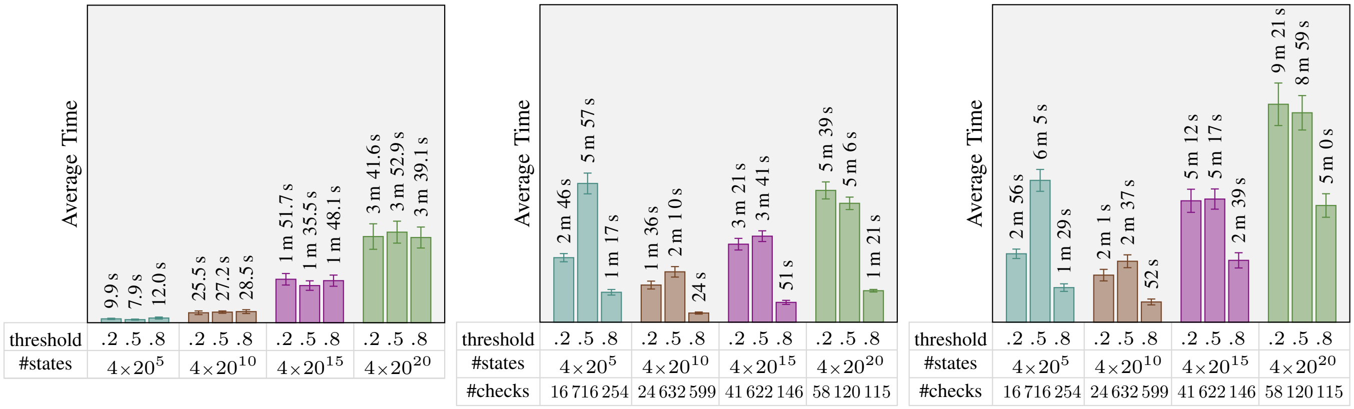
Our implementation is in Scala, and we used the Apache Commons Mathematics Library [56] to find eigenvalues of a matrix. Our simulations are carried out on Ubuntu 18.04 with i7-8700 CPU 3.2GHz and 16GB memory. We ran each test 50 times and here we report average running time as well as 95% confidence interval for different tests. Figure 5 shows the results for the case T is bounded (1000 and 10000), and Figure 6 shows the results for the case T is set to ∞. 'Threshold' is the value of p in our desired property. '#states' is the number of states in CTMC, and '#checks' is the number of check points the algorithm uses as it discretizes the time. Note that this number does not tell how many steps the algorithm takes to simulate the system for T units of time (or until it reaches the invariant distribution). It is the number of points in time for which we examine distribution of system at those times. When time is unbounded (*i.e.* $T = \infty$ in Figure 6), the algorithm firsts finds a time at which the system in known



(a) Time bound 1000

(b) Time bound 10000

Fig. 5: Bounded Time



(a) Time to find time-horizon

(b) Time to verify the property

(c) Total Time

Fig. 6: Unbounded Time

to be in the invariant distribution. It is easy to see that in the invariant distribution, our example is reduces to a birth–death process for which one can compute the invariant distribution analytically. Figure 6a shows the average amount of time our algorithm spent to find a time in which the distribution is known to be invariant. Figure 6b shows the average amount of time the algorithm uses to verify the property after a time horizon is fixed (note that our property of interest does not hold at the invariant distribution). Figure 6c show the sum of previous averages.

Our experimental results show that the running time of our algorithm is very robust to number of states, as, in worst case, the time increases logarithmically with respect to number of states. This is as one expects. Since we use statistical model checking, number of steps that the algorithm takes does not depend on the number of states. However, increasing number of states in a model, increases the time that each step of simulation takes. This is precisely why our algorithm logarithmically slows down as the number of states increases; in our implementation, complexity of simulating every step is linear to dimension of our morel (n).

VII. Conclusion

In this work, we proposed a method of verifying temporal logic formulas on stochastic hybrid systems via model reduction in both continuous-time and discrete-time. Specifically, we reduced the stochastic hybrid systems into Markov chains respectively by partitioning the state space and derive an upper bound the error. In addition, we propose stochastic algorithms to verify the temporal logic formulas on the Markov chains with arbitrarily high confidence.

References

[1] X. Jin, J. V. Deshmukh, J. Kapinski, K. Ueda, and K. Butts, “Benchmarks for model transformations and conformance checking,” in *1st International Workshop on Applied Verification for Continuous and Hybrid Systems (ARCH)*, 2014.

[2] I. Daniele, F. Alessandro, H. Marianne, B. Axel, and P. Maria, “A smart grid energy management problem for data-driven design with probabilistic reachability guarantees,” in *4th International Workshop on Applied Verification of Continuous and Hybrid Systems*, 2017, pp. 2–19.

[3] R. R. Rajkumar, I. Lee, L. Sha, and J. Stankovic, "Cyber-physical systems: the next computing revolution," in *Proceedings of the 47th design automation conference*. ACM, 2010, pp. 731–736.

[4] B. Liu, D. Hsu, and P. S. Thiagarajan, "Probabilistic approximations of ODEs based bio-pathway dynamics," *Theoretical Computer Science*, vol. 412, no. 21, pp. 2188–2206, May 2011.

[5] B. Liu, A. Hagiescu, S. K. Palaniappan, B. Chattopadhyay, Z. Cui, W.-F. Wong, and P. S. Thiagarajan, "Approximate probabilistic analysis of biopathway dynamics," *Bioinformatics*, vol. 28, no. 11, pp. 1508–1516, Jun. 2012.

[6] P. Zuliani, "Statistical model checking for biological applications," *STTT*, pp. 1–10, Aug. 2014.

[7] B. M. Gyori, B. Liu, S. Paul, R. Ramanathan, and P. Thiagarajan, "Approximate probabilistic verification of hybrid systems," in *Hybrid Systems Biology*. Springer, 2015, pp. 96–116.

[8] T. Henzinger, P. Kopke, A. Puri, and P. Varaiya, "What's decidable about hybrid automata?" *Journal of Computer and System Sciences*, vol. 57, no. 1, pp. 94–124, 1998.

[9] E. Clarke, A. Fehnker, Z. Han, B. Krogh, J. Ouaknine, O. Stursberg, and M. Theobald, "Abstraction and Counterexample-Guided Refinement in Model Checking of Hybrid Systems," *JFCS*, vol. 14, no. 4, pp. 583–604, 2003.

[10] R. Alur, T. Dang, and F. Ivancic, "Counter-Example Guided Predicate Abstraction of Hybrid Systems," in *TACAS 2003*, 2003, pp. 208–223.

[11] N. Roohi, P. Prabhakar, and M. Viswanathan, "HARE: A Hybrid Abstraction Refinement Engine for verifying non-linear hybrid automata," in *Proceedings of TACAS*, 2017, pp. 573–588.

[12] P. Tabuada and G. Pappas, "Linear time logic control of discrete-time linear systems," *IEEE Transactions on Automatic Control*, vol. 51, no. 12, pp. 1862–1877, Dec. 2006.

[13] M. Kloetzer and C. Belta, "A fully automated framework for control of linear systems from temporal logic specifications," *IEEE Transactions on Automatic Control*, vol. 53, no. 1, pp. 287–297, Feb. 2008.

[14] T. Wongpiromsarn, U. Topcu, and R. M. Murray, "Receding horizon control for temporal logic specifications," in *Proceedings of the 13th ACM International Conference on Hybrid Systems: Computation and Control*, ser. HSCC '10. New York, NY, USA: ACM, 2010, pp. 101–110.

[15] J. Liu, N. Ozay, U. Topcu, and R. M. Murray, "Synthesis of reactive switching protocols from temporal logic specifications," *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1771–1785, 2013.

[16] R. Chadha and M. Viswanathan, "A Counterexample Guided Abstraction-Refinement Framework for Markov Decision Processes," *ACM Transactions on Computational Logic*, vol. 12, no. 1, pp. 1:1–1:49, 2010.

[17] I. Tkachev and A. Abate, "Formula-free finite abstractions for linear temporal verification of stochastic hybrid systems," in *Proceedings of the 16th international conference on Hybrid Systems: Computation and Control*. ACM, 2013, pp. 283–292.

[18] I. Tkachev, A. Mereacre, J.-P. Katoen, and A. Abate, "Quantitative automata-based controller synthesis for non-autonomous stochastic hybrid systems," in *Proceedings of the 16th international conference on Hybrid Systems: Computation and Control*. ACM, 2013, pp. 293–302.

[19] Y. Kwon and G. Agha, "Linear inequality ltl (iltl): A model checker for discrete time markov chains," in *Formal Methods and Software Engineering*, ser. Lecture Notes in Computer Science, J. Davies, W. Schulte, and M. Barnett, Eds. Springer Berlin Heidelberg, 2004, vol. 3308, pp. 194–208.

[20] R. Alur, T. Feder, and T. A. Henzinger, "The benefits of relaxing punctuality," *J. ACM*, vol. 43, no. 1, pp. 116–146, 1996.

[21] A. J. Chorin, O. H. Hald, and R. Kupferman, "Optimal prediction and the mori-zwanzig representation of irreversible processes," *Proceedings of the National Academy of Sciences*, vol. 97, no. 7, pp. 2968–2973, Mar. 2000.

[22] C. Beck, S. Lall, T. Liang, and M. West, "Model reduction, optimal prediction, and the mori-zwanzig representation of markov chains," in *CDC/CCC*, 2009, pp. 3282–3287.

[23] A. A. Julius and G. J. Pappas, "Approximations of Stochastic Hybrid Systems," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1193–1203, 2009.

[24] A. Abate, J.-P. Katoen, J. Lygeros, and M. Prandini, "Approximate Model Checking of Stochastic Hybrid Systems," *European Journal of Control*, vol. 16, no. 6, pp. 624–641, 2010.

[25] A. Abate, A. D'Innocenzo, and M. D. D. Benedetto, "Approximate Abstractions of Stochastic Hybrid Systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 11, pp. 2688–2694, 2011.

[26] G. Pola, A. Girard, and P. Tabuada, "Approximately bisimilar symbolic models for nonlinear control systems," *Automatica*, vol. 44, no. 10, pp. 2508 – 2516, 2008.

[27] A. Girard, G. Pola, and P. Tabuada, "Approximately bisimilar symbolic models for incrementally stable switched systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 116–126, 2010.

[28] M. Zamani, G. Pola, M. Mazo, and P. Tabuada, "Symbolic models for nonlinear control systems without stability assumptions," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1804–1809, 2012.

[29] M. Zamani, P. M. Esfahani, R. Majumdar, A. Abate, and J. Lygeros, "Symbolic control of stochastic systems via approximately bisimilar finite abstractions," *IEEE Transactions on Automatic Control*, vol. 59, no. 12, pp. 3135–3150, 2014.

[30] H. L. S. Younes and R. G. Simmons, "Statistical probabilistic model checking with a focus on time-bounded properties," *Information and Computation*, vol. 204, no. 9, pp. 1368–1409, Sep. 2006.

[31] K. Sen, M. Viswanathan, and G. Agha, "On statistical model checking of stochastic systems," in *Computer Aided Verification*, ser. Lecture Notes in Computer Science, K. Etessami and S. K. Rajamani, Eds. Springer Berlin Heidelberg, Jan. 2005, no. 3576, pp. 266–280.

[32] H. L. S. Younes, "Error control for probabilistic model checking," in *Verification, Model Checking, and Abstract Interpretation, 7th International Conference, VMCAI 2006, Charleston, SC, USA, January 8-10, 2006, Proceedings*, 2006, pp. 142–156.

[33] N. Roohi, Y. Wang, M. West, G. Dullerud, and M. Viswanathan, "Statistical verification of the Toyota powertrain control verification benchmark," in *Proceedings of HSCC*, 2017, pp. 65–70.

[34] Y. Wang, N. Roohi, M. West, M. Viswanathan, and G. E. Dullerud, "Verifying continuous-time stochastic hybrid systems via mori-zwanzig model reduction," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE, 2016, pp. 3012–3017.

[35] ——, "Statistical verification of dynamical systems using set oriented methods," in *HSCC*, 2015, pp. 169–178.

[36] ——, "Statistical verification of dynamical systems using set oriented methods," in *Proceedings of the 18th International Conference on Hybrid Systems: Computation and Control*. ACM, 2015, pp. 169–178.

[37] A. R. Teel, A. Subbaraman, and A. Sferlazza, "Stability analysis for stochastic hybrid systems: A survey," *Automatica*, vol. 50, no. 10, pp. 2435 – 2456, 2014.

[38] A. R. Teel and J. P. Hespanha, "Stochastic hybrid systems: a modeling and stability theory tutorial," in *Decision and Control (CDC), 2015 IEEE 54th Annual Conference on*. IEEE, 2015, pp. 3116–3136.

[39] A. R. Teel, *Recent Developments in Stability Theory for Stochastic Hybrid Inclusions*. Cham: Springer International Publishing, 2017, pp. 329–354.

[40] A. Subbaraman and A. R. Teel, "Robust global recurrence for a class of stochastic hybrid systems," *Nonlinear Analysis: Hybrid Systems*, vol. 25, pp. 283 – 297, 2017.

[41] I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus*. Springer Science & Business Media, 2012, vol. 113.

[42] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*. Springer Science & Business Media, 2013, vol. 293.

[43] F. B. Hanson, "Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis and Computation," p. 29, 2007.

[44] O. Maler and D. Nickovic, *Monitoring Temporal Properties of Continuous Signals*, 2004, pp. 152–166.

[45] J. V. Deshmukh, A. Donzé, S. Ghosh, X. Jin, G. Juniwal, and S. A. Seshia, *Robust Online Monitoring of Signal Temporal Logic*, 2015, pp. 55–70.

[46] A. Donzé, T. Ferrère, and O. Maler, *Efficient Robust Monitoring for STL*, 2013, pp. 264–279.

[47] R. Alur and D. L. Dill, "A theory of timed automata," *Theor. Comput. Sci.*, vol. 126, no. 2, pp. 183–235, Apr. 1994.

[48] W. Rudin, "Functional analysis," 1973.

[49] T. Batu, L. Fortnow, R. Rubinfeld, W. D. Smith, and P. White, "Testing closeness of discrete distributions," *J. ACM*, vol. 60, no. 1, pp. 4:1–4:25, Feb. 2013.

[50] P. Gastin and D. Oddoux, "Fast ltl to büchi automata translation," in *Proceedings of the 13th International Conference on Computer Aided Verification*, ser. CAV '01. London, UK, UK: Springer-Verlag, 2001, pp. 53–65.

[51] A. Duret-Lutz, "Ltl translation improvements in spot," in *Proceedings of the Fifth International Conference on Verification and Evaluation of Computer and Communication Systems*, ser. VECoS'11. Swinton, UK, UK: British Computer Society, 2011, pp. 72–83.

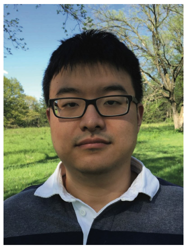
[52] A. Duret-Lutz and D. Poitrenaud, "Spot: an extensible model checking library using transition-based generalized büchi automata," in *IN PROC. OF MASCOTS'04*. IEEE Computer Society, 2004, pp. 76–83.

[53] M. Dellnitz and O. Junge, "On the approximation of complicated dynamical behavior," *SIAM Journal on Numerical Analysis*, vol. 36, no. 2, pp. 491–515, Jan. 1999.

[54] A. Wald, "Sequential tests of statistical hypotheses," *The Annals of Mathematical Statistics*, vol. 16, no. 2, pp. pp. 117–186, 1945.

[55] Y. S. Chow and H. Robbins, "On the asymptotic theory of fixed-width sequential confidence intervals for the mean," *The Annals of Mathematical Statistics*, vol. 36, no. 2, pp. 457–462, 04 1965.

[56] "Commons Math: The Apache Commons Mathematics Library," https://commons.apache.org/proper/commons-math, accessed: 2019-06-10.



Yu Wang is a Ph.D. candidate at the University of Illinois at Urbana-Champaign, majoring in Mechanical Engineering. His research focuses on the security and privacy of cyber-physical systems. He received his M.S. degrees in Mechanical Engineering, Mathematics, and Statistics from the same school, in 2014, 2016, and 2017, respectively, and his B.E. degree in Engineering Mechanics from Tsinghua University in 2012.

Nima Roohi is a postdoctoral researcher in Department of Computer and Information Science at the University of Pennsylvania. His research focuses on developing modeling and verification techniques for Cyber-Physical Systems. He received his Ph.D. degree from the Department of Computer Science and his M.S. degree in Mathematics at the University of Illinois Urbana-Champaign. Before joining UIUC, he received his M.S. in Software Engineering from the Sharif University of Technology and his B.S. in Software Engineering from the Amirkabir University of Technology.

Matthew West received the B.Sc. in pure and applied mathematics from the University of Western Australia, Crawley WA, Australia, in 1996, and the Ph.D. in control and dynamical systems from the California Institute of Technology, Pasadena, CA, USA, in 2004.

He was an Assistant Professor in the Department of Mathematics at the University of California, Davis, CA, USA, from 2003 to 2004, and an Assistant Professor in the Department of Aeronautics and Astronautics at Stanford University, Stanford, CA, USA, from 2004 to 2007. He joined the Department of Mechanical Science and Engineering at the University of Illinois at Urbana-Champaign, Urbana, IL, USA, as an Assistant Professor in 2008, and he is currently an Associate Professor in this department. His research interests include time integration of deterministic and stochastic systems, distributed computation, and particle methods for simulation and estimation.

Dr. West is a recipient of the NSF CAREER award

Mahesh Viswanathan received his Bachelor of Technology degree in Computer Science and Engineering from the Indian Institute of Technology, Kanpur, India in 1995, and his Ph.D. in Computer and Information Science from the University of Pennsylvania in 2000. He was a post-doctoral fellow at the Center for Discrete Mathematics (DIMACS), Rutgers University, with a joint appointment at Telcordia Technologies in the academic year 2000-01.

Since 2001, he has been on the faculty at the Department of Computer Science at the University of Illinois, Urbana-Champaign where he is currently a Professor. He was a visiting faculty fellow at the Courant Institute of Mathematics, New York University during the academic 2009–10, and a visiting faculty at the ´ Ecole Normale Supérieure at Cachan in the summer of 2011. His research interests are in the core areas of logic, automata theory, and algorithm design with its applications to automated verification of computing systems, including hybrid systems.

Geir E. Dullerud (F'08) received the B.A.Sc. degree in engineering science and the M.A.Sc. degree in electrical engineering both from the University of Toronto, Toronto, ON, Canada, in 1988 and 1990, respectively, and the Ph.D. degree in engineering from the University of Cambridge, Cambridge, U.K., in 1994.

Since 1998, he has been a Faculty Member with the Mechanical Science and Engineering Department, University of Illinois, Urbana-Champaign (UIUC), Urbana, IL, USA, where he is currently a W. Grafton and Lillian B. Wilkins Professor. He is the Director of the Decision and Control Laboratory of the Coordinated Science Laboratory. He has held visiting positions in electrical engineering at KTH, Stockholm, Sweden, in 2013, and in aeronautics and astronautics, Stanford University, Stanford, CA, USA, during 2005-2006. From 1996 to 1998, he was an Assistant Professor of applied mathematics at the University of Waterloo, Waterloo, ON, Canada. He was a Research Fellow and Lecturer in the Control and Dynamical Systems Department, California Institute of Technology, Pasadena, CA, USA, in 1994 and 1995. He has published two books: A Course in Robust Control Theory (Springer, 2000) and Control of Uncertain Sampled-Data Systems (Birkhauser, 1996). His current research interests include games and networked control, robotic vehicles, hybrid dynamical systems, and cyber-physical systems security.

Dr. Dullerud is currently Associate Editor of the SIAM Journal on Control and Optimization, and served in a similar role for Automatica. He received the National Science Foundation CAREER Award in 1999, and the Xerox Faculty Research Award at UIUC in 2005. He became an ASME Fellow in 2011. He was an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.