Reachability Computation for Switching Diffusions: Finite Abstractions with Certifiable and Tuneable Precision

Alessandro Abate
Department of Computer Science
University of Oxford
alessandro.abate@cs.ox.ac.uk

Luca Bortolussi
Department of Mathematics and Geosciences
University of Trieste
luca@dmi.units.it

Marta Kwiatkowska
Department of Computer Science
University of Oxford
marta.kwiatkowska@cs.ox.ac.uk

Luca Cardelli
Microsoft Research &
University of Oxford
luca@microsoft.com

Milan Ceska
Faculty of Information Technology
Brno University of Technology
ceskam@fit.vutbr.cz

Luca Laurenti
Department of Computer Science
University of Oxford
luca.laurenti@cs.ox.ac.uk

ABSTRACT

We consider continuous time stochastic hybrid systems with no resets and continuous dynamics described by linear stochastic differential equations – models also known as switching diffusions. We show that for this class of models reachability (and dually, safety) properties can be studied on an abstraction defined in terms of a discrete time and finite space Markov chain (DTMC), with provable error bounds. The technical contribution of the paper is a characterization of the uniform convergence of the time discretization of such stochastic processes with respect to safety properties. This allows us to newly provide a complete and sound numerical procedure for reachability and safety computation over switching diffusions.

Keywords

Switching diffusions; stochastic hybrid models; reachability and safety analysis; finite abstractions; time and space discretisation; numerical computations

1. INTRODUCTION

Hybrid models are natural in the context of cyber-physical systems applications, where continuous dynamics of physical variables are interleaved with discrete updates of finite-state models. Furthermore, in many engineering and natural systems, noise or uncertainty structured via probabilistic laws are relevant, which leads to stochastic models. In this context stochastic hybrid models encompass all these features, and their properties have been recently investigated [17, 10, 18, 3].

In this work we consider switching diffusions [27, 38, 6], models that are characterised by dynamics over a hybrid state space: continuous-time flows are determined by the solution of a mode-dependent linear diffusion process, whereas mode updates (over finitely many locations) hinge on events triggered by Poisson processes, with rates that depend on the continuous variables. As such, switching diffusions can be regarded as special instances of stochastic hybrid models, the latter dealing also with probabilistic resets between discrete-mode commutations. The models considered in this work are also fully observable and not subject to any form of non-determinism (such as control inputs, as discussed in [27, 6]).

This contribution investigates the problem of reachability analysis for switching diffusions, a central problem due to the duality between reachability and safety problems, and its role in the verification of many other specifications (thanks to product constructions). Whilst this is a widely investigated problem, contributions in the literature have been limited to the characterisation of this problem, with computational aspects that have been relegated to the use of approximation techniques often resorting to state-space gridding with no guarantees.

Contribution

This work provides a formal computational procedure for the reachability analysis problem over switching diffusions. As such, we address an open problem also for the special case of linear stochastic differential equations. More precisely, we provide approximation algorithms with certificates on their precision, which reduce the problem to the computation over a finite-state Markov chain. In other words, we show that probabilistic reachability can be formally computed over finite abstractions, obtained by discretising the continuous components of the models (time and space).

Related work

Stochastic hybrid models (SHS) are broadly discussed in [18, 10], and switching diffusions investigated in [27, 38, 1, 6].

The characterisation of probabilistic reachability for SHS is elaborated in [13] by means of a number of techniques, but not under the lens of computations. [13] leverages and extends theory developed for piece-wise deterministic Markov processes in [20]. Further, [29] has characterised probabilis-

DOI: 10.1145/1235

tic reachability for SHS as a solution of a PDE (HJI partial differential equation), but only provided weak convergence results for its computation, based on the approximation theory in [30]. A similar approach has been pursued in [34], but again with no numerical scheme with certifiable errors. It appears that the application of numerical schemes for time discretisation of SDE [28] are not of help. [22] has extended the characterisation to constrained reachability problems. In [11, 16, 12] numerical algorithms for verification of linear SDE, obtained for Markov population processes in the limit of high population, have been given with just weak convergence results.

On the other hand, for discrete-time stochastic hybrid models probabilistic reachability (and safety) have been fully, characterised [3], connected with verification procedures [2, 37], formally computed via software tools [23] leveraging finite abstractions, and indeed extended to general specifications [36].

An alternative approach towards formal, finite approximations of continuous-time stochastic models is discussed in [40] and extended in [39] to switching diffusions. Noteworthy are also techniques and tools for verification of related probabilistic models based on abstractions [41], measurability conditions [24], and SMT technology [25] approaches. These techniques, alongside that of this work, are clearly set apart from statistical model checking approaches [15].

2. STOCHASTIC HYBRID PROCESSES

We consider the following class of continuous time stochastic hybrid systems with no guards or resets, which are also commonly denoted as switching diffusions. We refer the reader to [10, 18] for technical details on the measure theoretical aspects underlying these processes.

DEFINITION 1. *A switching diffusion $\mathcal{H}$ is a tuple $\mathcal{H} = (\mathcal{Q}, K, F, G, W, \Lambda)$, where*

- $\mathcal{Q} = \{q_1, ..., q_{|\mathcal{Q}|}\}$ *is the set of* discrete modes
- $K \subseteq \mathbb{R}^m$, *for $m > 0$, is the state space of the* continuous dynamics. *The hybrid state space is defined as $\mathcal{D} = \cup_{q\in\mathcal{Q}}\{q\} \times K$*
- $F : \mathcal{Q} \to \mathbb{R}^{m\times m}$ *is the* drift *term for the continuous dynamics*
- $G : \mathcal{Q} \to \mathbb{R}^{m\times q}$ *is the* diffusion *associated to the continuous dynamics*
- *W is a $q-$dimensional Wiener process*
- $\Lambda : \mathcal{D} \times \mathcal{Q} \to \mathbb{R}_{\geq 0}$ *is an* intensity function, *where for $(q_i, x) \in \mathcal{D}, q_j \in \mathcal{Q}$, we define $\Lambda((q_i, x), q_j) = \lambda_{i,j}(x)$*

Let W be defined in the probability space $(\Omega, \mathcal{F}, P)$ with filtration $\mathcal{F}_t$, where a filtration is a family of $\sigma-$algebras representing the information available at time t. Then, given $\mathcal{H}$ and an initial condition $y_0 = (x_0, q_0) \in \mathcal{D}$, the stochastic process $Y = (X, \alpha)$, defined on the hybrid state space $\mathcal{D} = \cup_{q\in\mathcal{Q}}\{q\} \times K$ is a solution of $\mathcal{H}$ if it satisfies

$$dX(t) = F(\alpha(t)) \cdot X(t)dt + G(\alpha(t)) \cdot dW(t), \qquad (1)$$

and for $i \neq j$

$$P\left(\alpha(t + \Delta t) = q_j | Y(t) = (q_i, x)\right) = \lambda_{i,j}(x)\Delta t + o(\Delta t) \quad (2)$$

with $(X(0), \alpha(0)) = (x_0, q_0)$.

The discrete dynamics of Y, described by variable α, evolves as a jump process over the discrete state space $\mathcal{Q}$, with jump rate dependent on the continuous part. The continuous dynamics of Y evolves according to a linear diffusion. That is, when the discrete system is in a particular state, X evolves according to a linear SDE driven by a Wiener process. Then, when the discrete systems hits a change in its state, X continues evolving according to a different SDE without resetting its state.

ASSUMPTION 1. *We introduce the following assumptions:*

- *$\lambda_{i,j}(x)$ is a bounded and locally Lipschitz continuous function in x, for all $q_i, q_j \in \mathcal{Q}$*
- *$|F(q)x| + |G(q)| \leq C(1 + |x|)$ forall $q \in \mathcal{Q}$, for some constant C where $|G(q)| = \sum_{i,j} |G(q)(i, j)|$*
- *$|F(q)x - F(q)x'| \leq D|x - x'|$ forall $q \in \mathcal{Q}$, for some constant D*

The first condition guarantees that over any finite time interval, α almost surely jumps only a finite number of times, thus excluding Zeno behaviours. The second and third conditions guarantee that the continuous solution X exists and is unique, and that it remains bounded over a finite time interval [33].

EXAMPLE 1. *Consider the stochastic process X described by the following SDE*

$$dX(t) = F \cdot X(t)dt + G \cdot dW(t) \qquad (3)$$

with initial condition $X(0) = x_0 \in \mathbb{R}^m$. That is, X is the solution of a hybrid process $\mathcal{H}$ with a singleton discrete state space ($\mathcal{Q} = \{q\}$). It is well known that the evolution of the probability distribution of the solution of a SDE over time satisfies the following Fokker-Planck equation [26]

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial x_i} \left[(F(t) \cdot x)_i p(\mathbf{x}, t)\right] + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial^2}{\partial x_i \partial x_j} \left[D_{ij} p(\mathbf{x}, t)\right],$$

with diffusion tensor $D_{ij} = \sum_{k=1}^{q} G_{ik}G_{jk}$.

The following lemma guarantees that X, process solution of Equation 3 is a Gaussian process.

LEMMA 1. *[7, 35] Let $X(0)$ be a normally distributed random variable with expected value $E[X(0)] = E_{x_0}$ and covariance matrix $C_X(0) = E[X(0)X(0)] = C_{x_0}$. Then, X, as defined in* (3)*, is a Gaussian Markov process with expected value and covariance matrix given by*

$$\begin{cases} \frac{dE[X(t)]}{dt} & = FE[X(t)] \\ E[X(0)] & = E_{x_0}, \end{cases} \qquad (4)$$

$$\begin{cases} \frac{dC_X(t)}{dt} & = FC_X(t) + C_X(t)^T F + G(G^T) \\ dC_X(0) & = C_{x_0}. \end{cases} \qquad (5)$$

Lemma 1 allows us to derive the analytical solution for the expectation and variance of the solution of a linear SDE as

$$E[X(t)] = e^{Ft}E_{x_0},$$

$$C_X(t) = e^{Ft}C_{x_0}(e^{Ft})^T + \int_0^t (e^{F(t-s)})GG^T(e^{F(t-s)})^T ds.$$

3. PROBLEM DEFINITION

Given a stochastic process Y with state space $\mathcal{D}$, a *target set* $\mathcal{S} \subseteq \mathcal{D}$, which is assumed to be measurable, and a time interval $I \subseteq \mathbb{R}_{\geq 0}$, the reachability problem is defined as the search for the characterisation and computation of the probability that Y will reach $\mathcal{S}$ during I. This problem is dual to the safety problem, that is, computing the probability that the system will remain in a given, measurable safe region, over a given time interval. The characterisation of the two problems is thus interchangeable [2]. Reachability analysis is one of the fundamental problems in the quantitative analysis of models, and it is likewise key for the analysis of stochastic hybrid processes [14]. Model checking of Continuous Stochastic Logic (CSL) [8] reduces in computing reachability problems. Likewise for discrete-time stochastic hybrid systems, reachability and safety play a pivotal role for model checking PCTL formulae [31], and more complex properties via the product construction [36].

PROBLEM 1. *(Probabilistic Reachability) Let $\mathcal{H}$ be a hybrid process, and $Y = (X, \alpha)$ its solution with state space $\mathcal{D}$. Let $\mathcal{S} \subseteq \mathcal{D}$ be a measurable set and $I = [t_1, t_2]$ a time interval. The reachability probability for Y to reach $\mathcal{S}$ in I is defined as*

$$P_{reach}(Y, \mathcal{S}, I) = Prob\{\exists t \in I \, s.t. \, Y(t) \in \mathcal{S}\}. \quad (6)$$

The safety problem is introduced as

$$P_{safe}(Y, \mathcal{S}, I) = Prob\{\forall t \in I, Y(t) \in \mathcal{S}\}$$

and is the dual of the reachability problem, namely

$$P_{safe}(Y, \mathcal{S}, I) = 1 - P_{reach}(Y, \mathcal{S}^c, I)$$

where $\mathcal{S}^c$ is the complement of set $\mathcal{S}$.

Analytic solutions of Problem 1 for the class of hybrid systems we consider are in general infeasible, as they would be tantamount to viscosity solutions of systems of Hamilton-Jacobi-Bellman equations [29]. In this work we instead introduce a numerical algorithm that employs time- and space discretization to solve Problem 1 – in particular the time discretisation part of the scheme is new. We further show that for the class of processes considered in this paper, the safety value computed on the discrete time and finite space Markov chain (DTMC) abstraction, obtained from the overall procedure, converges uniformly to the safety value associated to the given (continuous) switching diffusion as the discretisation parameters become zero. We also offer explicit error bounds quantifying this approximation level.

In the following illustrative example, we consider a simple SDE model, for which analytical solutions to the probabilistic safety problem exist.

EXAMPLE 2. *Consider the stochastic process X described by the following SDE*

$$dX(t) = GdW(t),$$

where W is a uni-dimensional Brownian motion and $G \in \mathbb{R}_{\geq 0}$. Assume there is no discrete switching (and a single discrete location): then the probability density function of process X can be described by the following diffusion equation

$$\frac{\partial p(x, t \mid x_0)}{\partial t} = \frac{G^2}{2} \frac{\partial^2 p(x, t \mid x_0)}{\partial x^2},$$

with initial condition $p(x, 0 \mid x_0) = \delta(x - x_0)$. For $\bar{x} \in \mathbb{R}$, we consider the safe set $\mathcal{S}_{\bar{X}} = \{x \in \mathbb{R} : x \leq \bar{x}\}$. In order to solve $P_{safe}(X, S_{\bar{X}}, [0, 1])$, we need to integrate $p(x, t; x_0)$ with the boundary condition $P(\bar{x}, t) = 0$: this leads to the following density function for, $x < \bar{x}$,

$$p(x, t; x_0, \bar{x}) = \frac{1}{\sqrt{2\pi G t}} (e^{-\frac{(x - x_0)^2}{2Gt}} - e^{-\frac{(x - (2\bar{x} - x_0))^2}{2Gt}}).$$

We then obtain

$$P_{safe}(X, S_{\bar{X}}, [0, 1]) = \int_{-\infty}^{\bar{x}} p(x, 1; x_0, \bar{x}) dx = erf(\frac{\bar{x} - x_0}{\sqrt{2G}}),$$

where $erf()$ is the Gaussian error function.

4. TIME DISCRETIZATION

Given a hibrid system $\mathcal{H}$, its solution, $Y = (X, \alpha)$, is a continuous time Markov process defined on the hybrid space $\mathcal{D} = \cup_{q \in \mathcal{Q}} \{q\} \times K$, where $K \subseteq \mathbb{R}^m, m > 0$. By sampling Y with a fixed interval $h > 0$, we obtain a discrete time Markov process $\bar{Y} = (\bar{X}, \bar{\alpha})$ defined on the same hybrid state space $\mathcal{D}$ and such that $\bar{Y}(k) = Y(h \cdot k), k \in \mathbb{N}$.

DEFINITION 2. *The discrete time Markov process (DTMP) $(\bar{Y}(k) = (\bar{X}(k), \bar{\alpha}(k)), k \in \mathbb{N}$ is a time homogeneous hybrid model, uniquely defined by a quadruple $(\mathcal{D}, \sigma, T^c, T^d)$, where $(\mathcal{D}, \sigma)$ is the measurable space inherited from $\mathcal{H}$; $T^c : A \times \mathcal{D} \to [0, 1]$, for $A \subseteq \mathbb{R}^m$, is a continuous transition kernel; and $T^d : \mathcal{Q} \times \mathcal{D} \to [0, 1]$ is a discrete transition kernel.*

T^c and T^d describe the probability that the continuous and discrete components of the process transition onto a measurable set at the next discrete step, given the current state of the process. More precisely, for state $(q, x) \in \mathcal{D}$ and Borel-measurable set $(q', A) \subseteq \mathcal{D}$, we have that

$$T^c(A, x, q) = Prob(\bar{X}(k+1) \in A | \bar{X}(k) = x, \bar{\alpha}(k) = q)$$

$$T^d(q', x, q) = Prob(\bar{\alpha}(k+1) = q' | \bar{X}(k) = x, \bar{\alpha}(k) = q).$$

T^c and T^d fully characterize $\bar{Y} = (\bar{X}, \bar{\alpha})$. In the following proposition we derive an analytical form for such kernels. To keep the presentation simpler, only for the following theorem, we make the further simplification that the jump rates do not depend on the continuous state $x \in K$: $\lambda_{ij}(x) = \lambda_{ij}$. This assumption allows us to have a simpler form of the kernel. To deal with more general rate functions, the simplest way to proceed is to assume that they are piecewise constant in each considered interval of time, fixing the value of λ_{ij} to the one in the initial state. As rate functions are locally Lipschitz, the distance between the true rate and λ_{ij} will be bounded by a term of order $O(h)$, which can be lifted at the kernel level.

THEOREM 1. *Let $\mathcal{H} = (\mathcal{Q}, K, F, G, W, \Lambda)$ be a hybrid process and $Y = (X, \alpha)$ its solution. Assume the jump rates do not depend on the continuous state. Let $h > 0$ be a sampling time and $\bar{Y} = (\bar{X}, \bar{\alpha})$ the resulting DTMP. Call $\mathcal{N}(\bar{x} | E, C)$ the normal distribution with mean E and covariance C. Introduce terms*

$$\Gamma(i, t) = \int_0^t (e^{F(q_i)(t-m)}) G(q_i) G(q_i)^T (e^{F(q_i)(t-m)})^T dm,$$

$$\Omega_{\lambda_i, \lambda_j, t}(s) = (\lambda_j - \lambda_i) \frac{e^{(\lambda_j s - \lambda_j t - \lambda_i s)}}{e^{(-\lambda_i t)} - e^{(-\lambda_j t)}},$$

and for $x \in \mathbb{R}^m$ *define* $\lambda_i(x) = \sum_{j \neq i} \lambda_{i,j}(x)$. *Then, given* $(q, x) \in \mathcal{D}$ *and* (q', A) *a measurable set, it holds that*

$$T^c(A, x, q_i) = \int_A \mathcal{N}(\bar{x}|e^{F(q_i)\cdot h}x, \Gamma(i,h))d\bar{x} \cdot e^{-\lambda_i h} + \sum_{q_j \neq q_i} \int_A (\int_0^h \mathcal{N}(\bar{x}|E^{\mathcal{H}}_{q_i,x}(s), C^{\mathcal{H}}_{q_i,x}(s)) \cdot \Omega_{\lambda_i,\lambda_j,h}(s)ds) \cdot \frac{\lambda_{ij}}{\lambda_i} \cdot \lambda_i h \cdot e^{-\lambda_i h}\, d\bar{x} + \epsilon$$

and

$$T^d(q_j, x, q_i) = \begin{cases} e^{-\lambda_i h} + \epsilon & \text{if } q_i = q_j \\ \lambda_i h \cdot e^{-\lambda_i h} \cdot \frac{\lambda_{ij}}{\lambda_i} + \epsilon & \text{if } q_i \neq q_j \end{cases},$$

where

$$E^{\mathcal{H}}_{q_i,x}(s) = e^{F(q_i)s}e^{F(q_i)(h-s)}x,$$
$$C^{\mathcal{H}}_{q_i,x,s} = e^{F(q_i)s}\Gamma(i,s)(e^{F(q_i)s})^T + \Gamma(j, h-s),$$
$$0 \leq \epsilon \leq 1 - e^{-\lambda_i h} - \lambda_i h \cdot e^{-\lambda_i h}.$$

The full derivation of the continuous kernel, $T^c(A, x, q_i)$, is shown in the Appendix. Each integral over A quantifies the probability that the continuous component of the model enters set A, conditional on the discrete part of the process performing either 0 or 1 jumps during the sampling interval h. Assuming to be in the discrete location q_i, the probability of these events is respectively $e^{-\lambda_i h}$ and $\lambda_i h \cdot e^{-\lambda_i h}$ [19], where x is the state at time kh. If the discrete system makes no jumps within $[0, h]$, then, because of the memoryless property of the SDE, during this interval X evolves according to equations as in Lemma 1 and specific to location q_i. As a consequence, at time h, X is normally distributed, with mean $e^{F(q_i)\cdot h}x$ and variance $\Gamma(i, h)$. If instead the system jumps once within $[0, h]$, after marginalizing over the jump time and the state where this event happens, we end up with a *linear Gaussian model* [9]. This process is still Gaussian with mean and covariance matrix that can be derived from the equations in Lemma 1. Finally, parameter ϵ takes into account the probability associated to paths with more than one jump within $[kh, (k+1)h]$: based on the provided upper bound on ϵ, it is clear that the probability of such event becomes negligible as h gets small enough.

The discrete kernel $T^d(q_j, x, q_i)$, has a much simpler derivation. If we assume that the system makes at most one jump during h, then the probability that $q_j = q_i$ amounts to the probability that the system does not jump within $[0, h]$. Instead, for the condition $q_j \neq q_i$ the resulting probability is obtained as the probability of making a jump once, multiplied by the probability of jumping to the specific state q_j.

From T^c and T^d, for $(x, q_i) \in \mathcal{D}$ and for a measurable set $(A, q_j) \subseteq \mathcal{D}$, it is easy to calculate the following transition kernel

$$T((A,q_j),(x,q_i)) = Prob(Y((k+1)h) \in (A, x_j) | Y((k)h) = (q, x_i), k \in \mathbb{N}).$$

In fact, from Theorem 1, we have

$$T((A, q_j), (x, q_i)) = \begin{cases} \int_A \mathcal{N}(\bar{x}|e^{F(q_i)\cdot h}x, \Gamma(i,h))d\bar{x} \cdot e^{-\lambda_i h} & \text{if } q_i = q_j \\ \int_A (\int_0^h \mathcal{N}(\bar{x}|E^{\mathcal{H}}_{q_i,x}(s), C^{\mathcal{H}}_{q_i,x}(s)) \cdot \Omega_{\lambda_i,\lambda_j,h}(s)ds) \cdot d\bar{x} \cdot \frac{\lambda_{ij}}{\lambda_i}\lambda_i h e^{-\lambda_i h} & \text{if } q_i \neq q_j \end{cases}$$

Note that the derived kernels are time homogeneous: from a numerical point of view this is a key property that facilitates the practical computation of the resulting DTMP, which is also time homogeneous.

4.1 Error Bounds for Time Discretization

In this section we quantify the approximation level introduced by the discretisation procedure. More precisely, we characterize the error associated to the computation of reachability properties with the DTMP over a discrete set of sampling points, with sampling time $h > 0$: by deriving formal error bounds, we show uniform convergence as $h \to 0$.

ASSUMPTION 2. *Assume that the target set* $\mathcal{S}$ *is independent of the locations, namely select* $S \subseteq \mathbb{R}^m$ *so that* $\mathcal{S} = \cup_{q \in \mathcal{Q}}\{q\} \times S$.

Let $\mathcal{H} = (\mathcal{Q}, K, F, G, W, \Lambda)$ be a hybrid system and $Y = (X, \alpha)$ its solution. Let $I \subseteq \mathbb{R}_{\geq 0}$ be a finite time interval. For any $q \in \mathcal{Q}$, call X_q the solution of the SDE

$$dX_q(t) = F(q)X_q(t)dt + G(q)dW(t).$$

In this section we assume that X_q is a uni-dimensional, zero mean Gaussian process (GP). In the next subsection, we show how to generalize the results derived here for general GPs and multi-dimensional processes.

X_q is almost surely bounded within the interval I by Assumption 1. Set $h = min\{\frac{2^{-n}}{2\sqrt{2}K^2K_d}, 2^{-n}\}$ and $\epsilon_n = 2^{-\frac{n}{2}}$, where $n \in \mathbb{N}$, and K_d is a constant such that for any $t_1, t_2 \in I$,

$$\max_{q \in \mathcal{Q}}\{d_q(t_1, t_2)\} \leq K_d \cdot |t_2 - t_1|,$$

where d_q is a pseudometric defined as

$$d_q(t_1, t_2) = \sqrt{E[(X_q(t_2) - X_q(t_1))^2]},$$

and $K \geq 12$ is the universal constant in the Dudley's metric entropy integral [32]. It is possible to show that under mild assumptions the almost sure bound for X_q guarantees that K_d is finite [5]. Fix a set of sampling times $\Sigma = \{t_1, ..., t_{|\Sigma_n|}\}$, with step distance h. Call

$$S^{\epsilon_n} = \{x \in S \ : \ |x - \partial S| \geq \epsilon_n\},$$

where ∂S is the boundary of S and $|\cdot|$ is the Euclidean distance metrics between a point and a set. Define the events

$$\mathcal{A}^n = \{\forall t_i \in \Sigma,\, X(t_i) \in S^{\epsilon_n}\}$$

and

$$\mathcal{B} = \{\exists t \in [0, T] \ s.t. \ X(t) \notin S\}.$$

As $S \subseteq \mathbb{R}$, we have that $P_{safe}(Y, \mathcal{S}, I) = P_{safe}(X, S, I)$, where $P_{safe}(X, S, I)$ is the probability that the continuous

component X of Y, stays in S during I. It is easy to see that

$$P_{safe}(X,S,I) = \lim_{n\to\infty} P(\mathcal{A}^n \wedge \mathcal{B}^c).$$

For a finite $n > 0$, $P(\mathcal{A}^n \wedge \mathcal{B}^c)$ is a lower bound for the safety probability computed on $\mathcal{S}$. This is because it requires the systems to be inside $S^{\epsilon_n} \subseteq S$ at sampling times in Σ. Notice that for n big enough S and S^{ϵ_n} become indistinguishable. As a consequence, we can compute the reachability on S^{ϵ_n} instead of S, as for n big enough the two sets becomes indistinguishable. The dependence of S^{ϵ_n} from n is a key aspect of our approach. Let us define as $P(\mathcal{A}^n)$ the reachability probability computed considering only the discrete times in Σ.

THEOREM 2. *Under Assumption 1, it holds that for $n \geq 3$ and over a finite time interval $I \subseteq \mathbb{R}_{\geq 0}$*

$$P(\mathcal{A}^n) \geq P(\mathcal{A}^n \wedge \mathcal{B}^c) \geq P(\mathcal{A}^n) \cdot (1 - \frac{I}{h} \exp^{-(2^n - 2^{\frac{n}{2}}+1)}),$$

where $h = min\{\frac{2^{-n}}{2\sqrt{2}K^2 K_d}, 2^{-n}\}$.

COROLLARY 1. *Under the Assumption 1, it holds that*

$$\lim_{n\to\infty} P(\mathcal{A}^n \wedge \mathcal{B}^c) = \lim_{n\to\infty} P(\mathcal{A}^n).$$

Theorem 2 guarantees that for any $n \geq 3$, we obtain

$$|P(\mathcal{A}^n) - P(\mathcal{A}^n \wedge \mathcal{B}^c)| \leq \frac{I}{h} \exp^{-(2^n - 2^{\frac{n}{2}}+1)}.$$

This enables choosing, a priori, a sampling interval h that guarantees meeting a chosen error on the precision. The proof of Theorem 2 is given in the Appendix. Here, we explain the main ideas. The proof of Theorem 2 is based on the fact that, for any $q \in \mathcal{Q}$, X_q is a Gaussian process, which is almost surely bounded in T. It is possible to show that the supremum of X_q is still distributed as a Gaussian [4]. Then, the use of the entropy Dudley's integral [21] allows to bound the probability that each X_q stays in a ϵ_n-neighbourhood between two sampling points. The fact that S^{ϵ_n} depends on the sampling interval concludes the proof. Note that a key feature enabling this approach is the absence of resets of the continuous state upon mode change. As a consequence, we can simply assume to find constants for the "worst behaving X_q" in a particular interval, without worrying about the discrete mode changes.

4.1.1 Discussion on the error and extensions

In the derivation of Theorem 2 we assumed that the continuous component of Y, solution of $\mathcal{H}$, is zero mean and uni-dimensional. This is not a limitation: Lemma 1 guarantees that for any $q \in \mathcal{Q}$, the variance of the solution X_q is independent of the particular continuous location, depending exclusively on time. Moreover, given a set $S \subseteq \mathbb{R}$ and $h > 0$, for $E[X(0)] = x \in S$ from Equation 4, it is possible to derive a constant $K^m_{h,S}$ such that

$$sup_{q\in\mathbb{Q},t_1,t_2\in[0,h]}\{|E[X_q(t_2) - E[X_q(t_1)]]\} \leq K^m_{h,S} \cdot h.$$

Then, we can simply consider as target for the continuous components the set $S' = S \cup \{x \in \mathbb{R} - S : |x - \partial S| \leq K^m_{h,S} \cdot h\}$. As such, the bound computed for $X - E[X]$ on $\mathcal{S} = \cup_{q\in\mathcal{Q}} q \times S$ still holds for X on $\mathcal{S}' = \cup_{q\in\mathcal{Q}} q \times S'$.

One of the key properties of a multivariate Gaussian Process (mGP) is that each of its components is itself a Gaussian process. Moreover, the euclidean metric distance for X at time t can be defined as

$$|X(t)| = \sqrt{\sum_{i=1}^{m} |X_i(t)|^2},$$

where X_i is the $i-th$ component of X. As a consequence,

$$P(|X(t)| > \epsilon) \leq P(sup_{i\in[1,n]}(|X_i(t)| < \sqrt{\frac{\epsilon^2}{n}})).$$

These observations allow us to derive the following theorem, which generalize Theorem 2 to multi-dimensional continuous components.

THEOREM 3. *Let $\mathcal{H}$ be a hybrid process and $Y = (X, \alpha)$ its solution, with X m-dimensional process, for $m > 0$. Define $K_{d,i}$ the K_d constant relative to X_i as introduced in Section 4.1. Then, it holds that for $n \geq 3$ and over a finite time interval $I \subseteq \mathbb{R}_{\geq 0}$*

$$P(\mathcal{A}^n) \geq P(\mathcal{A}^n \wedge \mathcal{B}^c) \geq P(\mathcal{A}^n) \cdot (1 - \frac{I}{h} \exp^{-(2^n - 2^{\frac{n}{2}}+1)}),$$

where $h = min\{\frac{2^{-n}}{2\sqrt{2}K^2 \bar{K}_d}, 2^{-n}\}$, for $K \geq 12$ and $\bar{K}_d = sup_{i\in 1,\ldots,m}(K_{d,i})$.

OBSERVATION 1. *Consider an hybrid system $\mathcal{H}$ with solution $Y = (X, \alpha)$, where X takes values in $\mathbb{R}^m$, and α takes values in a finite set of discrete states $\mathcal{Q}$. Given a measurable set $\mathcal{S} = \cup_{q_i\in\mathcal{Q}}(q_i, S_i) \subseteq \mathcal{D}$, we can define $S' = \cap_{q_i\in\mathcal{Q}} S_i$ and $\mathcal{S}' = \cup_{q_i\in\mathcal{Q}}(q_i, S'$. Then, we have that for a general time interval I, $P_{safe}(Y,\mathcal{S},I) \geq P_{safe}(Y,\mathcal{S}',I)$. That is, if we need to compute probabilistic safety on a set that is discrete mode dependent, then, we can always compute a lower bound of this safety considering a location independent target set.*

As explained in Observation 1, Theorem 3 can still be used to get lower bounds of cases where the target set is continuous mode dependent. However, the bounds we obtain can be quite conservative if the target sets corresponding to different modes greatly differs.

5. STATE SPACE DISCRETIZATION

In order to complete the procedure leading to a model where we can numerically compute safety or reachability properties, we introduce a numerical scheme inspired by the results of [2, 23]. The numerical scheme is based on a Discrete Time Markov Chain (DTMC) approximation of the DTMP that results from the time discretization of the original switching diffusion process $\mathcal{H}$. We discuss convergence results and relative error bounds both of this second (state space) approximation step, and of the combined (time- and state approximation) procedure.

Let $\mathcal{S} = \cup_{q\in\mathcal{Q}}\{q\} \times A_q$ be the safe set, where $A_q \subseteq \mathbb{R}^m$. We assume $\mathcal{S}$ to be measurable and compact. Given $dx \in \mathbb{R}_{\geq 0}$, we define the grid $\mathcal{G}_{dx} = \cup_{q\in\mathbb{Q}} \cup_{i\in m_q} \{q\} \times A_{i,q}$, where $A_{i,q}$ are pairwise disjoint measurable sets, such that for $q \in \mathcal{Q}$ $\cup_{i\in m_q} A_{i,q} = A_q$, for $i \neq j$ $A_{i,q} \cap A_{j,q} = \emptyset$, and $A_{i,q} = \{x, x' \in A_q : |x - x'| \leq dx\}$. In other words, $\mathcal{G}_{dx}$ is a partition of $\mathcal{S}$ in sets of diameter dx. For each $(q, A_{i,q}) \in \mathcal{G}_{dx}$, we consider a representative point $(q, x_i) \in \{q\} \times A_{i,q}$. The set of representative points $\mathcal{S}_{dx} = \{(q, x_i), i \in \{1, \ldots m_q\}, q \in \mathcal{Q}\}$

makes up the finite state space of the DTMC, a discrete version of the set $\mathcal{S}$. Let us introduce $\xi : \mathcal{S} \to \mathcal{S}_{dx}$, a map that associates to any $(q, x) \in \mathcal{S}$ the corresponding representative point. Similarly, the set-valued map $\Xi : \mathcal{S} \to \mathcal{G}_{dx}$ relates any representative point to the concrete $A_{i,q}$ partition.

We define the discrete state space $\mathcal{Z}_{dx} = \mathcal{S}_{dx} \cup \phi$, where ϕ is a discrete state modeling all the states outside $\mathcal{S}$. Note that the compactness of $\mathcal{S}$ guarantees that $\mathcal{Z}$ is finite. The resulting DTMC is completely characterized by its transition kernel $T_{dx} : \mathcal{Z}_{dx} \times \mathcal{Z}_{dx} \to \mathbb{R}_{\geq 0}$, such that for $z_1 = (x_1, q_1), z_2 = (x_2, q_2) \in \mathcal{Z}_{dx}$, $T_{dx}(z_1, z_2)$ describes the probability of going in z_1 in the next discrete step, being in z_2 at the current time. T_{dx} can be easily computed from kernel T presented in Section 4 as

$$T_{dx}(z_1, z_2) = \begin{cases} T(z_1, z_2), & \text{if } z_1, z_2 \in \mathcal{S}_{dx} \\ 1 - \sum_{z_j \in \mathcal{S}_{dx} r} T(z_1, \Xi(z_j)), & \text{if } z_1 \in \mathcal{S}_{dx}, z_2 \in \phi \\ 1, & \text{if } z_1, z_2 \in \phi \\ 0, & \text{if } z_1 \in \phi, z_2 \in \mathcal{S}_{dx}. \end{cases}$$

5.1 Error Bounds for Space Discretization

Let $\bar{Y}$ the discrete time continuous space hybrid process derived through time discretization of Y, solution of the hybrid process $\mathcal{H}$, with initial condition $(x, q) \in \mathcal{S}$. Call Y^D the approximated DTMC with state space $\mathcal{Z}_{dx}$ and initial condition $(x^D, q^D) = \xi((x, q))$. We show that, for $I \subseteq \mathbb{R}_{\geq 0}$, under Assumption 1, the property $P_{safe}(Y^D, \mathcal{S}_{dx}, I)$ converges uniformly to $P_{safe}(\bar{Y}, S, I)$, which also allows us to derive uniform convergence on the original continuous time stochastic process, and to derive error bounds on the global approximation procedure.

Definition 3. *Let us introduce the following Lipschitz constants $h_1, h_2 \in \mathbb{R}_{\geq 0}$, which are such that*

$$|T^d(q', x_1, q) - T^d(q', x_2, q)| \leq h_1 \cdot |x_2 - x_1|,$$
$$\text{for all } (q, x_1), (q, x_2) \in \mathcal{S}, q' \in \mathcal{Q},$$
$$|t^c(x', x_1, q) - t^c(x', x_2, q)| \leq h_2 \cdot |x_2 - x_1|,$$
$$\text{for all } (q, x_1), (q, x_2) \in \mathcal{S}, x' \in K \cap \mathcal{S},$$

where t^c is the density function of the continuous kernel T^c.

Theorem 4. *[2] Let $\bar{Y}$ be the discrete time continuous space hybrid process with initial condition $(x, q) \in \mathcal{S}$, where $\mathcal{S}$ is a measurable set. Call Y^D the approximated DTMC with state space $\mathcal{Z}_{dx}$, where $dx > 0$ is the discretization parameter, and initial condition $(x^D, q^D) = \xi(x, q)$. Then, given $[0, N] \subseteq \mathbb{N}$, it holds that*

$$|P_{safe}(Y^D, \mathcal{S}_{dx}, N) - P_{safe}(\bar{Y}, \mathcal{S}, N)| \leq N \cdot \mathcal{K} \cdot dx,$$

where $\mathcal{K} = mh_1 + Lh_2$, with L the Lebesque measure of the continuous set $\mathcal{S}$, and m cardinality of the discrete set $\mathcal{Q}$.

Notice that, as $dx \downarrow 0$, the two probabilities collapse.

6. GLOBAL ALGORITHM AND ERRORS

Using the results in Theorem 4, we can derive the uniform convergence between $P_{safe}(Y^D, \mathcal{S}_{dx}, N)$ and $P_{safe}(Y, S, I)$ for $h, dx \to 0$ and N discretized version of I.

Theorem 5. *Let Y be the solution of a switching diffusion process $\mathcal{H}$ with initial condition $(x, q) \in \mathcal{S}$. Call Y^D the approximated DTMC, with $h, dx > 0$ time and space discretization parameters, and with initial condition $(x^D, q^D) = \xi((x, q))$. Then, given $I = [0, t] \subseteq \mathbb{R}_{\geq 0}$, it holds that*

$$|P_{safe}(Y^D, \mathcal{S}_{dx}, \lceil \frac{I}{h} \rceil) - P_{safe}(Y, \mathcal{S}, I)| \leq$$

$$\frac{I}{h} \cdot (\mathcal{K} dx + e^{-(2^n - 2^{\frac{n}{2}} + 1)}),$$

where $h = min\{\frac{2^{-n}}{2\sqrt{2}K^2 K_d}, 2^{-n}\}$ for $n \geq 3$, with $K \geq 12$ and $\bar{K}_d$ constant introduced in Section 4.

Proof. By triangular inequality we have

$$|P_{safe}(Y^D, \mathcal{S}_{dx}, \lceil \frac{I}{h} \rceil) - P_{safe}(Y, \mathcal{S}, I)| \leq$$

$$|P_{safe}(Y^D, \mathcal{S}_{dx}, \lceil \frac{I}{h} \rceil) - P_{safe}(\bar{Y}, \mathcal{S}, \lceil \frac{I}{h} \rceil)| +$$

$$|P_{safe}(\bar{Y}, \mathcal{S}, I) - P_{safe}(Y, \mathcal{I}, I)|.$$

The proof results from the application of Theorem 4 and Theorem 2. □

Algorithm 1 Probabilistic safety computation by finite DTMC abstraction

Require: $Y = (X, \alpha)$ solution of $\mathcal{H}$ with initial condition (x, q), safe set $\mathcal{S}$, finite time interval $I = [0, t]$, and parameters $dx, h = min\{\frac{2^{-n}}{2\sqrt{2}K^2 K_d}, 2^{-n}\}$;

1: Select the partition $\mathcal{G}_{dx} = \cup_{q \in \mathcal{Q}} \cup_{i \in m_q} \{q\} \times A_{i,q}$;
2: Select the set of representative points, leading to $\mathcal{S}_{dx}$;
3: Define the DTMC Y^D with state space $\mathcal{Z}_{dx} = \mathcal{S}_{dx} \cup \phi$, initial condition z_0 equals to 1 for the entry corresponding to $\xi((x, q))$ and 0 otherwise, and transition matrix P_{dx} such that $P_{dx}(i, j) = T_{dx}(z_i, z_j)$;
4: Compute $z^t = z_0 \cdot P_{dx}^{(\lceil \frac{I}{h} \rceil)}$;
5: Return $P_{safe}(Y, \mathcal{S}, I) = 1 - z^t(\phi)$ with error $\frac{I}{h} \cdot (\mathcal{K} dx + e^{-(2^n - 2^{\frac{n}{2}} + 1)})$.

In Algorithm 1 we present a numerical routine to compute safety properties over continuous time hybrid systems. The inputs of the Algorithm are $Y = (X, \alpha)$, solution of the continuous time hybrid process $\mathcal{H}$ with a given initial condition, a finite time interval I, the sampling time h, the grid parameter dx and the target set $\mathcal{S}$. (In the case study presented in the next section we consider parameters $h = 0.1$, $dx = 0.2$ and $I = [0, 2]$.) Theorem 5 allows to compute a bound on the error as a function of parameters dx and h. Moreover, such parameters can be selected to meet a required precision error. That is, given the maximum error that is tolerated, Theorem 5 returns possible h and dx that guarantee such error. In Lines 1, 2, 3 the algorithm computes the DTMC abstraction from Y and $\mathcal{S}$, as described in the previous section: P_{dx} is the transition probability matrix of the resulting DTMC [31], namely $P_{dx}(i, j)$ describes the probability of going from the discrete state z_i to the discrete state z_j at the next time step. Line 4 computes the transient evolution of the DTMC Y^D. This is done by multiplying the initial state z_0 for P_{dx} $\lceil \frac{I}{h} \rceil$ times, where $\lceil \frac{I}{h} \rceil$ are the number of discrete steps: $P_{safe}(Y, \mathcal{S}, I)$ is just the probability of not being in the sink state ϕ. A bound on the error is computed using Theorem 5.

7. CASE STUDY

We consider the continuous time switching diffusion process studied in [1]. The discrete state space is composed of two locations $\mathcal{Q} = \{on, off\}$, and the continuous process takes values in $\mathbb{R}^2$, so that the hybrid state space is $\mathcal{D} = \mathcal{Q} \times \mathbb{R}^2$. The drift is given by the following two matrices

$$F(on) = \begin{pmatrix} -0.6 & 0.3 \\ -0.6 & 0.15 \end{pmatrix}, \quad F(off) = \begin{pmatrix} -0.35 & 0 \\ 0.1 & -0.25 \end{pmatrix}.$$

The continuous dynamics are further affected by a $1-$dimensional Wiener process scaled by matrices

$$G(on) = \begin{pmatrix} 0.2 \\ 0.2 \end{pmatrix}, \quad G(off) = \begin{pmatrix} 0.3 \\ 0.3 \end{pmatrix}.$$

The Poisson measures have rates $\lambda_{on,off} = 0.41$ and $\lambda_{off,on} = 0.38$, respectively, both of which are independent of the continuous component of the process. We consider the Borel sigma algebra over $\mathcal{D}$ and a measurable set A. As the rates are independent of the continuous components, for $A \subseteq \mathbb{R}^2, q_i, q_j \in \{on, off\}$ with $q_i \neq q_j$, $x \in \mathbb{R}^2$ and $h \in \mathbb{R}_{\geq 0}$ small enough, we have the following transition kernels (see Theorem 1):

$$T^c(A, x, q_i, k) =$$
$$\int_A \mathcal{N}(\bar{x}|e^{F(q_i)\cdot h}x, \Gamma(i,h))d\bar{x} \cdot e^{-\lambda_{i,j}h} +$$
$$\int_A \int_0^h \mathcal{N}(\bar{x}|e^{F(q_i)\cdot s}e^{F(q_i)(h-s)}x, e^{F(q_i)\cdot s}\Gamma(i,s)(e^{F(q_i)\cdot s})^T +$$
$$\Gamma(j, h-s)) \cdot \Omega_{i,j,h}(s)ds \cdot (\lambda_{i,j}\, h\, e^{-\lambda_{i,j}h})d\bar{x}$$
$$T^d(q_j, x, q_i, k) = \begin{cases} e^{-\lambda_{i,j}h} & \text{if } q_i = q_j \\ \lambda_{i,j}h \cdot e^{-\lambda_{i,j}h} & \text{if } q_i \neq q_j \end{cases}$$

In order to choose h, we need to compute constants K_d, h_1, h_2. As the rate coefficients are independent of the continuous components we have $h_1 = 0$. It can be further derived that $h_2 \leq \max_{x\in\mathbb{R}^2}\{\frac{\partial t^c(x'|x,q_i)}{\partial x}\}$, where t^c is the density function of the kernel T^C. Further, $\bar{K}_d$ can be computed as

$$\bar{K}_d = \max_{q_i\in\{on,off\}, j\in\{1,2\}}\{\sqrt{\Gamma(i,h)(j,j)}\},$$

where $\Gamma(i,h)(j,j)$ is the component (j,j) of matrix $\Gamma(i,h)$. Note that K_d is also independent of the continuous component of the process.

In order to show the soundness of our method we have implemented in Matlab Algorithm 1, and compared the numerical implementation with empirical results obtained by simulations. We consider the following safe region

$$\mathcal{S} = \{x \in \mathbb{R}^2 \text{ s.t. for } i \in \{1,2\}, -0.2 \leq x_i \leq 1\},$$

where x_i is the $i-$th component of vector x. We select time interval $I = [0, 2]$. Theorem 5 guarantees that for given h and dx, respectively sampling time and space discretization parameters, the difference between the safety computed on the abstracted DTMC and the safety of the original switched diffusion is smaller than a computable bound. Notice that this bound can be conservative: for instance, for a choice of $h = 0.1$ and $dx = 0.2$, the theoretical error is not significative, whilst we have found a maximum empirical error of 0.11. In the following table we have compare the outcomes for $x_2 = 0.7$ and different values of x_1, obtained via an abstract DTMC made up of 5184 states, and the empirical values. The empirical values are computed with respect to 1000 stochastic simulations.

x_1	*Simulations*	$h = 0.1, dx = 0.2$
-0.2	0	0
-0.1	0.02	0.002
0.1	0.2	0.1
0.3	0.42	0.35
0.5	0.65	0.56
0.65	0.78	0.74
0.75	0.65	0.54

In order to increase the precision we can always decrease h and dx at the price of more computational effort (a larger DTMC abstraction). In order to guarantee a theoretical error ≤ 0.1, we could choose $h = 0.03$, which guarantees a theoretical time discretization error $\leq 4 \cdot 10^{-4}$. However, for such small h, t^c has very small variance, rendering h_2 big. As a consequence, in order to keep the error small, we would need $dx < 10^{-3}$) and the resulting DTMC would be composed of $> 10^6$ states. This dimensionality issue arises mostly because we are considering a uniform grid (dx constant). As a consequence, we use the same space resolution both for states with no probability mass and for states with probability mass, which are the great minority for h small. In fact, as described in the next Section, the usage of adaptive grid techniques [23] would allow to meet the given precision with a much smaller resulting DTMC, and it is targeted as future work.

8. CONCLUSIONS

We have presented a novel and formal approach to compute probabilistic reachability (and dually safety) for continuous time hybrid processes with no guards and no resets, and with continuous dynamics that can be described by linear stochastic differential equations. We have considered an approach based on space and time discretization of the original process, and derived uniform convergence of the algorithm, as well as error bounds that can be used to tune and control the approximation error.

The main contributions of the paper are the characterization of the kernels for the time discretizion of such processes and the error bound for the time discretization process. Finding formal bounds for the time discretization of stochastic hybrid processes has been an open problem, and, to our knowledge, only limited to results of weak convergence of the approximation. We have first presented the bound for uni-dimensional target sets, then have shown how to extend it to multidimensional polytopes (intersections of linear combinations of the components of the process). We have also shown how the bound grows linearly with the size of the target set.

For the space discretization we have considered an approach based on uniform gridding of the state space, insipred by the work in [2]. Although formally correct, this approach in combination with time discretization may result in large DTMC abstractions. In fact, as shown in a case study, the diameter of each grid location tends to grow as the sampling time of the time discretization process decreases. A much better solution would be to consider adaptive gridding techniques [23]. These would be extremely beneficial, as, when the sampling time is small, the distribution of the continuous kernel has very small variance. As a consequence, only a

very small set of states has not negligible probability mass. This is exactly the scenario where adaptive techniques perform better. As a future work, we plan to merge our time discretization approach with adaptive gridding techniques and to release a tool based on that. We also plan to include in the error bound the error derived from treating rate functions as piecewise constant in the computation of the kernel.

9. REFERENCES

[1] A. Abate. Probabilistic bisimulations of switching and resetting diffusions. In *49th IEEE Conference on Decision and Control (CDC)*, pages 5918–5923. IEEE, 2010.

[2] A. Abate, J.-P. Katoen, J. Lygeros, and M. Prandini. Approximate model checking of stochastic hybrid systems. *European Journal of Control*, 6:624–641, 2010.

[3] A. Abate, M. Prandini, J. Lygeros, and S. Sastry. Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. *Automatica*, 44(11):2724–2734, November 2008.

[4] R. J. Adler. *The geometry of random fields*, volume 62. Siam, 2010.

[5] R. J. Adler and J. E. Taylor. *Random fields and geometry*. Springer Science & Business Media, 2009.

[6] A. Arapostathis, V. Borkar, and M. Ghosh. *Ergodic Control of Diffusion Processes*. Cambridge University Press, 2012.

[7] L. Arnold. Stochastic differential equations. *New York*, 1974.

[8] A. Aziz, K. Sanwal, V. Singhal, and R. Brayton. Model-checking continuous-time Markov chains. *ACM Transactions on Computational Logic*, 1(1):162–170, 2000.

[9] C. M. Bishop. *Neural networks for pattern recognition*. Oxford university press, 1995.

[10] H. Blom and J. Lygeros (Eds.). *Stochastic Hybrid Systems: Theory and Safety Critical Applications*. Number 337 in Lecture Notes in Control and Information Sciences. Springer Verlag, Berlin Heidelberg, 2006.

[11] L. Bortolussi, L. Cardelli, M. Kwiatkowska, and L. Laurenti. Approximation of probabilistic reachability for chemical reaction networks using the linear noise approximation. In *Quantitative Evaluation of Systems*, pages 72–88. Springer, 2016.

[12] L. Bortolussi and R. Lanciani. Model checking markov population models by central limit approximation. In *International Conference on Quantitative Evaluation of Systems*, pages 123–138. Springer, 2013.

[13] L. Bujorianu. *Stochastic Reachability Analysis of Hybrid Systems*. Springer-Verlag, London, 2012.

[14] L. M. Bujorianu. *Stochastic reachability analysis of hybrid systems*. Springer Science & Business Media, 2012.

[15] P. Bulychev, A. David, K. Guldstrand Larsen, A. Legay, M. Mikučionis, and D. Bøgsted Poulsen. *Checking and Distributing Statistical Model Checking*, pages 449–463. Springer Berlin Heidelberg, 2012.

[16] L. Cardelli, M. Kwiatkowska, and L. Laurenti. Stochastic analysis of chemical reaction networks using linear noise approximation. In *International Conference on Computational Methods in Systems Biology*, pages 64–76. Springer, 2015.

[17] L. Cardelli, M. Kwiatkowska, and L. Laurenti. A stochastic hybrid approximation for chemical kinetics based on the linear noise approximation. In *International Conference on Computational Methods in Systems Biology*, pages 147–167. Springer, 2016.

[18] C. Cassandras and J. Lygeros (Eds.). *Stochastic Hybrid Systems*. Number 24 in Control Engineering. CRC Press, Boca Raton, 2006.

[19] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes: volume II: general theory and structure*. Springer Science & Business Media, 2007.

[20] M. Davis. *Markov Models and Optimization*. Chapman & Hall/CRC Press, London, 1993.

[21] R. M. Dudley. The sizes of compact subsets of hilbert space and continuity of gaussian processes. *Journal of Functional Analysis*, 1(3):290–330, 1967.

[22] P. M. Esfahani, D. Chatterjee, and J. Lygeros. The stochastic reach-avoid problem and set characterization for diffusions. *Automatica*, 70:43–56, 2016.

[23] S. Esmaeil Zadeh Soudjani and A. Abate. Adaptive and sequential gridding procedures for the abstraction and verification of stochastic processes. *SIAM Journal on Applied Dynamical Systems*, 12(2):921–956, 2013.

[24] M. Fränzle, E. M. Hahn, H. Hermanns, N. Wolovick, and L. Zhang. Measurability and safety verification for stochastic hybrid systems. In *HSCC*, pages 43–52, 2011.

[25] M. Fränzle, H. Hermanns, and T. Teige. Stochastic satisfiability modulo theory: A novel technique for the analysis of probabilistic hybrid systems. In M. Egerstedt and B. Misra, editors, *Hybrid Systems: Computation and Control*, volume 4981 of *Lecture Notes in Computer Science*, pages 172–186. Springer Verlag, Berlin Heidelberg, 2008.

[26] C. W. Gardiner et al. *Handbook of stochastic methods*, volume 3.

[27] M. K. Ghosh, A. Arapostathis, and S. I. Marcus. Optimal control of switching diffusions with application to flexible manufacturing systems. *SIAM Journal on Control and Optimization*, 31(5):1183–1204, 1993.

[28] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer Science & Business Media, 2011.

[29] K. Koutsoukos and D. Riley. Computational methods for reachability analysis of stochastic hybrid systems. In J. Hespanha and A. Tiwari, editors, *Hybrid Systems: Computation and Control*, volume 3927 of *Lecture Notes in Computer Science*, pages 377–391. Springer Verlag, Berlin Heidelberg, 2006.

[30] H. J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag, New York, 2001.

[31] M. Kwiatkowska, G. Norman, and D. Parker. Stochastic model checking. pages 220–270, 2007.

[32] P. Massart. *Concentration inequalities and model*

selection, volume 6. Springer, 2007.

[33] B. Øksendal. *Stochastic differential equations*. Springer, 2003.

[34] M. Prandini and J. Hu. Stochastic reachability: Theory and numerical approximation. In C. Cassandras and J. Lygeros, editors, *Stochastic hybrid systems*, Automation and Control Engineering Series 24, pages 107–138. Taylor & Francis Group/CRC Press, 2006.

[35] S. Särkkä et al. *Recursive Bayesian inference on stochastic differential equations*. Helsinki University of Technology, 2006.

[36] 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 ACM international conference on Hybrid Systems: computation and control*, pages 293–302, 2013.

[37] Y. Wang, N. Roohi, M. West, M. Viswanathan, and G. E. Dullerud. Statistical verification of dynamical systems using set oriented methods. In *Proceedings of the 18th International Conference on Hybrid Systems: Computation and Control*, pages 169–178. ACM, 2015.

[38] G. G. Yin and C. Zhu. *Hybrid switching diffusions: properties and applications*, volume 63. Springer Science & Business Media, 2009.

[39] M. Zamani and A. Abate. Symbolic models for randomly switched stochastic systems. *Systems Control Letters*, 69:38–46, 2014.

[40] 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*, 59(12):2825–2830, 2014.

[41] L. Zhang, Z. She, S. Ratschan, H. Hermanns, and E. M. Hahn. *Safety Verification for Probabilistic Hybrid Systems*, pages 196–211. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

APPENDIX

A. PROOFS

THEOREM 3. *Under Assumption 1, It holds that for any* $n \in \mathbb{N}$

$$P(\mathcal{A}^n) \geq P(\mathcal{A}^n \wedge \mathcal{B}^c) \geq P(\mathcal{A}^n) \cdot (1 - \frac{I}{h} \exp^{-(2^n - 2^{\frac{n}{2}}+1)}),$$

where $h = min\{\frac{2^{-n}}{2\sqrt{2K^2K_d}}, 2^{-n}\}$ *and* $K \geq 12$.

PROOF. For each $q \in \mathcal{Q}$, X_q is a bounded and uniform continuous GP in I, time interval of interest. Under some mild assumptions, this guarantees the existence of a sequence $\{\delta_n\}$ with $\delta_n \to 0$ such that $\phi(\delta_n) \leq 2^{-n}$ (see Theorem 2.1.3 of [5]), where

$$\phi(\delta_n) = E[\max_{q \in \mathcal{Q}}\{\sup_{s,s' \in I : |s'-s| \leq \delta_n} (X(s) - X(s'))\}].$$

Set $h = min\{\delta_n, 2^{-n}\}$and $\epsilon_n = 2^{-\frac{n}{2}}$. For a set of a sampling times $\Sigma = \{t_1, ..., t_{|\Sigma_n|}\}$, with step distance $h > 0$, call $S \subseteq \mathbb{R}^n$ the *safe set* and,

$$S^{\epsilon_n} = \{x \in S \ : \ |x - \partial S| \geq \epsilon_n\},$$

where ∂S is the boundary of S, and $|\cdot|$ stands for euclidean metric distance. Define the events

$$\mathcal{A}^n = \{\forall t_i \in \Sigma_n, . \ X(t_i) \in S^{\epsilon_n}\}$$

and

$$\mathcal{B} = \{\exists t \in I \ s.t. \ X(t) \notin \mathcal{S}\}.$$

Using the rules of probability we get

$$P(\mathcal{A}^n \wedge \mathcal{B}^c) = P(\mathcal{A}^n) \cdot (1 - P(\mathcal{B}|\mathcal{A}^n))$$

By definition of probability we have

$$0 \leq P(\mathcal{B}|\mathcal{A}^n) \leq P(\exists t_i \in \Sigma_n \ s.t. \sup_{t \in [t_i, t_i+h]} (X(t) - X(i)) > \epsilon_n) \leq$$

$$P(\exists t_i \in \Sigma_n \ s.t. \sup_{t \in [t_i, t_i+h]} (X(t) - X(i)) > \epsilon_n)$$

$$\leq \sum_{i=1}^{|\Sigma_n|} P(\sup_{t \in [t_i, t_i+h]} (X(t) - X(i)) > \epsilon_n)$$

In order to bound $P(\sup_{t \in [t_i, t_i+h]}(X(t) - X(i)) > \epsilon_n)$ we need to take into account that during $[t_i, t_i + h]$ the discrete state may hit a transition. However, there is no reset for the continuous components. As a consequence, it is enough to assume that during $[t_i, t_i + h]$ X always evolves accoridng to the "worst" behaving X_q. As a consequence, being X_q a GP, we can make use of the Borell's bound [4, 5]. Given an interval I and a centered and bounded Guassian process X_q with $\sigma_I = \sup_{t \in I}(\sigma(t))$, supremum of the standard deviation of the process, the Borrell's bound guarantees that

$$Prob(\sup_{t \in I} X_q(t) > u) \leq \exp^{-(u - E[sup_{t \in I} X_q(t)])^2/(\sigma_I^2)}$$

Applying this result to our case for intervals of the type $[t_i, t_i + h]$, and for $u = \epsilon_n = 2^{-\frac{n}{2}}$, for $n \geq 3$ we have

$$\sum_{i}^{|\Sigma_n|} P(\sup_{t \in [t_i, t_i+h]} (X(t) - X(t_i)) > \epsilon_n) \leq$$

$$|\Sigma_n| \exp^{-\frac{(2^{-\frac{n}{2}} - 2^{-n})^2}{2^{-2n}}} \leq$$

$$|\Sigma_n| \exp^{-(2^n - 2^{\frac{n}{2}}+1)}$$

At this point, the last, and non trivial, step in order to derive our convergence results and relative error bounds is to show that

$$\lim_{n \to \infty} |\Sigma_n| \exp^{-(2^n - 2^{\frac{n}{2}}+1)} = 0.$$

In fact, as

$$0 \leq P(\mathcal{B}|A^n) \leq |\Sigma_n| \exp^{-(2^n - 2^{\frac{n}{2}}+1)},$$

this would guarantee that

$$P_{safe}(X, S, I) = \lim_{n \to \infty} P(\mathcal{A}^n \wedge \mathcal{B}^c) = \lim_{n \to \infty} P(\mathcal{A}^n).$$

In order to do that, it is sufficient to show that $h = \frac{2^{-n}}{C}$, for some constant C. In fact, this implies $|\Sigma_n| = \frac{I \cdot C}{2^{-n}}$.

Recall that we chose h such that for all $t_i \in \Sigma_n$,

$$E[\sup_{t \in [t_i, t_i+h]} (X(t) - X(t_i)) \leq 2^{-n}.$$

As a consequence, it is enough to take h as the greatest interval smaller than 2^{-n} such that this condition is verified. For $t_i \in \Sigma_n$ call $\bar{X}_i = X(t) - X(t_i)$. We can now make use of the *Dudley's integral (or entropy integral)* [4], which guarantees that for $t_i \in \Sigma_n$,

$$E[\sup_{t\in[t_i,t_i+h]}(\bar{X}_i)] \leq$$

$$K\int_0^{\frac{diam([t_i,t_i+h])}{2}} \sqrt{\ln(N([t_i,t_i+h],d,\epsilon))}d\epsilon,$$

where K is a constant and d is a pseudo-metric defined as

$$d(t,t+dt) = \sqrt{E[(X(t+dt)-X(t))^2]}.$$

$N([t_i,t_i+h],d,\epsilon)$ represents the smallest number of balls of radius ϵ, which covers $[t_i,t_i+h]$, under metric d. $diam([t_i,t_i+h])$ is defined as

$$diam([t_i,t_i+h]) = sup_{s',s\in[t_i,t_i+h]}d(s',s)$$

and with our assumptions, it is possible to show that there exists a constant K_d such that

$$d(t,t+h) \leq K_d \cdot h$$

Moreover, for $\bar{T}_i = [t_i,t_i+h] \subseteq \mathbb{R}_{\geq 0}$ we have

$$N(\bar{T}_i,d,\epsilon) \leq \frac{K_d h}{2\epsilon} + 1,$$

This can be easily understood thinking at the geometry of the problem. As a consequence, we have

$$E[\sup_{t\in\bar{T}_i}(\bar{X}_i(t))] \leq K\int_0^{\sqrt{2}\cdot 2^{-n-1}} \sqrt{ln(\frac{K_d h}{2\epsilon}+1))}\ d\epsilon$$

Now, our property is satisfied if we chose h such that

$$K\int_0^{\sqrt{2}\cdot 2^{-n}} \sqrt{\ln(\frac{K_d h}{2\epsilon}+1)}\ d\epsilon \leq 2^{-n}.$$

The integral inequality we need to solve cannot be solved analytically. However, as $K_d h > 0$, we can write

$$K\int_0^{\sqrt{2}\cdot 2^{-n}} \sqrt{\ln(\frac{K_d h}{2\epsilon}+1)}\ d\epsilon \leq K\int_0^{\sqrt{2}\cdot 2^{-n}} \sqrt{\frac{K_d h}{2\epsilon}}\ d\epsilon =$$

$$K\sqrt{\frac{K_d h}{2}}\int_0^{\sqrt{2}\cdot 2^{-n}} \sqrt{\frac{1}{\epsilon}}\ d\epsilon = K\sqrt{\frac{K_d h}{2}}2\sqrt{\sqrt{2}2^{-n}}.$$

Asking for this quantity to be smaller than 2^{-n}, we obtain the following bound for the sampling time h :

$$h \leq min\{\frac{2^{-n}}{2\sqrt{2}K^2K_d}, 2^{-n}\}.$$

It is also possible to show that $K \geq 12$ []. □

THEOREM 2. *Let $\mathcal{H} = (\mathcal{Q}, K, F, G, W, \Lambda)$ be a hybrid process and $Y = (X,\alpha)$ its solution. Let $h > 0$ be a sampling time and $\bar{Y} = (\bar{X},\bar{\alpha})$ the resulting DTMP. Call $\mathcal{N}(\bar{x}|E,C)$ the normal distribution with mean E and covariance C. Introduce terms*

$$\Gamma(i,t) = \int_0^t (e^{F(q_i)(t-m)})G(q_i)G(q_i)^T(e^{F(q_i)(t-m)})^T dm,$$

$$\Omega_{\lambda_i,\lambda_j,t}(s) = (\lambda_j-\lambda_i)\frac{e^{(\lambda_j s-\lambda_j t-\lambda_i s)}}{e^{(-\lambda_i t)}-e^{(-\lambda_j t)}},$$

and for $x \in \mathbb{R}^n$ define $\lambda_i(x) = \sum_{j\neq i}\lambda_{i,j}(x)$. Then, given $(q,x) \in \mathcal{D}$ and (q',A) a measurable set, it holds that

$$T^c(A,x,q_i) =$$

$$\int_A \mathcal{N}(\bar{x}|e^{F(q_i)\cdot h}x, \Gamma(i,h))d\bar{x}\cdot e^{-\lambda_i h} +$$

$$\sum_{q_j\neq q_i}\int_A\int_0^h \mathcal{N}(\bar{x}|e^{F(q_i)\cdot s}e^{F(q_i)(h-s)}x, e^{F(q_i)\cdot s}\Gamma(i,s)(e^{F(q_i)\cdot s})^T + \Gamma(j,h-s$$

$$\Omega_{\lambda_i,\lambda_j,h}(s)ds\ \cdot\ \frac{\lambda_{ij}}{\lambda_i}\cdot\lambda_i h\cdot e^{-\lambda_i h}\ +\ \epsilon)d\bar{x},$$

and

$$T^d(q_j,x,q_i) = \begin{cases} e^{-\lambda_i h}+\epsilon & \text{if } q_i = q_j \\ \lambda_i h\cdot e^{-\lambda_i h}\cdot\frac{\lambda_{ij}}{\lambda_i}+\epsilon & \text{if } q_i \neq q_j \end{cases}$$

where $\epsilon \leq 1 - e^{-\lambda_i h} - \lambda_i h\cdot e^{-\lambda_i h}$.

PROOF. We consider the continuous kernel, as the discrete kernel can be derived similarly.

$$T^c(A,x,q_i) = \int_A t^c(\bar{x}|x,q_i)d\bar{x}$$

where $t^c(\bar{x}|x,q_i)$ is the density function of $X(h)$, continuous component of Y, assuming $X(0) = x$ and $\alpha(0) = \lambda_i$. Note that the kernel is time homogeneous. By marginalizing with respect to the number of time that α jumps during $[0,h]$, we have

$$T^c(A,x,q_i) = \int_A t^c(\bar{x}|x,q_i)d\bar{x} =$$

$$\int_A (t^c(\bar{x}|x,q_i,0\ \alpha\text{ firings})\cdot$$

$$Prob(\alpha \text{ does not fire in } [0,h]|x,q_i) +$$

$$t^c(\bar{x}|x,q_i,1\ \alpha\text{ firings }|\text{x,q}\cdot$$

$$Prob(\alpha \text{ fires once in } [0,h]|x,q_i))d\bar{x} + \epsilon$$

where $\epsilon \leq Prob(\alpha$ fires more than once in $[0,h]|x,q_i) = 1 - \sum_{i\in\{0,1\}}Prob(\alpha$ jumps i times in$[0,h]|x,q_i)$.

$t^c(\bar{x}|x,q_i 0\ \alpha$ firings) is the normal distribution derived from solving the linear SDE corresponding to mode q_i from initial condition x for the interval $[0,h]$ because of the Markovianity of SDE.

Because of the properties of Poisson processes we have that $Prob(\alpha$ does not fire in $[0,h]|x,q_i) = e^{-\lambda_i h}$, and $Prob(\alpha$ fires once in $[0,h]|x,q_i) = \lambda_i h e^{-\lambda_i h}$.

Marginalizing $t^c(\bar{x}|x,q_i,1\ \alpha$ firings) with respect to the discrete location where we jump and respect the istant when α jumps we get:

$$\sum_{q_j\neq q_i}\int_0^h t^c(\bar{x}|x,q_i,0\ \alpha\text{ firings},\alpha\text{ jumps at time } s \text{ in } q_j)\cdot$$

$$f(s|x,q_i,\alpha\text{ jumps once in }[0,h])Prob(\alpha\text{ jumps in } q_j)ds$$

The first term can be shown to be a linear Gaussian model. This class of models has been extensively studied in literature [9]. More specifically, it has a Gaussian distribution whose variance and expectation can be derived from Lemma 3. As a conseqeunce, we have

$$t^c(\bar{x}|x,q_i,0\ \alpha\text{ firings},\alpha\text{ jumps at time } s \text{ in } q_j) =$$

$\mathcal{N}(\bar{x}|e^{F(q_i)\cdot s}e^{F(q_i)(h-s)}x, e^{F(q_i)\cdot s}\Gamma(i,s)(e^{F(q_i)\cdot s})^T+\Gamma(j,h-s)$

.

$Prob(\alpha$ jumps in $q_j)$ is the probability of jumping in q_j at the next jump. This is $\frac{\lambda_{ij}}{\lambda_i}$.

$f(s|x,q,\alpha$jumps once and in $q_j)$ is the density function of the jumping time conditioned on the fact that we jump in $[0,h]$. This can be derived from properties of Poisson processes as $\Omega_{\lambda_i,\lambda_j,h}(s) = (\lambda_j - \lambda_i)\frac{e^{(\lambda_j s-\lambda_j t-\lambda_i s)}}{e^{(-\lambda_i t)}-e^{(-\lambda_j t)}}$.

□