Approximate Abstractions of Stochastic Hybrid Systems

Alessandro Abate,[1] Alessandro D'Innocenzo,[2]
and Maria D. Di Benedetto[2]

***Abstract*—We present a constructive procedure for obtaining a finite approximate abstraction of a discrete-time stochastic hybrid system. The procedure consists of a partition of the state space of the system and depends on a controllable parameter. Given proper continuity assumptions on the model, the approximation errors introduced by the abstraction procedure are explicitly computed and it is shown that they can be tuned through the parameter of the partition. The abstraction is interpreted as a Markov set-Chain. We show that the enforcement of certain ergodic properties on the stochastic hybrid model implies the existence of a finite abstraction with finite error in time over the concrete model, and allows introducing a finite-time algorithm that computes the abstraction.**

***Index Terms*—Stochastic Hybrid Systems, Markov Chains.**

I. Introduction and related work

The study of complex, heterogeneous, and probabilistic models such as Stochastic Hybrid Systems poses challenges, both analytically (e.g., steady-state analysis, synthesis of optimal controllers [1]) and computationally (e.g., reachability and safety analysis [2]). An approach that is used to cope with this issue is that of approximate abstraction: a system with smaller (possibly finite) state space is obtained, which is approximately equivalent to the original system [3]. Unlike the exact concept of equivalence, which is usually defined by the notions of language equivalence and bisimulation [4] and as such is quite restrictive since it requires a perfect correspondence between the trajectories of the original system and those of its abstraction, approximate notions of system equivalence [3], [5], [6] are endowed with a proper metric quantifying the distance between the trajectories of the original system and those of the approximate abstraction. The research on abstraction techniques for dynamical systems has two general goals. The first objective is that of proving the existence of a finite abstraction [4], while the second goal is that of developing finite time and tunable abstraction algorithms [5].

Abstraction techniques have been adapted to probabilistic models, for instance to discrete-space, continuous-time models [7]. Notions of bisimulation, which naturally lead to abstracted models, have been developed for classical discrete Markov processes in [8], [9], [10], and for jump linear stochastic systems in [6].

Weak approximations of continuous-time probabilistic models as locally-consistent Markov Chains have been introduced by [11], and applied on hybrid models in [12], [13], whereas approximations of discrete-time Markov models can be alternatively studied via renewal theory, as in [14]. Notice that both approaches are different than the present work in that they derive no explicit approximation bound. Related to this work, recently [15] has proposed explicit error bounds on a time and space discretization of a Markov process with certain ergodic properties.

In this work we provide new results on approximate abstractions of discrete-time stochastic hybrid systems (DTSHS). DTSHS encompass a number of other classes of stochastic hybrid models in the literature [16]. The main contributions of this paper can be summarized as follows:

- We introduce a procedure to construct an approximate abstraction of a DTSHS, which involves the partition of the state space and the approximation of the transition laws of the DTSHS over the partition sets, with an explicit computation of the error.
- The abstraction is interpreted as a Markov set-Chain (MSC) [17]. MSC are useful as they comprise both stochastic and non-deterministic parts. In the present context, the stochastic quantities approximate the probability law of the original system, while the non deterministic behaviors are used to quantitatively take into account the error introduced by the abstraction procedure.
- By posing some continuity assumptions on the DTSHS model, we derive an explicit and tunable bound on the error between the probability distribution of the abstracted model (the MSC) and that of the original model (the DTSHS, considered over the partition sets), for each time instant (and, in particular, in steady-state). The tunable bound allows for refinements of the abstraction procedure.
- Based on the derived error bound in time and given proper assumptions on the ergodicity of the original DTSHS, the contribution proposes a finite time algorithm to construct an approximate abstraction with arbitrary positive precision. The precision is related to the distance between the steady-state distributions.

Using the concept of distance between the probability laws of the original system and those of the abstraction, the proposed procedure and the associated time-dependent bounds represent a step towards a formalization of the notion of *approximate stochastic bisimulation* for general probabilistic models [8], [9], [10]. Looking forward, our work addresses the following general verification purpose: given a DTSHS, verify the probabilistic properties of the original system on a finite-dimensional MSC abstraction with arbitrary precision.

Section II introduces the class of stochastic hybrid models under study, namely the DTSHS. Section III recalls some results on MSC to be utilized in the rest of the work. Section IV introduces the abstraction procedure as a partitioning of the state space. Section V delves into the derivation of the errors associated with the abstraction, which turns the original DTSHS into a MSC. We show that, under proper ergodic assumptions on the DTSHS, it is possible to construct an approximate abstraction with arbitrary precision over time. We also propose an algorithm for building in finite time an abstraction endowed with the property that its steady state is arbitrarily close to that of the original system. Section VI concludes the article. The appendix contains the proofs of the statements.

II. Discrete-Time Stochastic Hybrid Systems

Definition 1 (DTSHS): A discrete-time stochastic hybrid system is a tuple $\mathcal{H} = (\mathcal{S}, T_q, T_t, T_r)$, where

- $\mathcal{S} := \cup_{q_i \in \mathcal{Q}} \{q_i\} \times \mathcal{D}_i$, is the hybrid state space, which consists of a set of discrete modes $\mathcal{Q} := \{q_1, q_2, \dots, q_m\}$, for some finite $m \in \mathbb{N}$, and of a set of continuous domains, one for each mode $q_i \in \mathcal{Q}$, each of which is defined to be a compact set $\mathcal{D}_i \subset \mathbb{R}^{n(q_i)}$. The function $n : \mathcal{Q} \to \mathbb{N}$ assigns to each $q_i \in \mathcal{Q}$ the finite dimension of the continuous state space $\mathbb{R}^{n(q_i)}$;
- $T_q : \mathcal{Q} \times \mathcal{S} \to [0,1]$ is a discrete stochastic kernel (the "discrete transition kernel") on $\mathcal{Q}$, given $\mathcal{S}$, which assigns to each $s = (q,x) \in \mathcal{S}$ a discrete probability distribution $T_q(\cdot|s)$ over $\mathcal{Q}$;
- $T_t : \mathcal{B}(\mathcal{D}_{(\cdot)}) \times \mathcal{S} \to [0,1]$ is a Borel-measurable stochastic kernel (the "continuous transition kernel") on $\mathcal{D}_{(\cdot)}$, given $\mathcal{S}$, which assigns to each $s = (q,x) \in \mathcal{S}$ a probability measure $T_t(\cdot|s)$ on the Borel space $(\mathcal{D}_q, \mathcal{B}(\mathcal{D}_q))$;

[1] Department of Aeronautics and Astronautics. Stanford University, CA, USA, and Delft Center for Systems and Control, TU Delft, The Netherlands - email: a.abate@tudelft.nl

[2] Department of Electrical Engineering and Computer Science, Center of Excellence DEWS. University of L'Aquila, Italy - email: adinnoce@ing.univaq.it, dibenede@parades.rm.cnr.it

The first two authors equally contributed to this work, supported by the European Commission MoVeS project FP7-ICT-2009-257005, and Marie Curie grant MANTRAS 249295, and by the NWO VENI grant 016.103.020.

- $T_r : \mathcal{B}(\mathcal{D}_{(\cdot)}) \times \mathcal{S} \times \mathcal{Q} \to [0,1]$ is a Borel-measurable stochastic kernel (the "reset kernel") on $\mathcal{D}_{(\cdot)}$, given $\mathcal{S} \times \mathcal{Q}$, that assigns to each $s = (q,x) \in \mathcal{S}$, and $q' \in \mathcal{Q}, q' \neq q$, a probability measure $T_r(\cdot|s,q')$ on the Borel space $(\mathcal{D}_{(q')}, \mathcal{B}(\mathcal{D}_{(q')}))$. □

The system initialization at time $k = 0$ is specified by some probability measure $\pi_0 : \mathcal{B}(\mathcal{S}) \to [0,1]$ on the Borel space $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$. Here and in Definition 1, $\mathcal{B}(\mathcal{S})$ is the σ-field generated by the subsets of $\mathcal{S}$ of the form $\cup_{q \in \mathcal{Q}} \{q\} \times B_q$, with B_q denoting a Borel set in $\mathcal{D}_q$. To understand the semantics of the model, let us state the definition of a solution of the DTSHS $\mathcal{H}$.

Definition 2 (Execution of a DTSHS): Given a DTSHS $\mathcal{H}$ and an initial distribution π_0, an execution of $\mathcal{H}$ is a stochastic process[1] $\{\mathbf{s}(k) = (\mathbf{q}(k), \mathbf{x}(k)), \forall k = 0, \dots, N+1\}$, with values in $\mathcal{S}$, generated by the following algorithm:

extract a value $s_0 = (q_0, x_0)$ *on* $\mathcal{S}$ *for* $\mathbf{s}(0)$, *according to* π_0;

for $k = 0$ *to* N,

extract a value $q_{k+1} \in \mathcal{Q}$ *for* $\mathbf{q}(k+1)$, *according to* $T_q(\cdot|s_k)$;

if $q_{k+1} \neq q_k \in \mathcal{Q}$,

then *extract* $x_{k+1} \in \mathcal{D}_{q_{k+1}}$ *for* $\mathbf{x}(k+1)$ *from* $T_r(\cdot|s_k, q_{k+1})$;

else *extract* $x_{k+1} \in \mathcal{D}_{q_k}$ *for* $\mathbf{x}(k+1)$ *from* $T_t(\cdot|s_k)$;

end *(if);*

end *(for).*

□

It is understood that, when $N = \infty$, the algorithm does not terminate. For the sake of conciseness, we make use of the following shortened notation for the probability kernels:

$$T(ds|(q,x)) = \begin{cases} T_t(dx|(q,x)) T_q(q'|(q,x)), & \text{if } q = q', \\ T_r(dx|(q,x),q') T_q(q'|(q,x)), & \text{if } q \neq q', \end{cases} \quad (1)$$

where $q, q' \in \mathcal{Q}$, and $s = (q,x) \in \mathcal{S}$. Notice that, for semantical consistency, we avoid the definition of the reset kernel T_r on $q' = q$. The execution $\mathbf{s}(k)$ of $\mathcal{H}$ associated with $s_0 \in \mathcal{S}$ is a stochastic process with probability measure $p^k_{s_0}$, at time $k \in [1, N+1]$, which is uniquely defined by the transition kernel T and the initial condition s_0 [18, Proposition 7.45]. It is easy to show that the execution $\mathbf{s}(k)$ of $\mathcal{H}$ is a Markov process with one-step transition kernel T [16]. We refer the reader to the details contained in [16] for further insights on the model, the complete understanding of its properties, and the comparison with other models in the literature (for instance, the *random evolution process* [19], [20]).

While continuity of the probability kernels is not strictly required in Definitions 1 and 2, we now raise the following assumption and suppose that it holds true throughout this work. The assumption will be useful to prove certain bounds on the transition probability of the DTSHS $\mathcal{H}$.

Assumption 1 (Continuity of the Stochastic Kernels): Suppose that the continuous stochastic kernels T_t, T_r of the DTSHS $\mathcal{H}$ admit densities t, r. Assume that the following Lipschitz properties hold for T_q, t, r:

1) $|T_q(\bar{q}|s) - T_q(\bar{q}|s')| \leq L_q \|x - x'\|$, for all $s = (q,x), s' = (q,x') \in \mathcal{D}_q$, and $q, \bar{q} \in \mathcal{Q}$;
2) $|t(\bar{x}|s) - t(\bar{x}|s')| \leq L_t \|x - x'\|$, for all $s = (q,x), s' = (q,x') \in \mathcal{D}_q$, and $(q,\bar{x}) \in \mathcal{D}_q, q \in \mathcal{Q}$;
3) $|r(\bar{x}|s,\bar{q}) - r(\bar{x}|s',\bar{q})| \leq L_r \|x - x'\|$, for all $s = (q,x), s' = (q,x') \in \mathcal{D}_q$, $(\bar{q},\bar{x}) \in \mathcal{D}_{\bar{q}}$, and $q, \bar{q} \in \mathcal{Q}, \bar{q} \neq q$,

where L_q, L_t, L_r are finite positive constants, and $\| \cdot \|$ is the Euclidean norm on $\mathcal{D}_q, q \in \mathcal{Q}$. □

III. Markov set-Chains

We recall the concept of Markov set-Chain (MSC), which in this paper is later used to prove properties of the abstraction. The results are from [17] and references therein. The framework is also related to that of interval Markov chains [21].

Definition 3 (Transition Set): [17, Definition 2.5] Let $P, Q \in \mathbb{R}^{n \times n}$ be nonnegative matrices (not necessarily stochastic) with $P \leq Q$, where $\leq$ holds element-wise. A transition set is

$$[P,Q] = \{A = (a_{ij}) \in \mathbb{R}^{n \times n} : a_{ij} \geq 0, \sum_{j=1}^{n} a_{ij} = 1, P \leq A \leq Q\}.$$

□

In this article, we assume that the set $[P,Q] \neq \varnothing$. Whenever the "bounding matrices" of the MSC will be clear from the context, we will use the more compact notation $[\Pi]$ for the transition set characterized by the interval matrix Π. A MSC is defined as a discrete-time Markov Chain, where the transition probabilities vary non-deterministically within a compact transition set $[\Pi]$ at each time step. More formally,

Definition 4 (Markov set-Chain): [17, Definition 2.5] Let $[\Pi]$ be a transition set, i.e. a compact set of $n \times n$ stochastic matrices. Consider the set of all non-homogeneous Markov Chains having their transition matrices in $[\Pi]$. We call the sequence $[\Pi], [\Pi]^2, \dots$ a Markov set-Chain, where $[\Pi]^k$ is defined by induction as the set of all possible products $A_1 A_2 \cdots A_k$, such that $\forall i = 1, \dots, k, A_i \in [\Pi]$. □

Let $[\pi_0]$ be a compact set made up of stochastic vectors of dimension $1 \times n$, introduced as in Definition 3. We call $[\pi_0]$ the initial distribution set. The compact set $[\pi_k] = [\pi_0][\Pi]^k$ is the k-th distribution set and the structure $[\pi_0], [\pi_0][\Pi], [\pi_0][\Pi]^2, \dots$ is the Markov set-Chain with initial distribution set $[\pi_0]$.

Definition 5 (Coefficient of Ergodicity of a Stochastic Matrix): [17, Definition 1.2] For a stochastic matrix A, its coefficient of ergodicity is defined as follows:

$$\mathcal{T}(A) = \frac{1}{2} \max_{i,j} ||a_i - a_j||,$$

where a_i, a_j are the i–th, j–th rows of A, and $\| \cdot \|$ is the standard 1-norm over row vectors: $\|x\| = \sum_k |x_k|$. □

It can be shown that the condition $\mathcal{T}(A) < 1$, along with the condition of irreducibility of the chain, implies the existence of a unique limiting and invariant distribution for the associated Markov Chain [17]. The previous definition can be directly extended to Markov set-Chains:

Definition 6 (Coefficient of Ergodicity of a Transition Set): [17, Definition 3.1] For any transition set $[\Pi]$, its coefficient of ergodicity is defined over the stochastic matrices that define $[\Pi]$ as follows:

$$\mathcal{T}([\Pi]) = \max_{A \in [\Pi]} \mathcal{T}(A).$$

□

Since $\mathcal{T}(\cdot)$ is a continuous function and $[\Pi]$ a compact set, the argument of the max exists. Similar to the simpler case of Markov Chains, the quantity $\mathcal{T}([\Pi]) \in [0,1]$ provides a measure of the "contractive" nature of the Markov set-Chain: the smaller $\mathcal{T}([\Pi])$, the "more contractive" the MSC. This quality is related to the regularity properties of the stochastic matrices that build up the MSC and can be exploited when studying its asymptotics [17]. The exact value of $\mathcal{T}([\Pi])$ can be approximated, given any $A \in [\Pi]$, as in Proposition 1, on page 5. Let us define the diameter of a compact set (referred to either matrices or vectors) as $\Delta([\Pi]) = \max_{A,A' \in [\Pi]} ||A - A'||$. Proposition 2 provides an upper bound for the diameter of the transition set $[\Pi]^k, k > 0$.

[1] In this work bold symbols denote (stochastic) processes, while a regular typeset is used for sample values or points on the state space.

The derived bounds are not necessarily tight, however they are sufficient for the objectives of the study (finiteness of bounds, proof of convergence), and as such they will be used in the following. Tighter results can be obtained with more sophistication: in particular, the notion of coefficient of ergodicity can be relaxed by looking at the concept of *scrambling coefficient* [17, Definition 3.3], which is the minimum positive integer r such that $\mathcal{T}(A) < 1, A = \prod_{i=1}^{r} A_i, \forall A_i \in [\Pi]$. The focus of the article is not that of seeking a set of "optimal" bounds for the abstraction procedure proposed in this work, but rather that of showing that finite bounds exist and can be properly tuned. We thus delegate the search for improved bounds to future work, as mentioned in Section VI.

IV. Abstraction Procedure: State Space Partitioning

The abstraction proposed in this work involves a partitioning procedure described in this Section. The partitioning algorithm is inspired by the work in [22], where a similar procedure is introduced on a simplified dynamical model in order to solve a class of optimal control problems via dynamic programming. Based on this procedure and its related approximation error, we introduce a Markov set-Chain [17] as the abstraction of the original DTSHS (see Section V-A). It is desirable that, as the approximation error of the introduced abstraction goes to zero, the dynamical properties of the original system hold if they are true on the abstraction [3] (see Section V-B for the computation of error bounds and for a study of asymptotic properties). In order to achieve this, it is necessary to enforce some continuity on the dynamics of the DTSHS: let us then uphold Assumption 1.

We introduce a finite partition of the hybrid state space $\mathcal{S} = \cup_{q\in\mathcal{Q}}\{q\}\times\mathcal{D}_q$ of $\mathcal{H}$. Let us recall that each domain $\mathcal{D}_q \subset \mathbb{R}^{n(q)}, q \in \mathcal{Q}$, is required to be compact. A partition $\{\mathcal{D}_q\}_\delta = \{D_q^i, i = 1,\ldots,m_q^\delta\}$ (which depends on a parameter δ, to be defined shortly) of the domain $\mathcal{D}_q$ is a covering of $\mathcal{D}_q$ made up of m_q^δ non-overlapping convex sets such that $\mathcal{D}_q \subseteq \bigcup_{i=1}^{m_q^\delta} D_q^i$, and $D_q^i \cap D_q^j = \varnothing, i \neq j$. The sets D_q^i can have any convex shape (we will be simply interested on a parameter δ that characterizes them), which makes the procedure general and flexible. In particular, the covering can be selected to exactly coincide with $\mathcal{D}_q$. In this work, for the sake of simplicity, the partition $\{\mathcal{D}_q\}_\delta$ of domain $\mathcal{D}_q$ for mode $q \in \mathcal{Q}$ is characterized as follows: consider a uniform square grid of width $\delta/\sqrt{n(q)}$, defined on $\mathbb{R}^{n(q)}$ and centered around points in the set

$$\left\{\left(\frac{m_1\delta}{\sqrt{n(q)}},\frac{m_2\delta}{\sqrt{n(q)}},\ldots,\frac{m_{n(q)}\delta}{\sqrt{n(q)}}\right):(m_1,m_2,\ldots,m_{n(q)})\in\mathbb{Z}^{n(q)}\right\}. \tag{2}$$

Each spatial set $D_q^i \in \{\mathcal{D}_q\}_\delta$ is introduced as a hyper-cube centered around a point in (2) for a particular choice of $(m_1, m_2, \ldots, m_{n(q)}) \in \mathbb{Z}^{n(q)}$, and defined as $D_q^i \doteq \Big\{x \in \mathcal{D}_q : \left((m_1-\frac{1}{2})\frac{\delta}{\sqrt{n(q)}},(m_2-\frac{1}{2})\frac{\delta}{\sqrt{n(q)}},\ldots,(m_{n(q)}-\frac{1}{2})\frac{\delta}{\sqrt{n(q)}}\right) \leq x < \left((m_1+\frac{1}{2})\frac{\delta}{\sqrt{n(q)}},(m_2+\frac{1}{2})\frac{\delta}{\sqrt{n(q)}},\ldots,(m_{n(q)}+\frac{1}{2})\frac{\delta}{\sqrt{n(q)}}\right)\Big\}$. The partition $\{\mathcal{D}_q\}_\delta$ is formally defined as the smallest collection of partition sets D_q^i that contains $\mathcal{D}_q$. If the cardinality of $\{\mathcal{D}_q\}_\delta$ is m_q^δ, then $\mathcal{D}_q \subseteq \bigcup_{i=1}^{m_q^\delta} D_q^i$. In this case each partition set D_q^i has diameter δ, and in general the parameter δ is defined to be the diameter of the partition sets (i.e., the maximum distance between any two points in the same equivalence class), and influences its cardinality m_q^δ (see equation (3) and following lines). The set $\{\mathcal{S}\}_\delta = \{\{q\}\times D_q^i, i = 1,\ldots,m_q^\delta, q \in \mathcal{Q}\}$ is then a partition of the whole hybrid state space $\mathcal{S}$ with cardinality $\sum_{q\in\mathcal{Q}} m_q^\delta$, and is such that $\mathcal{S} \subseteq \bigcup_{q\in\mathcal{Q}}\{q\}\times\left(\cup_{i=1}^{m_q^\delta} D_q^i\right)$. For any point $s = (q,x) \in \mathcal{S}$ there exists an element of the partition $\{q\}\times D_q^i \in \{\mathcal{S}\}_\delta$ such that $x \in D_q^i$. Let us introduce a function $\langle\cdot\rangle : \mathcal{S} \to \{\mathcal{S}\}_\delta$, which associates to any hybrid point $s = (q,x) \in \mathcal{S}$ its partition set $\langle s\rangle = \{q\}\times D_q^i \in \{\mathcal{S}\}_\delta$ for a specific $i \in \{1,2,\ldots,m_q^\delta\}$. Furthermore, given a hybrid point $s = (q,x)$ and its partition set $\langle s\rangle$, let us select any point $\bar{s} = (q,\bar{x}) \in \langle s\rangle$ to be the *representative point* of the partition set $\langle s\rangle$. For instance, we may select its centroid as in (2): $\bar{x} = \left(m_1\frac{\delta}{\sqrt{n(q)}}, m_2\frac{\delta}{\sqrt{n(q)}},\ldots,m_{n(q)}\frac{\delta}{\sqrt{n(q)}}\right)$, for a particular choice of $(m_1,m_2,\ldots,m_{n(q)}) \in \mathbb{Z}^{n(q)}$. The following expression relates a point $s = (q,x) \in \mathcal{S}$ with its representative point $\bar{s} = (q,\bar{x})$, within their equivalence class $\langle s\rangle$:

$$\forall s \in \mathcal{S},\ \exists\langle s\rangle \in \{\mathcal{S}\}_\delta, \bar{s} \in \mathcal{S} : (s,\bar{s} \in \langle s\rangle) \wedge (\|x-\bar{x}\| \leq \delta), \tag{3}$$

where $\|\cdot\|$ is the Euclidean norm. As we just did above for the partition sets, given any $q \in \mathcal{Q}$ and any compact subset $W \subseteq \mathbb{R}^{n(q)}$, we define its diameter to be the largest distance between any two points in W: $\lambda_W = \sup\{\|x-y\|, x,y \in W\}$. Consider now mode $q \in \mathcal{Q}$, the parameter $\frac{\delta}{\sqrt{n(q)}}$ that characterizes the partition introduced for that domain, and the (finite) diameter $\lambda_{\mathcal{D}_q}$ of the associated domain $\mathcal{D}_q$. The domain is contained in a cube with side equal to its diameter. Select the integer quantity $\left\lceil\frac{\lambda_{\mathcal{D}_q}}{\delta/\sqrt{n(q)}}\right\rceil$. The cardinality of the partition $\{\mathcal{D}_q\}_\delta$ can be upper bounded as $m_q^\delta \leq \left\lceil\frac{\lambda_{\mathcal{D}_q}}{\delta/\sqrt{n(q)}}\right\rceil^{n(q)} \doteq \mathrm{k}(q,\delta)$. Thus the cardinality of the complete partition $\{\mathcal{S}\}_\delta$ can be upper bounded by $\sum_{q\in\mathcal{Q}} m_q^\delta \leq \sum_{q\in\mathcal{Q}} \mathrm{k}(q,\delta) \doteq \mathrm{k}(\delta)$. It increases as the continuous dimension n increases, as the size of the domains (related to its diameter $\lambda_{\mathcal{D}_{(\cdot)}}$) grows, and as the partition parameter δ is refined.

V. Analysis of the Abstraction and of its Precision

In this section, we quantify the precision of the abstraction by providing explicit bounds on the approximation distance between the transition probability for the DTSHS and that originating from the partition procedure (to be introduced shortly). The bounds are used to define intervals which, along with the approximated transition probabilities computed over the finite partition, characterize the abstraction as a MSC (see Definition 7). Furthermore, we investigate the actual dynamics in time of the approximation error: if the obtained abstraction is endowed with some ergodic property, it is shown that the error remains finite over time. We also show that, under conditions on the original DTSHS, it is possible to obtain an abstraction with arbitrary precision by tuning the parameter δ associated with the partition. This, in connection with the spectral properties of the MSC (see Section III), allows to introduce a finite-time algorithm that computes an abstraction, given a precision bound as an a-priori specification.

Let us recall that for a generic hybrid point $s = (q,x) \in \mathcal{S}$, its corresponding partition set is $\langle s\rangle \in \{\mathcal{S}\}_\delta$, and its representative point in $\langle s\rangle$ is $\bar{s} = (q,\bar{x}) \in \mathcal{S}$. Select a second partition set $\langle s'\rangle \in \{\mathcal{S}\}_\delta$ as a target set and any $k_0 \geq 0$. Let us approximate the one-step transition probability

$$p_s(\langle s'\rangle) = Prob(\mathbf{s}(k_0+1) \in \langle s'\rangle \mid \mathbf{s}(k_0) = s), \tag{4}$$

with a related quantity defined on a representative point as

$$p_{\bar{s}}(\langle s'\rangle) = Prob(\mathbf{s}(k_0+1) \in \langle s'\rangle \mid \mathbf{s}(k_0) = \bar{s}). \tag{5}$$

(More generally, we write $p_s^k(\langle s'\rangle) = Prob(\mathbf{s}(k_0+k) \in \langle s'\rangle \mid \mathbf{s}(k_0) = s)$, where $k_0, k \in \mathbb{N}$, and often omit the apex in p^k when

$k = 1$.) As explained in (1), the values in (4) and (5) depend on the kernel T, and can be obtained by marginalizing the probability distribution of the DTSHS over the set $\langle s' \rangle$. This approximation introduces an error, which depends on the parameter δ associated with the partition, and which we quantify in a closed form next.

A. Single step error and definition of the MSC abstraction

With reference to (4)-(5), $\forall \langle s' \rangle \subseteq \mathcal{S}$, where $s' = (q', x')$ and assuming that $q' = q \in \mathcal{Q}$, using equation (1) and Assumption 1, the following holds:

$$\begin{aligned} &|p_s(\langle s' \rangle) - p_{\bar{s}}(\langle s' \rangle)| \\ &= \left| \int_{\langle s' \rangle} T((q', dz)|(q, x)) - \int_{\langle s' \rangle} T((q', dz)|(q, \bar{x})) \right| \\ &\le \int_{\langle s' \rangle} \Big\{ \Big| T_t(dz|(q, x)) T_q(q'|(q, x)) - T_t(dz|(q, x)) T_q(q'|(q, \bar{x})) \Big| \\ &\quad + \Big| T_t(dz|(q, x)) T_q(q'|(q, \bar{x})) - T_t(dz|(q, \bar{x})) T_q(q'|(q, \bar{x})) \Big| \Big\} \\ &\le \mathcal{L}_{\langle s' \rangle} (L_t + L_q) \|x - \bar{x}\|, \end{aligned}$$

where $\mathcal{L}_A$ is the Lebesgue measure of the Borel set $A \in \mathcal{B}(\mathcal{D}_q)$, denoting the volume of set A. If instead $\langle s' \rangle \subseteq \mathcal{D}_{q'}, q' \neq q$, we can establish the following: $|p_s(\langle s' \rangle) - p_{\bar{s}}(\langle s' \rangle)| \le \mathcal{L}_{\langle s' \rangle}(L_r + L_q) \|x - \bar{x}\|$. Introducing the quantity $L \doteq \max\{L_t + L_q, L_r + L_q\}$ and the finite constant $n \doteq \max_{q \in \mathcal{Q}} n(q)$ it thus holds that, $\forall s, s' \in \mathcal{S}$:

$$|p_s(\langle s' \rangle) - p_{\bar{s}}(\langle s' \rangle)| \le \mathcal{L}_{\langle s' \rangle} L \|x - \bar{x}\| \le \delta^n L \delta \doteq \epsilon(\delta, n, L). \quad (6)$$

The quantity $\epsilon(\delta, n, L)$ denotes an upper bound on the error that depends on the continuity of the transition kernels (constant L), and which is an increasing function of the discretization diameter δ and of the continuous dimension n of the hybrid state space. By virtue of the state space partition procedure and of the error bound computed in (6), it is possible to associate to the DTSHS a MSC as follows:

Definition 7 (MSC Abstraction of a DTSHS): Given a DTSHS $\mathcal{H}$, let us introduce a partition $\{\mathcal{S}\}_\delta$ of its hybrid state space $\mathcal{S}$, parameterized by δ. An approximate abstraction of $\mathcal{H}$ is a Markov set-Chain $[\mathcal{M}]$ with state space coinciding with the quotient space of $\{\mathcal{S}\}_\delta$. The transition intervals of the MSC $[\mathcal{M}]$ are defined by:

1) computing the transition probabilities in (5) over the sets $\langle s \rangle \subseteq \{\mathcal{S}\}_\delta$ and their representative points $\bar{s} \in \mathcal{S}$;
2) introducing the error bound defined by (6).

For any $\langle s \rangle, \langle s' \rangle \in \{\mathcal{S}\}_\delta$, the elements of the MSC $[\mathcal{M}]$ are formally defined as: $[p_{\bar{s}}(\langle s' \rangle)] \doteq [\max\{0, p_{\bar{s}}(\langle s' \rangle) - \epsilon(\delta, n, L)\}, \min\{1, p_{\bar{s}}(\langle s' \rangle) + \epsilon(\delta, n, L)\}]$, and its k-th distribution set is denoted as $[p^k_{\bar{s}}(\langle s' \rangle)]$. The state cardinality of the MSC $[\mathcal{M}]$ is upper-bounded by the quantity $\mathrm{k}(\delta) = \sum_{q \in \mathcal{Q}} \left\lceil \frac{\lambda_{\mathcal{D}_q}}{\delta / \sqrt{n(q)}} \right\rceil^{n(q)}$. □

In general, we are interested in checking the validity of certain properties of the DTSHS $\mathcal{H}$ on the MSC $[\mathcal{M}]$. For instance, in Section V-C we will exploit some spectral properties of $[\mathcal{M}]$ to show related asymptotics of $\mathcal{H}$. To achieve this general goal, we first extend the study of the approximation error over time.

B. Error dynamics

In this Section we analyze the dynamics in time of the approximation error in equation (6) for the introduced MSC abstraction. Consider a point $s \in \mathcal{S}$, the representative point $\bar{s} \in \langle s \rangle$, any partition set $\langle s' \rangle \in \{\mathcal{S}\}_\delta$, and any $k > k_0 = 0$. Let us focus on the following two entities: $p^k_s(\langle s' \rangle) = Prob\big(\mathbf{s}(k) \in \langle s' \rangle \mid \mathbf{s}(0) = s\big)$ and $p^k_{\bar{s}}(\langle s' \rangle) = Prob\big(\mathbf{s}(k) \in \langle s' \rangle \mid \mathbf{s}(0) = \bar{s}\big)$. The distribution p^k_s of the DTSHS $\mathcal{H}$ over the sets of the partition $\{\mathcal{S}\}_\delta$ is derived, for any $k > 0$, by marginalization. This distribution can be associated with a non-homogeneous Markov Chain M which evolves on the quotient space $\{\mathcal{S}\}_\delta$. Additionally, the quantity $[p^k_{\bar{s}}]$ is the distribution set over the space of $\{\mathcal{S}\}_\delta$, and is generated by the Markov set-Chain $[\mathcal{M}]$, given an initial probability distribution concentrated on $\langle s \rangle$. Table I contains a summary of the different models used in this work.

model	state space	probability distribution/interval at $k > 0$ (conditional on s at k=0)
DTSHS $\mathcal{H}$	$\mathcal{S}$	$p^k_s(C), C \in \mathcal{B}(\mathcal{S})$
MC M	$\{\mathcal{S}\}_\delta$	$p^k_s(\langle s' \rangle), \langle s' \rangle \in \{\mathcal{S}\}_\delta$
MSC $[\mathcal{M}]$	$\{\mathcal{S}\}_\delta$	$[p^k_{\bar{s}}(\langle s' \rangle)], \langle s' \rangle \in \{\mathcal{S}\}_\delta$

TABLE I
SUMMARY OF NOTATIONS FOR THE MODELS UNDER STUDY

Let us introduce a function $f(\theta, n, k) : \mathbb{R}^+ \times \mathbb{N} \times \mathbb{N} \to \mathbb{R}^+$, inductively defined over time $k \ge 0$ as follows:

$$f(\theta, n, 0) = f_0, f_0 \ge 0, \; f(\theta, n, k{+}1) = (\theta n{+}1) f(\theta, n, k) {+} \theta. \quad (7)$$

The function f is monotonically increasing, as it is clear from its explicit form: $f(\theta, n, k) = (\theta n + 1)^k f_0 + \theta \sum_{l=1}^{k} (\theta n + 1)^{l-1}, \; k > 0$. The following result extends the calculation of the error bound in (6) over time, by looking at the distance between the probability distribution p^k_s of the MC M, (the chain obtained from $\mathcal{H}$ by marginalization over the partition sets at time k) and the distribution set $[p^k_{\bar{s}}]$ at time k induced by the MSC $[\mathcal{M}]$.

Theorem 1: Given a DTSHS $\mathcal{H}$, let us introduce a partition $\{\mathcal{S}\}_\delta$ of the hybrid state space $\mathcal{S}$, characterized by parameter δ. Assume that the corresponding Markov set-Chain abstraction $[\mathcal{M}]$ has coefficient of ergodicity $\mathcal{T}([\mathcal{M}])$. For any $s \in \mathcal{S}$, corresponding point $\bar{s}$, and any partition set $\langle s' \rangle \in \{\mathcal{S}\}_\delta$, the following holds, $\forall k > 0$:

$$\begin{aligned} &d_h(p^k_s(\langle s' \rangle), [p^k_{\bar{s}}(\langle s' \rangle)]) \le \\ &\min \left\{ f(\epsilon, \mathrm{k}(\delta), k), \quad \mathcal{T}([\mathcal{M}])^k + \mathrm{k}(\delta) \epsilon \sum_{l=0}^{k-1} \mathcal{T}([\mathcal{M}])^l \right\}, \end{aligned} \quad (8)$$

where d_h is obtained by the Hausdorff metric [17]; $\epsilon = \epsilon(\delta, n, L)$ is the error bound introduced by the abstraction procedure in equation (6); n is a finite upper-bound on the dimension of the continuous state of $\mathcal{H}$; $\mathrm{k}(\delta)$ is a finite upper-bound on the dimension of $[\mathcal{M}]$; L is a finite upper-bound on the Lipschitz constants of the probabilistic kernels of $\mathcal{H}$; and f has been introduced in (7) and is such that $f(\cdot, \cdot, 0) = 0$. In particular, if $[\mathcal{M}]$ is ergodic, i.e. if $\mathcal{T}([\mathcal{M}]) < 1$,

$$\begin{aligned} &d_h(p^k_s(\langle s' \rangle), [p^k_{\bar{s}}(\langle s' \rangle)]) \le \\ &\min \left\{ f(\epsilon, \mathrm{k}(\delta), k), \quad \mathcal{T}([\mathcal{M}])^k + \frac{\mathrm{k}(\delta) \epsilon}{1 - \mathcal{T}([\mathcal{M}])} \right\}, \end{aligned} \quad (9)$$

which is finite, for any $k > 0$. □

Equations (8)-(9) provide a time-dependent bound for the approximation error. The bound is finite in time if the MSC abstraction $[\mathcal{M}]$ induced by the partition parameter δ is ergodic (that is, if its coefficient of ergodicity $\mathcal{T}([\mathcal{M}])$ is strictly less than one). Next we prove that, under the following Assumption 2 on $\mathcal{H}$, there always exists a procedure (i.e., a choice of $\delta > 0$ for the partition) that yields an ergodic abstraction.

Let us recall a few notions. A DTSHS $\mathcal{H}$ is ψ-irreducible if there exists a measure ψ on $\mathcal{B}(\mathcal{S})$ such that, for all $C \in \mathcal{B}(\mathcal{S})$ with $\psi(C) > 0$, $\exists m < \infty : p^m_x(C) > 0$, for any $x \in \mathcal{S}$ [23, Sec. 4.2.1]. A set C is said to be a ν_m-small set for a non-trivial measure ν_m on $\mathcal{B}(\mathcal{S})$ if $\exists m > 0 : \forall x \in C, \forall B \in \mathcal{B}(\mathcal{S}),\, p^m_x(B) \ge \nu_m(B)$ [23, Sec. 5.2]. Suppose that DTSHS $\mathcal{H}$ is ψ-irreducible: $\mathcal{H}$ is strongly aperiodic if there exists a ν_1-small set $C \in \mathcal{B}(\mathcal{S})$ with $\nu_1(C) > 0$ [23, Sec. 5.4.3].

Assumption 2: The DTSHS $\mathcal{H}$ is ψ-irreducible and strongly aperiodic. □

Theorem 2: Consider a DTSHS $\mathcal{H}$. If Assumption 2 holds, it is possible to select a partition parameter $\delta > 0$, such that the induced abstraction $[\mathcal{M}]$ satisfies $\mathcal{T}([\mathcal{M}]) < 1$. □

C. Steady-State Computation with the Abstraction

Theorem 2 guarantees that, under some structural assumption on a DTSHS $\mathcal{H}$, it is possible to select a partition parameter $\delta > 0$ and to construct an approximate abstraction $[\mathcal{M}]$ which, by virtue of the bound in (9), is "close" over time to $\mathcal{H}$. In this section we propose an algorithm which, given a DTSHS $\mathcal{H}$ and a desired precision parameter $\phi > 0$, determines in a finite number of steps the steady-state behavior of $\mathcal{H}$ with precision ϕ by selecting a parameter $\delta(\phi)$ for the approximation procedure that generates the MSC $[\mathcal{M}]$ and by computing its steady-state interval.

More formally, let us consider a DTSHS $\mathcal{H}$ and let us uphold Assumption 2. By Theorem 2, there exists a positive parameter $\delta(\phi)$ and an associated approximation procedure on $\mathcal{H}$ that generates a MSC $[\mathcal{M}] : \mathcal{T}([\mathcal{M}]) < 1$. The steady state interval $[p^\infty]$ of $[\mathcal{M}]$ can be bounded by the diameter $\Delta([p^\infty])$. A sufficient condition to achieve the desired precision by the abstraction $[\mathcal{M}]$ is thus $\Delta([p^\infty]) \leq \phi$. Consider the one-step error $\epsilon(\delta(\phi), n, L)$ that characterizes the abstraction with cardinality $\mathrm{k}(\delta(\phi))$. Theorem 1 guarantees that a sufficient condition for finding a partition parameter $\delta(\phi)$ that induces an MSC abstraction with the required precision ϕ is $\frac{\mathrm{k}(\delta(\phi))\epsilon(\delta(\phi),n,L)}{1-\mathcal{T}([\mathcal{M}])} \leq \phi$. This bound holds if $\mathcal{T}([\mathcal{M}]) < 1$, which is enforced by Theorem 2.

The following algorithm introduces a sequence of partition procedures characterized by monotonically decreasing parameters $\{\delta_i\}_{i \geq 0}$, until the sufficient condition is satisfied. We define M_i to be the MC computed over the partition sets characterized by δ_i, $[\mathcal{M}_i]$ to be the associated MSC, and a variable τ_i to represents an upper bound for $\mathcal{T}([\mathcal{M}_i])$, as discussed in Section III.

Algorithm 1 (Compute steady state of $\mathcal{H}$ with precision $\phi > 0$):
set integer $i = 0$, real $\tau_i = 0$, and δ_i such that $\mathrm{k}(\delta_i)\epsilon(\delta_i, n, L) \leq \phi$;
***for** $i \geq 0$*
compute M_i according to the approximation with parameter δ_i;
set $\tau_i = \mathcal{T}(M_i) + \mathrm{k}(\delta_i)\epsilon(\delta_i, n, L)$;
***if** $\tau_i \geq \min\{1, \tau_{i-1}\}$, **then** set $\delta_{i+1} = a\delta_i$, for some $a < 1$;*
***else if** $\frac{\mathrm{k}(\delta_i)\epsilon(\delta_i,n,L)}{1-\tau_i} > \phi$, **then** set δ_{i+1} s.t. $\frac{\mathrm{k}(\delta_{i+1})\epsilon(\delta_{i+1},n,L)}{1-\tau_i} \leq \phi$;*
else exit;
***end** (if)*
set $i = i + 1$;
***end** (for)*
compute the steady state p_i^∞ of M_i.

Theorem 3: If Assumption 2 holds, Algorithm 1 terminates in a finite number of steps. □

VI. Conclusions

This work has introduced an abstraction procedure for discrete-time Stochastic Hybrid Systems (DTSHS). The approximate abstraction is interpreted as a Markov set-Chain (MSC). By raising some continuity assumptions on the stochastic kernels that characterize the DTSHS, we have derived an analytic formula relating the accuracy of the state space partition and the error of the approximate abstraction. Additionally, we have shown a bound in time for the distance between the transition probabilities of the abstract model (the MSC) and those of the original DTSHS. Under proper assumptions, the error bounds are finite over time and there exists a finite-time algorithm that computes the abstraction.

References

[1] C. Cassandras and J. Lygeros, *Stochastic hybrid systems*, ser. Automation and Control Engineering. Taylor & Francis Group/CRC Press, 2006.

[2] C. Courcoubetis and M. Yannakakis, "The complexity of probabilistic verification," *Journal of the ACM*, vol. 42, no. 4, pp. 857–907, 1995.

[3] A. Girard and G. J. Pappas, "Approximation metrics for discrete and continuous systems," *IEEE Trans. on Automatic Control*, vol. 52(5), pp. 782–798, 2007.

[4] R. Alur, T. Henzinger, G. Lafferriere, and G. Pappas, "Discrete abstractions of hybrid systems," *Proceedings of the IEEE*, vol. 88(2), pp. 971–984, July 2000.

[5] G. P. A. Girard 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.

[6] A.A. Julius and G.J. Pappas, "Approximations of stochastic hybrid systems," *IEEE Trans. Automatic Control*, vol. 54, no. 6, pp. 1193–1203, June 2009.

[7] C. Baier, J. Katoen, H. Hermanns, and V. Wolf, "Comparative branching-time semantics for Markov Chains," *Information and Computation*, vol. 200, no. 2, pp. 149–214, 2005.

[8] K. Larsen and A. Skou, "Bisimulation through probabilistic testing," *Information and Computation*, vol. 94, pp. 1–28, 1991.

[9] J. Desharnais, A. Edalat, and P. Panangaden, "Bisimulation for labeled Markov processes," *Information and Computation*, vol. 179, no. 2, pp. 163–193, 2002.

[10] H. Hermanns, *Interactive Markov Chains, and the Quest for Quantified Quality*, ser. Lecture Notes in Computer Science. Springer Verlag, 2002, vol. 2428.

[11] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic Systems Theory*. Cambridge, Massachussets: MIT Press, 1984.

[12] X. Koutsoukos, "Optimal control of stochastic hybrid systems based on locally consistent Markov decision processes," *International Journal of Hybrid Systems*, vol. 4, pp. 301–318, 2004.

[13] M. Prandini and J. Hu, "A numerical approximation scheme for reachability analysis of stochastic hybrid systems with state-dependent switchings," in *In Proc. IEEE Int. Conf. Decision and Control*, New Orleans, LA, Dec. 2007, pp. 4662–4667.

[14] C. Guihenneuc-Jouyaux and C. Robert, "Discretization of continuous Markov Chains and Markov Chain Monte Carlo convergence assessment," *Journal of the American Statistical Association*, vol. 93, no. 443, pp. 1055–1067, 1998.

[15] M. Sarich, F. Noe, and C. Schütte, "On the approximation quality of Markov state models," *Multiscale Modeling and Simulation*, vol. 8, no. 4, pp. 1154–1177, 2010.

[16] A. Abate, M. Prandini, J. Lygeros, and S. Sastry, "Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems," *Automatica*, vol. 44, no. 11, pp. 2724–2734, Nov 2008.

[17] H. J. Hartfiel, *Markov set-Chains*, ser. Lecture Notes in Mathematics. Springer-Verlag Berlin Heidelberg, 1998, vol. 1695.

[18] D. P. Bertsekas and S. E. Shreve, *Stochastic optimal control: the discrete-time case*. Athena Scientific, 1996.

[19] J. Cohen, "Random evolutions in discrete and continuous time," *Stochastic Processes and their Applications*, vol. 9, no. 4, pp. 245–251, 1979.

[20] M. Keepler, "Random evolutions processes induced by discrete time Markov Chains," *Port. Mathematica*, vol. 55, no. 4, pp. 391–400, 1998.

[21] I. Kozine and L. Utkin, "Interval-valued finite Markov Chains," *Reliable Computing*, vol. 8, pp. 97–113, 2002.

[22] D. P. Bertsekas, "Convergence of discretization procedures in dynamic programming," *IEEE Transactions on Automatic Control*, vol. AC-20, no. 3, pp. 415–419, 1975.

[23] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*. Springer Verlag, 1993.

Appendix

Proposition 1: [17, Theorem 3.1] Let $[\Pi]$ be the interval $[P, Q]$ and $A \in [\Pi]$, then: $|\mathcal{T}([\Pi]) - \mathcal{T}(A)| \leq ||Q - P||$. □
The used matrix norm is the induced 1-norm: $\|A\| = \max_{x \neq 0} \frac{\|xA\|}{\|x\|}$ [17, Appendix A.1].

Proposition 2: [17, Theorems 3.4, 3.11] Given a Markov set-Chain with transition set $[\Pi] = [P, Q]$, then

$$\Delta([\Pi]^k) \leq \mathcal{T}([\Pi])^k + (\mathcal{T}([\Pi])^{k-1} + \ldots + 1)\Delta([\Pi]).$$

In particular, if $\mathcal{T}([\Pi]) < 1$, given any initial distribution set $[\pi_0]$, there exists a unique limit set $[\pi_\infty]$ that is invariant, i.e. such that $[\pi_\infty][\Pi] = [\pi_\infty]$, and such that $\lim_{k\to\infty}[\pi_k] = \lim_{k\to\infty}[\pi_0][\Pi]^k = [\pi_\infty]$. The following holds: $\Delta([\pi_\infty]) \leq \frac{\Delta([\Pi])}{1-\mathcal{T}([\Pi])} \leq \frac{||Q-P||}{1-\mathcal{T}([\Pi])}$. □

The notion of limit of a vector interval hinges on the Hausdorff distance [17], which is a distance between sets.

Corollary 1: Given a Markov set-Chain with transition set $[\Pi] = [P, Q]$ such that $\mathcal{T}([\Pi]) < 1$, and any initial interval vector $[\pi_0]$, then for any $A \in [\Pi]$ such that $\mathcal{T}(A) + ||Q - P|| \leq 1$

$$\Delta([\pi_\infty]) \leq \frac{\Delta([\Pi])}{1-\mathcal{T}([\Pi])} \leq \frac{||Q-P||}{1-\mathcal{T}(A)-||Q-P||}.$$

□

Lemma 1: Consider a discrete-time homogeneous Markov Chain defined by an $n \times n$ stochastic matrix $P = \{p_{ij}\}$ and a legitimate initial probability distribution $p(0)$. Let $p(k), k \geq 0$, be the associated probability distribution vector. Given a real constant $\theta > 0$, consider the Markov set-Chain defined by the $n \times n$ stochastic interval matrix $\bar{P} = \{[\max\{0, p_{ij} - \theta\}, \min\{1, p_{ij} + \theta\}]\}$ and the initial probability distribution $\bar{p}(0) = p(0)$, and let $\bar{p}(k), k \geq 0$, be the associated probability interval vector. The following holds:

$$\forall k \geq 0, \qquad d_h\big(\bar{p}(k), p(k)\big) \leq f(\theta, n, k), \tag{10}$$

where d_h is the Hausdorff distance [17], and $f(\theta, n, k)$ is defined in equation (7) with $f(\theta, n, 0) = 0$. □

Proof: (By induction) Equation (10) is valid for $k = 0$, since $\bar{p}(0) = p(0)$. Let (10) hold for $k > 0$. Introducing function $\mu : \mathbb{R} \to [0, 1], \mu(x) = \min\{\max\{x, 0\}, 1\}$, and resorting to $P, \bar{P}$:

$$
\begin{aligned}
&d_h\big(\bar{p}(k+1), p(k+1)\big) \leq \\
&\max_{\substack{i=1,\ldots,n,\\ \alpha\in[-\theta,+\theta],\\ \beta\in[-f(\theta,n,k),+f(\theta,n,k)]}} \left\{ \begin{array}{l} |\mu(p_{i1}+\alpha)(p_1(k)+\beta)+\cdots \\ +\mu(p_{in}+\alpha)(p_n(k)+\beta)-p_i(k+1)| \end{array} \right\} \leq \\
&\max_{i=1,\ldots,n} |p_i(k+1) + \theta n f(\theta, n, k) + \theta + f(\theta, n, k) - p_i(k+1)|,
\end{aligned}
$$

which is $(\theta n + 1) f(\theta, n, k) + \theta = f(\theta, n, k + 1)$, and shows that (10) holds. The first inequality is set up by expressing the probability interval of the MSC $\bar{P}$ at time $k + 1$ according to its structure, whereas the second is derived by upper-bounding the multiplication of probabilities. If $d_h(\bar{p}(0), p(0)) \neq 0$, the proof is adapted by taking $f_0 = d_h(\bar{p}(0), p(0)) = \|\bar{p}(0) - p(0)\|$, where $\|\cdot\|$ is the 1-norm. ■

Proof of Theorem 1: The case $k = 1$ follows from equation (6). If, instead of a single point $s \in \mathcal{S}$, the system is initialized over a probability distribution π_0 over $\mathcal{S}$, then both M and $[\mathcal{M}]$ will be initialized on a marginalization of π_0 over the partition sets of $\{\mathcal{S}\}_\delta$. For any $k \geq 1$, one can show by direct calculation on the definition of the MSC $[\mathcal{M}]$ and by Lemma 1 that the approximation error can be upper-bounded as

$$d_h(p_s^k(\langle s'\rangle), [p_{\bar{s}}^k(\langle s'\rangle)]) \leq f(\epsilon, \mathrm{k}(\delta), k), \tag{11}$$

where f has been introduced in (7) and where we have inherited the initialization $f(\cdot, \cdot, 0) = 0$. This bound corresponds to the error growth that is obtained when the MSC $[\mathcal{M}]$ is elevated to the power of k. Being monotonically increasing, it can become conservative as time k grows. This leads to consider a second bound, to be combined with the first. Observe that the stochastic behavior $[p_{\bar{s}}^k]$ (generated by $[\mathcal{M}]$) is conservative with respect to p_s^k (generated by M). This allows to state that $d_h(p_s^k(\langle s'\rangle), [p_{\bar{s}}^k(\langle s'\rangle)]) \leq \Delta([\mathcal{M}]^k)$. Resorting to Proposition 2, it is possible to conclude that

$$
\begin{aligned}
&d_h(p_s^k(\langle s'\rangle), [p_{\bar{s}}^k(\langle s'\rangle)]) \leq \\
&\mathcal{T}([\mathcal{M}])^k + (\mathcal{T}([\mathcal{M}])^{k-1} + \ldots + 1)\mathrm{k}(\delta)\epsilon(\delta, n, L).
\end{aligned} \tag{12}
$$

By Corollary 1, if the MSC is ergodic with $\mathcal{T}([\mathcal{M}]) < 1$, then

$$d_h(p_s^k(\langle s'\rangle), [p_{\bar{s}}^k(\langle s'\rangle)]) \leq \mathcal{T}([\mathcal{M}])^k + \frac{\mathrm{k}(\delta)\epsilon(\delta, n, L)}{1 - \mathcal{T}([\mathcal{M}])}. \tag{13}$$

The right-hand side is made up of two terms: the first is finite and decreasing in k, whereas the second is fixed. As a result, the bound is finite in time. The inequalities in (8) and (9) in the statement of the theorem follow by considering, respectively, the pair of bounds (11)-(12) and (11)-(13). ■

Proof of Theorem 2: Pick a discretization parameter $\delta > 0$ for $\mathcal{H}$. The obtained $\mathrm{k}(\delta)$-dimensional MSC $[\mathcal{M}]$ is made up of the elements $[\max\{0, p_{\bar{s}}(\langle s'\rangle) - \epsilon(\delta, n, L)\}, \min\{1, p_{\bar{s}}(\langle s'\rangle) + \epsilon(\delta, n, L)\}]$ and, as per (6), $\epsilon(\delta, n, L) \leq \delta^{n+1}L, n = \max_{q\in\mathcal{Q}} n(q), L = \max\{L_t + L_q, L_r + L_q\}$. Select a generic stochastic matrix $A = (A_{ij}) \in [\mathcal{M}]$, where the element A_{ij} refers to the partition sets $\langle s_i\rangle, \langle s_j\rangle, i, j = 1, \ldots, \mathrm{k}(\delta)$. To claim the ergodicity of $[\mathcal{M}]$, we are interested in showing that the generic MC A extracted from $[\mathcal{M}]$ is such that $\mathcal{T}(A) < 1$. To achieve this, it is sufficient to show that matrix A is irreducible and aperiodic [17].

Let us denote with $b(\epsilon, \mathrm{k}(\delta), m)$ the bound on the RHS of (8) from Theorem 1 at time $m > 0$. With reference to the partition sets $\langle s_i\rangle, \langle s_j\rangle, i, j = 1, \ldots, \mathrm{k}(\delta)$, generic point $s_i \in \langle s_i\rangle$, we have that $(A^m)_{i,j} \geq \max\{p_{s_i}^m(\langle s_j\rangle) - b(\epsilon, \mathrm{k}(\delta), m), 0\}$. At any $m > 0$ and as the discretization parameter $\delta \downarrow 0$, the term $p_{s_i}^m(\langle s_j\rangle)$ - whenever positive - converges to zero as a function of δ^n (which is due to the volume of $\langle s_j\rangle$), whereas $b(\epsilon, \mathrm{k}(\delta), m)$ converges to zero at least as a function of δ^{n+1} (cfr. (6), (7), and bound in (8) – the convergence can be made faster if local Lipschitz constants rather than L are considered in (6)). The continuity of the transition kernels of $\mathcal{H}$ allows to select a partition parameter $\delta^* > 0$ small enough so that the following two conditions hold:

1 Choose $\delta^* > 0$ so that, if $\psi(\langle s_j\rangle) > 0$, then $p_{s_i}^m(\langle s_j\rangle) - b(\epsilon, \mathrm{k}(\delta^*), m) > 0$, for $i, j = 1, \ldots, \mathrm{k}(\delta^*)$, where $m = \max_{j=1,\ldots,\mathrm{k}(\delta^*)} m(\langle s_j\rangle)$ and $m(\langle s_j\rangle)$ is the finite ψ-irreducibility index of $\mathcal{H}$ over $\langle s_j\rangle$. This allows to obtain that $(A^m)_{i,j} > 0$.

2 Consider any ν_1-small set $C \in \mathcal{B}(\mathcal{S})$ with $\nu_1(C) > 0$, which is such that $\forall s \in C, p_x(C) > \nu_1(C) > 0$. Choose $\delta^* > 0$ so that $\exists i = 1, \ldots, \mathrm{k}(\delta^*) : \langle s_i\rangle \subseteq C \wedge \nu_1(\langle s_i\rangle) > 0$, and $p_{s_i}(\langle s_i\rangle) - b(\epsilon, \mathrm{k}(\delta^*), 1) > 0$. (Notice that the partition set $\langle s_i\rangle$ is related to a positive probability of self-loop $p_{s_i}(\langle s_i\rangle) > \nu_1(\langle s_i\rangle) > 0$.) Then this yields $A_{i,i} > 0$.

This choice of δ^* renders matrix $A \in [\mathcal{M}]$ irreducible (over the ψ-irreducibility classes of $\mathcal{H}$) in view of condition 1, as well as aperiodic thanks to condition 2. Thus, $\mathcal{T}(A) < 1$. The conclusion that $\mathcal{T}([\mathcal{M}]) < 1$ is drawn based on the generic choice of $A \in [\mathcal{M}]$ and given the compactness of the interval matrix $[\mathcal{M}]$ and the continuity of $\mathcal{T}(A)$ as a function of its entries in A. ■

Proof of Theorem 3: The updates on the parameter δ_i are contractive, since the multiplication $\mathrm{k}(x)\epsilon(x, \cdot, \cdot)$ is monotonically increasing with respect to x. Furthermore, notice that the "else if" condition cannot be true for two consecutive choices of the parameter. Therefore we have that $\lim_{i\to+\infty} \delta_i = 0$, which says that $\lim_{i\to+\infty} \Delta([\mathcal{M}_i]) = 0$ (see Section III) and that $\lim_{i\to+\infty} \mathcal{T}([\mathcal{M}_i]) < 1$ (by Theorem 2). The above two considerations imply that $\lim_{i\to+\infty} \tau_i = \lim_{i\to+\infty} \mathcal{T}(\mathcal{M}_i) < 1$. Thus, there exists a finite index i^* such that for any $i > i^* : \frac{\mathrm{k}(\delta_i)\epsilon(\delta_i, n, L)}{1-\tau_i} \leq \phi$. According to this condition, the steady state p_i^∞ of M_i is an estimate of the steady state of $\mathcal{H}$ with precision ϕ. Notice that the condition provided by Assumption 2 is sufficient, but not necessary. ■