

# Measuring Performance of Continuous-Time Stochastic Processes using Timed Automata

Tomáš Brázdil  
brazdil@fi.muni.cz

Jan Krčál  
krccal@fi.muni.cz

Jan Křetínský  
jan.kretinsky@fi.muni.cz

Antonín Kučera  
kucera@fi.muni.cz

Vojtěch Řehák  
rehak@fi.muni.cz

Faculty of Informatics, Masaryk University  
Botanická 68a, 60200 Brno  
Czech Republic

## ABSTRACT

We propose deterministic timed automata (DTA) as a model-independent language for specifying performance and dependability measures over continuous-time stochastic processes. Technically, these measures are defined as limit frequencies of locations (control states) of a DTA that observes computations of a given stochastic process. Then, we study the properties of DTA measures over semi-Markov processes in greater detail. We show that DTA measures over semi-Markov processes are well-defined with probability one, and there are only finitely many values that can be assumed by these measures with positive probability. We also give an algorithm which approximates these values and the associated probabilities up to an arbitrarily small given precision. Thus, we obtain a general and effective framework for analysing DTA measures over semi-Markov processes.

## Categories and Subject Descriptors

G.3 [Mathematics of Computing]: Probability and statistics—*Stochastic processes*

## General Terms

Performance, Theory

## Keywords

semi-Markov processes, timed automata, performance analysis, general state space Markov chains, stochastic stability

## 1. INTRODUCTION

Continuous-time stochastic processes, such as continuous-time Markov chains, semi-Markov processes, or generalized semi-Markov processes [21, 6, 19, 17], have been widely used in practice to determine performance and dependability characteristics of real-world systems. The desired behaviour

of such systems is specified by various measures such as mean response time, throughput, expected frequency of errors, etc. These measures are often formulated just semi-formally and chosen specifically for the system under study in a somewhat ad hoc manner. One example of a rigorous and model-independent specification language for performance and dependability properties is Continuous Stochastic Logic (CSL) [3, 5] which allows to specify both steady state and transient measures over the underlying stochastic process. The syntax and semantics of CSL is inspired by the well-known non-probabilistic logic CTL [14]. The syntax of CSL defines *state* and *path* formulae, interpreted over the states and runs of a given stochastic process  $\mathcal{M}$ . In particular, there are two probabilistic operators,  $\mathcal{P}_{\bowtie \varrho}(\cdot)$  and  $\mathcal{S}_{\bowtie \varrho}(\cdot)$ , which refer to the transient and steady state behaviour of  $\mathcal{M}$ , respectively. Here  $\bowtie$  is a numerical comparison (such as  $\leq$ ) and  $\varrho \in [0, 1]$  is a rational constant. If  $\varphi$  is a path formula<sup>1</sup> (which is either valid or invalid for every run of  $\mathcal{M}$ ), then  $\mathcal{P}_{\geq 0.7}(\varphi)$  is a state formula which says “the probability of all runs satisfying  $\varphi$  is at least 0.7”. If  $\Phi$  is a state formula, i.e.,  $\Phi$  is either valid or invalid in every state, then  $\mathcal{S}_{\geq 0.5}(\Phi)$  is also a state formula which says “the  $\pi$ -weighted sum over all states where  $\Phi$  holds is at least 0.5”. Here  $\pi$  is the steady-state distribution of  $\mathcal{M}$ . The logic CSL can express quite complicated properties and the corresponding model-checking problem over continuous-time Markov chains is decidable. However, there are also several disadvantages.

- (a) The semantics of steady state probabilistic operator  $\mathcal{S}_{\bowtie \varrho}(\cdot)$  assumes the existence of invariant distribution which is not guaranteed to exist for all types of stochastic processes with continuous time (the existing works mainly consider CSL as a specification language for ergodic continuous-time Markov chains).
- (b) In CSL formulae, all measures are explicitly quantified, and the model-checking algorithm just verifies constraints over these measures. Alternatively, we might wish to *compute* certain measures up to a given precision.

In this paper, we propose *deterministic timed automata (DTA)* [2] as a model-independent specification language

<sup>1</sup>In CSL,  $\varphi$  can be of the form  $\mathcal{X}_I\Phi$  or  $\Phi_1\mathcal{U}_I\Phi_2$  where  $\Phi, \Phi_1, \Phi_2$  are state formulae, and  $\mathcal{X}_I, \mathcal{U}_I$  are the modal connectives of CTL parametrized by an interval  $I$ . Boolean connectives can be used to combine just state formulae.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

HSCC'11, April 12–14, 2011, Chicago, Illinois, USA.

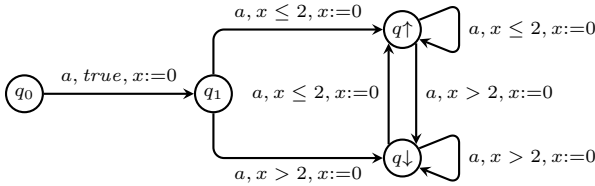
Copyright 2011 ACM 978-1-4503-0629-4/11/04 ...\$10.00.

for performance and dependability measures of continuous-time stochastic processes. The “language” of DTA can be interpreted over *arbitrary* stochastic processes that generate timed words, and their expressive power appears sufficiently rich to capture many interesting run-time properties (although we do not relate the expressiveness of CSL and DTA formally, they are surely incomparable because of different “nature” of the two formalisms). Roughly speaking, a DTA  $\mathcal{A}$  “observes” runs of a given stochastic process  $\mathcal{M}$  and “remembers” certain information in its control states (which are called *locations*). Since  $\mathcal{A}$  is deterministic, for every run  $\sigma$  of  $\mathcal{M}$  there is a unique computation  $\mathcal{A}(\sigma)$  of  $\mathcal{A}$ , which determines a unique tuple of “frequencies” of visits to the individual locations of  $\mathcal{A}$  along  $\sigma$ . These frequencies are the values of “performance measures” defined by  $\mathcal{A}$  (in fact, we consider *discrete* and *timed* frequencies which are based on the same concept but defined somewhat differently).

Let us explain the idea in more detail. Consider some stochastic process  $\mathcal{M}$  whose computations (or runs) are infinite sequences of the form  $\sigma = s_0 t_0 s_1 t_1 \dots$  where all  $s_i$  are “states” and  $t_i$  is the time spent by performing the transition from  $s_i$  to  $s_{i+1}$ . Also assume a suitable probability space defined over the runs of  $\mathcal{M}$ . Let  $\Sigma$  by a finite alphabet and  $L$  a labelling which assigns a unique letter  $L(s) \in \Sigma$  to every state  $s$  of  $\mathcal{M}$ . Intuitively, the letters of  $\Sigma$  correspond to collections of predicates that are valid in a given state. Thus, every run  $\sigma = s_0 t_0 s_1 t_1 \dots$  of  $\mathcal{M}$  determines a unique *timed word*  $w_\sigma = L(s_0) t_0 L(s_1) t_1 \dots$  over  $\Sigma$ .

A DTA over  $\Sigma$  is a finite-state automaton  $\mathcal{A}$  equipped with finitely many internal clocks. Each control state (or location)  $q$  of  $\mathcal{A}$  has finitely many outgoing edges  $q \rightarrow q'$  labeled by triples  $(a, g, X)$ , where  $a \in \Sigma$ ,  $g$  is a “guard” (a constraint on the current clock values), and  $X$  is a subset of clocks that are reset to zero after performing the edge. A configuration of  $\mathcal{A}$  is a pair  $(q, \nu)$ , where  $q$  and  $\nu$  are the current location and the current clock valuation, respectively. Every timed word  $w = c_0 c_1 c_2 c_3 \dots$  over  $\Sigma$  (where  $c_i \in \Sigma$  iff  $i$  is even) then determines a unique run  $\mathcal{A}(w) = (q_0, \nu_0) (q_1, \nu_1) (q_2, \nu_2) \dots$  of  $\mathcal{A}$  where  $q_0$  is an *initial* location,  $\nu_0$  assigns zero to every clock, and  $(q_{i+1}, \nu_{i+1})$  is obtained from  $(q_i, \nu_i)$  either by performing the only enabled edge  $q_i \rightarrow q_{i+1}$  labeled by  $(c_i, g, X)$  if  $i$  is even, or by simultaneously increasing all clocks by  $c_i$  if  $i$  is odd.

As a simple example, consider the following DTA  $\hat{\mathcal{A}}$  over the alphabet  $\{a\}$  with one clock  $x$  and the initial location  $q_0$ :



Intuitively,  $\hat{\mathcal{A}}$  observes time stamps in a given timed word and enters either  $q_\uparrow$  or  $q_\downarrow$  depending on whether a given stamp is bounded by 2 or not, respectively. For example, a word  $w = a \ 0.2 \ a \ 2.4 \ a \ 2.1 \dots$  determines the run  $\hat{\mathcal{A}}(w) = (q_0, 0) (q_1, 0) (q_1, 0.2) (q_\uparrow, 0) (q_\uparrow, 2.4) (q_\downarrow, 0) (q_\downarrow, 2.1) \dots$

Let  $w = a_0 t_0 a_1 t_1 \dots$  be a timed word over  $\Sigma$  and  $q$  a location of  $\mathcal{A}$ . For every  $i \in \mathbb{N}_0$ , let  $T^i(w)$  be the stamp  $t_i$  of  $w$ , and  $Q^i(w)$  the location of  $\mathcal{A}$  entered after reading the finite prefix  $a_0 t_0 \dots a_i$  of  $w$ . Further, let  $1_q^i(w)$  be either 1 or 0 depending on whether  $Q^i(w) = q$  or not, respectively. We define the *discrete* and *timed* frequency of visits to  $q$  along

$\mathcal{A}(w)$ , denoted by  $\mathbf{d}_q^{\mathcal{A}}(w)$  and  $\mathbf{c}_q^{\mathcal{A}}(w)$ , in the following way (the ‘ $\mathcal{A}$ ’ index is omitted when it is clear from the context):

$$\mathbf{d}_q^{\mathcal{A}}(w) = \limsup_{n \rightarrow \infty} \frac{\sum_{i=1}^n 1_q^i(w)}{n}$$

$$\mathbf{c}_q^{\mathcal{A}}(w) = \limsup_{n \rightarrow \infty} \frac{\sum_{i=1}^n T^i(w) \cdot 1_q^i(w)}{\sum_{i=1}^n T^i(w)}$$

Thus, every timed word  $w$  determines the tuple  $\mathbf{d}^{\mathcal{A}}(w) = (\mathbf{d}_q^{\mathcal{A}}(w))_{q \in Q}$  and the tuple  $\mathbf{c}^{\mathcal{A}}(w) = (\mathbf{c}_q^{\mathcal{A}}(w))_{q \in Q}$  of *discrete* and *timed*  $\mathcal{A}$ -measures, respectively.

DTA measures can encode various performance and dependability properties of stochastic systems with continuous time. For example, consider again the DTA  $\hat{\mathcal{A}}$  above and assume that all states of a given stochastic process  $\mathcal{M}$  are labeled with  $a$ . Then, the fraction

$$\frac{\mathbf{d}_{q_\uparrow}(w_\sigma)}{\mathbf{d}_{q_\uparrow}(w_\sigma) + \mathbf{d}_{q_\downarrow}(w_\sigma)}$$

corresponds to the percentage of transitions of  $\mathcal{M}$  that are performed within 2 seconds along a run  $\sigma$ . If  $\mathcal{M}$  is an ergodic continuous-time Markov chain, then the above fraction takes the *same* value for almost all runs  $\sigma$  of  $\mathcal{M}$ . However, it makes sense to consider this fraction also for non-ergodic processes. For example, we may be interested in the expected value of  $\mathbf{d}_{q_\uparrow}/(\mathbf{d}_{q_\uparrow} + \mathbf{d}_{q_\downarrow})$ , or in the probability of all runs  $\sigma$  such that the fraction is at least 0.5.

One general trouble with DTA measures is that  $\mathbf{d}_q^{\mathcal{A}}(w)$  and  $\mathbf{c}_q^{\mathcal{A}}(w)$  faithfully capture the frequency of visits to  $q$  along  $w$  only if the limits

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n 1_q^i(w)}{n} \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n T^i(w) \cdot 1_q^i(w)}{\sum_{i=1}^n T^i(w)}$$

exist, in which case we say that  $\mathbf{d}^{\mathcal{A}}$  and  $\mathbf{c}^{\mathcal{A}}$  are *well-defined* for  $w$ , respectively. So, one general question that should be answered when analyzing the properties of DTA measures over a particular class of stochastic processes is whether  $\mathbf{d}^{\mathcal{A}}$  and  $\mathbf{c}^{\mathcal{A}}$  are well-defined for almost all runs. If the answer is negative, we might either try to re-design our DTA or accept the fact that the limit frequency of the considered event simply does not exist (and stick to lim sup).

In this paper, we study DTA measures over semi-Markov processes (SMPs). An SMP is essentially a discrete-time Markov chain where each transition is assigned (apart of its discrete probability) a *delay density*, which defines the distribution of time needed to perform the transition. A computation (run) of an SMP  $\mathcal{M}$  is initiated in some state  $s_0$ , which is also chosen randomly according to a fixed initial distribution over the state space of  $\mathcal{M}$ . The next transition is selected according to the fixed transition probabilities, and the selected transition takes time chosen randomly according to the density associated to the transition. Hence, each run of  $\mathcal{M}$  is an infinite sequence  $s_0 t_0 s_1 t_1 \dots$ , where all  $s_i$  are states of  $\mathcal{M}$  and  $t_i$  are time stamps. The probability of (certain) subsets of runs in  $\mathcal{M}$  is measured in the standard way (see Section 2).

The *main contribution* of this paper are general results about DTA measures over semi-Markov processes, which are valid for all SMPs where the employed density functions are bounded from zero on every closed subinterval (see Section 2). Under this assumption, we prove that for every SMP  $\mathcal{M}$  and every DTA  $\mathcal{A}$  we have the following:

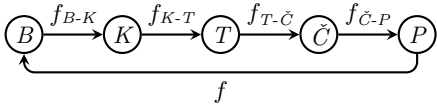
- (1) Both discrete and timed  $\mathcal{A}$ -measures are well defined for almost all runs of  $\mathcal{M}$ .

- (2) Almost all runs of  $\mathcal{M}$  can be divided into finitely many pairwise disjoint subsets  $\mathcal{R}_1, \dots, \mathcal{R}_k$  so that  $\mathbf{d}^{\mathcal{A}}(w)$  takes the same value for almost all  $w \in \mathcal{R}_j$ , where  $1 \leq j \leq k$ . The same result holds also for  $\mathbf{c}^{\mathcal{A}}$ . (Let us note that  $k$  can be larger than 1 even if  $\mathcal{M}$  is strongly connected.)
- (3) The observations behind the results of (1) and (2) can be used to compute the  $k$  and effectively approximate the probability of all  $\mathcal{R}_j$  together with the associated values of discrete or timed  $\mathcal{A}$ -measures up to an arbitrarily small given precision. More precisely, we show that these quantities are expressible using the  $m$ -step transition kernel  $P^m$  of the product process  $\mathcal{M} \times \mathcal{A}$  defined for  $\mathcal{M}$  and  $\mathcal{A}$  (see Section 3.2), and we give generic bounds on the number of steps  $m$  that is sufficient to achieve the required precision. The  $m$ -step transition kernel is defined by nested integrals (see Section 3.1) and can be approximated by numerical methods (see, e.g., [16, 9]). This makes the whole framework effective. The design of more efficient algorithms as well as more detailed analysis applicable to concrete subclasses of SMP are left for future work.

To get some intuition about potential applicability of our results (and about the actual power of DTA which is hidden mainly in their ability to accumulate the total time of several transitions in internal clocks), let us start with a simple example. Consider the following itinerary for travelling between Brno and Prague:

	Brno	Kuřim	Tišnov	Čáslav	Prague
<i>arrival</i>		1:15	2:30	3:30	4:50
<i>departure</i>	0:00	1:20	2:40	3:35	

A traveller has to change a train at each of the three intermediate stops, and she needs at least 3 minutes to walk between the platforms. Assume that all trains depart on time, but can be delayed. Further, assume that travelling time between  $X$  and  $Y$  has density  $f_{X-Y}$ . We wonder what is the chance that a traveller reaches Prague from Brno without missing any train and at most 5 minutes after the scheduled arrival. Answering this question “by hand” is not simple (though still possible). However, it is almost trivial to rephrase this question in terms of DTA measures. The itinerary can be modeled by the following semi-Markov process, where the density  $f$  is irrelevant and  $\Sigma = \{B, K, T, \check{C}, P\}$ .



The property of “reaching Prague from Brno without missing any train and at most 5 minutes after the scheduled arrival” is encoded by the DTA  $\bar{\mathcal{A}}$  of Figure 1. The automaton uses just one clock  $x$  to measure the total elapsed time, and the guards reflect the required timing constraints. Starting in location *init*, the automaton eventually reaches either the location  $p\uparrow$  or  $p\downarrow$ , which corresponds to satisfaction or violation of the above property, and then it is “restarted”. Hence, we are interested in the relative frequency of visits to  $p\uparrow$  among the visits to  $p\uparrow$  or  $p\downarrow$ . Using our results, it follows that  $\mathbf{d}^{\mathcal{A}}$  is well-defined and takes the same value for almost all runs of  $\mathcal{M}$ . Hence, the random variable  $\mathbf{d}_{p\uparrow}/(\mathbf{d}_{p\uparrow} + \mathbf{d}_{p\downarrow})$  also takes the same value with probability one, and this (unique) value is the quantity of our interest.

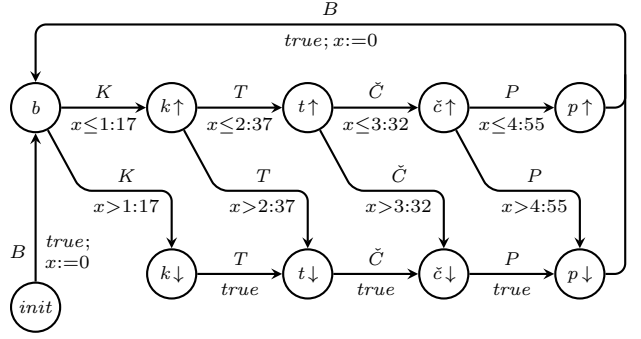


Figure 1: A deterministic timed automaton  $\bar{\mathcal{A}}$ .

Now imagine we wish to model and analyse the flow of passengers in London metro at rush hours. The SMP states then correspond to stations, transition probabilities encode the percentage of passengers traveling in a given direction, and the densities encode the distribution of travelling time. A DTA can be used to monitor a complex list of timing restrictions such as “there is enough time to change a train”, “travelling between important stations does not take more than 30 minutes if one the given routes is used”, “trains do not arrive more than 2 minutes later than scheduled”, etc. For this we already need several internal clocks. Apart of some auxiliary locations, the constructed DTA would also have special locations used to encode satisfaction/violation of a given restriction (in the DTA  $\bar{\mathcal{A}}$  of Figure 1,  $(p, \uparrow)$  and  $(p, \downarrow)$  are such special locations). Using the results presented in this paper, one may not only study the overall satisfaction of these restrictions, but also estimate the impact of changes in the underlying model (for example, if a given line becomes slower due to some repairs, one may evaluate the decrease in various dependability measures without changing the constructed DTA).

*Proof techniques.* For a given SMP  $\mathcal{M}$  and a given DTA  $\mathcal{A}$  we first construct their synchronized product  $\mathcal{M} \times \mathcal{A}$ , which is another stochastic process. In fact, it turns out that  $\mathcal{M} \times \mathcal{A}$  is a discrete-time Markov chain with uncountable state-space. Then, we apply a variant of the standard region construction [2] and thus partition the state-space of  $\mathcal{M} \times \mathcal{A}$  into finitely many equivalence classes. At the very core of our paper there are several non-trivial observations about the structure of  $\mathcal{M} \times \mathcal{A}$  and its region graph which establish a powerful link to the well-developed ergodic theory of Markov chains with general state-space (see, e.g., [18, 20]). In this way, we obtain the results of items (1) and (2) mentioned above. Some additional work is required to analyze the algorithm presented in Section 4 (whose properties are summarized in item (3) above). Due to space constraints, most of the proofs are omitted and can be found in [10].

*Related work.* There is a vast literature on continuous-time Markov chains, semi-Markov processes, or even more general stochastic models such as generalized semi-Markov processes (we refer to, e.g., [21, 6, 19, 17]). In the computer science context, most works on continuous-time stochastic models concern model-checking against a given class of temporal properties [3, 5]. The usefulness of CSL model-checking for dependability analysis is advocated in [15]. Timed automata [2] have been originally used as a model of (non-stochastic) real-time systems. Probabilistic semantics of

timed automata is proposed in [4, 7]. The idea of using timed automata as a specification language for continuous-time stochastic processes is relatively recent. In [13], the model-checking problem for continuous-time Markov chains and linear-time properties represented by timed automata is considered (the task is to determine the probability of all timed words that are accepted by a given timed automaton). A more general model of two-player games over generalized semi-Markov processes with qualitative reachability objectives specified by deterministic timed automata is studied in [11].

## 2. PRELIMINARIES

In this paper, the sets of all positive integers, non-negative integers, real numbers, positive real numbers, and non-negative real numbers are denoted by  $\mathbb{N}$ ,  $\mathbb{N}_0$ ,  $\mathbb{R}$ ,  $\mathbb{R}_{>0}$ , and  $\mathbb{R}_{\geq 0}$ , respectively.

Let  $A$  be a finite or countably infinite set. A *discrete probability distribution* on  $A$  is a function  $\alpha : A \rightarrow \mathbb{R}_{\geq 0}$  such that  $\sum_{a \in A} \alpha(a) = 1$ . We say that  $\alpha$  is *rational* if  $\alpha(a)$  is rational for every  $a \in A$ . The set of all distributions on  $A$  is denoted by  $\mathcal{D}(A)$ . A  $\sigma$ -field over a set  $\Omega$  is a set  $\mathcal{F} \subseteq 2^\Omega$  that includes  $\Omega$  and is closed under complement and countable union. A *measurable space* is a pair  $(\Omega, \mathcal{F})$  where  $\Omega$  is a set called *sample space* and  $\mathcal{F}$  is a  $\sigma$ -field over  $\Omega$  whose elements are called *measurable sets*. A *probability measure* over a measurable space  $(\Omega, \mathcal{F})$  is a function  $\mathcal{P} : \mathcal{F} \rightarrow \mathbb{R}_{\geq 0}$  such that, for each countable collection  $\{X_i\}_{i \in I}$  of pairwise disjoint elements of  $\mathcal{F}$ ,  $\mathcal{P}(\bigcup_{i \in I} X_i) = \sum_{i \in I} \mathcal{P}(X_i)$ , and moreover  $\mathcal{P}(\Omega) = 1$ . A *probability space* is a triple  $(\Omega, \mathcal{F}, \mathcal{P})$ , where  $(\Omega, \mathcal{F})$  is a measurable space and  $\mathcal{P}$  is a probability measure over  $(\Omega, \mathcal{F})$ . We say that a property  $A \subseteq \Omega$  holds *for almost all* elements of a measurable set  $Y$  if  $\mathcal{P}(Y) > 0$ ,  $A \cap Y \in \mathcal{F}$ , and  $\mathcal{P}(A | Y) = 1$ .

All of the integrals used in this paper should be understood as Lebesgue integrals, although we use Riemann-like notation when appropriate.

### 2.1 Semi-Markov processes

A semi-Markov process (see, e.g., [21]) can be seen as discrete-time Markov chains where each transition is equipped with a density function specifying the distribution of time needed to perform the transition. Formally, let  $\mathfrak{D}$  be a set of delay densities, i.e., measurable functions  $f : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$  satisfying  $\int_0^\infty f(t) dt = 1$  where  $f(t) = 0$  for every  $t < 0$ . Moreover, for technical reasons, we assume that each  $f \in \mathfrak{D}$  satisfies the following: There is an interval  $I$  either of the form  $[\ell, u]$  with  $\ell, u \in \mathbb{N}_0$ ,  $\ell < u$ , or  $[\ell, \infty)$  with  $\ell \in \mathbb{N}_0$ , such that

- for all  $t \in \mathbb{R} \setminus I$  we have that  $f(t) = 0$ ,
- for all  $[c, d] \subseteq I$  there is  $b > 0$  such that for all  $t \in [c, d]$  we have that  $f(t) \geq b$ .

The assumption that  $\ell, u$  are natural numbers is adopted only for the sake of simplicity. Our results can easily be generalized to the setting where  $I$  is an interval with rational bounds or even a finite union of such intervals.

**DEFINITION 2.1.** A semi-Markov process (SMP) is a tuple  $\mathcal{M} = (S, \mathbf{P}, \mathbf{D}, \alpha_0)$ , where  $S$  is a finite set of states,  $\mathbf{P} : S \rightarrow \mathcal{D}(S)$  is a transition probability function,  $\mathbf{D} : S \times S \rightarrow \mathfrak{D}$  is a delay function which to each transition assigns its delay density, and  $\alpha_0 \in \mathcal{D}(S)$  is an initial distribution.

A computation (run) of a SMP  $\mathcal{M}$  is initiated in some state  $s_0$ , which is chosen randomly according to  $\alpha_0$ . In the current state  $s_i$ , the next state  $s_{i+1}$  is selected randomly according to the distribution  $\mathbf{P}(s_i)$ , and the selected transition  $(s_i, s_{i+1})$  takes a random time  $t_i$  chosen according to the density  $\mathbf{D}(s_i, s_{i+1})$ . Hence, each run of  $\mathcal{M}$  is an infinite timed word  $s_0 t_0 s_1 t_1 \dots$ , where  $s_i \in S$  and  $t_i \in \mathbb{R}_{\geq 0}$  for all  $i \in \mathbb{N}_0$ . We use  $\mathcal{R}_{\mathcal{M}}$  to denote the set of all runs of  $\mathcal{M}$ .

Now we define a probability space  $(\mathcal{R}_{\mathcal{M}}, \mathcal{F}_{\mathcal{M}}, \mathcal{P}_{\mathcal{M}})$  over the runs of  $\mathcal{M}$  (we often omit the index  $\mathcal{M}$  if it is clear from the context). A *template* is a finite sequence of the form  $B = s_0 I_0 s_1 I_1 \dots s_{n+1}$  such that  $n \geq 0$  and  $I_i$  is an interval in  $\mathbb{R}_{\geq 0}$  for every  $0 \leq i \leq n$ . Each such  $B$  determines the corresponding *cylinder*  $\mathcal{R}(B) \subseteq \mathcal{R}$  consisting of all runs of the form  $\hat{s}_0 t_0 \hat{s}_1 t_1 \dots$ , where  $\hat{s}_i = s_i$  for all  $0 \leq i \leq n+1$ , and  $t_i \in I_i$  for all  $0 \leq i \leq n$ . The  $\sigma$ -field  $\mathcal{F}$  is the Borel  $\sigma$ -field generated by all cylinders. For each template  $B = s_0 I_0 s_1 I_1 \dots s_{n+1}$ , let  $p_i = \mathbf{P}(s_i)(s_{i+1})$  and  $f_i = \mathbf{D}(s_i, s_{i+1})$  for all  $0 \leq i \leq n$ . The probability  $\mathcal{P}(\mathcal{R}(B))$  is defined as follows:

$$\alpha_0(s_0) \cdot \prod_{i=0}^n p_i \cdot \int_{t_i \in I_i} f_i(t_i) dt_i$$

Then,  $\mathcal{P}$  is extended to  $\mathcal{F}$  (in the unique way) by applying the extension theorem (see, e.g., [8]).

### 2.2 Deterministic timed automata

Let  $\mathcal{X}$  be a finite set of *clocks*. A *valuation* is a function  $\nu : \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$ . For every valuation  $\nu$  and every subset  $X \subseteq \mathcal{X}$  of clocks, we use  $\nu[X := \mathbf{0}]$  to denote the unique valuation such that  $\nu[X := \mathbf{0}](x)$  is equal either to 0 or  $\nu(x)$ , depending on whether  $x \in X$  or not, respectively. Further, for every valuation  $\nu$  and every  $\delta \in \mathbb{R}_{\geq 0}$ , the symbol  $\nu + \delta$  denotes the unique valuation such that  $(\nu + \delta)(x) = \nu(x) + \delta$  for all  $x \in \mathcal{X}$ . Sometimes we assume an implicate linear ordering on clocks and slightly abuse our notation by identifying a valuation  $\nu$  with the associated vector of reals.

A *clock constraint* (or *guard*) is a finite conjunction of basic constraints of the form  $x \bowtie c$ , where  $x \in \mathcal{X}$ ,  $\bowtie \in \{<, \leq, >, \geq\}$ , and  $c \in \mathbb{N}_0$ . For every valuation  $\nu$  and every clock constraint  $g$  we have that  $\nu$  either does or does not satisfy  $g$ , written  $\nu \models g$  or  $\nu \not\models g$ , respectively (the satisfaction relation is defined in the expected way). Sometimes we identify a guard  $g$  with the set of all valuations that satisfy  $g$  and write, e.g.,  $g \cap g'$ . The set of all guards over  $\mathcal{X}$  is denoted by  $\mathcal{B}(\mathcal{X})$ .

**DEFINITION 2.2.** A deterministic timed automaton (DTA) is a tuple  $\mathcal{A} = (Q, \Sigma, \mathcal{X}, \longrightarrow, q_0)$ , where  $Q$  is a nonempty finite set of locations,  $\Sigma$  is a finite alphabet,  $\mathcal{X}$  is a finite set of clocks,  $q_0 \in Q$  is an initial location, and  $\longrightarrow \subseteq Q \times \Sigma \times \mathcal{B}(\mathcal{X}) \times 2^{\mathcal{X}} \times Q$  is an edge relation such that for all  $q \in Q$  and  $a \in \Sigma$  we have the following:

1. the guards are deterministic, i.e., for all edges of the form  $(q, a, g_1, X_1, q_1)$  and  $(q, a, g_2, X_2, q_2)$  such that  $g_1 \cap g_2 \neq \emptyset$  we have that  $g_1 = g_2$ ,  $X_1 = X_2$ , and  $q_1 = q_2$ ;
2. the guards are total, i.e., for all  $q \in Q$ ,  $a \in \Sigma$ , and every valuation  $\nu$  there is an edge  $(q, a, g, X, q')$  such that  $\nu \models g$ .

A *configuration* of  $\mathcal{A}$  is a pair  $(q, \nu)$ , where  $q \in Q$  and  $\nu$  is a valuation. An *infinite timed word* over  $\Sigma$  is an infinite sequence  $w = c_0 c_1 c_2 c_3 \dots$ , where  $c_i \in \Sigma$  when  $i$  is even, and

$c_i \in \mathbb{R}_{\geq 0}$  when  $i$  is odd. The *run* of  $\mathcal{A}$  on  $w$  is the unique infinite sequence of configurations  $\mathcal{A}(w) = (q_0, \nu_0)(q_1, \nu_1) \cdots$  such that  $q_0$  is the initial location of  $\mathcal{A}$ ,  $\nu_0(x) = 0$  for all  $x \in \mathcal{X}$ , and for each  $i \in \mathbb{N}_0$  we have that

- if  $c_i$  is a time stamp, then  $q_{i+1} = q_i$  and  $\nu_{i+1} = \nu_i + c_i$ ;
- if  $c_i$  is a letter of  $\Sigma$ , then there is a unique edge  $(q_i, c_i, g, X, q)$  such that  $\nu_i \models g$ , and we require that  $q_{i+1} = q$  and  $\nu_{i+1} = \nu_i[X := \mathbf{0}]$ .

Notice that we do not define any acceptance condition for DTA. Instead, we understand DTA as finite-state *observers* that analyze timed words and report about certain events by entering designated locations. The “frequency” of these events is formally captured by the quantities  $d_q$  and  $c_q$  defined below.

Let  $\mathcal{A} = (Q, \Sigma, \mathcal{X}, \longrightarrow, q_0)$  be a DTA,  $q \in Q$  some location, and  $w = a_0 t_0 a_1 t_1 \cdots$  a timed word over  $\Sigma$ . For every  $i \in \mathbb{N}_0$ , let  $T^i(w)$  be the stamp  $t_i$  of  $w$ , and  $Q^i(w)$  the unique location of  $\mathcal{A}$  entered after reading the finite prefix  $a_0 t_0 \cdots a_i$  of  $w$ . Further, let  $1_q^i(w)$  be either 1 or 0 depending on whether  $Q^i(w) = q$  or not, respectively. The *discrete* and *timed* frequency of visits to  $q$  along  $\mathcal{A}(w)$ , denoted by  $\mathbf{d}_q^{\mathcal{A}}(w)$  and  $\mathbf{c}_q^{\mathcal{A}}(w)$ , are defined in the following way (if  $\mathcal{A}$  is clear, it is omitted):

$$\begin{aligned} \mathbf{d}_q^{\mathcal{A}}(w) &= \limsup_{n \rightarrow \infty} \frac{\sum_{i=1}^n 1_q^i(w)}{n} \\ \mathbf{c}_q^{\mathcal{A}}(w) &= \limsup_{n \rightarrow \infty} \frac{\sum_{i=1}^n T^i(w) \cdot 1_q^i(w)}{\sum_{i=1}^n T^i(w)} \end{aligned}$$

Hence, every timed word  $w$  determines the tuple  $\mathbf{d}^{\mathcal{A}} = (\mathbf{d}_q^{\mathcal{A}}(w))_{q \in Q}$  and the tuple  $\mathbf{c}^{\mathcal{A}} = (\mathbf{c}_q^{\mathcal{A}}(w))_{q \in Q}$  of *discrete* and *timed*  $\mathcal{A}$ -measures, respectively. The  $\mathcal{A}$ -measures were defined using limsup, because the corresponding limits may not exist in general. If  $\lim_{n \rightarrow \infty} \sum_{i=1}^n 1_q^i(w)/n$  exists for all  $q \in Q$ , we say that  $\mathbf{d}^{\mathcal{A}}$  is *well-defined* for  $w$ . Similarly, if  $\lim_{n \rightarrow \infty} (\sum_{i=1}^n T^i(w) \cdot 1_q^i(w)) / (\sum_{i=1}^n T^i(w))$  exists for all  $q$ , we say that  $\mathbf{c}^{\mathcal{A}}$  is *well-defined* for  $w$ .

As we already noted in Section 1, a DTA  $\mathcal{A}$  can be used to observe runs in a given SMP  $\mathcal{M}$  after labeling all states of  $\mathcal{M}$  with the letters of  $\Sigma$  by a suitable  $L : S \rightarrow \Sigma$ . Then, every run  $\sigma = s_0 t_0 s_1 t_1 \cdots$  of  $\mathcal{M}$  determines a unique timed word  $w_\sigma = L(s_0) t_0 L(s_1) t_1 \cdots$ , and one can easily show that for every timed word  $w$ , the set  $\{\sigma \in \mathcal{R} \mid w_\sigma = w\}$  is measurable in  $(\mathcal{R}, \mathcal{F}, \mathcal{P})$ .

### 3. DTA MEASURES OVER SMPS

Throughout this section we fix an SMP  $\mathcal{M} = (S, \mathbf{P}, \mathbf{D}, \alpha_0)$  and a DTA  $\mathcal{A} = (Q, \Sigma, \mathcal{X}, \longrightarrow, q_0)$  where  $\mathcal{X} = \{x_1, \dots, x_n\}$ . To simplify our notation, we assume that  $\Sigma = S$ , i.e., every run  $\sigma$  of  $\mathcal{M}$  is a timed word over  $\Sigma$  (hence, we do not need to introduce any labeling  $L : S \rightarrow \Sigma$ ). This technical assumption does not affect the generality of our results (all of our arguments and proofs work exactly as they are, we only need to rewrite them using less readable notation). Our goal is to prove the following:

THEOREM 3.1.

1.  $\mathbf{d}^{\mathcal{A}}$  is well-defined for almost all runs of  $\mathcal{M}$ .
2. There are pairwise disjoint sets  $\mathcal{R}_1, \dots, \mathcal{R}_k$  of runs in  $\mathcal{M}$  such that  $\mathcal{P}(\mathcal{R}_1 \cup \dots \cup \mathcal{R}_k) = 1$ , and for every  $1 \leq j \leq k$  there is a tuple  $D_j$  such that  $\mathbf{d}^{\mathcal{A}}(\sigma) =$

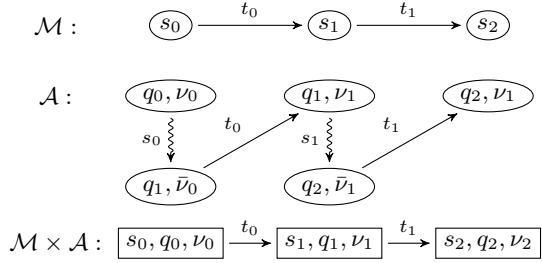


Figure 2: Synchronizing  $\mathcal{M}$  and  $\mathcal{A}$  in  $\mathcal{M} \times \mathcal{A}$ . Notice that  $\nu_0 = \bar{\nu}_0 = \mathbf{0}$  and  $\nu_{i+1} = \bar{\nu}_i + t_i$ .

$D_j$  for almost all  $\sigma \in \mathcal{R}_j$  (we use  $D_{j,q}$  to denote the  $q$ -component of  $D_j$ ).

In Section 4, we show how to compute the  $k$  and approximate  $\mathcal{P}(\mathcal{R}_j)$  and  $D_j$  up to an arbitrarily small given precision.

An immediate corollary of Theorem 3.1 is an analogous result for  $\mathbf{c}^{\mathcal{A}}$ .

COROLLARY 3.2.  $\mathbf{c}^{\mathcal{A}}$  is well-defined for almost all runs of  $\mathcal{M}$ . Further, there are pairwise disjoint sets  $\mathcal{R}_1, \dots, \mathcal{R}_K$  of runs in  $\mathcal{M}$  such that  $\mathcal{P}(\mathcal{R}_1 \cup \dots \cup \mathcal{R}_K) = 1$ , and for every  $1 \leq j \leq K$  there is a tuple  $C_j$  such that  $\mathbf{c}^{\mathcal{A}}(\sigma) = C_j$  for almost all  $\sigma \in \mathcal{R}_j$ .

Corollary 3.2 follows from Theorem 3.1 simply by considering the discrete  $\mathbf{d}^{S \times \mathcal{A}}$  measure, where the DTA  $S \times \mathcal{A}$  is obtained from  $\mathcal{A}$  in the following way: the set of locations of  $S \times \mathcal{A}$  is  $\{q_0\} \cup (S \times Q)$ , and for every transition  $(q_0, s, g, X, q')$  of  $\mathcal{A}$  we add a transition  $(q_0, s, g, X, (s, q'))$  to  $S \times \mathcal{A}$  and for every transition  $(q, s, g, X, q')$  and every  $s' \in S$  we add a transition  $((s', q), s, g, X, (s, q'))$  to  $S \times \mathcal{A}$ . The initial location of  $S \times \mathcal{A}$  is  $q_0$ . Intuitively,  $S \times \mathcal{A}$  is the same as  $\mathcal{A}$  but it explicitly “remembers” the letter which was used to enter the current location. Let  $k$  and  $D_j$  be the constants of Theorem 3.1 constructed for  $\mathcal{M}$  and  $S \times \mathcal{A}$ . Observe that the expected time of performing a transition from a given  $s \in S$ , denoted by  $E_s$ , is given by  $E_s = \sum_{s' \in S} \mathbf{P}(s)(s') \cdot E_{s,s'}$ , where  $E_{s,s'}$  is the expectation of a random variable with the density  $\mathbf{D}(s, s')$ . From this we easily obtain that

$$C_{j,q} = \frac{\sum_{s \in S} E_s \cdot D_{j,(s,q)}}{\sum_{p \in Q} \sum_{s \in S} E_s \cdot D_{j,(s,p)}} \quad (1)$$

for all  $q \in Q$  and  $1 \leq j \leq k$ . The details are given in [10]. Hence, we can also compute the constant  $K$  and approximate  $\mathcal{P}(\mathcal{R}_j)$  and  $C_j$  for every  $1 \leq j \leq K$  using Equation (1).

It remains to prove Theorem 3.1. Let us start by sketching the overall structure of our proof. First, we construct a synchronous product  $\mathcal{M} \times \mathcal{A}$  of  $\mathcal{M}$  and  $\mathcal{A}$ , which is a Markov chain with an uncountable state space  $\Gamma_{\mathcal{M} \times \mathcal{A}} = S \times Q \times (\mathbb{R}_{\geq 0})^n$ . Intuitively,  $\mathcal{M} \times \mathcal{A}$  behaves in the same way as  $\mathcal{M}$  and simulates the computation of  $\mathcal{A}$  on-the-fly (see Figure 2). Then, we construct a finite region graph  $G_{\mathcal{M} \times \mathcal{A}}$  over the product  $\mathcal{M} \times \mathcal{A}$ . The nodes of  $G_{\mathcal{M} \times \mathcal{A}}$  are the sets of states that, roughly speaking, satisfy the same guards of  $\mathcal{A}$ . Edges are induced by transitions of the product (note that if two states satisfy the same guards, the sets of enabled outgoing transitions are the same). By relying on arguments presented in [1, 11], we show that almost all runs reach a node of a bottom strongly connected component (BSCC)  $\mathcal{C}$  of  $G_{\mathcal{M} \times \mathcal{A}}$  (by definition, each run which enters  $\mathcal{C}$  remains in  $\mathcal{C}$ ). This gives us the partition of the set of runs

of  $\mathcal{M}$  into the sets  $\mathcal{R}_1, \dots, \mathcal{R}_k$  (each  $\mathcal{R}_j$  corresponds to one of the BSCCs of  $G_{\mathcal{M} \times \mathcal{A}}$ ).

Subsequently, we concentrate on a fixed BSCC  $\mathcal{C}$ , and prove that almost all runs that reach  $\mathcal{C}$  have the same frequency of visits to a given  $q \in Q$  (this gives us the constant  $D_{j,q}$ ). Here we employ several deep results from the theory of general state space Markov chains (see Theorem 3.6). To apply these results, we prove that assuming aperiodicity of  $G_{\mathcal{M} \times \mathcal{A}}$  (see Definition 3.10), the state space of the product  $\mathcal{M} \times \mathcal{A}$  is *small* (see Definition 3.5 and Lemma 3.11 below). This is perhaps the most demanding part of our proof. Roughly speaking, we show that there is a distinguished subset of states reachable from each state in a fixed number of steps with probability bounded from 0. By applying Theorem 3.6, we obtain a complete invariant distribution on the product, i.e., in principle, we obtain a constant frequency of any non-trivial subset of states. From this we derive our results in a straightforward way. If  $G_{\mathcal{M} \times \mathcal{A}}$  is periodic, we use standard techniques for removing periodicity and then basically follow the same stream of arguments as in the aperiodic case.

### 3.1 General state space Markov chains

We start by recalling the definition of “ordinary” discrete-time Markov chains with discrete state space (DTMC). A DTMC is given by a finite or countably infinite state space  $S$ , an initial probability distribution over  $S$ , and a one-step transition matrix  $P$  which defines the probability  $P(s, s')$  of every transition  $(s, s') \in S \times S$  so that  $\sum_{s' \in S} P(s, s') = 1$  for every  $s \in S$ . In the setting of uncountable state spaces, transition probabilities cannot be specified by a transition matrix. Instead, one defines the probabilities of moving from a given state  $s$  to a given measurable subset  $X$  of states. Hence, the concept of transition matrix is replaced with a more general notion of *transition kernel* defined below.

**DEFINITION 3.3.** *A transition kernel over a measurable space  $(\Gamma, \mathcal{G})$  is a function  $P : \Gamma \times \mathcal{G} \rightarrow [0, 1]$  such that*

1.  $P(z, \cdot)$  is a probability measure over  $(\Gamma, \mathcal{G})$  for each  $z \in \Gamma$ ;
2.  $P(\cdot, A)$  is a measurable function for each  $A \in \mathcal{G}$  (i.e., for every  $c \in \mathbb{R}$ , the set of all  $z \in \Gamma$  satisfying  $P(z, A) \geq c$  belongs to  $\mathcal{G}$ ).

A transition kernel is the core of the following definition.

**DEFINITION 3.4.** *A general state space Markov chain (GSSMC) with a state space  $(\Gamma, \mathcal{G})$ , a transition kernel  $P$  and an initial probability measure  $\mu$  is a stochastic process  $\Phi = \Phi_1, \Phi_2, \dots$  such that each  $\Phi_i$  is a random variable over a probability space  $(\Omega_\Phi, \mathcal{F}_\Phi, \mathcal{P}_\Phi)$  where*

- $\Omega_\Phi$  is a set of runs, i.e., infinite words over  $\Gamma$ .
- $\mathcal{F}_\Phi$  is the product  $\sigma$ -field  $\bigotimes_{i=0}^\infty \mathcal{G}$ .
- $\mathcal{P}_\Phi$  is the unique probability measure over  $(\Omega_\Phi, \mathcal{F}_\Phi)$  such that for every finite sequence  $A_0, \dots, A_n \in \mathcal{F}_\Phi$  we have that  $\mathcal{P}_\Phi(\Phi_0 \in A_0, \dots, \Phi_n \in A_n)$  is equal to

$$\int_{y_0 \in A_0} \dots \int_{y_{n-1} \in A_{n-1}} \mu(dy_0) \cdot P(y_0, dy_1) \dots P(y_{n-1}, A_n). \quad (2)$$

- Each  $\Phi_i$  is the projection of elements of  $\Omega_\Phi$  onto the  $i$ -th component.

A *path* is a finite sequence  $z_1 \dots z_n$  of states from  $\Gamma$ . From Equation (2) we get that  $\Phi$  also satisfies the following properties which will be used to show several results about the chain  $\Phi$  by working with the transition kernel only.

1.  $\mathcal{P}_\Phi(\Phi_0 \in A_0) = \mu(A_0)$ ,
2.  $\mathcal{P}_\Phi(\Phi_{n+1} \in A \mid \Phi_n, \dots, \Phi_0) = \mathcal{P}_\Phi(\Phi_{n+1} \in A \mid \Phi_n) = P(\Phi_n, A)$  almost surely,
3.  $\mathcal{P}_\Phi(\Phi_{n+m} \in A \mid \Phi_n) = P^m(\Phi_n, A)$  almost surely,

where the  $m$ -step transition kernel  $P^m$  is defined as follows:

$$P^1(z, A) = P(z, A)$$

$$P^{i+1}(z, A) = \int_{\Gamma} P(z, dy) \cdot P^i(y, A).$$

Notice that the transition kernel and the  $m$ -step transition kernel are analogous counterparts to the transition matrix and the  $k$ -step transition matrix of a DTMC.

As we mentioned above, our proof of Theorem 3.1 employs several results of GSSMC theory. In particular, we make use of the notion of *smallness* of the state space defined as follows.

**DEFINITION 3.5.** *Let  $m \in \mathbb{N}$ ,  $\varepsilon > 0$ , and  $\nu$  be a probability measure on  $\mathcal{G}$ . A set  $C \in \mathcal{G}$  is  $(m, \varepsilon, \nu)$ -small if for all  $x \in C$  and  $B \in \mathcal{G}$  we have that  $P^m(x, B) \geq \varepsilon \cdot \nu(B)$ .*

GSSMCs where the whole state space is small have many nice properties, and the relevant ones are summarized in the following theorem.

**THEOREM 3.6.** *If  $\Gamma$  is  $(m, \varepsilon, \nu)$ -small, then*

1. **[Existence of invariant measure]** *There exists a unique probability measure  $\pi$  such that for all  $A \in \mathcal{G}$  we have that*

$$\pi(A) = \int_{\Gamma} \pi(dx) P(x, A)$$

2. **[Strong law of large numbers]** *If  $h : \Gamma \rightarrow \mathbb{R}$  satisfies  $\int_{\Gamma} h(x) \pi(dx) < \infty$ , then almost surely*

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n h(\Phi_i)}{n} = \int_{\Gamma} h(x) \pi(dx)$$

3. **[Uniform ergodicity]** *For all  $x \in \Gamma$ ,  $A \in \mathcal{G}$ , and all  $n \in \mathbb{N}$ ,*

$$\sup_{A \in \mathcal{G}} |P^n(x, A) - \pi(A)| \leq (1 - \varepsilon)^{\lfloor n/m \rfloor}$$

**PROOF.** The theorem is a consequence of standard results for GSSMCs. Since  $\Gamma$  is  $(m, \varepsilon, \nu)$ -small, we have

- (i)  $\Phi$  is by definition  $\varphi$ -irreducible for  $\varphi = \nu$ , and thus also  $\psi$ -irreducible by [18, Proposition 4.2.2];
- (ii)  $\Gamma$  is by definition also  $(a, \varepsilon, \nu)$ -petite (see [18, Section 5.5.2]), where  $a$  is the Dirac distribution on  $\mathbb{N}_0$  with  $a(m) = 1$ ,  $a(n) = 0$  for  $n \neq m$ ;
- (iii) the first return time to  $\Gamma$  is trivially 1.

ad 1. By (iii),  $\Gamma$  is not uniformly transient, hence by (i), (ii) and [18, Theorem 8.0.2],  $\Phi$  is recurrent. Thus by [18, Theorem 10.0.1], there exists a unique invariant probability measure  $\pi$ .

ad 2. By (i)-(iii) and [18, Theorem 10.4.10 (ii)],  $\Phi$  is positive Harris. Therefore, we may apply [18, Theorem 17.0.1 (i)] and obtain the desired result.

ad 3. This follows immediately from [20, Theorem 8].  $\square$

### 3.2 The product process

The *product process* of  $\mathcal{M}$  and  $\mathcal{A}$ , denoted by  $\mathcal{M} \times \mathcal{A}$ , is a GSSMC with the state space  $\Gamma_{\mathcal{M} \times \mathcal{A}} = S \times Q \times (\mathbb{R}_{\geq 0})^n$ , where  $n = |\mathcal{X}|$  is the number of clocks of  $\mathcal{A}$ . The  $\sigma$ -field over  $\Gamma_{\mathcal{M} \times \mathcal{A}}$  is the product  $\sigma$ -field  $\mathcal{G}_{\mathcal{M} \times \mathcal{A}} = 2^S \otimes 2^Q \otimes \mathfrak{B}^n$  where  $\mathfrak{B}^n$  is the Borel  $\sigma$ -field over the set  $(\mathbb{R}_{\geq 0})^n$ . For each  $A \in \mathcal{G}_{\mathcal{M} \times \mathcal{A}}$ , the initial probability  $\mu_{\mathcal{M} \times \mathcal{A}}(A)$  is equal to  $\sum_{(s, q_0, \mathbf{0}) \in A} \alpha_0(s)$  (recall that  $\alpha_0$  is the initial distribution of  $\mathcal{M}$ ).

The behavior of  $\mathcal{M} \times \mathcal{A}$  is depicted in Figure 2. Each step of the product process corresponds to one step of  $\mathcal{M}$  and two steps of  $\mathcal{A}$ . The step of the product starts by simulating the discrete step of  $\mathcal{A}$  that reads the current state of  $\mathcal{M}$  and possibly resets some clocks, followed by simulating simultaneously the step of  $\mathcal{M}$  that takes time  $t$  and the corresponding step of  $\mathcal{A}$  which reads the time stamp  $t$ .

Now we define the transition kernel  $P_{\mathcal{M} \times \mathcal{A}}$  of the product process. Let  $z = (s, q, \nu)$  be a state of  $\Gamma_{\mathcal{M} \times \mathcal{A}}$ , and let  $(\bar{q}, \bar{\nu})$  be the configuration of  $\mathcal{A}$  entered from the configuration  $(q, \nu)$  after reading  $s$  (note that  $\bar{\nu}$  is not necessarily the same as  $\nu$  because  $\mathcal{A}$  may reset some clocks). It suffices to define  $P_{\mathcal{M} \times \mathcal{A}}(z, \cdot)$  only for generators of  $\mathcal{G}_{\mathcal{M} \times \mathcal{A}}$  and then apply the extension theorem (see, e.g., [8]) to obtain a unique probability measure  $P_{\mathcal{M} \times \mathcal{A}}(z, \cdot)$  over  $(\Gamma_{\mathcal{M} \times \mathcal{A}}, \mathcal{G}_{\mathcal{M} \times \mathcal{A}})$ . Generators of  $\mathcal{G}_{\mathcal{M} \times \mathcal{A}}$  are sets of the form  $\{s'\} \times \{q'\} \times \mathbf{I}$  where  $s' \in S$ ,  $q' \in Q$  and  $\mathbf{I}$  is the product  $I_1 \times \dots \times I_n$  of intervals  $I_i$  in  $\mathbb{R}_{\geq 0}$ . If  $q' \neq \bar{q}$ , then we define  $P_{\mathcal{M} \times \mathcal{A}}(z, \{s'\} \times \{q'\} \times \mathbf{I}) = 0$ . Otherwise, we define

$$P_{\mathcal{M} \times \mathcal{A}}(z, \{s'\} \times \{q'\} \times \mathbf{I}) = \mathbf{P}(s)(s') \cdot \int_0^\infty f(t) \cdot \mathbf{1}_{\mathbf{I}}(\bar{\nu} + t) dt$$

Here  $f = \mathbf{D}(s, s')$  and  $\mathbf{1}_{\mathbf{I}}$  is the indicator function of the set  $\mathbf{I}$ .

Since  $P_{\mathcal{M} \times \mathcal{A}}(z, \cdot)$  is by definition a probability measure over  $(\Gamma_{\mathcal{M} \times \mathcal{A}}, \mathcal{G}_{\mathcal{M} \times \mathcal{A}})$ , it remains to check the second condition of Definition 3.3.

LEMMA 3.7. *Let  $A \in \mathcal{G}_{\mathcal{M} \times \mathcal{A}}$ . Then  $P_{\mathcal{M} \times \mathcal{A}}(\cdot, A)$  is a measurable function, i.e.,  $\mathcal{M} \times \mathcal{A}$  is a GSSMC.*

A proof of this lemma can be found in [10]. Recall that by Definition 3.4,  $\mathcal{P}_{\mathcal{M} \times \mathcal{A}}$  is the unique probability measure on the product  $\sigma$ -field  $\mathcal{F}_{\mathcal{M} \times \mathcal{A}} = \bigotimes_{i=0}^\infty \mathcal{G}_{\mathcal{M} \times \mathcal{A}}$  induced by  $P_{\mathcal{M} \times \mathcal{A}}$  and the initial probability measure  $\mu_{\mathcal{M} \times \mathcal{A}}$ .

#### 3.2.1 The correspondence between $\mathcal{M} \times \mathcal{A}$ and $\mathcal{M}$

In this subsection we show that  $\mathcal{M} \times \mathcal{A}$  correctly reflects the behaviour of  $\mathcal{M}$ . First, we define the  $\mathbf{d}^{\mathcal{A}}$  measure for  $\mathcal{M} \times \mathcal{A}$ . (As the DTA  $\mathcal{A}$  is fixed, we omit them and write  $\mathbf{d}$  and  $\mathbf{d}_q$  instead of  $\mathbf{d}^{\mathcal{A}}$  and  $\mathbf{d}_q^{\mathcal{A}}$ , respectively.) Let  $\sigma = (s_0, q_0, \nu_0)(s_1, q_1, \nu_1) \dots$  be a run of  $\mathcal{M} \times \mathcal{A}$  and  $q \in Q$  a location. For every  $i \in \mathbb{N}_0$ , let  $\mathbf{1}_q^i(\sigma)$  be either 1 or 0 depending on whether if  $q_i = q$  or not, respectively. We put

$$\mathbf{d}_q(\sigma) = \limsup_{n \rightarrow \infty} \frac{\sum_{i=1}^n \mathbf{1}_q^i(\sigma)}{n}$$

LEMMA 3.8. *There is a measurable one-to-one mapping  $\xi$  from the set of runs of  $\mathcal{M}$  to the set of runs of  $\mathcal{M} \times \mathcal{A}$  such that*

- $\xi$  preserves measure, i.e., for every measurable set  $X$  of runs of  $\mathcal{M}$  we have that  $\xi(X)$  is also measurable and  $\mathcal{P}_{\mathcal{M}}(X) = \mathcal{P}_{\mathcal{M} \times \mathcal{A}}(\xi(X))$ ;

- $\xi$  preserves  $\mathbf{d}$ , i.e., for every run  $\sigma$  of  $\mathcal{M}$  and every  $q \in Q$  we have that  $\mathbf{d}_q(\sigma)$  is well-defined iff  $\mathbf{d}_q(\xi(\sigma))$  is well-defined, and  $\mathbf{d}_q(\sigma) = \mathbf{d}_q(\xi(\sigma))$ .

A formal proof of Lemma 3.8 is given in [10].

#### 3.2.2 The region graph of $\mathcal{M} \times \mathcal{A}$

Although the state-space  $\Gamma_{\mathcal{M} \times \mathcal{A}}$  is uncountable, we can define the standard *region relation*  $\sim$  [2] over  $\Gamma_{\mathcal{M} \times \mathcal{A}}$  with finite index, and then work with finitely many *regions*. For a given  $a \in \mathbb{R}$ , we use  $\text{frac}(a)$  to denote the fractional part of  $a$ , and  $\text{int}(a)$  to denote the integral part of  $a$ . For  $a, b \in \mathbb{R}$ , we say that  $a$  and  $b$  agree on integral part if  $\text{int}(a) = \text{int}(b)$  and neither or both  $a, b$  are integers.

We denote by  $B_{\max}$  the maximal constant that appears in the guards of  $\mathcal{A}$  and say that a clock  $x \in \mathcal{X}$  is *relevant* for  $\nu$  if  $\nu(x) \leq B_{\max}$ . Finally, we put  $(s_1, q_1, \nu_1) \sim (s_2, q_2, \nu_2)$  if

- $s_1 = s_2$  and  $q_1 = q_2$ ;
- for all relevant  $x \in \mathcal{X}$  we have that  $\nu_1(x)$  and  $\nu_2(x)$  agree on integral parts;
- for all relevant  $x, y \in \mathcal{X}$  we have that  $\text{frac}(\nu_1(x)) \leq \text{frac}(\nu_1(y))$  iff  $\text{frac}(\nu_2(x)) \leq \text{frac}(\nu_2(y))$ .

Note that  $\sim$  is an equivalence with finite index. The equivalence classes of  $\sim$  are called *regions*. Observe that states in the same region have the same behavior with respect to qualitative reachability. This is formalized in the following lemma.

LEMMA 3.9. *Let  $R$  and  $T$  be regions and  $z, z' \in R$ . Then  $P_{\mathcal{M} \times \mathcal{A}}(z, T) > 0$  iff  $P_{\mathcal{M} \times \mathcal{A}}(z', T) > 0$ .*

A proof of Lemma 3.9 can be found in [11]. Further, we define a finite *region graph*  $G_{\mathcal{M} \times \mathcal{A}} = (V, E)$  where the set of vertices  $V$  is the set of regions and for every pair of regions  $R, R'$  there is an edge  $(R, R') \in E$  iff  $P_{\mathcal{M} \times \mathcal{A}}(z, R') > 0$  for some  $z \in R$  (due to Lemma 3.9, the concrete choice of  $z$  is irrelevant). For technical reasons, we assume that  $V$  contains only regions reachable with positive probability in  $\mathcal{M} \times \mathcal{A}$ .

### 3.3 Finishing the proof of Theorem 3.1

Our proof is divided into three parts. In the first part we consider a general region graph which is not necessarily strongly connected, and show that we can actually concentrate just on its BSCCs. In the second part we study a given BSCC under the aperiodicity assumption. Finally, in the last part we consider a general BSCC which may be periodic. (The second part is included mainly for the sake of readability.)

#### Non-strongly connected region graph

Let  $\mathcal{C}_1, \dots, \mathcal{C}_k$  be the BSCCs of the region graph. The set  $\mathcal{R}_i$  consists of all runs  $\sigma$  of  $\mathcal{M}$  such that  $\xi(\omega)$  visits (a configuration in a region of)  $\mathcal{C}_i$ , where  $\xi$  is the mapping of Lemma 3.8. By applying the arguments of [1, 11], it follows that almost runs in  $\mathcal{M} \times \mathcal{A}$  visit a configuration of a BSCC. By Lemma 3.8,  $\xi$  preserves  $\mathbf{d}$  and the probability  $\mathcal{P}_{\mathcal{M}}(\mathcal{R}_i)$  is equal to the probability of visiting  $\mathcal{C}_i$  in  $\mathcal{M} \times \mathcal{A}$ . Further, since the value of  $\mathbf{d}$  does not depend on a finite prefix of a run, we may safely assume that  $\mathcal{M} \times \mathcal{A}$  is initialized in  $\mathcal{C}_i$  in such a way that the initial distribution corresponds to the conditional distribution of the first visit to  $\mathcal{C}_i$  conditioned on visiting  $\mathcal{C}_i$ .

In a BSCC  $\mathcal{C}_i$ , there may be some *growing* clocks that are never reset. Since the values of growing clocks are just constantly increasing, the product process never returns to a state it has visited before. Therefore, there is no invariant distribution. Observe that all runs initiated in  $\mathcal{C}_i$  eventually reach a configuration where the values of all growing clocks are larger than the maximal constant  $B_{\max}$  employed in the guards of  $\mathcal{A}$ . This means that  $\mathcal{C}_i$  actually consists *only* of regions where all growing clocks are irrelevant (see Section 3.2.2), because  $\mathcal{C}_i$  would not be strongly connected otherwise. Hence, we can safely *remove* every growing clock  $x$  from  $\mathcal{C}_i$ , replacing all guards of the form  $x > c$  or  $x \geq c$  with *true* and all guards of the form  $x < c$  or  $x \leq c$  with *false*. So, from now on we assume that there are no growing clocks in  $\mathcal{C}_i$ .

### Strongly connected & aperiodic region graph

In this part we consider a given BSCC  $\mathcal{C}_i$  of the region graph  $G_{\mathcal{M} \times \mathcal{A}}$ . This is equivalent to assuming that  $G_{\mathcal{M} \times \mathcal{A}}$  is strongly connected and  $\Gamma_{\mathcal{M} \times \mathcal{A}}$  is equal to the union of all regions of  $G_{\mathcal{M} \times \mathcal{A}}$  (recall that  $G_{\mathcal{M} \times \mathcal{A}}$  consists just of regions reachable with positive probability in  $\mathcal{M} \times \mathcal{A}$ ). We also assume that there are no growing clocks (see the previous part). Further, in this subsection we assume that  $G_{\mathcal{M} \times \mathcal{A}}$  is aperiodic in the following sense.

**DEFINITION 3.10.** *A period  $p$  of the region graph  $G_{\mathcal{M} \times \mathcal{A}}$  is the greatest common divisor of lengths of all cycles in  $G_{\mathcal{M} \times \mathcal{A}}$ . The region graph  $G_{\mathcal{M} \times \mathcal{A}}$  is aperiodic if  $p = 1$ .*

The key to proving Theorem 3.1 in the current restricted setting is to show that the state space of  $\mathcal{M} \times \mathcal{A}$  is small (recall Definition 3.5) and then apply Theorem 3.6 (1) and (2) to obtain the required characterization of the long-run behavior of  $\mathcal{M} \times \mathcal{A}$ .

**PROPOSITION 3.11.** *Assume that  $G_{\mathcal{M} \times \mathcal{A}}$  is strongly connected and aperiodic. Then there exist a region  $R$ , a measurable subset  $S \subseteq R$ ,  $n \in \mathbb{N}$ ,  $b > 0$ , and a probability measure  $\kappa$  such that  $\kappa(S) = 1$  and for all measurable  $T \subseteq S$  and  $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$  we have that  $P_{\mathcal{M} \times \mathcal{A}}^n(z, T) > b \cdot \kappa(T)$ . In other words, the set  $\Gamma_{\mathcal{M} \times \mathcal{A}}$  of all states of the GSSMC  $\mathcal{M} \times \mathcal{A}$  is  $(n, b, \kappa)$ -small.*

**PROOF SKETCH.** We show that there exist  $z^* \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ ,  $n \in \mathbb{N}$ , and  $\gamma > 0$  such that for an arbitrary starting state  $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$  there is a path from  $z$  to  $z^*$  of length exactly  $n$  that is  $\gamma$ -wide in the sense that the waiting time of any transition in the path can be changed by  $\pm\gamma$  without ending up in a different region in the end. The target set  $S$  then corresponds to a “neighbourhood” of  $z^*$  within the region of  $z^*$ . Any small enough sub-neighbourhood of  $z^*$  is visited by a set of runs that follow the  $\gamma$ -wide path closely enough. The probability of this set of runs then depends linearly on the size of the sub-neighbourhood when measured by  $\kappa$ , where  $\kappa$  is essentially the Lebesgue measure restricted to  $S$ .

So, it remains to find suitable  $z^*$ ,  $n$ , and  $\gamma$ . For a given starting state  $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ , we construct a path of fixed length  $n$  (independent of  $z$ ) that always ends in the same state  $z^*$ . Further, the path is  $\gamma$ -wide for some  $\gamma > 0$  independent of  $z$ . Technically, the path is obtained by concatenating five sub-paths each of which has a fixed length independent of  $z$ . These sub-paths are described in greater detail below.

In the first sub-path, we move to a  $\delta$ -separated state for some fixed  $\delta > 0$  independent of  $z$ . A state is  $\delta$ -separated if the fractional parts of all relevant clocks are approximately

equally distributed on the  $[0, 1]$  line segment (each two of them have distance at least  $\delta$ ). We can easily build the first sub-path so that it is  $\delta$ -wide.

For the second sub-path, we first fix some region  $R_1$ . Since  $G_{\mathcal{M} \times \mathcal{A}}$  is strongly connected and aperiodic, there is a fixed  $n'$  such that  $R_1$  is reachable from an arbitrary state of  $\Gamma_{\mathcal{M} \times \mathcal{A}}$  in exactly  $n'$  transitions. The second sub-path is chosen as a  $(\delta/n')$ -wide path of length  $n'$  that leads to a  $(\delta/n')$ -separated state of  $R_1$  (we show that such a sub-path is guaranteed to exist; intuitively, the reason why the separation and wideness may decrease proportionally to  $n'$  is that the fractional parts of relevant clock may be forced to move closer and closer to each other by the resets performed along the sub-path).

In the third sub-path, we squeeze the fractional parts of all relevant clocks close to 0. We go through a fixed region path  $R_1 \cdots R_k$  (independent of  $z$ ) so that in each step we shift the time by an integral value minus a small constant  $c$  (note that the fractional parts of clocks reset during this path have fixed relative distances). Thus, we reach a state  $z'_k$  that is “almost fixed” in the sense that the values of all relevant clocks in  $z'_k$  are the same for every starting state  $z$ . Note that the third sub-path is  $c$ -wide. At this point, we should note that if we defined the product process somewhat differently by identifying all states differing only in the values of irrelevant clocks (which does not lead to any technical complications), we would be done, i.e., we could put  $z^* = z'_k$ . We have neglected this possibility mainly for presentation reasons. So, we need two more sub-paths to fix the values of irrelevant clocks.

In the fourth sub-path, we act similarly as in the first sub-path and prepare ourselves for the final sub-path. We reach a  $\delta$ -separated state that is almost equal to a fixed state  $z_\ell \in R_\ell$ . Again, we do it by a  $\delta$ -wide path of a fixed length.

In the fifth sub-path, we follow a fixed region path  $R_\ell \cdots R_{\ell+m}$  such that each clock not relevant in  $R_\ell$  is reset along this path, and hence we reach a fixed state  $z^* \in R_{\ell+m}$ . Here we use our assumption that every clock can be reset to zero (i.e., there are no growing clocks).  $\square$

Now we may finish the proof of Theorem 3.1. By Theorem 3.6 (1), there is a unique invariant distribution  $\pi$  on  $\Gamma_{\mathcal{M} \times \mathcal{A}}$ . For every  $q \in Q$ , we denote by  $A_q$  the set of all states of  $\mathcal{M} \times \mathcal{A}$  of the form  $(s, q, \nu) \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ . By Theorem 3.6 (2), for almost all runs  $\sigma$  of  $\mathcal{M} \times \mathcal{A}$  we have that  $\mathbf{d}(\sigma)$  is well-defined and  $\mathbf{d}_q(\sigma) = \sum \pi(A_q)$ . By Lemma 3.8, we obtain the same for almost all runs of  $\mathcal{M}$ .

### Strongly connected & periodic region graph

Now we consider a general BSCC  $\mathcal{C}_i$  of the region graph  $G_{\mathcal{M} \times \mathcal{A}}$ . Technically, we adopt the same setup as the previous part but remove the aperiodicity condition. That is, we assume that  $G_{\mathcal{M} \times \mathcal{A}}$  is strongly connected,  $\Gamma_{\mathcal{M} \times \mathcal{A}}$  is equal to the union of all regions of  $G_{\mathcal{M} \times \mathcal{A}}$ , and there are no growing clocks.

Let  $p$  be the period of  $G_{\mathcal{M} \times \mathcal{A}}$ . In this case,  $\mathcal{M} \times \mathcal{A}$  is not necessarily small in the sense of Definition 3.5. By employing standard methods for periodic Markov chains, we decompose  $\mathcal{M} \times \mathcal{A}$  into  $p$  stochastic processes  $\Phi_0, \dots, \Phi_{p-1}$  where each  $\Phi_k$  makes steps corresponding to  $p$  steps of the original process  $\mathcal{M} \times \mathcal{A}$  (except for the first step which corresponds just to  $k$  steps of  $\mathcal{M} \times \mathcal{A}$ ). Each  $\Phi_k$  is aperiodic and hence small (this follows by slightly generalizing the arguments of the previous part; see Proposition 3.13). Thus, we can apply Theorem 3.6 to each  $\Phi_k$  separately and express the frequency of visits to  $q$  in  $\Phi_k$  in terms of a unique invari-



ant distribution  $\pi_k$  for  $\Phi_k$ . Finally, we obtain the frequency of visits to  $q$  in  $\mathcal{M} \times \mathcal{A}$  as an average of the corresponding frequencies in  $\Phi_k$ .

Let us start by decomposing the set of nodes  $V$  of  $G_{\mathcal{M} \times \mathcal{A}}$  into  $p$  classes that constitute a cyclic structure (see e.g. [12, Theorem 4.1]).

**LEMMA 3.12.** *There are disjoint sets  $V_0, \dots, V_{p-1} \subseteq V$  such that  $V = \bigcup_{k=0}^{p-1} V_k$  and for all  $u, v \in V$  we have that  $(u, v) \in E$  iff there is  $k \in \{0, \dots, p-1\}$  satisfying  $u \in V_k$  and  $v \in V_j$  where  $j = (k+1) \bmod p$ .*

For each  $k \in \{0, \dots, p-1\}$  we construct a GSSMC  $\Phi_k$  with state space  $\Gamma_{\mathcal{M} \times \mathcal{A}}^k = \bigcup_{R \in V_k} R$ , a transition kernel  $P^k(\cdot, \cdot)$  restricted to  $\Gamma_{\mathcal{M} \times \mathcal{A}}^k$ , and an initial probability measure  $\mu_k$  defined by  $\mu_k(A) = \int_{z \in \Gamma_{\mathcal{M} \times \mathcal{A}}^k} \mu(dz) \cdot P^k(z, A)$ . For each  $k$ , we define the discrete frequency  $\mathbf{d}_q^k$  of visits  $q$  in the process  $\Phi_k$ . Then we show that if  $\mathbf{d}^k$  is well-defined in  $\Phi_k$ , we can express the frequency  $\mathbf{d}_q$  in  $\mathcal{M} \times \mathcal{A}$ .

Note that for every run  $z_0 z_1 \dots$  of  $\mathcal{M} \times \mathcal{A}$ , the word  $z_k z_{p+k} z_{2p+k} \dots$  is a run of  $\Phi_k$ . For a run  $\sigma = (s_0, q_0, \nu_0)(s_1, q_1, \nu_1) \dots$ ,  $k \in \{0, \dots, p-1\}$ , and a location  $q \in Q$ , let define  $1_q^{i,k}(\sigma)$  to be either 1 or 0 depending on whether  $q_{ip+k} = q$  or not, respectively. Further, we put

$$\mathbf{d}_q^k(\sigma) = \limsup_{n \rightarrow \infty} \frac{\sum_{i=1}^n 1_q^{i,k}(\sigma)}{n}$$

Assuming that each  $\mathbf{d}^k$  is well-defined, for almost all runs  $\sigma$  of  $\mathcal{M} \times \mathcal{A}$  we have the following:

$$\begin{aligned} \mathbf{d}_q(\sigma) &= \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n 1_q^i(\sigma)}{n} = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \sum_{k=0}^{p-1} 1_q^{i,k}(\sigma)}{np} \\ &= \frac{1}{p} \sum_{k=0}^{p-1} \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n 1_q^{i,k}(\sigma)}{n} = \frac{1}{p} \sum_{k=0}^{p-1} \mathbf{d}_q^k(\sigma) \end{aligned}$$

So, it suffices to concentrate on  $\mathbf{d}_q^k$ . The following proposition is a generalization of Proposition 3.11 to periodic processes.

**PROPOSITION 3.13.** *Assume that  $G_{\mathcal{M} \times \mathcal{A}}$  is strongly connected and has a period  $p$ . For every  $k \in \{0, \dots, p-1\}$  there exist a region  $R_k \in V_k$ , a measurable  $S_k \subset R_k$ ,  $n_k \in \mathbb{N}$ ,  $b_k > 0$ , and a probability measure  $\kappa_k$  such that  $\kappa_k(S_k) = 1$  and for every measurable  $T \subseteq S_k$  and  $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}^k$  we have  $P_{\mathcal{M} \times \mathcal{A}}^{n_k \cdot p}(z, T) > b_k \cdot \kappa_k(T)$ . In other words,  $\Phi_k$  is  $(n_k, b_k, \kappa_k)$ -small.*

By Theorem 3.6 (1), for every  $k \in \{0, \dots, p-1\}$ , there is a unique invariant distribution  $\pi_k$  on  $\Gamma_{\mathcal{M} \times \mathcal{A}}$  for the process  $\Phi_k$ . By Theorem 3.6 (2), each  $\mathbf{d}^k$  is well-defined and for almost all runs  $\sigma$  we have that  $\mathbf{d}_q^k(\sigma) = \pi_k(A_q)$ . Thus, we obtain

$$\mathbf{d}_q(\sigma) = \frac{1}{p} \sum_{k=0}^{p-1} \pi_k(A_q)$$

## 4. APPROXIMATING DTA MEASURES

In this section we show how to approximate the DTA measures for SMPs using the  $m$ -step transition kernel  $P_{\mathcal{M} \times \mathcal{A}}^m$  of  $\mathcal{M} \times \mathcal{A}$ . The procedure for computing  $P_{\mathcal{M} \times \mathcal{A}}^m$  up to a sufficient precision is taken as a ‘‘black box’’ part of the algorithm, we concentrate just on developing generic bounds on  $m$  that are sufficient to achieve the required precision.

For simplicity, we assume that the initial distribution  $\alpha_0$  of  $\mathcal{M}$  assigns 1 to some  $s_0 \in S$  (all of the results presented in

this section can easily be generalized to an arbitrary initial distribution). The initial state in  $\mathcal{M} \times \mathcal{A}$  is  $z_0 = (s_0, q_0, \mathbf{0})$ .

As we already noted in the previous section, the constant  $k$  of Theorem 3.1 is the number of BSCCs of  $G_{\mathcal{M} \times \mathcal{A}}$ . For the rest of this section, we fix some  $1 \leq j \leq k$ , and write just  $\mathcal{C}$ ,  $\mathcal{R}$  and  $D$  instead of  $\mathcal{C}_j$ ,  $\mathcal{R}_j$  and  $D_j$ , respectively. We slightly abuse our notation by using  $\mathcal{C}$  to denote also the set of configurations that belong to some region of  $\mathcal{C}$  (particularly in expressions such as  $P_{\mathcal{M} \times \mathcal{A}}(z, \mathcal{C})$ ).

The probability  $\mathcal{P}_{\mathcal{M}}(\mathcal{R})$  is equal to the probability of visiting  $\mathcal{C}$  in  $\mathcal{M} \times \mathcal{A}$ . Observe that

$$\mathcal{P}_{\mathcal{M}}(\mathcal{R}) = \lim_{i \rightarrow \infty} P_{\mathcal{M} \times \mathcal{A}}^i(z_0, \mathcal{C})$$

Let us analyze the speed of this approximation. First, we need to introduce several parameters. Let  $p_{\min}$  be the smallest transition probability in  $\mathcal{M}$ , and  $\mathfrak{D}(\mathcal{M})$  the set of delay densities used in  $\mathcal{M}$ , i.e.,  $\mathfrak{D}(\mathcal{M}) = \{\mathbf{D}(s, s') \mid s, s' \in S\}$ . Let  $|V|$  be the number of vertices (regions) of  $G_{\mathcal{M} \times \mathcal{A}}$ . Due to our assumptions imposed on delay densities, there is a fixed bound  $c_{\mathfrak{D}} > 0$  such that, for all  $f \in \mathfrak{D}(\mathcal{M})$  and  $x \in [0, B_{\max}]$ , either  $f(x) > c_{\mathfrak{D}}$  or  $f(x) = 0$ . Further,  $\int_{B_{\max}}^\infty f(x) dx$  is either larger than  $c_{\mathfrak{D}}$  or equal to 0.

**THEOREM 4.1.** *For every  $i \in \mathbb{N}$  we have that*

$$\mathcal{P}_{\mathcal{M}}(\mathcal{R}) - P_{\mathcal{M} \times \mathcal{A}}^i(z_0, \mathcal{C}) \leq \left(1 - \left(\frac{p_{\min} \cdot c_{\mathfrak{D}}}{c}\right)^c\right)^{\lfloor i/c \rfloor}$$

where  $c = 4 \cdot |V|$ .

**PROOF SKETCH.** We denote by  $B$  the union of all regions that belong to BSCCs of  $G_{\mathcal{M} \times \mathcal{A}}$ . We show that for  $c = 4 \cdot |V|$  there is a lower bound  $p_{\text{bound}} = (p_{\min} \cdot c_{\mathfrak{D}} \cdot 1/c)^c$  on the probability of reaching  $B$  in at most  $c$  steps from any state  $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ . Note that then the probability of not hitting  $B$  after  $i = m \cdot c$  steps is at most  $(1 - p_{\text{bound}})^m$ . However, this means that  $P_{\mathcal{M} \times \mathcal{A}}^i(z, \mathcal{C})$  cannot differ from the probability of reaching  $\mathcal{C}$  (and thus also from  $\mathcal{P}_{\mathcal{M}}(\mathcal{R})$ ) by more than  $(1 - p_{\text{bound}})^m$  because  $\mathcal{C} \subseteq B$  and the probability of reaching  $\mathcal{C}$  from  $B \setminus \mathcal{C}$  is 0.

The bound  $p_{\text{bound}}$  is provided by arguments similar to the proof of Proposition 3.11. From any state  $z$  we build a  $\delta$ -wide path to a state in  $B$  that has length bounded by  $4 \cdot |V|$  such that  $\delta = p_{\min} \cdot c_{\mathfrak{D}} \cdot 1/c$ . The paths that follow this  $\delta$ -wide path closely enough (hence, reach  $B$ ) have probability  $p_{\text{bound}}$ .  $\square$

Now let us concentrate on approximating the tuple  $D$ . This can be done by considering just the BSCC  $\mathcal{C}$ . Similarly as in Section 3, from now on we assume that  $\mathcal{C}$  is the set of nodes of  $G_{\mathcal{M} \times \mathcal{A}}$  (i.e.,  $G_{\mathcal{M} \times \mathcal{A}}$  is strongly-connected) and that  $\Gamma_{\mathcal{M} \times \mathcal{A}}$  is equal to the union of all regions of  $\mathcal{C}$ .

As in Section 3, we start with the aperiodic case. Then, Theorem 3.6 (3.) implies that each  $D_q$  can be approximated using  $P_{\mathcal{M} \times \mathcal{A}}^i(u, A_q)$  where  $u$  is an arbitrary state of  $\Gamma_{\mathcal{M} \times \mathcal{A}}$  and  $A_q$  is the set of all states of  $\mathcal{M} \times \mathcal{A}$  of the form  $(s, q, \nu)$ . More precisely, we obtain the following:

**THEOREM 4.2.** *Assume that  $G_{\mathcal{M} \times \mathcal{A}}$  is strongly connected and aperiodic. Then for all  $i \in \mathbb{N}$ ,  $u \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ , and  $q \in Q$*

$$\left| D_q - P_{\mathcal{M} \times \mathcal{A}}^i(u, A_q) \right| \leq \left(1 - \left(\frac{p_{\min} \cdot c_{\mathfrak{D}}}{r}\right)^r\right)^{\lfloor i/r \rfloor}$$

where  $r = \lfloor |V|^{4 \ln |V|} \rfloor$ .

**PROOF.** From the proof of Proposition 3.13 (for details see [10]), we obtain that  $\Gamma_{\mathcal{M} \times \mathcal{A}}$  is  $(m, \varepsilon, \kappa)$ -small with  $m \leq r$  and  $\varepsilon = \left(\frac{p_{\min} \cdot c_{\mathfrak{D}}}{r}\right)^r$ , and the result follows from Theorem 3.6 (3.).  $\square$

Now let us consider the general (periodic) case. We adopt the same notation as in Section 3, i.e., the period of  $G_{\mathcal{M} \times \mathcal{A}}$  is denoted by  $p$ , the decomposition of the set  $V$  by  $V_0, \dots, V_{p-1}$  (see Lemma 3.12), and  $\Gamma_{\mathcal{M} \times \mathcal{A}}^k$  denotes the set  $\bigcup_{R \in V_k} R$  for every  $k \in \{0, \dots, p-1\}$ .

**THEOREM 4.3.** *For every  $i \in \mathbb{N}$  we have that*

$$\left| D_q - \frac{1}{p} \cdot \sum_{k=0}^{p-1} P_{\mathcal{M} \times \mathcal{A}}^{i,p}(u_k, A_q) \right| \leq \left( 1 - \left( \frac{p_{\min} \cdot c_{\mathfrak{D}}}{r} \right)^r \right)^{\lfloor i/r \rfloor}$$

where  $u_k \in \Gamma_{\mathcal{M} \times \mathcal{A}}^k$  and  $r = \lfloor |V|^{4 \ln |V|} \rfloor$ .

**PROOF.** Due to the results of Section 3 we have that  $D_q = \frac{1}{p} \cdot \sum_{k=0}^{p-1} \pi_k(A_q)$ , where  $\pi_k$  is the invariant measure for the  $k$ -th aperiodic decomposition  $\Phi_k$  of the product process  $\mathcal{M} \times \mathcal{A}$  (i.e.  $\pi_k$  is a measure over  $\Gamma_{\mathcal{M} \times \mathcal{A}}^k$ ). From the proof of Proposition 3.13 (for details see [10]),  $\Gamma_{\mathcal{M} \times \mathcal{A}}^k$  is  $(m, \varepsilon, \kappa)$ -small with  $m \leq r$  and  $\varepsilon = \left( \frac{p_{\min} c_{\mathfrak{D}}}{r} \right)^r$ , and the result follows from Theorem 3.6 (3.) applied to each  $\Gamma_{\mathcal{M} \times \mathcal{A}}^k$  separately.  $\square$

## 5. CONCLUSIONS

We have shown that DTA measures over semi-Markov processes are well-defined for almost all runs and assume only finitely many values with positive probability. We also indicated how to approximate DTA measures and the associated probabilities up to an arbitrarily small given precision.

Our approximation algorithm is quite naive and there is a lot of space for further improvement. An interesting open question is whether one can design more efficient algorithms with low complexity in the size of SMP (the size of DTA specifications should stay relatively small in most applications, and hence the (inevitable) exponential blowup in the size of DTA is actually not so problematic).

Another interesting question is whether the results presented in this paper can be extended to more general stochastic models such as generalized semi-Markov processes.

### Acknowledgement

The authors thank Petr Slovák for his many useful comments. The work has been supported by the Institute for Theoretical Computer Science, project No. 1M0545, the Czech Science Foundation, grant No. P202/10/1469 (T. Brázdil, A. Kučera), No. 201/08/P459 (V. Řehák), and No. 102/09/H042 (J. Krčál), and Brno PhD Talent Financial Aid (J. Křetínský).

## 6. REFERENCES

- [1] R. Alur, C. Courcoubetis, and D.L. Dill. Verifying automata specifications of probabilistic real-time systems. In *Real-Time: Theory in Practice*, volume 600 of *Lecture Notes in Computer Science*, pages 28–44. Springer, 1992.
- [2] R. Alur and D. Dill. A theory of timed automata. *Theoretical Computer Science*, 126(2):183–235, 1994. Fundamental Study.
- [3] A. Aziz, K. Sanwal, V. Singhal, and R.K. Brayton. Model-checking continuous-time Markov chains. *ACM Transactions on Computational Logic*, 1(1):162–170, 2000.
- [4] C. Baier, N. Bertrand, P. Bouyer, T. Brihaye, and M. Größer. Almost-sure model checking of infinite paths in one-clock timed automata. In *Proceedings of LICS 2008*, pages 217–226. IEEE Computer Society Press, 2008.
- [5] C. Baier, B.R. Haverkort, H. Hermanns, and J.-P. Katoen. Model-checking algorithms for continuous-time Markov chains. *IEEE Transactions on Software Engineering*, 29(6):524–541, 2003.
- [6] V.L. Barbu and N. Limnios. *Semi-Markov Chains and Hidden Semi-Markov Models toward Applications*. Springer, 2008.
- [7] N. Bertrand, P. Bouyer, T. Brihaye, and N. Markey. Quantitative model-checking of one-clock timed automata under probabilistic semantics. In *Proceedings of 5th Int. Conf. on Quantitative Evaluation of Systems (QEST'08)*, pages 55–64. IEEE Computer Society Press, 2008.
- [8] P. Billingsley. *Probability and Measure*. Wiley, 1995.
- [9] D. Bini, G. Latouche, and B. Meini. *Numerical methods for Structured Markov Chains*. Oxford University Press, 2005.
- [10] T. Brázdil, J. Krčál, J. Křetínský, A. Kučera, and V. Řehák. Measuring performance of continuous-time stochastic processes using timed automata. *CoRR*, abs/1101.4204, 2011.
- [11] T. Brázdil, J. Krčál, J. Křetínský, A. Kučera, and V. Řehák. Stochastic real-time games with qualitative timed automata objectives. In *Proceedings of CONCUR 2010*, volume 6269 of *Lecture Notes in Computer Science*, pages 207–221. Springer, 2010.
- [12] P. Brémaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*. Springer, 1998.
- [13] T. Chen, T. Han, J.-P. Katoen, and A. Mereacre. Quantitative model checking of continuous-time Markov chains against timed automata specifications. In *Proceedings of LICS 2009*, pages 309–318. IEEE Computer Society Press, 2009.
- [14] E.A. Emerson and E.M. Clark. Using branching time temporal logic to synthesize synchronization skeletons. *Science of Computer Programming*, 2:241–266, 1982.
- [15] B.R. Haverkort, H. Hermanns, and J.-P. Katoen. On the use of model checking techniques for quantitative dependability evaluation. In *Proceedings of IEEE Symp. Reliable Distributed Systems*, pages 228–238, 2000.
- [16] R. Kress. *Numerical Analysis*. Springer, 1998.
- [17] K. Matthes. Zur Theorie der Bedienungsprozesse. *Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions, Random Processes*, pages 513–528, 1962.
- [18] S. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge University Press, 2009.
- [19] J.R. Norris. *Markov Chains*. Cambridge University Press, 1998.
- [20] G.O. Roberts and J.S. Rosenthal. General state space Markov chains and MCMC algorithms. *Probability Surveys*, 1:20–71, 2004.
- [21] S.M. Ross. *Stochastic Processes*. Wiley, 1996.