

# Transient analysis of non-Markovian models using stochastic state classes

András Horváth<sup>a,\*</sup>, Marco Paolieri<sup>b</sup>, Lorenzo Ridi<sup>b</sup>, Enrico Vicario<sup>b</sup>

<sup>a</sup> Dipartimento di Informatica, Università di Torino, Italy

<sup>b</sup> Dipartimento Sistemi e Informatica, Università di Firenze, Italy

## ARTICLE INFO

### Article history:

Received 21 February 2011

Accepted 23 November 2011

Available online 16 January 2012

### Keywords:

Generalised Semi Markov Process  
Markov Regenerative Process  
Transient analysis  
Non-Markovian Stochastic Petri Net  
Stochastic Time Petri Net  
Stochastic state class

## ABSTRACT

The method of stochastic state classes approaches the analysis of Generalised Semi Markov Processes (GSMPs) through the symbolic derivation of probability density functions over supports described by Difference Bounds Matrix (DBM) zones. This makes steady state analysis viable, provided that at least one regeneration point is visited by every cyclic behaviour of the model.

We extend the approach providing a way to derive transient probabilities. To this end, stochastic state classes are extended with a supplementary timer that enables the symbolic derivation of the distribution of time at which a class can be entered. The approach is amenable to efficient implementation when model timings are given by expolynomial distributions, and it can be applied to perform transient analysis of GSMPs within any given time bound. In the special case of models underlying a Markov Regenerative Process (MRGP), the method can also be applied to the symbolic derivation of local and global kernels, which in turn provide transient probabilities through numerical integration of generalised renewal equations. Since much of the complexity of this analysis is due to the local kernel, we propose a selective derivation of its entries depending on the specific transient measure targeted by the analysis.

## 1. Introduction

Discrete event systems encountered in real-time applications usually involve multiple concurrent generally distributed timers, often supported over finite time domains. Quantitative modelling of this kind of systems is conveniently supported by non-Markovian Stochastic Petri Nets but, in the general case, the underlying stochastic behaviour is described by a Generalised Semi-Markov Process (GSMP) [1,2] and simulation is the only viable approach to quantitative evaluation [3].

Analytical treatment becomes viable under restrictions on the degree of concurrency of non-exponential timers. In particular, most existing techniques develop on the so-called *enabling restriction*, which assumes that at most one generally-distributed transition is enabled in any reachable tangible marking, so that activity cycles of generally distributed transitions never overlap. In this case, the model underlies a Markov Regenerative Process (MRGP) which regenerates at every change in the enabling status of non-exponential timed transitions and behaves as a (subordinated) continuous time Markov chain (CTMC) between any two regeneration points [4,2,5,6]. Conditional probabilities evolve according to a set of generalised Markov renewal equations, formulated as a system of Volterra integral equations of the second kind [7] defined by a local and a global kernel which capture the transient probability of the states before the first regeneration point and the cumulative distribution of the time to reach the states at the first regeneration point, respectively. Under enabling restriction, both

\* Corresponding author.

E-mail address: [horvath@di.unito.it](mailto:horvath@di.unito.it) (A. Horváth).

kernels can be expressed in closed-form in terms of the exponential of the matrix describing the subordinated CTMC [8] and evaluated numerically through uniformisation. Generalised Markov renewal equations can then be solved by numerical approaches in the time domain or through Laplace–Stieltjes transform.

In principle, the analysis of a model with multiple concurrent generally distributed timers can be formulated through the theory of *supplementary variables* [9,6,10], extending the logical state (the marking) with the vector of ages of generally distributed enabled transitions. However, this results in a set of integro-differential equations whose practical solution is limited to one or two concurrently enabled non-exponential distributions, thus falling again within the limits of the enabling restriction [11]. The limit of the enabling restriction is overcome in [12,13] but only in the case that all timed transitions are either exponential or deterministic. In this case, sampling the process at equidistant time points yields a General State Space Markov Chain whose kernel of transition probabilities is derived through numerical integration. Limiting and transient probabilities of tangible markings are derived by numerical integration of steady-state and transient equations taking the form of a set of Volterra differential equations [12] and a set of Fredholm integral equations [13], respectively.

More recently, [14–16] proposed a new analytical approach that manages the case of multiple concurrent generally distributed timers with possibly overlapping activity cycles, provided that every cyclic behaviour that changes the enabling status of generally distributed transitions visits at least one regeneration point [15]. The approach relies on the stochastic expansion of non-deterministic state-classes described by Difference Bounds Matrix (DBM) zones [17–19] which are commonly employed in qualitative verification tools. This yields so-called stochastic state classes characterising sojourn time distributions which are then organised in a graph that explicitly describes transitions among classes. The stochastic state class graph abstracts the behaviour of a GSMP into an embedded Discrete Time Markov Chain (DTMC) that samples the process after each transition and thus allows the evaluation of discrete time transient and steady state measures referred to the number of fired transitions. In addition, the distribution associated with stochastic classes allows the derivation of average sojourn times between subsequent transitions and thus enables the evaluation of steady state probabilities of the overall continuous time process [16]. However, for what pertains to continuous time transient behaviour, the approach only supports the derivation of the cumulative distribution of the time elapsed along a selected firing sequence that starts from a regenerative class where the process loses memory of its past history [15,20]. This comprises a major limit for the application in design and verification of real time systems where non-functional requirements are natively concerned with the behaviour within the short-term of deadlines or periods.

In this paper, we extend the approach of stochastic state classes so as to support the derivation of continuous time transient probabilities. To this end, stochastic state classes are extended with a supplementary timer that keeps track of the time elapsed from the initial state class. This makes class density functions dependent upon the entering time, and thus enables a symbolic derivation of the distribution of times at which the states of a zone can be reached. The approach is amenable to efficient implementation when temporal parameters in the model are given by expolynomial distributions, and it can be applied to the transient analysis of GSMPs within any given time bound and regardless of the existence of regeneration points. In the special case of models underlying MRGPs that always reach a regeneration in a finite number of steps, the method can be applied to the symbolic derivation of local and global kernels, even when regeneration periods break the limit of a subordinated CTMC and rather evolve according to complex stochastic structures with multiple concurrent non-exponential transitions. Transient probabilities can thus be derived through numerical integration of classical generalised Markov renewal equations. In doing so, the local kernel turns out to be the major source of complexity, both for the number of non-null entries and for the fact that its entries are derived from multivariate distributions whenever multiple generally distributed transitions are concurrently enabled. We show that this complexity can be reduced through a structured approach to the derivation of transient probabilities of specific states, minimising the size of the local kernel for the specific transient measure targeted by the analysis. Finally, we introduce an approximation technique, based on a partial enumeration of regeneration periods, attaining a twofold result: on the one hand, it enables the analysis of MRGP processes that include infinite series of transitions without regeneration points; on the other hand, it simplifies the enumeration for complex models whose dimension would make an exact analysis computationally prohibitive. The impact of approximation on the evaluated transient probabilities is estimated and a solution for its mitigation is devised.

The rest of the paper is organised as follows. In Section 2, we recall syntax and semantics of a variant of non-Markovian Stochastic Petri Nets, which we call stochastic Time Petri Nets (sTPNs) to emphasise the interest on concurrent generally distributed transitions with possibly bounded supports. In Section 3, we introduce transient stochastic state classes, develop the calculus for their derivation and show how they support the derivation of transient probabilities. In Section 4, we discuss the application of the approach to MRGPs showing how the symbolic forms of local and global kernels can be derived. We illustrate how the complexity of the local kernel can be reduced considering the desired transient measure and how the analysis can be simplified, or in some cases even made viable, by means of approximation. Conclusions are drawn in Section 5.

## 2. Stochastic Time Petri Nets

We formulate the analysis problem with reference to a variant of (non-Markovian) Stochastic Petri Nets (SPNs), which we call stochastic Time Petri Nets (sTPNs) [15,16]. As suggested by the name, an sTPN is here regarded as a non-deterministic Time Petri Net (TPN) [21,18,17] extended with a stochastic characterisation of timers and choices: while the TPN identifies a set of feasible behaviours, the stochastic extension associates them with a measure of probability.

## 2.1. Syntax

A stochastic Time Petri Net [15,16] is a tuple:

$$sTPN = \langle P, T, A^-, A^+, A^\bullet, m_0, EFT, LFT, \mathcal{F}, \mathcal{C} \rangle. \quad (1)$$

As in Time Petri Nets [21,18,17]:  $P$  is a set of places;  $T$  a set of transitions;  $A^- \subseteq P \times T$ ,  $A^+ \subseteq T \times P$  and  $A^\bullet \subseteq P \times T$  are the usual pre-condition, post-condition, and inhibitor arc sets, respectively;  $m_0$  is the initial marking.  $EFT: T \rightarrow \mathbb{Q}_{\geq 0}$  and  $LFT: T \rightarrow \mathbb{Q}_{\geq 0} \cup \{+\infty\}$  associate each transition with a firing interval  $[EFT(t), LFT(t)]$  defined by an earliest firing time and a latest firing time. Note that earliest and latest firing times are restrained to be rational values, which takes some theoretical relevance for the termination of the analysis process, as explained in Section 3.1.

In addition,  $\mathcal{C}: T \rightarrow \mathbb{R}_{>0}$  associates each transition  $t \in T$  with a positive weight  $\mathcal{C}(t)$  and  $\mathcal{F}: T \rightarrow CDF$  associates each  $t \in T$  with a Cumulative Distribution Function (CDF)  $F_t: [EFT(t), LFT(t)] \rightarrow [0, 1]$ .

## 2.2. Semantics

The state of an sTPN is a pair  $s = \langle m, \tau \rangle$ , where  $m: P \rightarrow \mathbb{N}$  is a marking and  $\tau: T \rightarrow \mathbb{R}_{\geq 0}$  associates each transition with a real-valued time to fire (remaining time). A transition  $t_0 \in T$  is *enabled* if each of its input places contains at least one token and none of its inhibiting places contains any token, and it is *firable* if it is enabled and its time to fire  $\tau(t_0)$  is not higher than that of any other enabled transition. When multiple transitions are firable, the choice is resolved through a random switch determined by the weights  $\mathcal{C}$ :

$$Prob\{t_0 \text{ is selected}\} = \frac{\mathcal{C}(t_0)}{\sum_{t_i \in T^f(s)} \mathcal{C}(t_i)}$$

where  $T^f(s)$  is the set of transitions that are firable in the state  $s$ .

When a transition  $t_0$  fires, the state  $s = \langle m, \tau \rangle$  is replaced by  $s' = \langle m', \tau' \rangle$ , which we write as  $s \xrightarrow{t_0} s'$ . Marking  $m'$  is derived from  $m$  by removing a token from each input place of  $t_0$  and by adding a token to each output place of  $t_0$ , as usual in Petri Nets:

$$\begin{aligned} m_{tmp}(p) &= \begin{cases} m(p) - 1 & \text{if } \langle p, t_0 \rangle \in A^- \\ m(p) & \text{otherwise,} \end{cases} \\ m'(p) &= \begin{cases} m_{tmp}(p) + 1 & \text{if } \langle t_0, p \rangle \in A^+ \\ m_{tmp}(p) & \text{otherwise.} \end{cases} \end{aligned} \quad (2)$$

Transitions that are enabled both by the intermediate marking  $m_{tmp}$  and by the final marking  $m'$  are said *persistent*, while those that are enabled by  $m'$  but not by  $m_{tmp}$  are said *newly enabled*. A transition  $t_0$  enabled after its own firing is always regarded as newly enabled, as usual in Time Petri Nets [18,17]. After the firing of  $t_0$ , the time to fire  $\tau'(t_i)$  of each persistent transition  $t_i$  is reduced by the time elapsed in the previous state, i.e.

$$\tau'(t_i) = \tau(t_i) - \tau(t_0), \quad (3)$$

whereas the time to fire  $\tau'(t_a)$  of each newly enabled transition  $t_a$  is sampled in the firing interval according to the CDF  $\mathcal{F}(t_a) = F_{t_a}$ :

$$\begin{aligned} EFT(t_a) &\leq \tau'(t_a) \leq LFT(t_a), \\ Prob\{\tau'(t_a) \leq x\} &= F_{t_a}(x). \end{aligned} \quad (4)$$

For every transition  $t \in T$  with  $EFT(t) < LFT(t)$ , we assume that the CDF  $F_t$  can be expressed as the integral function of a Probability Density Function  $f_t$  (PDF):

$$F_t(x) = \int_0^x f_t(y) dy. \quad (5)$$

Following the usual terminology of quantitative evaluation, a transition  $t$  is called *immediate* (IMM) if  $EFT(t) = LFT(t) = 0$  or *timed* if it is not immediate. A timed transition is *generally distributed* (GEN) if it is not distributed as a negative exponential over  $[0, \infty)$  (EXP). As a special case, a GEN transition  $t$  is *deterministic* (DET) if  $EFT(t) = LFT(t) > 0$ . For simplicity, we maintain the notation of Eq. (5) also for IMM and DET transitions, using the Dirac  $\delta$  function as a PDF. This abuse of notation could be avoided by resorting to a (more tedious) formulation based on the partitioned form described in [15].

### 2.3. Underlying stochastic process

The evolution over time of the marking of an sTPN identifies an underlying continuous time stochastic process [2]. Due to transitions with GEN distribution, the process can accumulate memory over time, producing different classes of stochastic processes depending on the presence of *regeneration points*, i.e. states where all GEN timers lose their memory [2,4]. This basically depends on the conditions of persistence of GEN transitions.

If GEN transitions never persist at any firing, the process regenerates at each step, and thus falls in the class of Semi Markov Processes (SMPs) [22].

If GEN transitions never persist at the firing or newly enabling of other GEN transitions (enabling restriction), then the process is a special case of Markov regenerative process [22] that behaves as a sequence of continuous time Markov chains subordinated to the activity periods of GEN transitions. Analysis methods for this class of models have been finely developed relying on Markov renewal theory [4,5] and on the method of supplementary variables [9,6,10].

When a GEN transition can persist at changes of the enabling status of other GEN transitions, GEN activity periods can overlap. If this occurs in a cycle, the model can exhibit infinite behaviours without a regeneration point. If the set of such behaviours has a null measure, the process always guarantees with probability 1 to reach a next regeneration and still belongs to the class of MRGPs [22]. Whereas, when there is a non-null probability that a regeneration point is never reached, then the process becomes a generalised semi-Markov process.

In principle, the overlap of GEN activity periods can occur in any model with at least two GEN transitions, e.g. in the parallel composition of two (possibly interacting) sTPNs each behaving as an SMP. However, the most complex and general case occurs when two GEN transitions persist at the firing of a third GEN transition, so that their remaining times to fire become mutually dependent variables, possibly distributed over a non-rectangular support.

### 3. Transient analysis through stochastic state classes

The method of stochastic state classes [14,15] approaches the analysis of an sTPN model (with any number of concurrent generally distributed transitions) through the construction of an embedded chain that samples the state of the underlying stochastic process after each transition firing. Each state in the embedded chain is called *stochastic state class*, and it is encoded as a triple made up of the marking, the support of times to fire and their distribution. A succession relation among stochastic classes identifies a so-called *stochastic class graph* which captures the relation of reachability among subsequent samples in the embedded chain.

Since times to fire are real-valued, the embedded chain is a Discrete Time Continuous Space Stochastic Process. However, the set of reachable stochastic state classes is countable and turns out to be finite under fairly general conditions. The stochastic class graph can thus be enumerated and managed as a finite Discrete Time Markov Chain, opening the way to *discrete time* evaluation, both in *transient* and in *steady state* regime [15]. In addition, the distribution of times to fire observed at the entrance in a class supports the derivation of the average sojourn time within each class, thus enabling the reconstruction of *continuous time steady state* probabilities [16].

*Transient stochastic classes* extend stochastic state classes through the introduction of an additional timer accumulating the time elapsed since the initial state of the process [23]. The distribution of the elapsed time provides a means to recover the distribution of probability of the absolute time at which the class can be reached, thus opening the way to the evaluation of transient probabilities also in continuous time.

#### 3.1. Non-deterministic state classes and DBM zones

We recall here the concepts of state class, state class graph, and Difference Bounds Matrix (DBM) zone, which are usually employed in symbolic state space analysis of non-deterministic models [18,19,17,24], and which will here comprise the ground for the identification of supports of multi-variate distributions of times to fire of persistent GEN transitions.

A state class is a set of states with the same marking, the same set of newly enabled transitions, and different values for the vector of times to fire of enabled transitions.

**Definition 3.1** (*State Class*). A state class is a triple  $S = \langle m, New, D \rangle$  where  $m$  is a marking;  $D \subseteq (\mathbb{R}_{\geq 0})^N$  is a time domain, i.e. a set of values for the vector  $\underline{\tau} = \langle \tau_0, \tau_1, \dots, \tau_{N-1} \rangle$  of times to fire of transitions enabled by  $m$ , with  $\tau_i$  denoting the time to fire of transition  $t_i$ ;  $New$  is the set of transitions that are newly enabled when the class is entered.

Definition 3.1 extends the usual definition of state class [18,19,17,24] with the addition of the set  $New$ , which distinguishes two classes with the same marking and the same time domain but different conditions of newly-enabling and persistence. This is in fact not relevant for the identification of feasible behaviours in non-deterministic analysis, but it takes a major relevance in stochastic evaluation, particularly for what concerns the identification of regeneration points.

**Definition 3.2** (*State Class Succession*). The state class  $S_2 = \langle m_2, New_2, D_2 \rangle$  is the successor of the state class  $S_1 = \langle m_1, New_1, D_1 \rangle$  through the firing of a transition  $t_0$ , which we write as  $S_1 \xrightarrow{t_0} S_2$ , iff:  $t_0$  is fireable in some state contained in  $S_1$ ;  $S_2$  contains all and only the states that can be reached from some state in  $S_1$ ;  $New_2$  is the set of transitions that are newly enabled when marking  $m_2$  is reached from marking  $m_1$  through the firing of  $t_0$ .

**Definition 3.3** (State Class Graph). Given an initial state class  $S_0$ , the relation  $\xrightarrow{t_0}$  identifies a *state class graph*  $(V, E)$  where the set of vertices  $V$  is the set of state classes reachable from  $S_0$  through the transitive closure of  $\xrightarrow{t_0}$  and the set of edges is  $E = \{ \langle S_1, t_0, S_2 \rangle \in V \times T \times V \mid S_1 \xrightarrow{t_0} S_2 \}$ .

As in most exploitations, we address here the case that time domains in the state class graph are DBM zones.

**Definition 3.4** (DBM Zone). A *Difference Bounds Matrix* zone for a vector  $\langle \tau_0, \tau_1, \dots, \tau_{N-1} \rangle$  is the set  $D \subseteq (\mathbb{R}_{\geq 0})^N$  of solutions of a set of linear inequalities of the form:

$$D = \left\{ \begin{array}{l} \tau_i - \tau_j \leq b_{ij} \\ \tau_* = 0 \\ \forall i \neq j \in \{*, 0, 1, \dots, N-1\} \end{array} \right\}$$

where coefficients  $b_{ij} \in \mathbb{R} \cup \{+\infty\}$  determine the boundaries of variation of the vector  $\langle \tau_0, \tau_1, \dots, \tau_{N-1} \rangle$ , and  $\tau_*$  is a fictitious ground variable which permits to encompass simple inequalities of the form  $\tau_i \leq b_{i*}$  or  $-\tau_j \leq b_{*j}$ .

When some of the inequalities in the set  $D$  are not effective (i.e. they define constraints that are implied by other inequalities), different values of coefficients  $b_{ij}$  may yield the same set of solutions, and thus encode the same DBM zone. To disambiguate the representation, and also to support efficient manipulation, a DBM is said to be in *normal form* when each coefficient  $b_{ij}$  is equal to the maximum value attained by the difference  $\tau_i - \tau_j$  in any solution of the set. For any non-empty DBM zone, the normal form exists, is unique, it is univocally identified by the triangular inequality

$$b_{ij} \leq b_{ih} + b_{hj} \quad \forall i, j, h \in \{*, 0, 1, \dots, N-1\} \quad \text{with } i \neq j, i \neq h, h \neq j,$$

and it can be derived in polynomial time as the solution of an all-shortest-path problem [17].

Given the normal form of a time domain for a state class  $S_1$ , the normal form of the time domain of the class  $S_2$  reached from  $S_1$  through the firing of  $t_0$  can be derived in time  $O(N^2)$ , with  $N$  denoting the number of transitions enabled in  $S_1$ , using a warm-restart reduction of the Floyd–Warshall algorithm [17]. As an immediate consequence of the steps of that derivation, the DBM encoding can be shown to enjoy various relevant closure properties.

**Lemma 3.1** (DBM Closure). *If the time domain of the initial class is a DBM zone, then (i) time domains of all classes in the state class graph are DBM zones; (ii) finite coefficients in the normal form of the DBM of any reachable class can be expressed as summation over the elements of a finite set  $B$  composed of the coefficients  $b_{ij}$  in the normal form of the initial class, the zero value, and the latest firing time and opposite of the earliest firing time of any transition in the sTPN model; (iii) finite coefficients in the normal form of the DBM of any reachable class are bounded to fall within the minimum and the maximum value in  $B$ .*

By relying on these results, the state class graph can be shown to be finite, provided that the sTPN can reach a finite number of markings, that finite coefficients in the normal form of the initial class are rational values, and that this also holds for the endpoints of firing intervals of transitions in the sTPN model:

**Lemma 3.2** (State Class Graph Finiteness). *If the number of markings that the sTPN model can reach is finite, if  $EFT(t) \in \mathbb{Q}_{\geq 0}$  and  $LFT(t) \in \mathbb{Q}_{\geq 0} \cup \{+\infty\}$  for every transition  $t$ , if the time domain of the initial class is a DBM and every coefficient of its normal form belongs to  $\mathbb{Q} \cup \{+\infty\}$ , then the state class graph is finite.*

**Proof.** If *ab absurdo* the state class graph is not finite, then it includes an infinite path  $\rho$  that visits an unbounded number of classes with the same marking  $m^*$  but with always different firing domains. Since these classes share the same marking, their time domains are built on the same set of enabled transitions. There must thus be at least some coefficient, say  $b_{ij}^*$ , that takes an infinite number of different *finite* values in the classes visited by  $\rho$ .

Let  $B$  be the set composed of the *finite* values of the coefficients  $b_{ij}^0$  in the initial class, the latest firing times and the opposite of earliest firing times of transitions in the model, plus the value 0:

$$B = \bigcup_{i,j,t} \{ b_{ij}^0, LFT(t), -EFT(t), 0 \}.$$

Since all the elements of  $B$  are rational values, they can be represented as:

$$B = \bigcup_{i,j,t} \left\{ \frac{\beta_{ij}^0}{\beta}, \frac{\lambda(t)}{\beta}, -\frac{\epsilon(t)}{\beta}, 0 \right\}$$

where  $\beta, \beta_{ij}^0, \lambda(t)$ , and  $\epsilon(t)$  are natural values.

By Lemma 3.1, any finite value for  $b_{ij}^*$  can be expressed as a summation over values in  $B$ , and it is constrained to fall within a bounded interval  $\left[ -\frac{M}{\beta}, \frac{M}{\beta} \right]$  with  $M = \max_{i,j,t} \{ \beta_{ij}^0, \lambda(t), \epsilon(t), 0 \}$ , which implies that the number of different values for  $b_{ij}^*$  is finite.  $\square$



### 3.2. Transient stochastic state classes

Transient stochastic state classes extend non-deterministic state classes in a twofold manner.

On the one hand, they add a timer, called hereinafter *age*, that accounts for the time elapsed since the beginning of the behaviour at the entrance in the initial class. Actually, for a subtle yet not conceptual reason, the age timer represents the opposite of the elapsed time: it starts from 0 in the initial class and decreases towards negative values as time passes; in doing so, the age variable evolves with the same slope as that of times to fire, so that the joint domain of times to fire and age is maintained in the shape of a DBM zone.

On the other hand, transient stochastic state classes add a joint probability density function for the age and the vector of times to fire of enabled transitions, that accounts for the measure of probability induced by the stochastic ingredients,  $\mathcal{F}$  and  $\mathcal{C}$ , of the sTPN model over the space of feasible behaviours.

**Definition 3.5** (*Transient Stochastic State Class*). A *transient stochastic state class* (transient class for short) is a quadruple  $\langle m, New, D, f_{\langle \tau_{age}, \underline{\tau} \rangle} \rangle$  where  $m$  is a marking;  $New$  identifies the set of newly enabled transitions;  $\langle \tau_{age}, \underline{\tau} \rangle$  is a random variable called *time vector* made up of the scalar variable  $\tau_{age}$  representing the opposite of the elapsed time together with the vector  $\underline{\tau} = \langle \tau_0, \tau_1, \dots, \tau_{N-1} \rangle$  representing the times to fire of transitions enabled by  $m$ ;  $f_{\langle \tau_{age}, \underline{\tau} \rangle}$  is the PDF of  $\langle \tau_{age}, \underline{\tau} \rangle$  over the support  $D$ .

**Definition 3.6** (*Transient Stochastic State Class Succession*). Given two transient stochastic classes  $\Sigma = \langle m, New, D, f_{\langle \tau_{age}, \underline{\tau} \rangle} \rangle$  and  $\Sigma' = \langle m', New', D', f_{\langle \tau'_{age}, \underline{\tau}' \rangle} \rangle$ , we say that  $\Sigma'$  is the successor of  $\Sigma$  through the firing of a transition  $t_0$ , and we write  $\Sigma \xrightarrow{t_0} \Sigma'$ , iff the following property holds: if the marking of the net is  $m$  and the time vector is a random variable  $\langle \tau_{age}, \underline{\tau} \rangle$  with support  $D$  and probability density function  $f_{\langle \tau_{age}, \underline{\tau} \rangle}$ , then  $t_0$  has a non-null probability  $\mu_0$  to fire, and if it fires then the model reaches a new marking  $m'$  and a new time vector  $\langle \tau'_{age}, \underline{\tau}' \rangle$  with newly enabled transitions identified by  $New'$ , support  $D'$ , and probability density function  $f_{\langle \tau'_{age}, \underline{\tau}' \rangle}$ .

### 3.3. Calculus of successor classes

The transient stochastic class  $\Sigma' = \langle m', New', D', f_{\langle \tau'_{age}, \underline{\tau}' \rangle} \rangle$  reached from  $\Sigma = \langle m, New, D, f_{\langle \tau_{age}, \underline{\tau} \rangle} \rangle$  through the firing of  $t_0$  can be derived by following the steps of [15], with slight changes accounting for the special role played by  $\tau_{age}$ : such a variable is decreased in the same way as that of all the other variables as time passes, but it does not restrict the firability of transitions and is never reset after a firing. In the derivation, we use the two following notational conventions: (i) given a vector  $\underline{x} = \langle x_0, \dots, x_{N-1} \rangle \in \mathbb{R}^N$  and a scalar  $\delta \in \mathbb{R}$ ,  $\underline{x} + \delta$  denotes the vector  $\langle x_0 + \delta, \dots, x_{N-1} + \delta \rangle$ ; (ii) if  $D \subseteq \mathbb{R}^N$  and  $n \in \{0, \dots, N-1\}$ , then  $D \downarrow_{\tau_n}$  denotes the projection of  $D$  that eliminates  $\tau_n$ :

$$D \downarrow_{\tau_n} = \{ \langle x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{N-1} \rangle \in \mathbb{R}^{N-1} \mid \exists x_n \in \mathbb{R} \text{ such that } \langle x_0, \dots, x_{n-1}, x_n, x_{n+1}, \dots, x_{N-1} \rangle \in D \}.$$

(0) *Successor detection and probability*  $\mu_0$ : Let  $\langle \tau_{age}, \underline{\tau} \rangle = \langle \tau_{age}, \tau_0, \tau_1, \dots, \tau_{N-1} \rangle$  be the time vector in the class  $\Sigma$ , and let it be distributed over  $D \subseteq \mathbb{R}_{\leq 0} \times \mathbb{R}_{\geq 0}^N$  according to the PDF  $f_{\langle \tau_{age}, \underline{\tau} \rangle}(x_{age}, \underline{x})$  with  $\underline{x} = \langle x_0, x_1, \dots, x_{N-1} \rangle$ . According to the semantics of sTPNs, the firing of a transition  $t_0 \in T$  is possible iff the restricted domain  $D^0 = D \cap \{ \tau_0 \leq \tau_n, n = 1, \dots, N-1 \}$  is not empty and the integral of  $f_{\langle \tau_{age}, \underline{\tau} \rangle}$  over  $D^0$  is non-null:

$$\mu_0 \stackrel{\text{def}}{=} \int_{D^0} f_{\langle \tau_{age}, \underline{\tau} \rangle}(x_{age}, \underline{x}) \, dx_{age} \, d\underline{x} > 0. \quad (6)$$

Note that in the restricted domain  $D^0$  the variable  $\tau_0$  is required to be less than or equal to any other time to fire variable (except for the  $\tau_{age}$  variable) and that the integral of  $f_{\langle \tau_{age}, \underline{\tau} \rangle}$  over  $D^0$  corresponds to the probability  $\mu_0$  that the transition  $t_0$  will be the one that fires in  $\Sigma$ .

If  $\mu_0 > 0$ , the marking  $m'$  and the set of newly enabled transitions  $New'$  are derived according to the semantics of the underlying PN model. Besides, the support  $D$  of the time vector after the firing of  $t_0$  is derived according to the underlying TPN model regarding  $\tau_{age}$  as a persistent time to fire. Finally, the PDF  $f_{\langle \tau'_{age}, \underline{\tau}' \rangle}$  is derived through the following four steps.

(1) *Precedence conditioning*: The assumption that  $t_0$  fires before any other transition conditions the time vector and thus yields a new random variable  $\langle \tau'_{age}, \underline{\tau}' \rangle = \langle \tau_{age}, \underline{\tau} \mid \tau_0 \leq \tau_n, n = 1, \dots, N-1 \rangle$  distributed over  $D^0$  according to:

$$f_{\langle \tau'_{age}, \underline{\tau}' \rangle}(x_{age}, \underline{x}) = \frac{f_{\langle \tau_{age}, \underline{\tau} \rangle}(x_{age}, \underline{x})}{\mu_0}. \quad (7)$$

(2) *Time advancement and projection*: When  $t_0$  fires, the age and the times to fire of all the other transitions are reduced by the time to fire value of  $t_0$  given by  $\tau_0^a$ , and  $\tau_0^a$  itself is eliminated through a projection. This yields a variable  $\langle \tau'_{age}, \underline{\tau}' \rangle = \langle \tau_{age}^a - \tau_0^a, \tau_1^a - \tau_0^a, \dots, \tau_{N-1}^a - \tau_0^a \rangle$ , distributed over  $D^b = D^0 \downarrow_{\tau_0}$  according to:

$$f_{\langle \tau'_{age}, \underline{\tau}' \rangle}(x_{age}, \underline{x}) = \int_{\text{Min}^0(x_{age}, \underline{x})}^{\text{Max}^0(x_{age}, \underline{x})} f_{\langle \tau_{age}, \underline{\tau} \rangle}(x_{age} + x_0, x_0, \underline{x} + x_0) \, dx_0 \quad (8)$$

where  $\underline{x} = \langle x_1, \dots, x_{N-1} \rangle$ , and the interval  $[\text{Min}^0(x_{\text{age}}, \underline{x}), \text{Max}^0(x_{\text{age}}, \underline{x})]$  provides the support representing all possible values of  $\tau_0$ .

(3) *Disabling*: If a single transition, say  $t_1$ , is disabled at the firing of  $t_0$ , its elimination yields a new variable  $\langle \tau_{\text{age}}^c, \underline{\tau}^c \rangle = \langle \tau_{\text{age}}^b, \tau_2^b, \dots, \tau_{N-1}^b \rangle$ , distributed over  $D^c = D^b \downarrow_{\tau_1}$  according to:

$$f_{(\tau_{\text{age}}^c, \underline{\tau}^c)}(x_{\text{age}}, \underline{x}) = \int_{\text{Min}^1(x_{\text{age}}, \underline{x})}^{\text{Max}^1(x_{\text{age}}, \underline{x})} f_{(\tau_{\text{age}}^b, \underline{\tau}^b)}(x_{\text{age}}, x_1, \underline{x}) dx_1 \quad (9)$$

where  $\underline{x} = \langle x_2, \dots, x_{N-1} \rangle$ . When multiple transitions are disabled at the firing of  $t_0$ , the same step is repeated for each one of them.

(4) *Newly enabling*: If a single transition  $t_N$  with static firing density function  $f_N(x_N)$  over the support  $[EFT_N, LFT_N]$  is newly enabled at the firing of  $t_0$ , then the time vector  $\langle \tau'_{\text{age}}, \underline{\tau}' \rangle = \langle \tau_{\text{age}}^c, \tau_2^c, \dots, \tau_{N-1}^c, \tau_N \rangle$  of the successor class  $\Sigma'$  is distributed over  $D' = D^c \times [EFT_N, LFT_N]$  according to:

$$f_{(\tau_{\text{age}}^d, \underline{\tau}^d)}(x_{\text{age}}, \underline{x}, x_N) = f_{(\tau_{\text{age}}^c, \underline{\tau}^c)}(x_{\text{age}}, \underline{x}) \cdot f_N(x_N) \quad (10)$$

where  $\underline{x} = \langle x_2, \dots, x_{N-1} \rangle$ . When multiple transitions are newly enabled, the product form is extended with a factor for each one of them.

If the vector of times to fire  $\underline{\tau}$  in the initial class  $\Sigma_0$  is supported over a DBM domain  $D$ , then by Lemma 3.1 the domains of all reached classes are still in DBM form. Moreover, in the special case that the static density function  $f_t$  associated with each timed transition  $t$  in the model is a truncated exponential function, i.e.

$$f_t(y) = \begin{cases} \sum_{k=1}^K c_k y^{a_k} e^{-\lambda_k y} & EFT(t) \leq y \leq LFT(t) \\ 0 & \text{otherwise} \end{cases}$$

with  $a_k \in \mathbb{N}$ ,  $c_k \in \mathbb{R}$ , and  $\lambda_k \in \mathbb{R}_{\geq 0}$  [15,2,25], then density functions of reachable transient classes accept a closed-form, which is efficiently computed in the Oris tool [26,27] through repeated application of the calculus within a conventional forward enumeration (semi-)algorithm. Note that this class of functions includes uniform, deterministic and phase type distributions.

Finally note that, according to the use intended in this paper, in the initial class  $\Sigma_0$ , the age  $\tau_{\text{age}}$  is a deterministic variable with value 0. However, the calculus also accepts the case that  $\tau_{\text{age}}$  is associated with any type of distribution supported over a subinterval of the set of non-positive real numbers. This can take relevance when the analysis technique is embedded in some compositional or incremental approach.

### 3.4. Transient analysis through enumeration of classes

The succession relation  $\xrightarrow{t_0}$  identifies a *transient stochastic graph*, which extends the state class graph by associating each vertex with a joint probability distribution of the age and the times to fire of enabled transitions. The extension is one-to-many, as the same state class can be reached with different joint distributions of age and times to fire.

For the purposes of the present treatment, the stochastic graph is more conveniently encoded into a *tree*: since the age is never reset, the graph includes a limited number of confluences, so that the size of the tree is basically the same as that of the graph; besides, the tree associates each node with a univocal history, which becomes useful in the evaluation of the probability that a node is reached.

**Definition 3.7** (*Transient Stochastic Tree*). Given an initial transient class  $\Sigma_0$ , the relation  $\xrightarrow{t_0}$  identifies a *transient stochastic tree*  $\langle V, E \rangle$  where each node  $v \in V$  is associated with a transient stochastic class  $\Sigma(v)$  and each arc is associated with a transition of the sTPN model; the root node is associated with the initial class  $\Sigma_0$ ; the node  $v_1$  has a child  $v_2$  reached through an arc associated with  $t_0$  iff  $\Sigma(v_1) \xrightarrow{t_0} \Sigma(v_2)$ .

For any node  $v$  in the transient stochastic tree, the path  $\rho_v = \langle u_0, u_1, \dots, u_K \rangle$  from the root  $u_0$  to  $v = u_K$  univocally identifies a feasible execution sequence for the sTPN model. The transient class  $\Sigma(v)$  samples the state of the underlying process of the sTPN model at the time when such sequence is completed and allows a direct reconstruction of transient behaviour of the sTPN model in both discrete and continuous time.

For any transition  $t_i$  associated with an arc connecting a node  $u_i$  with its successor  $u_{i+1}$  on the path  $\rho_v$ , the probability  $\mu_i$  that  $t_i$  fires in  $\Sigma(u_i) = \langle m, \text{New}, D, f \rangle$  before any other enabled transition can be computed as the measure of the restricted time domain  $D^i$  according to the probability density function  $f$ , as shown in Eq. (6).

The probability  $\eta_v$  that the execution proceeds along  $\rho_v$  can thus be derived as the product of transition probabilities of the arcs traversed by  $\rho_v$ :

$$\eta_v = \prod_{i=0}^{K-1} \mu_i. \quad (11)$$

The probability that the node  $v$  is reached within some time  $t$  can be derived from the probability  $\eta_v$  and the transient stochastic state class  $\Sigma(v) = \langle m_v, D_v, New_v, f_v \rangle$  as

$$\psi_v(t) = \eta_v \cdot \int_{D_v \cap \{0 \leq -x_{age} \leq t\}} f_v(x_{age}, \underline{x}) \, dx_{age} \, d\underline{x}. \quad (12)$$

Similarly, the probability  $\pi_v(t)$  that  $v$  is the last node reached within some time  $t$  can be expressed as the probability that  $v$  has been reached at some time  $\tau \in [0, t]$  and the sojourn time in  $v$  is higher than  $t - \tau$ :

$$\pi_v(t) = \eta_v \cdot \int_{D_v \cap \{0 \leq -x_{age} \leq t\} \cap \{\min(\underline{x}) > t + x_{age}\}} f_v(x_{age}, \underline{x}) \, dx_{age} \, d\underline{x}. \quad (13)$$

Finally, the probability density function  $f_{in,v}$  of the entering time for the node  $v$  can be computed as

$$f_{in,v}(t) = \int_{D_v \downarrow \tau_{age}} f_v(-t, \underline{x}) \, d\underline{x}, \quad (14)$$

i.e. from the marginalisation of  $f_v(x_{age}, \underline{x})$  with respect to the vector  $\underline{x}$  of times to fire of all enabled transitions. Note that measures obtained from the transient stochastic state class  $\Sigma(v)$  are conditioned to the assumption that  $v$  has been reached, which justifies the factor  $\eta_v$  in Eqs. (12) and (13).

Probabilities  $\pi_v(t)$  can be aggregated according to different abstractions attaining different levels of granularity. Each such abstraction results in a different concept of state to which the analysis can be referred.

In the usual approach, the analysis refers to the marking process  $\mathbb{M} = \{M(t), t \geq 0\}$ , and transient probabilities are obtained by adding up the probabilities of nodes that share a common marking:

$$\text{Prob}\{M(t) = m\} = \sum_{v | \Sigma(v) \text{ has marking } m} \pi_v(t). \quad (15)$$

While directly connected to the model interpretation, this level of aggregation neglects much of the inherent structure of the state space, as the same marking can be reached with different supports and distributions for times to fire of enabled transitions. A finer abstraction distinguishing also the condition of newly enabling will be later considered in Section 4.

### 3.5. Termination

The results of Section 3.4 enable a straight approach to the evaluation of transient probabilities of any sTPN, regardless of the class of its underlying stochastic process, provided that the number of transient classes that can be reached within the time horizon of interest is bounded.

The following result provides the basis that permits to characterise the problem of termination with reference to the state class graph of the underlying TPN model.

**Lemma 3.3.** *If the transient stochastic tree includes a vertex  $v$  with  $\Sigma(v) = \langle M_v, New_v, D_v, f_v \rangle$ , then the class graph includes a vertex  $S = \langle M_v, New_v, D_v \downarrow \tau_{age} \rangle$ . Moreover, if  $\Sigma(v) \xrightarrow{t} \Sigma'$  with  $\Sigma' = \langle M', New', D', f' \rangle$ , then  $S \xrightarrow{t} S'$  with  $S' = \langle M', New', D' \downarrow \tau_{age} \rangle$ .*

**Proof.** The first part of the statement is satisfied by construction for the root node of the transient stochastic tree and the initial class of the state class graph. By induction, it will thus be sufficient showing the second part of the statement.

According to Eq. (6),  $\Sigma(v) \xrightarrow{t} \Sigma'$  implies that  $D_v^0 \downarrow \tau_{age}$  is not empty, which in turn implies that  $t$  is a feasible outgoing event from  $S$ . Let then  $\hat{S} = \langle \hat{M}, \hat{New}, \hat{D} \rangle$  be the state class reached from  $S$  through  $t$  (i.e.  $S \xrightarrow{t} \hat{S}$ ). The marking  $\hat{M}$  and the set of newly enabled transitions  $\hat{New}$  are equal to those of  $\Sigma'$ , since they are derived through the same steps defined by the semantics of the underlying TPN. Besides,  $\hat{D} = D' \downarrow \tau_{age}$  can be proven by resorting to the general properties of the normal form of DBM zones by leveraging on the assumption that  $\tau_{age}$  is not considered in the restriction that determines  $D^0$ .  $\square$

Lemma 3.3 permits to regard the state class graph as a kind of non-deterministic projection of the transient stochastic tree, or, vice versa, regard the transient stochastic tree as a stochastic and aged expansion of the class graph. Note that, while Lemma 3.3 guarantees that each transient stochastic class has a unique projection in the state class graph, the inverse relation is one-to-many as each state class can be the projection of multiple (and possibly infinite) transient stochastic classes. If the state class graph is finite (which is guaranteed under fairly general conditions by Lemma 3.2), the number of transient classes reached within any predefined time is certainly finite provided that the age is advanced by a positive quantity in every cyclic behaviour:

**Lemma 3.4.** *If the state class graph is finite, and every cycle that it contains traverses at least one edge labelled by a transition  $t$  with non-null static earliest firing time (i.e.  $EFT(t) > 0$ ), then, for any time bound  $x \in \mathbb{R}_{>0}$ , the number of transient stochastic classes including states with  $(-\tau_{age}) \leq x$  is finite.*

**Proof.** Ab absurdo, let  $\rho$  be an infinite path in the transient stochastic tree and let each class visited by  $\rho$  include at least one state satisfying  $(-\tau_{age}) \leq x$ . By transitive closure of Definition 3.6, a transient class includes a state  $s$  satisfying  $(-\tau_{age}) \leq x$



```

BUILD-TRANSIENT-TREE( $\Sigma_0, \epsilon, x$ )
1   $V = \emptyset$ 
2   $E = \emptyset$ 
3   $v_0 = \text{NEW-NODE}(\Sigma_0)$ 
4   $\eta(v_0) = 1$ 
5   $Q = \{(\cdot, v_0)\}$ 
6  while  $\sum_{(u,v) \in Q} \psi_v(x) > \epsilon$ 
7       $(u, v) = \arg \max_{(i,j) \in Q} \psi_j(x)$ 
8       $Q = Q \setminus \{(u, v)\}$ 
9       $V = V \cup \{v\}$ 
10     if  $v \neq v_0$ 
11          $E = E \cup \{\text{NEW-ARC}(u, v)\}$ 
12     for each successor  $\Sigma' = \langle m', \text{New}', D', f_{\langle \tau'_{age}, \mathcal{I}' \rangle} \rangle$  of  $\Sigma(v)$  with probability  $\mu > 0$ 
13          $z = \text{NEW-NODE}(\Sigma')$ 
14          $\eta(z) = \eta(v) \cdot \mu$ 
15          $Q = Q \cup \{(v, z)\}$ 
16 return  $\langle V, E, \eta \rangle$ 

```

**Fig. 1.** The pseudo-code describes the construction of the transient stochastic tree rooted in  $\Sigma_0$  with approximation  $\epsilon \geq 0$  within time  $x$ . The next node added to the tree is the one with highest reaching probability.

iff the model accepts a behaviour that can reach  $s$  within a time not higher than  $x$ . According to this,  $\rho$  accepts a timing that never exceeds the bound  $x$ .

Since the state class graph is finite, some state class  $S_*$  is visited infinitely often by the projection of  $\rho$ . The trace  $\rho$  can thus be decomposed into an infinite sequence of sub-traces  $\{\rho_i\}_{i=0}^\infty$ , and the projection of each sub-trace  $\rho_i$  on the class graph is a cycle that starts and terminates on class  $S_*$ . By hypothesis, each such cycle includes the firing of at least one transition with non-null static earliest firing time. Since transitions are finite, there is at least one transition with a non-null earliest firing time which is fired infinitely often along  $\rho$ , which contradicts the assumption that  $\rho$  accepts a timing that never diverges beyond  $x$ .  $\square$

**Lemma 3.4** does not manage the (frequent) case of models allowing cyclic behaviours that *can* be completed in zero time. In this case, termination can be guaranteed in probability, provided that the model does not include cycles that *must* be completed in zero time.

**Lemma 3.5.** *If the class graph is finite and every cycle that it contains traverses at least one non-immediate transition, then for every time limit  $x \in \mathbb{R}_{>0}$  and every probability threshold  $\epsilon \in \mathbb{R}_{>0}$  there exists  $\bar{N}_{x,\epsilon} \in \mathbb{N}$  such that  $\forall N > \bar{N}_{x,\epsilon}$ , the probability that after  $N$  steps the age has not yet exceeded  $x$  is lower than  $\epsilon$ .*

**Proof.** Assume, *ab absurdo*, that such  $\bar{N}_{x,\epsilon} \in \mathbb{N}$  does not exist. This implies that there exists an infinite sequence  $r$  for which  $\text{Prob}\{r \text{ performed in } [0, x]\} > 0$ . By the assumptions on the class graph, such  $r$  has to contain an infinite number of non-immediate transitions. Since there are a finite number of non-immediate transitions in the model, there must be in  $r$  at least one non-immediate transition  $t$  that is fired infinitely many times. Denote by  $X_i$ ,  $i = 1, 2, \dots$ , the consecutive firing times of  $t$  in  $r$ . Then, in order to have  $\text{Prob}\{r \text{ performed in } [0, x]\} > 0$ , we must have  $\text{Prob}\{\sum_{i=1}^\infty X_i < x\} > 0$ , which is absurd.  $\square$

**Fig. 1** reports a high level pseudo-code of a possible algorithm that builds up a finite subset of the transient stochastic tree that, with probability not lower than  $1 - \epsilon$ , includes every behaviour within time  $x$ . The set  $Q$  contains arcs  $(u, v)$  each representing a feasible successor  $v$  of a leaf node  $u$ . The next node added to the tree is the successor  $v$  with highest reaching probability  $\eta(v)$ . The enumeration is halted when the total reaching probability (within  $x$ ) of the discarded successors in  $Q$  is lower than the error threshold  $\epsilon$ .

**Table 1** summarises the symbols and notation adopted in the introduction of transient stochastic state classes.

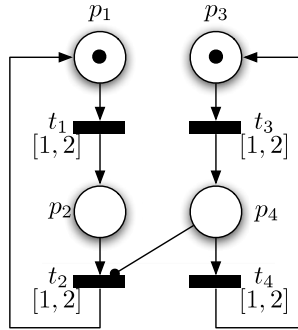
### 3.6. Computational experience

**Fig. 2** depicts the sTPN model of the  $G/G/1/2/2$  preemptive queue. This is a closed queue with two clients, GEN arrival times (transitions  $t_1$  and  $t_3$ ) and GEN service times (transitions  $t_2$  and  $t_4$ ). Service follows a Preemptive Repeat Different policy: the service of client 1 is interrupted by the arrival of client 2 and restarted from crash when the service of client 2 is completed. All transitions have firing times distributed uniformly on the interval  $[1, 2]$ .

The simpler case of the  $M/G/1/2/2$  preemptive queue has been analysed in several papers, by leveraging on the fact that the model never shows a GEN transition that persists at the change of the enabling status of other GEN transitions, so that the underlying stochastic process can always be characterised as a CTMC subordinated to a GEN activity period. On the contrary, in the  $G/G/1/2/2$  preemptive queue, activity cycles of GEN transitions can overlap. Due to the specific firing

**Table 1**  
Symbols and notation of transient stochastic state classes.

|  |  |
|--|--|
| $F_t f_t$  | CDF and PDF of the time to fire of transition $t \in T$  |
| $S = \langle m, New, D \rangle$  | State class  |
| $m$  | Marking of the sTPN  |
| $New$  | Set of newly enabled transitions   |
| $D \subseteq (\mathbb{R}_{\geq 0})^N$                                    | Time domain for the vector $\underline{\tau} = \langle \tau_0, \tau_1, \dots, \tau_{N-1} \rangle$ of times to fire |
| $S_1 \xrightarrow{t_0} S_2$  | The state class $S_2$ is the successor of $S_1$ through the firing of $t_0 \in T$                                  |
| $\Sigma = \langle m, New, D, f_{(\tau_{age}, \underline{\tau})} \rangle$ | Transient stochastic state class   |
| $\tau_{age}$   | Opposite of the elapsed time   |
| $f_{(\tau_{age}, \underline{\tau})}$                                     | PDF of the random variable $\langle \tau_{age}, \underline{\tau} \rangle$  |
| $D \subseteq \mathbb{R}_{\leq 0} \times \mathbb{R}_{\geq 0}^N$           | Time domain over which the random variable $\langle \tau_{age}, \underline{\tau} \rangle$ is distributed           |
| $\Sigma \xrightarrow{t_0} \Sigma'$                                       | $\Sigma'$ is the successor of $\Sigma$ through the firing of $t_0 \in T$   |
| $D^0$  | Time domain $D$ restricted to the hypothesis $\tau_0 \leq \tau_n$ for $n = 1, \dots, N - 1$                        |
| $\mu_0$  | Probability that transition $t_0 \in T$ will fire in the transient class $\Sigma$                                  |
| $\langle V, E \rangle$   | Transient stochastic tree  |
| $\Sigma(v)$  | Transient stochastic class associated with node $v \in V$  |
| $\rho_v$   | Path from the root to node $v \in V$ in the transient stochastic tree  |
| $\eta_v$   | Probability that the execution proceeds along the path $\rho_v$  |
| $\psi_v(t)$  | Probability that node $v \in V$ is reached within time $t$   |
| $\pi_v(t)$   | Probability that node $v \in V$ is the last node reached within time $t$   |
| $f_{in,v}(t)$  | PDF of the entering time for node $v \in V$  |



**Fig. 2.** Petri net representation of the preemptive  $G/G/1/2/2$  queue. All transitions have firing times distributed uniformly on the interval  $[1, 2]$ .

intervals assumed, overlaps can produce behaviours with unbounded duration and unbounded number of events between subsequent regeneration points, casting the model in a class for which no analysis approach has been proposed yet.

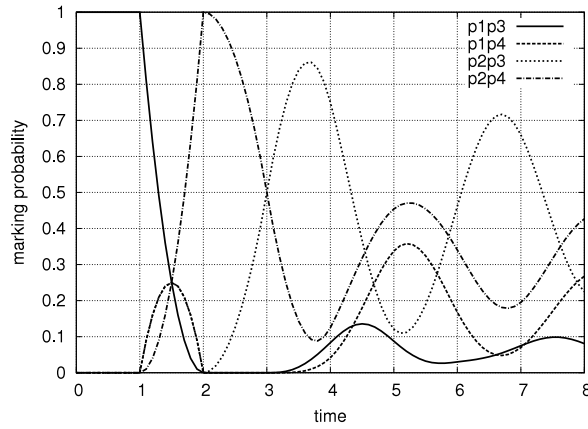
Using the proposed method, transient analysis in the interval  $[0, x]$  is reduced to the enumeration of transient classes whose support  $D$  has a non-null intersection with the constraint  $(-\tau_{age}) > x$ . Termination is guaranteed by Lemma 3.4: since the queue is closed, the number of reachable markings is finite and thus, according to Lemma 3.2, the graph of non-deterministic classes is finite; in addition, each cycle in the state class graph includes the firing of at least one transition with  $EFT > 0$  (actually, in the model, all transitions have  $EFT = 1$ ). The lapse of time up to  $x = 8$  is completely covered with 125 classes and the calculations require about four minutes with a standard portable computer. Resulting transient probabilities are depicted in Fig. 3.

If the temporal parameters of the model are changed by enlarging all firing intervals to  $[0, 2]$ , then the state class graph is still finite but includes cycles that can be executed in arbitrarily short time. In this case, the number of classes satisfying the constraint  $(-\tau_{age}) > x$  is unbounded. Termination can be guaranteed in probability according to Lemma 3.5 under the assumption of some approximation threshold on the total unallocated probability.

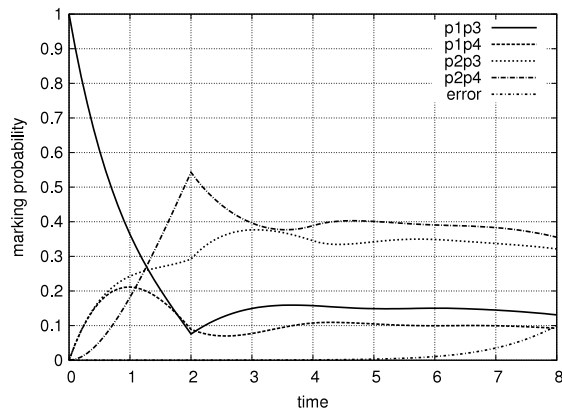
Fig. 4 shows transient probabilities of the 4 reachable markings within  $x = 8$  using threshold  $\epsilon = 0.1$ , while Fig. 5 shows transient probability of marking  $p1p3$  for different values of allowed approximation, along with the number of enumerated classes required to complete the analysis. Note that, by construction, truncation yields under-approximated probabilities and the threshold  $\epsilon$  comprises an upper bound of the approximation error. The number of enumerated classes and the computational time depend on the error threshold  $\epsilon$ , and correspond to 5689 enumerated classes and 13 h for  $\epsilon = 0.1$ , 2592 enumerated classes and 4 h for  $\epsilon = 0.2$ , 1512 enumerated classes and 2.5 h for  $\epsilon = 0.3$ .

#### 4. Exploiting regeneration points

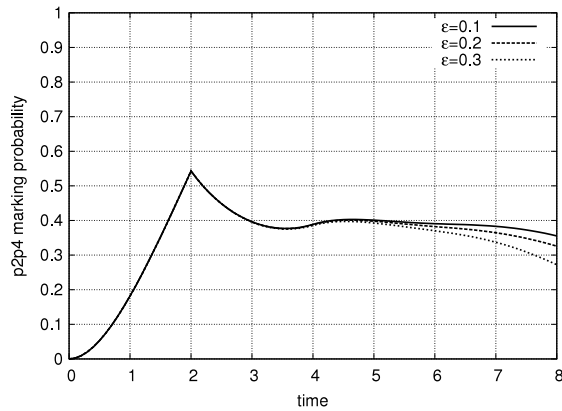
Straight enumeration of transient state classes provides a terminating algorithm for the transient analysis of any sTPN model within a bounded transient horizon and with a given limit of approximation. In principle, this allows the transient analysis of any generalised semi-Markov process with timers progressing with equal speed [3]. However, a practical



**Fig. 3.** Transient probabilities of the 4 markings reachable in the preemptive  $G/G/1/2/2$  queue.



**Fig. 4.** Transient probabilities of the 4 markings reachable in the preemptive  $G/G/1/2/2$  queue with firing intervals enlarged to  $[0, 2]$ , using threshold  $\epsilon = 0.1$ .



**Fig. 5.** Transient probabilities of marking  $p2p4$  in the preemptive  $G/G/1/2/2$  queue with firing intervals enlarged to  $[0, 2]$  using thresholds 0.1, 0.2 and 0.3.

limitation may arise from the number of classes needed for the analysis when the transient horizon grows with respect to the average sojourn time in each class.

When the process underlying the sTPN model exhibits regeneration points, this complexity may be circumvented by leveraging on the theory of Markov regenerative processes. Following this approach, the enumeration of transient stochastic classes can be limited to the first epoch before a regeneration and repeated from every regenerative state, deriving the local and global kernels that fully characterise the behaviour of the MRGP. Transient probabilities at any given time can then be computed as the numerical solution of a set of generalised Markov renewal equations [22].

#### 4.1. Regenerative classes

**Definition 4.1** (*Regenerative Class*). We call *regenerative* a transient stochastic class where all times to fire are either newly enabled, or exponentially distributed, or deterministic, or bounded to take a deterministic delay with respect to a time to fire satisfying any of the previous conditions.

In a regenerative class, the distribution of times to fire of enabled transitions is uniquely determined, so that it is not possible that two different regenerative classes have the same underlying non-deterministic class. In a different perspective, this means that when a regenerative class is entered, the future behaviour of the process is not conditioned by the history through which the class has been reached. The state reached when the class is entered is thus a regenerative state, as usually intended in Markov renewal theory.

It is relevant to note here that the same marking can be reached both in regenerative and non-regenerative classes. This is for instance the case of the marking  $p_2p_4$  in the  $G/G/1/2/2$  queue of Fig. 3: the marking determines that  $t_4$  is the only enabled transition; however,  $t_4$  is newly enabled and the state is regenerative if and only if the marking is reached through the firing of  $t_3$ .

According to this remark, in order to exploit regeneration conditions, a proper abstraction of the state shall distinguish not only the marking but also the condition of regeneration. We will thus refer the analysis to a refinement of the marking process that associates each marking  $m$  with two different states distinguishing whether  $m$  is reached under a condition of regeneration:  $\bar{M} = \{\bar{M}(t), t \in \mathbb{R}_{\geq 0}\}$ , with  $\bar{M}(t) = \langle m, \text{reg} \rangle$  and  $\text{reg} \in \{\text{restart}, \text{continue}\}$  meaning that at time  $t$  the marking is  $m$  and it has been reached in a condition of regeneration ( $\text{reg} = \text{restart}$ ) or not ( $\text{reg} = \text{continue}$ ). For example, in the  $G/G/1/2/2$  queue, the marking  $p_1p_3$  identifies a single non-regenerative state  $\langle p_1p_3, \text{continue} \rangle$ , while the marking  $p_2p_4$  identifies a regenerative state  $\langle p_2p_4, \text{restart} \rangle$  and a non-regenerative state  $\langle p_2p_4, \text{continue} \rangle$ .

#### 4.2. Analysis based on regeneration points

If  $\bar{M}$  embeds an infinite sequence of regeneration points at times  $S_0, S_1, \dots$  with  $S_0 = 0$  and  $\lim_{n \rightarrow \infty} S_n = \infty$ , then  $\bar{M}$  is an MRGP and can be completely characterised through a local kernel  $L_{ij}(t)$  and a global kernel  $G_{ij}(t)$  [22].

The local kernel captures the evolution in the period between two subsequent regeneration points, providing the probability that, starting from state  $i$  at time 0, no regeneration occurs within  $t$  and the state at time  $t$  is  $j$ :

$$L_{ij}(t) = \text{Prob} \{ \bar{M}(t) = j, S_1 > t \mid \bar{M}(S_0) = i \}. \quad (16)$$

The global kernel characterises the state where the process ends up at the first regeneration, providing the probability that, starting from state  $i$  at time 0, the first regeneration occurs at some state  $j$  within time  $t$ :

$$G_{ij}(t) = \text{Prob} \{ S_1 \leq t, \bar{M}(S_1) = j \mid \bar{M}(S_0) = i \}. \quad (17)$$

Note that

$$\sum_j L_{ij}(t) + G_{ij}(t) = 1, \quad \forall i, t. \quad (18)$$

For any pair of states  $i$  and  $j$ , conditional transient probabilities  $\pi_{ij}(t)$  can be expressed as the solution of a set of generalised Markov renewal equations [7]:

$$\pi_{ij}(t) \stackrel{\text{def}}{=} \text{Prob} \{ \bar{M}(t) = j \mid \bar{M}(0) = i \} \quad (19)$$

$$= L_{ij}(t) + \sum_k \int_0^t \frac{dG_{ik}(x)}{dx} \pi_{kj}(t-x) dx. \quad (20)$$

Eq. (19) has the form of a set of Volterra integral equations of the second kind, which can be solved through various numerical techniques widely experimented in both transient and steady state analysis of MRGPs [28]. Derivation of transition probabilities  $\pi_{ij}(t)$  can thus be reduced to the computation of the local and global kernels  $L_{ij}(t)$  and  $G_{ij}(t)$ .

For each state  $i$ , the kernels can be derived from the stochastic classes of the transient tree constructed from  $i$  in such a way that successors of regenerative classes are not enumerated. In this tree, the leaf nodes represent regenerative classes (characterising the global kernel) and non-leaf nodes represent non-regenerative classes (characterising the local kernel).

Specifically, for any node  $v$  associated with state  $j$  and reached through regeneration, let  $\eta_v$  be the probability that  $v$  is eventually reached, let  $f_v(x_{\text{age}}, \underline{x})$  be the PDF of the class in the node, and let  $f_{\text{in},v}(t)$  be probability density function of the entering time conditioned to the assumption that the node is eventually reached. The analytical form of the derivative of the  $ij$  entry of the global kernel can thus be obtained as:

$$\frac{dG_{ij}(x)}{dx} = \sum_{v \mid \bar{M}(v)=j} \eta_v \cdot f_{\text{in},v}(x). \quad (21)$$

Note that, since  $j$  is a regeneration class,  $f_v(x_{\text{age}}, \underline{x})$  is in product form and thus permits a straightforward derivation of the analytical form of  $f_{\text{in},v}(t)$ .

Besides, for any node  $v$  with a non-regenerative class associated with state  $l$ , the probability that at time  $t$  the process is in the node  $v$  is expressed according to Eq. (13). The derivation of  $\pi_v(t)$  requires a multidimensional integral over a DBM

```

COMPUTE-KERNELS( $\Sigma_0, \epsilon, x$ )
1   $S = \{m_0\}$ 
2   $S^* = \emptyset$ 
3  while  $S \neq \emptyset$ 
4      select marking  $i$  from  $S$ 
5       $S = S \setminus \{i\}$ 
6       $S^* = S^* \cup \{i\}$ 
7       $V = \emptyset$ 
8       $E = \emptyset$ 
9       $v_i = \text{NEW-NODE}(\Sigma_i)$ 
10      $\eta(v_i) = 1$ 
11      $Q = \{(\cdot, v_i)\}$ 
12     while  $\sum_{(u,v) \in Q} \psi_v(x) > \epsilon$ 
13          $(u, v) = \arg \max_{(i,j) \in Q} \psi_j(x)$ 
14          $Q = Q \setminus \{(u, v)\}$ 
15          $V = V \cup \{v\}$ 
16         if  $v \neq v_i$ 
17              $E = E \cup \{\text{NEW-ARC}(u, v)\}$ 
18         if  $\Sigma(v)$  is not regenerative
19             for each successor  $\Sigma' = \langle m', \text{New}', D', f_{\langle \tau'_{age}, \mathcal{I}' \rangle} \rangle$  of  $\Sigma(v)$  with probability  $\mu > 0$ 
20                  $z = \text{NEW-NODE}(\Sigma')$ 
21                  $\eta(z) = \eta(v) \cdot \mu$ 
22                  $Q = Q \cup \{(v, z)\}$ 
23                 if  $m' \notin S^*$ 
24                      $S = S \cup \{m'\}$ 
25      $L_{ij}(x) = \sum_{v \in V : \Sigma(v) \text{ non-regenerative with marking } j} \eta(v) \cdot \pi_v(x) \quad \forall j$ 
26      $G'_{ij}(x) = \sum_{v \in V : \Sigma(v) \text{ regenerative with marking } j} \eta(v) \cdot f_{in,v}(x) \quad \forall j$ 

```

**Fig. 6.** The pseudo-code describes the computation of local and global kernel from an initial class  $\Sigma_0$  with approximation  $\epsilon \geq 0$  within time  $x$ .

domain that depends on  $t$ . Derivation of the analytical form of  $\pi_v(t)$  is thus quite hard and results in a complex piecewise representation; however, its computation can be effectively accomplished for any selected value of  $t$ , which is sufficient for the numerical integration of Volterra equations. The numerical value of the  $il$  entry of the local kernel can thus be obtained as:

$$L_{il}(x) = \sum_{v|M(v)=l} \eta_v \cdot \pi_v(x). \quad (22)$$

All the entries of the local and global kernels are thus truncated expolynomial functions, provided that the static density functions associated with timed transitions are also truncated expolynomial functions.

**Fig. 6** reports a high level pseudo-code of a possible algorithm computing the global and local kernels through the construction of the transient stochastic trees rooted in every different regenerative class and limited to the first regenerative epoch (with an error threshold  $\epsilon$  for behaviours within time  $x$ ). In [Appendix](#), we use a simple model to illustrate the calculations step by step.

#### 4.3. Limiting the complexity of the local kernel

As mentioned above, derivation of the entries of the local and global kernels present different levels of difficulty. We discuss here two different general cases where the number of entries of the local kernel needed for the analysis can be limited.

In the first case, assume that we want to compute the distribution of the time to reach a given marking for the first time, i.e., our goal is to determine a first passage time distribution. This problem can be faced without computing any entry of the local kernel as follows. We modify the model in such a way that the considered marking becomes absorbing, i.e., the model gets blocked as soon as the marking is reached. This way, when the marking is reached the process is in a regeneration point. Consequently, it is not necessary to describe what happens within a regeneration period and the global kernel of the modified model is sufficient to characterise the desired first passage time distribution. Having computed the entries of the global kernel,  $G_{ij}(t)$ , the fact that the process is not described inside a regeneration period can be formally introduced by setting the entries of the local kernel as

$$L_{ij}(t) = \begin{cases} 1 - \sum_k G_{ik}(t) & i = j \\ 0 & i \neq j \end{cases}$$

which implies that the process remains in the same marking for the whole regeneration period. Note that this way we simplified the original MRGP into a semi-Markov process and hence the computation itself can be carried out through techniques for SMPs.

In the second case, assume that we are interested in the transient probabilities of a single marking, in particular, marking 1. In order to compute the desired probability, we need to compute all entries of the global kernel,  $G_{ij}(t)$ , and those entries of the local kernel that lead to marking 1,  $L_{i1}(t)$ . The remaining entries of the local kernel can be simply set to

$$L_{ij}(t) = \begin{cases} 1 - \sum_k G_{ik}(t) - L_{i1}(t) & i = j \\ 0 & i \neq j \end{cases} \quad j \neq 1$$

because inside a regeneration period all we have to know is if the process is in marking 1 or somewhere else.

By performing the analysis in the above described manner we decrease the number of local kernel entries that have to be computed. The size of the kernels can also be reduced if there exist markings that are not involved in the analysis objectives and do not play a role in the global kernel.

#### 4.4. Termination

A condition sufficient for guaranteeing termination in the enumeration of transient stochastic classes reached within the first regeneration period can be conveniently characterised in terms of cyclic behaviours occurring in the graph of underlying state classes.

**Lemma 4.1.** *If the class graph is finite and every cycle that it contains visits at least one regenerative state class, then the number of transient classes visited before reaching the first regenerative transient stochastic class is finite.*

**Proof.** Ab absurdo, let  $\rho$  be an infinite path in the transient stochastic tree that never reaches a node associated with a regeneration class. Following the same steps of the proof of Lemma 3.4,  $\rho$  includes a finite prefix  $\rho_{pre} \rightarrow \rho_{cycle}$  such that the projection of  $\rho_{cycle}$  on the class graph is a cycle. By hypothesis, this cycle visits a regeneration state class  $S_r$ , which implies that within the (finite) termination of  $\rho_{pre} \rightarrow \rho_{cycle}$  the enumeration has visited a transient stochastic regeneration class.  $\square$

Following an aggregation approach such as that proposed for steady state analysis in [16], this condition can be relaxed so as to encompass cycles of immediate or exponential transitions that do not change the enabling status of generally distributed transitions. However, when the state class graph accepts cyclic behaviours that always maintain some GEN transition persistent, then the model accepts behaviours that never reach a regeneration and thus results in an unbounded number of transient stochastic classes. In this case, termination can still be guaranteed by Lemma 3.5 under the assumption of some non-null threshold  $\epsilon$  and upper time bound  $x$ .

#### 4.5. Approximate regenerative analysis

If an MRGP can exhibit an infinite sequence of transitions without regeneration points, then the enumeration of the classes that are needed to characterise a regeneration period does not terminate. Even in this case, however, it is possible to perform an analysis based on an approximate description of the regeneration periods. This approximate approach can be also applied to MRGPs with regeneration periods characterised by finite transient stochastic trees to reduce the number of nodes needed for the analysis.

The approximate analysis is based on a partial characterisation of the regeneration periods which can be obtained by a partial enumeration of those classes that can appear in the regeneration periods. The resulting transient tree, unlike the tree resulting from a complete enumeration, contains leaves representing non-regenerative classes. Let us denote by  $\hat{L}_{ij}(t)$  and  $\hat{G}_{ij}(t)$  the entries of the kernels calculated from the classes contained in this partial tree according to Eqs. (21)–(22). It is easy to show that the following relations hold between  $\hat{L}_{ij}(t)$ ,  $\hat{G}_{ij}(t)$  and the exact kernels  $L_{ij}(t)$ ,  $G_{ij}(t)$  of the MRGP:

$$\hat{L}_{ij}(t) \leq L_{ij}(t), \quad \hat{G}_{ij}(t) \leq G_{ij}(t).$$

Moreover, as opposed to (18), we have

$$\lim_{t \rightarrow \infty} \sum_j \hat{L}_{ij}(t) + \hat{G}_{ij}(t) < 1.$$

It follows that calculating  $\hat{\pi}_{ij}(t)$  from the defective kernels  $\hat{L}_{ij}(t)$  and  $\hat{G}_{ij}(t)$  in the fashion of (19) leads to probabilities

$$\hat{\pi}_{ij}(t) \leq \pi_{ij}(t)$$



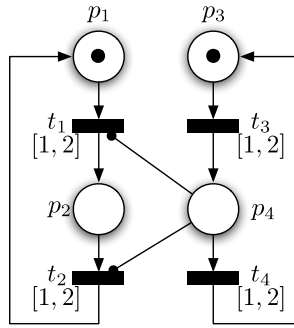


Fig. 7. Petri net representation of the modified  $G/G/1/2/2$  preemptive queue.

providing a lower bound of the real transient probabilities. A generally less tight upper bound can also be given based on  $\hat{\pi}_{ij}(t)$  as

$$\pi_{ij}(t) \leq 1 - \sum_{j|j \neq i} \hat{\pi}_{ij}(t).$$

In general, less defective kernels result in more precise approximated probabilities  $\hat{\pi}_{ij}(t)$ . Unfortunately, however, the error due to defective kernels grows as time advances and if all states are recurrent we have

$$\lim_{t \rightarrow \infty} \hat{\pi}_{ij}(t) = 0.$$

Consequently, the above approach can provide useful bounds of the transient probabilities only for a finite time horizon. In order to obtain an approximation for a long time horizon or for steady state probabilities, one can apply the following renormalised kernels

$$\check{L}_{ij}(t) = \hat{L}_{ij}(t), \quad \check{G}_{ij}(t) = \hat{G}_{ij}(t) \frac{\left(1 - \sum_k \hat{L}_{ij}(t)\right)}{\sum_k \hat{G}_{ik}(t)}$$

which satisfy Eq. (18). In Section 4.6, we will illustrate the use of both  $\hat{L}_{ij}(t)$ ,  $\hat{G}_{ij}(t)$  and  $\check{L}_{ij}(t)$ ,  $\check{G}_{ij}(t)$ .

#### 4.6. Computational experience

In principle, the original  $G/G/1/2/2$  preemptive queue of Fig. 2 cannot be analysed through the method based on regenerative points as the underlying process can visit an infinite sequence of non-regenerative classes. We consider here a simplified version of the  $G/G/1/2/2$  queue in which an additional inhibitor arc is added from place  $p_4$  to transition  $t_1$ , so that the arrival of client 2 entails both service interruption and arrival inhibition of client 1 (Fig. 7). This way, the corresponding process always reaches a regeneration in a finite number of steps and becomes amenable to the exact analysis based on regeneration points. Solving for  $x = 100$  requires the construction of 4 transient stochastic trees limited to the first regeneration epoch, for a total of 12 enumerated classes (1 s computation); results are shown in Fig. 8 for two different time scales.

The original  $G/G/1/2/2$  preemptive queue of Fig. 2 can be analysed through the approximated approach of Section 4.5. Fig. 9 depicts the results for different values of the approximation threshold  $\epsilon$ . Marking probabilities are defective and the curve denoted as “error” shows the missing probability which grows with the approximation threshold and with the transient time. It is worth noting that, unlike the approach of Section 3, in this case the approximation threshold does not represent an upper bound for the defectiveness of the actual evaluated solution. The computation requires about four minutes with  $\epsilon = 0.01$  and about 17 min with  $\epsilon = 0.0001$ .

In Fig. 10, we illustrate the results obtained with the non-defective approximate kernels,  $\check{L}_{ij}(t)$  and  $\check{G}_{ij}(t)$ . In particular, the figures report the transient probabilities of two markings calculated with defective and non-defective kernels (using the same approximation threshold,  $\epsilon = 0.01$ ) and the exact transient probabilities. The exact values were obtained by the supplementary variable approach [9] which requires for this model to memorise vectors with a total of about  $10^7$  entries and took about 30 h of computation. We performed the calculations with several values of  $\epsilon$  and found that for this model the maximal absolute error of the transient probabilities obtained by the non-defective kernels is smaller than the approximation threshold  $\epsilon$ , but this cannot be guaranteed in general. It is worth to note that the approach leads to very good approximation of the steady state probabilities.

Fig. 11 depicts a more complex model, introduced in [15] to illustrate steady state analysis of stochastic state classes; although the distribution of the length of the regeneration periods still has finite support, the intrinsic complexity of its

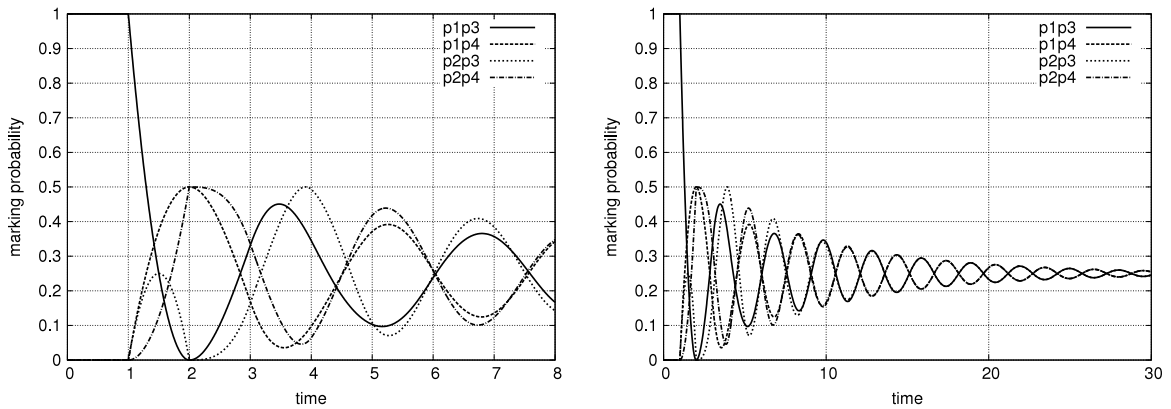


Fig. 8. Transient probabilities of the 4 markings reachable in the modified  $G/G/1/2/2$  queue for two different time scales.

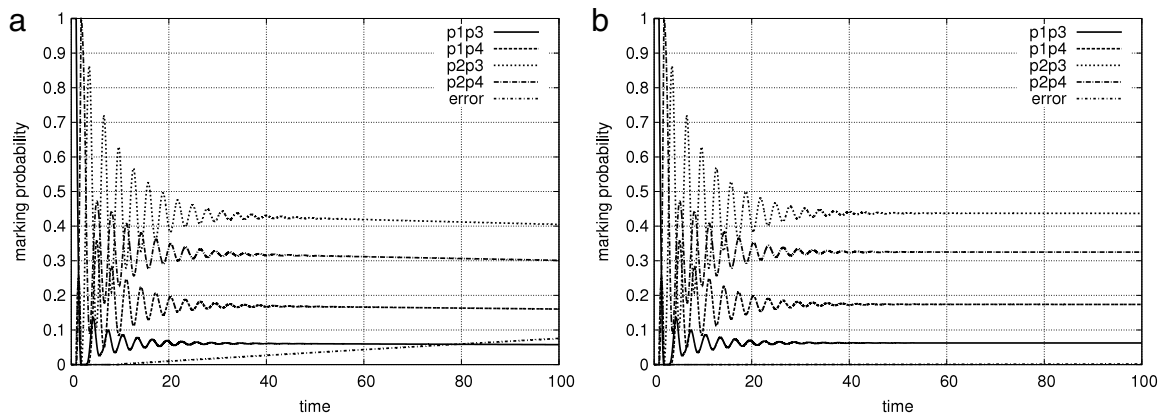


Fig. 9. Transient marking probabilities of  $G/G/1/2/2$  preemptive queue, computed with approximate kernels,  $\hat{L}_{ij}(t)$  and  $\hat{G}_{ij}(t)$ , for different approximation thresholds: (a)  $\epsilon = 0.01$ , 25 enumerated classes; (b)  $\epsilon = 0.0001$ , 36 enumerated classes.

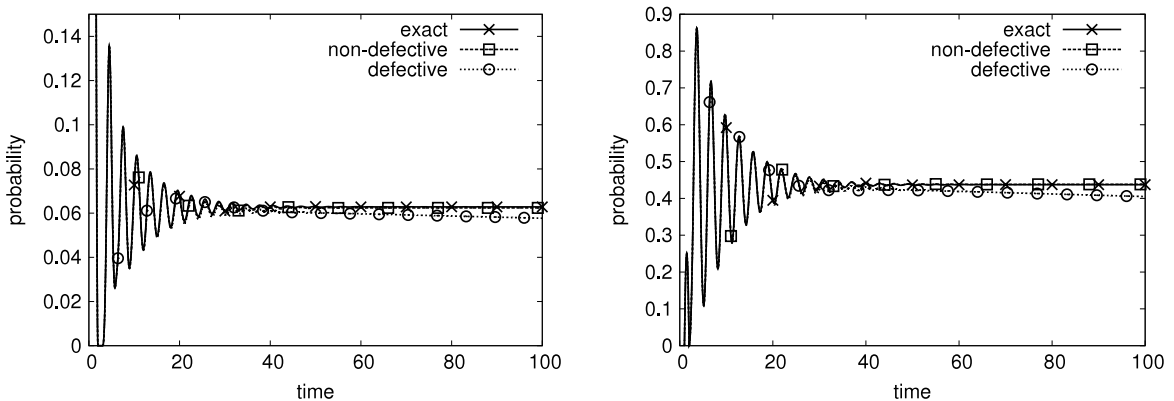
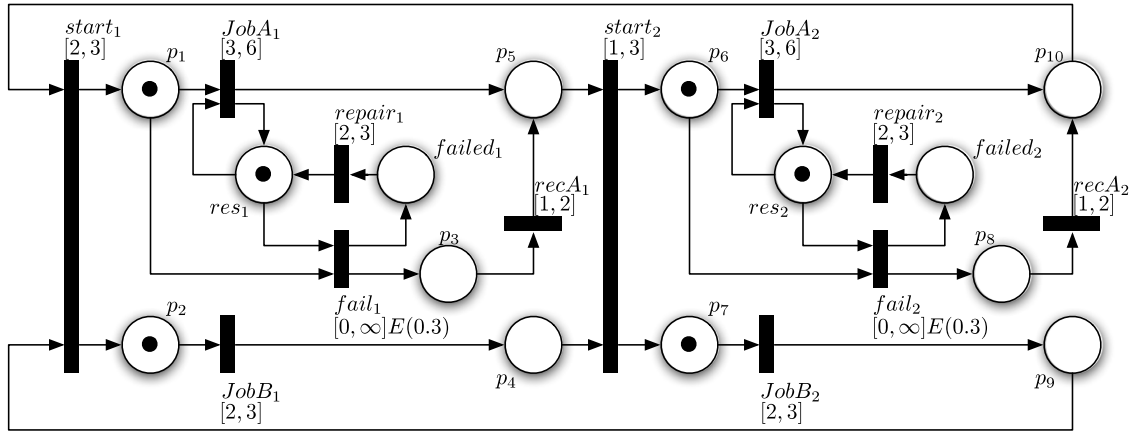
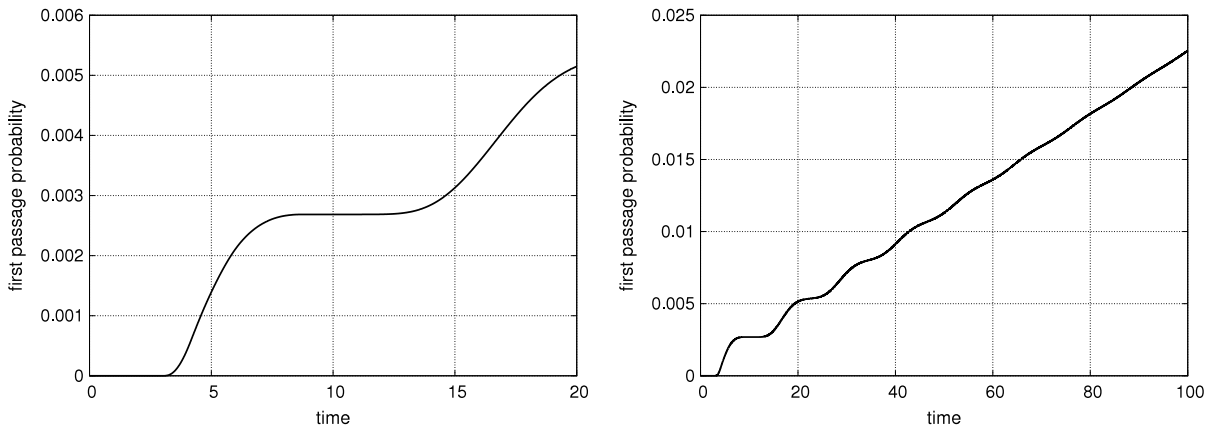


Fig. 10. Transient probabilities of markings  $p1p3$  (left) and  $p2p3$  (right): exact values, approximation with non-defective kernels ( $\epsilon = 0.01$ ) and approximation with defective kernels ( $\epsilon = 0.01$ ).

structure makes it more suitable than the previous simpler examples for a real benchmark of the proposed approach. This model represents a system composed of two production cells alternating through the passage of control to each other (transitions *start1* and *start2*). The two cells are identical with the single exception that transitions *start1* and *start2* have different earliest firing times. Both cells are composed of two parallel activities, named *JobA* and *JobB*, which are associated with uniform duration. In both cells, *JobA* requires a resource called *res* which may fail during the usage according to an exponential distribution with rate parameter 0.3. If failure occurs then *JobA* is interrupted and a recovery activity (*recA*) together with a repair action is started. The recovery and the repair actions both require a duration with uniform distribution.



**Fig. 11.** Petri net modelling a system with two production cells.  $E(0.3)$  denotes an exponential transition with rate 0.3 and all other transitions are uniform on the indicated interval.



**Fig. 12.** Probability of having reached double failure at least once as function of time for two different time scales.

We concentrate on the situation when both resources are in the failed state (i.e. both places  $failed_1$  and  $failed_2$  are marked) and compute two related performance indices.

First, we consider the distribution of the time to reach such double failure for the first time. As described in Section 4.3, this kind of measure can be computed without computing any entry of the local kernel. Originally there are 6 markings in the model that can be reached through regeneration. We stop the model as soon as double failure is reached and this adds a 7th regenerative marking (an absorbing one) to the model. Accordingly, the global kernel is 7 by 7. The enumeration of the classes has to be performed starting from the 6 different non-absorbing markings. In total, the 6 enumerations result in 93 classes. About 30 classes are regenerative and the global kernel is built based on these classes. Once the global kernel is defined, MRGP or SMP transient analysis techniques can be used to compute the desired probability. The computation takes about one hour and the results are depicted in Fig. 12.

Second, we consider the probability that at time  $t$  the system is in such a marking that both resources are failed. As we are interested in a particular situation only, once again we can perform the computations based on a selective analysis of the kernels as described in Section 4.3. The enumeration has to be performed for the 6 different regenerative markings which results in a total of about 130 classes. The number is higher than before because the model is not stopped anymore when double failure is reached. About 40 of the classes are regenerative and these are needed to construct the global kernel. For what concerns the local kernel, we need only those entries that corresponds to double failure. There are only 3 such non regenerative classes in which both failed places are marked. The resulting kernels are 7 by 7 as before because we need the 6 markings that can be reached through regeneration and a 7th state where we keep track of being in double failed state. This 7th state is reached only by the local kernel. The results of the computation (obtained in about one hour), which were validated by simulation, are depicted in Fig. 13.

As a measure of comparison between standard and regenerative analysis of this example, while the former needs to enumerate about 200 classes to cover the transient interval up to  $T = 8$ , the latter obtains global and local kernels performing partial enumerations from 6 different regenerative markings, for a total of about 130 classes.

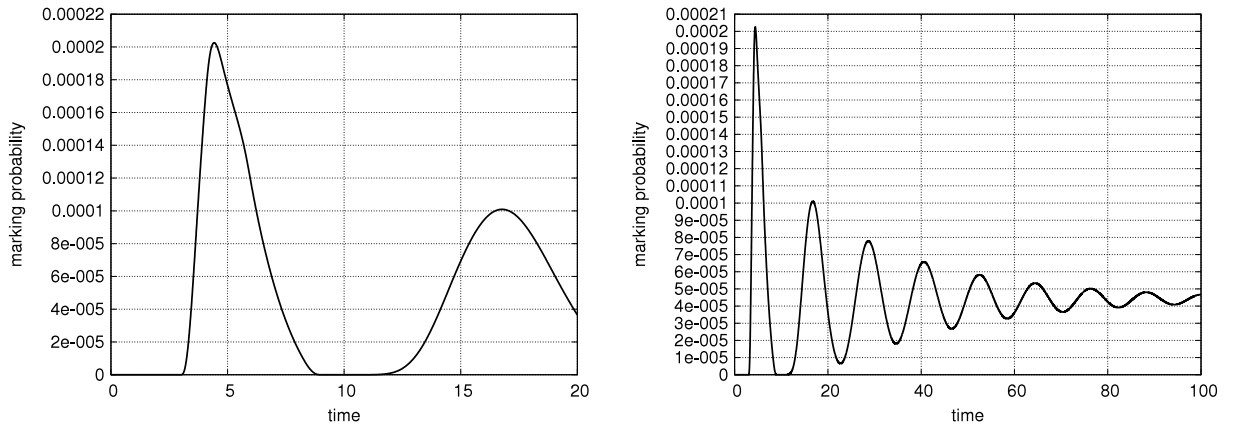


Fig. 13. Probability of being in double failure as function of time for two different time scales.

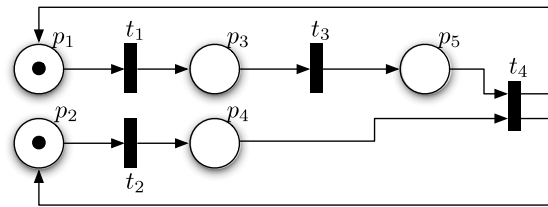


Fig. 14. A trivial model to illustrate the calculations.

## 5. Conclusions

In this paper, a technique for the transient analysis of generalised semi-Markov processes has been proposed. This technique, which is an extension of the method of stochastic state classes, is based on keeping track of the elapsed time in the state density function of the classes. The approach is applicable, in principle, to any generalised semi-Markov process where all the active timers proceed with the same speed, and it results in the possibility of symbolic derivation of transient measures in the case when all timings are expolynomial.

The time and memory complexity of the approach depend on many factors: the number of states of the model, the level of concurrency of the timers, and the time to fire distributions. A detailed complexity analysis was out of scope for this paper and will be studied in the future. It has been shown however that, if the model always visits regeneration points within a finite number of steps, the approach can be used to characterise the regeneration period even if the involved stochastic process contains multiple concurrent non-exponential activities; moreover, an approximation method has been developed in order to enable the analysis also for those processes allowing infinite sequences of steps without regeneration points and to limit the computational complexity when dealing with large and complex models. Based on a Java implementation in the Oris tool, the method has been tested and resulted to be applicable to models whose transient analysis has not been carried out before.

## Appendix

We consider the model depicted in Fig. 14 and assume that the firing time distribution of transitions  $t_1$  and  $t_2$  is uniform with support interval  $[0, 3]$  and the firing time distribution of  $t_3$  is given by the PDF  $f_3(x) = 0.5x$  on the support  $[0, 2]$ . We perform the calculations that lead to the description of a regeneration period starting in the marking  $p_1p_2$ . Accordingly, the enumeration of the transient classes has to be started in the class corresponding to this marking and carried out in such a way that successors of classes corresponding to regeneration points are not generated. Once the enumeration is finished, the entries of the local and global kernel are derived from the transient classes.

A transient class is given by a marking and the joint PDF of the opposite of the elapsed time and the time to fire values of the enabled transitions. In the initial marking the elapsed time is zero and the time to fire values are determined simply by the static firing time distribution of the enabled transitions. Therefore, the density associated with the initial class  $\Sigma_0$  is

$$f_0(a, x_1, x_2) = \begin{cases} \frac{1}{9}\delta(a) & 0 \leq x_1 \leq 3, 0 \leq x_2 \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

where  $\delta(\cdot)$  denotes the Dirac delta function. (Both in the expression above and in the following we provide the domain explicitly together with the state density function.) We consider first the firing of  $t_1$  leading to class  $\Sigma_1$  where the marking

is  $p_2p_3$ . As the firing time distributions of  $t_1$  and  $t_2$  are identical the integral in the denominator of (7) corresponding to the probability that  $t_1$  fires first yields  $1/2$ . Next we have to perform the time advancement and projection step according to (8) as

$$\int_0^{x_2} 2f_0(a + x_1, x_1, x_2 + x_1)dx_1 = \begin{cases} \frac{2}{9} & -3 \leq a \leq 0, 0 \leq x_2 \leq 3, 0 \leq x_2 - a \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

where the constraint  $0 \leq x_2 - a \leq 3$  is due to the fact that the elapsed time,  $-a$ , plus the time to fire of  $t_2$ ,  $x_2$ , has to be between 0 and 3. As no disabling happens by the firing of  $t_1$ , the last step in computing the density associated with class  $\Sigma_1$  is to consider the enabling of  $t_3$  applying (10). This leads to

$$f_1(a, x_2, x_3) = \begin{cases} \frac{1}{9}x_3 & -3 \leq a \leq 0, 0 \leq x_2 \leq 3, \\ & 0 \leq x_3 \leq 2, 0 \leq x_2 - a \leq 3 \\ 0 & \text{otherwise.} \end{cases}$$

The firing of  $t_2$  in class  $\Sigma_0$  can be considered the same way with the difference that no enabling takes place. The resulting class,  $\Sigma_2$ , where the marking is  $p_1p_4$  has associated density

$$f_2(a, x_1) = \begin{cases} \frac{2}{9} & -3 \leq a \leq 0, 0 \leq x_1 \leq 3, 0 \leq x_1 - a \leq 3 \\ 0 & \text{otherwise.} \end{cases}$$

Next we consider the firing of transition  $t_2$  leaving class  $\Sigma_1$  leading to the new class  $\Sigma_3$  with marking  $p_3p_4$ . This happens with probability

$$\int_{a=0}^{\infty} \int_{x_3=0}^{\infty} \int_{x_2=0}^{x_3} f_1(a, x_2, x_3)dx_2dx_3da = 2/3.$$

Since neither enabling nor disabling happens due to the firing of  $t_2$ , the time advancement and projection step according to (8) yields directly the density of the class  $\Sigma_3$

$$f_3(a, x_3) = \int_{x_2=0}^{x_3} \frac{3}{2}f_1(a + x_2, x_2, x_3 + x_2)dx_2 = \begin{cases} \frac{1}{12}a(a - 2x_3) & 0 \leq x_3 \leq 2, -2 \leq a \leq 0, 0 \leq x_3 - a \leq 2 \\ \frac{1}{12}(4 - x_3^2) & 0 \leq x_3 \leq 2, -3 \leq a \leq 0, 2 \leq x_3 - a \leq 5 \\ 0 & \text{otherwise} \end{cases}$$

where once again we have inequalities regarding the sum of the elapsed time,  $-a$ , and the time to fire of  $x_3$ . The other possibility in  $\Sigma_1$  is the firing of  $t_3$  which happens with probability  $1/3$ . The resulting class  $\Sigma_4$  is with marking  $p_3p_4$  and its density can be computed by time advancement and projection (8) which leads to

$$f_4(a, x_2) = \int_{x_3=0}^{x_2} 3f_1(a + x_3, x_2 + x_3, x_3)dx_3 = \begin{cases} \frac{2}{3} & 0 \leq x_2 \leq 1, -3 \leq a \leq -2, 0 \leq x_2 - a \leq 3 \\ \frac{a^2}{6} & 0 \leq x_2 \leq 3, -2 \leq a \leq 0, 0 \leq x_2 - a \leq 3 \\ 0 & \text{otherwise.} \end{cases}$$

From  $\Sigma_2$ , by the firing of the only enabled transition  $t_1$ , a new class,  $\Sigma_5$ , is reached with probability 1. Due to the firing of  $t_1$ ,  $t_3$  would get enabled but, as the firing of  $t_1$  constitutes a regeneration point and our aim is to characterise only the regeneration period, the enabling step given in (10) is not performed. Accordingly, the density associated with  $\Sigma_5$  is obtained by the usual time advancement and projection step according to (8) which yields

$$f_5(a) = \int_0^{\infty} f_2(a + x_1, x_1)dx_1 = \begin{cases} -\frac{2a}{9} & -3 \leq a \leq 0 \\ 0 & \text{otherwise} \end{cases}$$

where the inequality  $-3 \leq a \leq 0$  means that the elapsed time having reached this class is between 0 and 3. From  $\Sigma_3$  as well a single outgoing event is possible, namely, the firing of  $t_3$  which leads to a new class,  $\Sigma_6$ , through a regeneration point. Accordingly, the enabling of  $t_4$  is not performed and the time advancement and projection step according to (8) is enough to characterise  $\Sigma_6$  where the marking is  $p_4p_5$ . We have

$$f_6(a) = \int_0^{\infty} f_3(a + x_3, x_3)dx_3 = \begin{cases} -\frac{(-1+a)(5+a)^2}{36} & -5 \leq a \leq -3 \\ \frac{4}{9} & -3 \leq a \leq -2 \\ -\frac{a^3}{18} & -2 \leq a \leq 0 \\ 0 & \text{otherwise} \end{cases}$$

where the inequalities indicate that the elapsed time having reached this class is between 0 and 5. The last outgoing event to consider is the firing of  $t_2$  in class  $\Sigma_4$  which leads to a new class,  $\Sigma_7$ , by regeneration where the marking is  $p_2p_5$ . The density associated with  $\Sigma_7$  is

$$f_7(a) = \int_0^\infty f_2(a + x_2, x_2) dx_2 = \begin{cases} -\frac{2}{9}(4 + 3a) & -3 \leq a \leq -2 \\ -\frac{a^3}{18} & -2 \leq a \leq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Having enumerated the classes from the initial marking  $p_1p_2$  up to the regeneration points, the entries of the kernels can be determined by the application of (13). Classes reached by regeneration contribute to the global kernel while the others to the local kernel. Note that even in this simple example there is a marking, namely,  $p_3p_4$  that can be reached both through a regenerative and a non-regenerative event. It is also possible that the same marking appears in different regenerative (non-regenerative) classes; in this case these classes have to be aggregated to the same global (local) kernel entry.

Hereinafter, we compute some entries of the local and global kernels. The states in the kernels will be referred to according to the table

|         |          |          |          |          |          |          |
|---------|----------|----------|----------|----------|----------|----------|
| marking | $p_1p_2$ | $p_2p_3$ | $p_1p_4$ | $p_3p_4$ | $p_2p_5$ | $p_4p_5$ |
| state   | 0        | 1        | 2        | 3        | 4        | 5        |

As an example for a local kernel entry, we consider  $L_{12}(t)$  which corresponds to the probability that the regeneration period started at time 0 in state 0 (marking  $p_1p_2$ ), it is not yet finished at time  $t$  and the process is in state 1 (marking  $p_2p_3$ ) at time  $t$ . This probability can be calculated by (13) considering the density in class  $\Sigma_1$  and the fact that this class is passed through with probability 1/2. As a result we have

$$L_{12}(t) = \frac{1}{2} \int_{a=0}^t \int_{x_2=t-a}^\infty \int_{x_3=t-a}^\infty f_1(-a, x_2, x_3) da dx_2 dx_3 = \begin{cases} \frac{1}{108} t(36 - 12t - 3t^2 + t^3) & 0 \leq t \leq 2 \\ \frac{4}{27} (3 - t) & 2 \leq t \leq 3 \\ 0 & \text{otherwise.} \end{cases}$$

Next we calculate the global kernel entry,  $G_{05}(t)$ , which gives the probability that the regeneration period started at time 0 in state 0 (marking  $p_1p_2$ ), it has been finished before time  $t$  and the regeneration leads to state 5 (marking  $p_4p_5$ ). Both  $\Sigma_6$  and  $\Sigma_7$  are with marking  $p_4p_5$  and hence they will both contribute to  $G_{05}(t)$ . Class  $\Sigma_6$  is reached with probability 1/3 while  $\Sigma_7$  with probability 1/6. Applying (13) we have

$$G_{05}(t) = \frac{1}{3} \int_0^t f_6(-a) da + \frac{1}{6} \int_0^t f_7(-a) da = \begin{cases} \frac{t^4}{144} & 0 \leq t \leq 2 \\ \frac{-2 + t^2}{-159 + 100t + 30t^2 - 12t^3 + t^4} & 2 \leq t \leq 3 \\ \frac{18}{432} & 3 \leq t \leq 5 \\ \frac{1}{2} & 5 \leq t. \end{cases}$$

It can be seen that the entry of the local kernel requires to deal with a more complicated integral than the one used for the global kernel.

Several of the functions derived above have different symbolic forms in different intervals. None of these have however jump discontinuities.

## References

- [1] P.W. Glynn, A GSMP formalism for discrete-event systems, *Proceedings of the IEEE* 77 (1989) 14–23.
- [2] G. Ciardo, R. German, C. Lindemann, A characterization of the stochastic process underlying a stochastic petri net, *IEEE Transactions on Software Engineering* 20 (7) (1994) 506–515.
- [3] P.J. Haas, G.S. Shedler, Stochastic petri net representation of discrete event simulations, *IEEE Transactions on Software Engineering* 15 (4) (1989) 381–393.
- [4] H. Choi, V.G. Kulkarni, K. Trivedi, Markov regenerative stochastic petri nets, *Performance Evaluation* 20 (1994) 337–357.
- [5] A. Bobbio, M. Telek, Markov regenerative SPN with non-overlapping activity cycles *Int. Computer Performance and Dependability Symp. - IPDS95*, pp. 124–133, 1995.
- [6] R. German, C. Lindemann, Analysis of stochastic petri nets by the method of supplementary variables, *Performance Evaluation* 20 (1994) 317–335.
- [7] E. Cinlar, *Introduction to Stochastic Processes*, Prentice Hall, Englewood Cliffs, 1975.
- [8] R. German, D. Logothetis, K.S. Trivedi, Transient analysis of markov regenerative stochastic petri nets: a comparison of approaches, in: *PNNP'95: Proceedings of the Sixth International Workshop on Petri Nets and Performance Models*, IEEE Computer Society, Washington, DC, USA, 1995, p. 103.
- [9] D. Cox, The analysis of non-markovian stochastic processes by the inclusion of supplementary variables, *Proceedings of the Cambridge Philosophical Society* 51 (1955) 433–440.
- [10] M. Telek, A. Horváth, Supplementary variable approach applied to the transient analysis of age-MRSPNs, in: *Proceedings of International Performance and Dependability Symposium*, 1998, pp. 44–51.
- [11] R. German, M. Telek, Formal relation of Markov renewal theory and supplementary variables in the analysis of Stochastic Petri Nets, in: *Proceedings of 8th International Workshop on Petri Nets and Performance Models*, pp. 64–73, 1999.



- [12] C. Lindemann, G.S. Shedler, Numerical analysis of deterministic and stochastic petri nets with concurrent deterministic transitions, *Performance Evaluation* 27–28 (1996) 565–582.
- [13] C. Lindemann, A. Thiemmer, Transient analysis of deterministic and stochastic petri nets with concurrent deterministic transitions, *Performance Evaluation* 36–37 (1999) 35–54.
- [14] G. Bucci, R. Piovosi, L. Sassoli, E. Vicario, Introducing probability within state class analysis of dense time dependent systems, in: *Proceedings of the 2nd International Conference on the Quantitative Evaluation of Systems (QEST)*, 2005.
- [15] L. Carnevali, L. Sassoli, E. Vicario, Using stochastic state classes in quantitative evaluation of dense-time reactive systems, *IEEE Transactions on Software Engineering* (2009).
- [16] A. Horváth, E. Vicario, Aggregated stochastic state classes in quantitative evaluation of non-markovian stochastic petri nets, in: *Proceedings of 6th International Conference on the Quantitative Evaluation of Systems (QEST)*, Budapest, Hungary, September 2009.
- [17] E. Vicario, Static analysis and dynamic steering of time dependent systems using time petri nets, *IEEE Transactions on Software Engineering* (2001).
- [18] B. Berthomieu, M. Diaz, Modeling and verification of time dependent systems using time petri nets, *IEEE Transactions on Software Engineering* 17 (3) (1991).
- [19] D. Dill, Timing assumptions and verification of finite-state concurrent systems, in: *Proceedings of Workshop on Computer Aided Verification Methods for Finite State Systems*, 1989.
- [20] L. Carnevali, L. Ridi, E. Vicario, Partial stochastic characterization of timed runs over dbm domains, in: *Proceedings of the 9th International Workshop on Performability Modeling of Computer and Communication Systems (PMCCS 2009)*, Eger, Hungary, September 2009.
- [21] P. Merlin, D.J. Farber, Recoverability of communication protocols, *IEEE Transactions on Communications* 24 (9) (1976).
- [22] V.G. Kulkarni, *Modeling and Analysis of Stochastic Systems*, Chapman and Hall, 1995.
- [23] A. Horváth, L. Ridi, E. Vicario, Transient analysis of generalised semi-Markov processes using transient stochastic state classes, *Proceedings of the 7th International Conference on Quantitative Evaluation of Systems (QEST'10)* (2010).
- [24] J. Bengtsson, W. Yi, Timed automata: semantics, algorithms and tools, in: *Concurrency and Petri Nets*. LNCS, 3098, 2004.
- [25] R. Sahner, K. Trivedi, Reliability modeling using SHARPE, *Reliability*, *IEEE Transactions on* 36 (2) (2009) 186–193.
- [26] G. Bucci, L. Carnevali, L. Ridi, E. Vicario, Oris: a tool for modeling, verification and evaluation of real-time systems, *STTT*, *International Journal on Software Tools for Technology Transfer* 12 (5) (2010) 391–403.
- [27] L. Carnevali, L. Ridi, E. Vicario, A framework for simulation and symbolic state space analysis of non-markovian models, *SAFECOMP* (2011) 409–422.
- [28] D. Logothetis, K. Trivedi, A. Puliafito, Markov regenerative models, in: *Proceedings of the International Computer Performance and Dependability Symposium on Computer Performance and Dependability Symposium (IPDS'95)*, IEEE Computer Society, Washington, DC, USA, 1995.



**András Horváth** was born in 1974 in Budapest where he received the M.Sc. degree in Computer Science from the University of Technology and Economics. From 1998 to 2002 he was a Ph.D. student supervised by Miklós Telek at the same university. From 2003 he is a researcher at the University of Turin (Italy). His research interests are in the area of stochastic processes including performance analysis of non-Markovian systems and modeling issues of communication networks.



**Marco Paolieri** is currently a Ph.D. student in Computer Engineering, Multimedia and Telecommunications at the University of Florence, where he received his B.S. and M.S. in Computer Engineering (2011). His research interests focus on the quantitative evaluation of performance and dependability of concurrent real-time systems, stochastic modeling and software verification.



**Lorenzo Ridi** received the bachelor degree in Computer Engineering in 2005 and the master degree in Computer Engineering in 2007. He is now a student in the Ph.D. program in Computer Engineering, Multimedia and Telecommunications at the University of Florence. His research activity mainly focuses on formal techniques for the specification, the qualitative verification and the quantitative validation of stochastic time-dependent systems, and their integration in the development lifecycle of real-time software. He is a member of the Software Technologies Laboratory (STLab) of the University of Florence.



**Enrico Vicario** is a full professor of computer science at the engineering School of the University of Florence. His scientific and technological activity is presently focused on quantitative evaluation of concurrent non-Markovian models, formal methods for correctness verification and model-driven development of real-time systems, software engineering methods and software architectures.