Markov Regenerative Stochastic Petri Nets with Age Type General Transitions

Miklós Telek[1] and Andrea Bobbio[2]

[1] Department of Telecommunications
Technical University of Budapest, 1521 Budapest, Hungary

[2] Dipartimento di Elettronica per l'Automazione
Università di Brescia, 25123 Brescia, Italy

Abstract. Markov Regenerative Stochastic Petri Nets *(MRSPN)* have been recently introduced in the literature with the aim of combining exponential and non-exponential firing times into a single model. However, the realizations of the general *MRSPN* model, so far discussed, require that at most a single non-exponential transition is enabled in each marking and that its associated memory policy is of enabling type. The present paper extends the previous models by allowing the memory policy to be of age type and by allowing multiple general transitions to be simultaneously enabled, provided that their enabling intervals do not overlap. A final completely developed example, that couldn't have been considered in previous formulations, derives the closed form expressions for the transient state probabilities for a queueing system with *preemptive resume (prs)* service policy.

Key words: Markov regenerative processes, Stochastic Petri Nets, Queueing systems with preemptive resume service, Transient analysis.

1 Introduction

Markov Regenerative Stochastic Petri Nets are defined as the class of Stochastic Petri Nets (*SPN*) whose marking process is mapped into a Markov Regenerative Process (*MRGP*) [11, 8]. The concept of *MRSPN* was first proposed by Choi et al. in [7], when they recognized that the *Deterministic and Stochastic PN (DSPN)* model, defined by Ajmone and Chiola in [2], could be considered as a member of this class.

In the *DSPN* [2], at most one deterministic transition is enabled in each marking, and the deterministic transitions are assigned an enabling memory policy (after the taxonomy in [1]). The steady state solution algorithm, provided in [2], was then revisited in [16] and some structural extensions were proposed in [10]. Choi et al. [7] developed the transient analysis of the same *DSPN* model, based on the transient equations of the underlying Markov regenerative process. In [8, 13], deterministic transitions were replaced by generally distributed transitions, while in [9], the case of multiple deterministic transitions of enabling memory type activated in the same marking was considered.

The analysis technique developed for this class of models, consists in identifying a sequence of regeneration points and by analysing the behavior of the marking process between any two successive regeneration points. The restriction of the marking process between two successive regeneration points is called the subordinated process [16]. All the models discussed in the mentioned references require that the generally distributed (or deterministic) transitions are assigned a firing policy of enabling memory type [1]. The enabling memory policy means that each time the transition becomes enabled its firing time is resampled from the original distribution and the time spent without firing in prior enabling periods is lost. In [2, 16, 7, 8], the subordinated processes are restricted to be Continuous Time Markov Chains *(CTMC)*, while the steady state analysis of semi-Markov subordinated processes has been investigated in [9].

The aim of this paper is to introduce a new class of models, called *AgeMRSPN*, characterized by the fact that generally distributed transitions have an age memory policy, and multiple general transitions can be simultaneously enabled provided that a dominant transition exists whose enabling period determines the occurrence of two successive regeneration time points. It will be shown that the above assumptions entail that the subordinated processes can be reward semi-Markov processes. The age memory policy means that each time the transition becomes enabled its firing time is resumed from the previously attained value, so that the time possibly spent without firing in prior enabling periods is not lost. The age memory policy needs to be invoked to model *preemptive resume (prs)* service strategies, where the server is able to recover the execution of an interrupted job by keeping memory of the work already performed so that, upon restart, only the residual service needs to be completed.

A general closed form analytical solution for the transient state probabilities is derived in the Laplace transform domain. For the special case, in which the generally distributed transitions have an exponential polynomial *(EP)* firing time, an effective algorithm is developed. The numerical computation requires a combination of symbolic and numerical steps and is, in the present state of developement, restricted to small case examples.

After introducing the notation and the definition of an *AgeMRSPN* in Section 2, an analytical procedure for deriving the closed form transient equation for the transition probability matrix is presented in Section 3. Section 4 is devoted to illustrate a detailed derivation of the transient probabilities in a M/G/1/2/2 queuing system with *prs* service. This example revisits the case already studied in [2, 7, 8], but introduces modeling features that couldn't have been considered in the framework of the previous methodologies.

2 Markov Regenerative Stochastic Petri Nets

The untimed model is a marked Petri Net (*PN*) represented by a tuple $PN = (P, T, I, O, H, M)$, where P is the set of places, T the set of transitions, I, O and H the input, output and inhibitor functions respectively, and M is the marking. The reachability set $\mathcal{R}(M_0)$ is the set of all the markings that can be generated from an initial marking M_0. The marking process $\mathcal{M}(x)$ denotes the marking occupied by the *PN* at time x.

It is shown in [1] that, when a transition is assigned a non-exponential firing time, the nature of the marking process $\mathcal{M}(x)$ is univocally identified if a *memory policy* is attached to each transition. The memory policy specifies how the process is conditioned upon the past. Following [1], the memory policy is realized through a memory variable a_k, associated to each transition t_k. The memory variable is a functional that depends on the time during which t_k has been enabled according to the following three alternatives [1]:

- *Resampling policy* - The memory variable a_k is reset to zero at any change of marking.
- *Enabling memory policy* - The memory variable a_k accounts for the work performed by the activity corresponding to t_k from the last epoch in which t_k has been enabled. When transition t_k is disabled (even without firing) a_k is reset.
- *Age memory policy* - The memory variable a_k accounts for the work performed by the activity corresponding to t_k from its last firing up to the current epoch and is reset only when t_k fires.

At the entrance in a new marking, the residual firing time is computed for each enabled timed transition given its memory variable, so that the next marking is determined by the minimal residual firing time among the enabled transitions (*race policy* [1]). Since the three mentioned policies are equivalent for an exponential distribution, due to the memoryless property, the corresponding memory variable can be assumed identically zero. The set of transitions can be partitioned into a subset of exponential transitions (EXP) and a subset of generally distributed transitions (GEN).

A regeneration time point in a time homogeneous stochastic process is the epoch of entrance in a state in which the Markov property holds (i.e. the future evolution does not depend on the past history but only on the present state). A stochastic process for which a sequence of regeneration time points can be identified is called a Markov Regenerative Process [8, 11].

Definition 1. According to the semantics in [1], a regeneration time point in the marking process $\mathcal{M}(x)$ is the epoch of entrance in a marking $M_{(n)}$ in which all the memory variables are equal to 0. A SPN whose marking process $\mathcal{M}(x)$ is a Markov Regenerative Process is called a Markov Regenerative SPN (MRSPN).

The portion of the marking process confined between any two successive regeneration time points is called the subordinated process [16]. The subclass *MRSPN**, defined in [8], is obtained by restricting Definition 1 according to the following specifications: *i)* in each marking, at most a single GEN transition is enabled being all the other transitions EXP; *ii)* the memory policy associated to every GEN transition is of enabling memory type. As a consequence of the above specifications all the subordinated processes are *CTMC*'s. In order to remove the above restrictions, to some extent, the notion of active and dominant transition is introduced [6].

Definition 2. A transition is active when its memory variable is greater than zero; the activity cycle of a transition is the period of time in which the transition is active. A transition is dominant with respect to a subordinated process if its activity cycle determines the two successive regeneration time points in which the subordinated process is confined.

It has been shown in [6], that a solvable class of *MRSPN* corresponds to models in which the activity cycles of the GEN transitions do not overlap, and the subordinated processes are semi-Markovian.

Definition 3. An AgeMRSPN is a MRSPN in which:

i The set T is partitioned into EXP and GEN transitions;
ii To any GEN transition t_g a generally distributed random variable γ_g, with Cumulative Distribution Function $G_g(y)$, and a memory variable a_g with age memory policy is associated.
iii The regeneration intervals between any two successive regeneration time points are dominated by a single age memory GEN transition and the subordinated processes are semi-Markov.

A single realization of the marking process $\mathcal{M}(x)$ can be represented by the following timed execution sequence:

$$\mathcal{T}_E = \{ (\tau_0^*, M_{(0)}) ; (\tau_1^*, M_{(1)}) ; \ldots ; (\tau_i^*, M_{(i)}) ; \ldots \} \tag{1}$$

where τ_i^* represents a regeneration time point and $M_{(i)}$ the entered marking. By Definition 1, τ_i^* is such that at the entrance in $M_{(i)}$ all the memory variables are zero. The successive regeneration time point τ_{i+1}^* is derived from τ_i^* as follows:

1. If no GEN transition is enabled in marking $M_{(i)}$, τ_{i+1}^* is the first time after τ_i^* that a state change occurs.
2. If an age memory GEN transition t_g starts its activity cycle in marking $M_{(i)}$ and the subordinated process is dominated by t_g, τ_{i+1}^* is the firing time of t_g.

In the case 1) above, the subordinated process between two consecutive regeneration time points is a single step *CTMC* since only EXP transitions are enabled and any firing provides the next regeneration point.

In the case 2) above, during $[\tau_i^*, \tau_{i+1}^*)$, the *PN* can evolve in the subset of $\mathcal{R}(M_0)$ reachable from $M_{(i)}$, during the activity cycle of the dominant GEN transition t_g and the subordinated process inside this interval is semi-Markov.

Definition 3 has two major implications. Since the subordinated process is semi-Markov, multiple general transitions can be simultaneously enabled inside the firing process of t_g, provided that their activity cycles do not overlap [6]. The second implication is that, during the subordinated process, the dominant age memory GEN transition needs not to be continuously enabled; in fact, the associated memory variable is not reset even if the transition is disabled before firing. In order to track the enabling/disabling condition of the dominant GEN transition t_g, we introduce a reward (indicator) variable which is equal to 1 in those markings in which t_g is enabled and equal to 0 in those markings in which t_g is not enabled. The binary reward variables are then grouped into a reward vector and the subordinated processes are formulated in terms of semi-Markov reward models [17, 3]. The memory variable a_g corresponding to the dominant GEN transition is computed as the accumulated reward in the semi-Markov reward subordinated process and the successive regeneration time point (the firing epoch of t_g) occurs when the memory variable a_g accumulates a time equal to the firing time γ_g of the corresponding transition. Resorting to the computational properties of stochastic reward models [3], the cdf of the successive regeneration time point is evaluated as the first time at which the functional a_g hits an absorbing barrier of height γ_g.

The firing of the dominant GEN transition t_g in the subordinated process starting in the regeneration marking i, can only occur in a state k in which the reward variable is equal to one (t_g is enabled). After the firing of t_g in state k, the successor marking ℓ is determined by the branching probability matrix $\mathbf{\Delta}^{(g)} = [\Delta^{(g)}_{k\ell}]$ [7, 9], where:

$$\Delta^{(g)}_{k\ell} = Pr\{\,\text{next marking is } \ell \,|\, \text{current marking is } k,\ t_g \text{ fires}\,\} \qquad (2)$$

By virtue of the time homogeneity, and without loss of generality, any two successive regeneration time points can be supposed to be $x = \tau_0^* = 0$ and $x = \tau_1^*$. Let us define the following matrix valued functions [8, 11]:

$$\begin{array}{lll} \mathbf{V}(x) = [V_{ij}(x)] & \text{such that} & V_{ij}(x) = Pr\{\mathcal{M}(x) = j \,|\, \mathcal{M}(\tau_0^*) = i\} \\ \mathbf{K}(x) = [K_{ij}(x)] & \quad " & K_{ij}(x) = Pr\{M_{(1)} = j\,,\, \tau_1^* \le x |\, \mathcal{M}(\tau_0^*) = i\} \\ \mathbf{E}(x) = [E_{ij}(x)] & \quad " & E_{ij}(x) = Pr\{\mathcal{M}(x) = j\,,\, \tau_1^* > x |\, \mathcal{M}(\tau_0^*) = i\} \end{array} \qquad (3)$$

Matrix $\mathbf{V}(x)$ is the transition probability matrix and provides the probability that the stochastic process $\mathcal{M}(x)$ is in marking j at time x given it was in i at $x = 0$. The matrix $\mathbf{K}(x)$ is the *global kernel* of the *MRGP* and provides the cdf of the event that the next regeneration marking is $M_{(1)} = j$ at time τ_1^*, given marking i at $\tau_0^* = 0$. Finally, the matrix $\mathbf{E}(x)$ is the *local kernel* since describes the behavior of the marking process $\mathcal{M}(x)$ inside two consecutive regeneration time points. The generic element $E_{ij}(x)$ provides the probability that the process is in state j at x starting from i at $\tau_0^* = 0$ before the next regeneration time point. From the above definitions:

$$\sum_j [K_{ij}(x) + E_{ij}(x)] = 1$$

As specified by (3), for each state $M_{(i)} = i$, the entries of the i-th row of the matrices $\mathbf{K}(x)$ and $\mathbf{E}(x)$ depend only on the behavior of the subordinated process starting from $M_{(i)}$, given that $M_{(i)}$ is a regeneration state. If $M_{(i)}$ cannot be a regeneration state, the corresponding entries are irrelevant. The transient behavior of the *MRSPN* can be evaluated by solving the following generalized Markov renewal equation [11, 8]:

$$\mathbf{V}(x) \;=\; \mathbf{E}(x) \;+\; \mathbf{K} \,*\, \mathbf{V}(x) \tag{4}$$

where $\mathbf{K} * \mathbf{V}(x)$ is a convolution matrix, whose (i, j)-th entry is:

$$[\mathbf{K} \,*\, \mathbf{V}(x)]_{ij} \;=\; \sum_k \int_0^x dK_{ik}(y)\, V_{kj}(x-y) \tag{5}$$

By denoting the Laplace Stieltjes transform *(LST)* of a function $F(x)$ by $F^{\sim}(s) = \int_0^\infty e^{-sx}\, dF(x)$, Equation (4) becomes:

$$\mathbf{V}^{\sim}(s) \;=\; \mathbf{E}^{\sim}(s) \;+\; \mathbf{K}^{\sim}(s)\; \mathbf{V}^{\sim}(s) \tag{6}$$

whose solution is:

$$\mathbf{V}^{\sim}(s) \;= [\mathbf{I} \,-\, \mathbf{K}^{\sim}(s)]^{-1}\; \mathbf{E}^{\sim}(s) \tag{7}$$

The steady state solution can be evaluated as $\lim_{s\to 0} \mathbf{V}^{\sim}(s)$.

3 Transient analysis of the subordinated process

Let $M_{(i)} = i$ be a regeneration marking according to Definition 1. In the *AgeMR-SPN* model, only two classes of subordinated processes can be encountered:

1. *Single step CTMC.*
2. *Reward Semi-Markov Process.*

3.1 Subordinated single step CTMC

In the regeneration marking i only EXP transitions are enabled. The next regeneration time point is the epoch of jump into any one of the immediately reachable states. The subordinated process starting from state i is a *CTMC* with a single transient state (state i with initial probability equal to 1) and a number of absorbing states equal to the number of immediately reachable states.

Let $T_e^{(i)}$ be the set of EXP transitions enabled in the regeneration marking i, λ_e the transition rate of transition $t_e \in T_e^{(i)}$, and $\lambda^i = \sum_{t_e \in T_e^{(i)}} \lambda_e$. The entry $K_{ij}(x)$ provides the probability of reaching the successive regeneration state j before time x. The entry $E_{ij}(x)$ gives the probability of being in state j at time x starting from i, before the next regeneration time point. Since, in this case, any firing provides a new regeneration time point, the only nonzero entry of the i-th row of matrix $\mathbf{E}(x)$ corresponds to $j = i$. In the *LST* domain, the following expressions hold:

$$K_{ij}^{\sim}(s) \;=\; \frac{\lambda_e}{\lambda^i + s}\, \Delta_{ij}^{(e)} \qquad\qquad E_{ij}^{\sim}(s) \;=\; \delta_{ij}\, \frac{s}{\lambda^i + s} \tag{8}$$

where δ_{ij} is the Kronecker delta.

3.2 Subordinated Reward Semi-Markov Process

At $x = \tau_0^* = 0$ the dominant age memory GEN transition t_g starts its firing process in the regeneration state i ($a_g = 0$). The successive regeneration time point τ_1^* is the epoch of firing of t_g and this event occurs as the accumulated reward (memory variable) a_g reaches the value γ_g for the first time.

Let $\Omega(i)$ be the subset of $\mathcal{R}(M_0)$ grouping the states of the subordinated process (i.e. the states reachable from i before firing t_g). For notational convenience we do not renumber the states in $\Omega(i)$ so that all the subsequent matrix functions have the dimensions $(\mathcal{N} \times \mathcal{N})$ (cardinality of $\mathcal{R}(M_0)$), but with the significant entries located in position (k, ℓ) only, with $k, \ell \in \Omega(i)$.

Let $Z^{(i)}(x)$ $(x \geq 0)$ be the semi-Markov process defined over $\Omega(i)$ and $\underline{r}^{(i)}$ the corresponding binary reward vector. With this notation, $r_k^{(i)} = 1$ (0) means that t_g is enabled (not enabled) in state k, and the memory variable a_g increases at a rate $r_k^{(i)}$ when $Z^{(i)}(x) = k$. The subordinated process coincides with $Z^{(i)}(x)$ when the initial state is state i with probability 1 ($Pr\{Z^{(i)}(0) = i\} = 1$).

Let $\mathbf{Q}^{(i)}(x) = [Q_{k\ell}^{(i)}(x)]$ be the kernel of the semi-Markov process $Z^{(i)}(x)$. The initial probability vector is $\underline{Q}_0^{(i)} = [0, 0, \ldots, 1_i, \ldots, 0]$ (a vector with all the entries equal to 0 but entry i equal to 1). We denote by H the time duration until the first embedded time point in the semi-Markov process starting from state k at time 0 ($Z^{(i)}(0) = k$). The generic element (for $k, \ell \in \Omega(i)$)

$$Q_{k\ell}^{(i)}(x) = Pr\left\{H \leq x, Z^{(i)}(H^+) = \ell \,|\, Z^{(i)}(0) = k\right\}$$

is the distribution of H supposed that a transition from state k to state ℓ took place at the embedded time point. If diagonal elements in $\mathbf{Q}^{(i)}(x)$ are nonzero the next embedded time point can be determined by a transition from state k to state k. The distribution of H is:

$$Q_k^{(i)}(x) = \sum_{\ell \in \Omega(i)} Q_{k\ell}^{(i)}(x) \qquad\qquad (k = 1, ..., n)$$

and, finally, the probability of jumping from state k to ℓ at time $H = x$ is:

$$\frac{dQ_{k\ell}^{(i)}(x)}{dQ_k^{(i)}(x)} = Pr\left\{Z^{(i)}(x^+) = \ell \,|\, H = x,\, Z^{(i)}(0) = k\right\}$$

Let us fix the value of the random firing time $\gamma_g = y$ and let us introduce two matrix functions: $\mathbf{F}^{(i)}(x, y)$ and $\mathbf{P}^{(i)}(x, y)$ so defined:

$$\begin{aligned} F_{k\ell}^{(i)}(x,y) &= Pr\{Z^{(i)}(\tau_1^{*\,-}) = \ell\,,\, \tau_1^* \leq x \,|\, Z^{(i)}(0) = k\,,\, \gamma_g = y\} \\ P_{k\ell}^{(i)}(x,y) &= Pr\{Z^{(i)}(x) = \ell\,,\, \tau_1^* > x \,|\, Z^{(i)}(0) = k\,,\, \gamma_g = y\} \end{aligned} \qquad (9)$$

– $P^{(i)}_{k\ell}(x,y)$ is the probability of being in state ℓ at time x before absorption at the barrier y, starting in state k at $x=0$, and being γ_g equal to a constant value y.
– $F^{(i)}_{k\ell}(x,y)$ is the probability that t_g fires from state ℓ (hitting the absorbing barrier y in ℓ) before x, starting in state k at $x=0$, and being γ_g equal to a constant value y.
– $\mathbf{\Delta}^{(g)}$ is the branching probability matrix and represents the successor marking ℓ that is reached by firing t_g in state k (the firing of t_g can only occur in a state k in which $r^{(i)}_k = 1$).

From (9), it follows:

$$\sum_{\ell} [\, F^{(i)}_{k\ell}(x,y) \; + \; P^{(i)}_{k\ell}(x,y) \,] \; = \; 1$$

Due to the particular structure of the initial probability vector $\underline{Q}^{(i)}_0$, the entries of the i-th row of the matrices $\mathbf{K}(x)$ and $\mathbf{E}(x)$ are related to $\mathbf{F}^{(i)}(x,y)$ and $\mathbf{P}^{(i)}(x,y)$ by the following expressions:

$$\begin{aligned} K_{ij}(x) &= \int_{y=0}^{\infty} \sum_{k} F^{(i)}_{ik}(x,y)\, \varDelta^{(g)}_{kj}\, dG_g(y) \\ E_{ij}(x) &= \int_{y=0}^{\infty} P^{(i)}_{ij}(x,y)\, dG_g(y) \end{aligned} \tag{10}$$

Evaluation of $F^{(i)}_{k\ell}(x,y)$ and $P^{(i)}_{k\ell}(x,y)$ can be inferred from [15,4]. We include the derivation for completeness. In order to avoid unnecessarily cumbersome notation in the following expressions, we neglect the explicit dependence on the particular subordinated process by eliminating the superscript. It is however tacitly intended, that all the quantities $\underline{r}$, $\mathbf{Q}(x)$, $\mathbf{F}(x,y)$, $\mathbf{P}(x,y)$, $\mathbf{\Delta}$ and Ω refer to the specific process subordinated to state i.

Derivation of $\mathbf{F}(x,y)$ Conditioning on $H = h$, let us define:

$$F_{k\ell}(x,y \,|\, H=h) \;=\; \begin{cases} \delta_{k\ell}\, U\left(x \;-\; \dfrac{y}{r_k}\right) & \text{if} : h\, r_k \geq y \\ \\ \displaystyle\sum_{u\in\Omega} \frac{dQ_{ku}(h)}{dQ_k(h)}\, F_{u\ell}(x-h, y-hr_k) & \text{if} : h\, r_k < y \end{cases} \tag{11}$$

where $U(x)$ is the unit step function. In (11), two mutually exclusive events are identified. If $r_k \neq 0$ and $h\, r_k \geq y$, a sojourn time equal to y is accumulated before leaving state k, so that the firing time (next regeneration time point) is $\tau_1^* = y/r_k$. If $h\, r_k < y$ then a transition occurs to state u with probability

$dQ_{ku}(h)/dQ_k(h)$ and the residual service $(y - hr_k)$ should be accomplished starting from state u at time $(x - h)$. Taking the LST transform of (11) with respect to x, we get:

$$F^{\sim}_{k\ell}(s, y \mid H = h) = \begin{cases} \delta_{k\ell}\, exp(-sy/r_k) & \text{if} : h\, r_k \geq y \\ exp(-sh) \displaystyle\sum_{u \in \Omega} \frac{dQ_{ku}(h)}{dQ_k(h)}\, F^{\sim}_{u\ell}(s, y - hr_k) & \text{if} : h\, r_k < y \end{cases} \tag{12}$$

Unconditioning with respect to h, (12) becomes:

$$F^{\sim}_{k\ell}(s, y) = \delta_{k\ell} \left[1 - Q_k\left(\frac{y}{r_k}\right)\right] exp(-sy/r_k) + \sum_{u \in \Omega} \int_{h=0}^{\frac{y}{r_k}} exp(-s\, h)\, F^{\sim}_{u\ell}(s, y - hr_k)\, dQ_{ku}(h) \tag{13}$$

Taking the Laplace transform *(LT)* with respect to y (denoting by w the transform variable), and evaluating the integrals we obtain, for the double LST-LT transform $F^{\sim *}_{k\ell}(s, w)$, the following expression:

$$F^{\sim *}_{k\ell}(s, w) = \delta_{k\ell}\, \frac{r_k\, [1 - Q^{\sim}_k(s + w\, r_k)]}{s + w\, r_k} + \sum_{u \in \Omega} Q^{\sim}_{ku}(s + w\, r_k)\, F^{\sim *}_{u\ell}(s, w) \tag{14}$$

Derivation of $\mathbf{P}(x, y)$ The derivation follows the same pattern as for the function $\mathbf{F}(x, y)$. Conditioning on $H = h$, let us define:

$$P_{k\ell}(x, y \mid H = h) = \begin{cases} \delta_{k\ell} \left[U(x) - U\left(x - \frac{y}{r_k}\right)\right] & \text{if} : h\, r_k \geq y \\ \delta_{k\ell}\, [U(x) - U(x - h)] + \displaystyle\sum_{u \in \Omega} \frac{dQ_{ku}(h)}{dQ_k(h)}\, P_{u\ell}(x - h, y - hr_k) & \text{if} : h\, r_k < y \end{cases} \tag{15}$$

In (15), two mutually exclusive events are identified. If $r_k \neq 0$ and $y \leq h\, r_k$, then the process spends all its time up to absorption in the initial state k. If $h\, r_k < y$ then a transition occurs to state u with probability $dQ_{ku}(h)/dQ_k(h)$ and then the process jumps to state ℓ in the remaining time $(x - h)$ before completing the residual work $(y - hr_k)$. Taking the LST transform of (15) with respect to

x, we get:

$$P^{\sim}_{k\ell}(s, y \,|\, H = h) \;=\; \begin{cases} \delta_{k\ell}\;[1\;-\;exp(-sy/r_k)] & \text{if}: h\,r_k\;\geq\;y \\ \delta_{k\ell}\;\left[1\;-\;e^{-sh}\right]\;+\;e^{-sh}\displaystyle\sum_{u\in\Omega}\frac{dQ_{ku}(h)}{dQ_k(h)}\;P^{\sim}_{u\ell}(s, y - hr_k) & \text{if}: h\,r_k\;<\;y \end{cases} \tag{16}$$

Unconditioning (16) with respect to h, taking the *LT* transform with respect to y (denoting w the transform variable), and finally evaluating the integrals we obtain that the double *LST-LT* transform $P^{\sim *}_{k\ell}(s, w)$ satisfies the following equation:

$$P^{\sim *}_{k\ell}(s, w) \;=\; \delta_{k\ell}\;\frac{s\;[1\;-\;Q^{\sim}_k(s\;+\;w\,r_k)]}{w(s\;+\;w\,r_k)}\;+\;\sum_{u\in\Omega}\;Q^{\sim}_{ku}(s\;+\;w\,r_k)\,P^{\sim *}_{u\ell}(s, w) \tag{17}$$

EP distributed firing time Let us define an exponential polynomial *(EP)* distribution $G_E(y)$ as a distribution with rational Laplace transform whose density can be expressed as:

$$g_E(y) = \sum_{p=1}^{n}\sum_{r=0}^{m-1}\;c_{pr}\;y^r\;e^{-\lambda_p y} \tag{18}$$

where n is the number of distinct eigenvalues (λ_p), m is the supremum of the eigenvalue multiplicities, and c_{pr} is a constant coefficient[1]. When the dominant GEN transition is associated with an *EP* random firing time, an efficient computational procedure can be envisaged for handling the Laplace inverse transformation with respect to w and the integration with respect to $G_E(y)$.

Theorem 1. *When the firing time is an* EP *r.v. with density function* $g_E(y)$ *(18), the entries of the kernel matrices can be evaluated as follows:*

$$E^{\sim}_{ij}(s) = \sum_{p=1}^{n}\sum_{r=0}^{m-1}(-1)^r c_{pr}\frac{d^r P^{\sim *}_{ij}(s, w)}{dw^r}\bigg|_{w=\lambda_p} \tag{19}$$

$$K^{\sim}_{ij}(s) = \sum_{p=1}^{n}\sum_{r=0}^{m-1}(-1)^r c_{pr}\frac{d^r \sum_k\;F^{\sim *}_{ik}(s, w)\,\Delta^{(i)}_{kj}}{dw^r}\bigg|_{w=\lambda_p} \tag{20}$$

where the derivative of order $r = 0$ *simply means the substitution of the value* $w = \lambda_p$ *in the r.h.s.*

[1] The definition of *EP* r.v. given here requires the Laplace transform to be rational and is more restrictive than the definition of expolynomial distributions proposed in [9] in connection with *MRSPN*.

Proof. When γ_g is an *EP* r.v. Equation (10) becomes:

$$
\begin{aligned}
& E^{\sim}_{ij}(s) \;=\; \int_{y=0}^{\infty} P^{\sim}_{ij}(s,y)\, dG_E(y) \;=\; \int_{y=0}^{\infty} g_E(y)\, P^{\sim}_{ij}(s,y)\, dy \;= \\
& \sum_{p=1}^{n}\sum_{r=0}^{m-1} c_{pr} \int_{y=0}^{\infty} y^r e^{-\lambda_p y}\, P^{\sim}_{ij}(s,y)\, dy = \\
& \sum_{p=1}^{n}\sum_{r=0}^{m-1} (-1)^r c_{pr} \int_{y=0}^{\infty} \frac{d^r}{d\lambda_p^r} e^{-\lambda_p y}\, P^{\sim}_{ij}(s,y)\, dy = \\
& \sum_{p=1}^{n}\sum_{r=0}^{m-1} (-1)^r c_{pr} \frac{d^r}{d\lambda_p^r} \int_{y=0}^{\infty} e^{-\lambda_p y}\, P^{\sim}_{ij}(s,y)\, dy = \\
& \sum_{p=1}^{n}\sum_{r=0}^{m-1} (-1)^r c_{pr} \frac{d^r\, P^{(i)\,\sim *}_{ij}(s,\lambda_p)}{d\lambda_p^r}
\end{aligned}
\qquad (21)
$$

from which the first part of the theorem (equation 19) follows. The proof for $K^{\sim}_{ij}(s)$ follows the same pattern.

This approach is very effective, when the multiplicity of the eigenvalues is equal to 1, since the inverse Laplace transformation and integration in (21) reduces to a simple substitution; otherwise the symbolic derivation is required. A wellknown and convenient subclass of *EP* distributions is the class of *PH* distributions arising from the time to absorption of *CTMC*'s with at least one absorbing state. When all the GEN firing times are *PH* random variables and the subordinated processes are *CTMC*'s, the transient state probabilities can be alternatively evaluated by expanding the state space $\mathcal{R}(M_0)$ taking into account all the possible stage combinations of each *PH* transition. A completely automated tool that implements the state space expansion technique is in [12].

3.3 Derivation of $\mathbf{V}(x)$

The evaluation of the entries of the state transition probability matrix $\mathbf{V}(x)$ requires the following steps to be performed:

- Derivation of the double Laplace transform matrix functions $F^{*\sim}_{k\ell}(s,w)$ and $P^{*\sim}_{k\ell}(s,w)$, according to Equations (14) and (17), respectively.
- Evaluation of the LST transforms $F^{\sim}_{k\ell}(s,y)$ and $P^{\sim}_{k\ell}(s,y)$ by symbolic inverse Laplace transformation with respect to the firing time variable w.
- Evaluation of the LST transforms $\mathbf{K}^{\sim}(s)$ and $\mathbf{E}^{\sim}(s)$ by unconditioning the results of the previous step with respect to the distribution of the firing time $G_g(y)$ (Equation 10).
- Symbolic matrix inversion and matrix multiplication by using a standard package (e.g. MATHEMATICA) in order to obtain $\mathbf{V}^{\sim}(s)$ (Equation 7).

– Time domain solution obtained by a numerical inversion of the entries of $\mathbf{V}^{\sim}(s)$, resorting to the Jagerman's method [14] (for the sake of uniformity, this step has been implemented in MATHEMATICA language).

When $G_g(y)$ is an *EP*, Theorem 1 can be applied instead of steps 2 and 3. In the particular case in which the subordinated process $Z(t)$ is a *CTMC*, all the sojourn time distributions become exponential and Equations (14) and (17) can be simplified accordingly [6]. Due to the required symbolic and numerical steps, the procedure outlined in the previous points is effective only for small values of the cardinality of the reachability set.

4 M/G/1/2/2 with Preemptive Resume Service

The M/D/1/2/2 queueing system has been considered as a benchmark example in the recent literature on non-Markovian *SPN*. The example has been introduced in [2], where the steady state solution was derived. The transient analysis for the same system was carried on in [7] and the model was extended by allowing GEN service times in [8]. The effect of different preemption policies has been studied in [5] and the analysis of the M/D/1/2/2 queueing system with *prs* service policy is in [6]. In the following, we apply the procedure developed in the previous Section to the case of *prs* service policy and generally distributed service time.

4.1 Model assumptions

Figure 1a shows a *PN* describing the M/G/1/2/2 system in which any new job preempts the job under service. We assume that the service policy is of *prs* type: a preempted job is resumed as soon as the server becomes idle, but the prior work is not lost and the residual service time needs to be completed. Place p_1 contains the customers thinking, while place p_2 contains the number of submitted jobs (including the one under service). Starting from the initial marking $s_1 = (2\ 0\ 0\ 1)$ (Figure 1b), t_1 is the only enabled transition. Firing of t_1 represents the submission of the first job and leads to state $s_2 = (1\ 1\ 1\ 0)$. In s_2 transitions t_2 and t_3 are competing. t_2 represents the service of the submitted job and its firing returns the system to the initial state s_1. t_3 represents the submission of the second job and its firing disables t_2 by removing one token from p_3 (the first job becomes dormant). In $s_3 = (0\ 2\ 0\ 1)$ one job is under service and one job is dormant, and the only enabled activity is the service of the active job. Firing of t_4 leads the system again in s_2, where the dormant job is recovered. Assuming the thinking time of both customers to be EXP with parameter λ, t_1 is associated an exponential firing rate equal to $(2\,\lambda)$ and t_3 a firing rate equal to λ. Transitions t_2 and t_4 are assigned a GEN service time with distribution $G_g(x)$ and an age memory policy.

Each time t_2 is disabled without firing (t_3 fires before t_2) the memory variable a_2 is not reset. Hence, as the second job completes (t_4 fires), the system returns

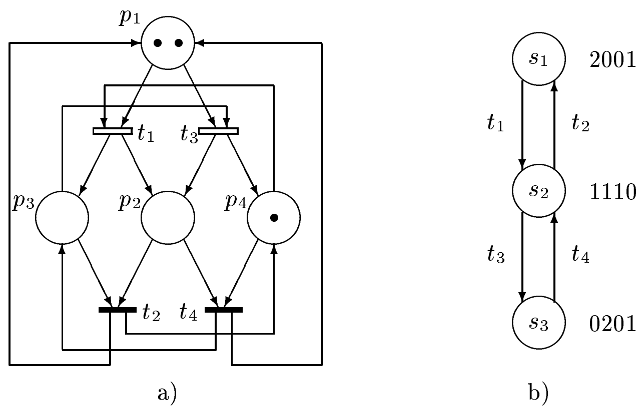

Figure 1 - Preemptive M/G/1/2/2 queue with identical customers

in s_2 remembering the value of a_2, so that the time to complete the interrupted job can be evaluated as the residual service time given a_2. a_2 counts the total time during which t_2 is enabled before firing, and is equal to the cumulative sojourn time in s_2. The assignment of the age memory policy to t_2 realizes a *prs* service mechanism.

The regeneration time points in the marking process $\mathcal{M}(x)$ correspond to the epochs of entrance in markings in which the memory variables associated to all the transitions are equal to zero. By inspecting Figure 1b), the regeneration time points result to be the epochs of entrance in s_1 and of entrance in s_2 from s_1. s_3 can never be a regeneration marking, since the memory variable a_2 is not reset at the entrance in s_3: the process can sojourn in s_3 only between two successive regeneration points (Figure 2).

The process subordinated to state s_1 is a single step *CTMC* (being EXP the only enabled transition t_1) and includes the only immediately reachable state s_2. The process subordinated to state s_2 is dominated by the GEN age memory transition t_2 and includes the states s_3 and s_2 reachable from s_2 before firing of t_2. Since s_2 is the only state in which t_2 is enabled, the corresponding reward rate vector is $\underline{r}^{(2)} = [0\ 1\ 0]$. Finally, the only relevant nonzero entry of the branching probability matrix is $\Delta_{21}^{(2)} = 1$, since firing of t_2 can only occur from state s_2 leading to state s_1.

A possible realization of the marking process subordinated to state s_2 is shown in Figure 2: the subordinated process is semi-Markov since t_4 is GEN. The memory variable a_2 grows whenever the process sojourns in state s_2, and the firing of t_2 is determined by the first passage time of a_2 across the absorbing barrier of height γ_2.

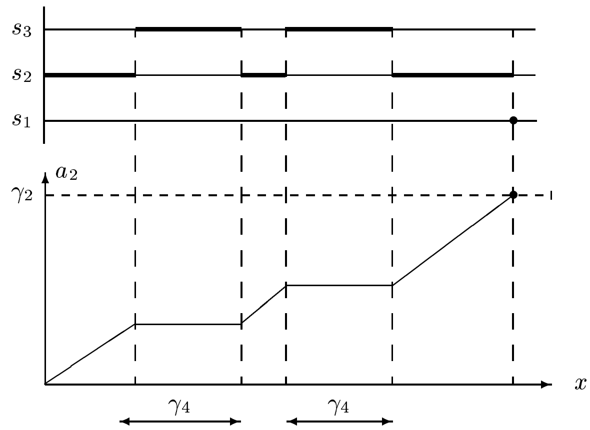

Figure 2 - A possible realization of the process subordinated to state s_2

4.2 Numerical Results

The closed form *LST* expressions of $\mathbf{K}(x)$ and $\mathbf{E}(x)$ for the *prs* M/G/1/2/2 queuing systems are derived in detail, considering two specific classes of GEN firing times (namely: the uniform and the *EP*). Let us build up the $\mathbf{K}^{\sim}(s)$ and $\mathbf{E}^{\sim}(s)$ matrices row by row by considering separately all the states that can be regeneration states and can originate a subordinated process. Since s_3 can never be a regeneration state the third row of the above matrices is irrelevant.
1) - The starting regeneration state is s_1 - No GEN transition is enabled in s_1 and the next regeneration state can only be state s_2. Applying (8) we obtain:

$$K^{\sim}_{11}(s) = 0 \qquad K^{\sim}_{12}(s) = \frac{2\lambda}{s + 2\lambda} \qquad K^{\sim}_{13}(s) = 0$$

and

$$E^{\sim}_{11}(s) = \frac{s}{s + 2\lambda} \qquad E^{\sim}_{12}(s) = 0 \qquad E^{\sim}_{13}(s) = 0 \tag{22}$$

2) - The starting regeneration state is s_2 - Transition t_2 is the dominant transition and the next regeneration time point is the epoch of firing of t_2. t_2 is an age memory GEN transition with Cdf $G_g(y)$, hence, the conditions of Section 3.2. are met. The subordinated process (Figure 2) comprises states s_2 and s_3 and is a semi-Markov process whose kernel is:

$$Q^{\sim}(s) = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & \dfrac{\lambda}{s + \lambda} \\ 0 & G^{\sim}_g(s) & 0 \end{vmatrix}$$

where $G^{\sim}_g(s)$ is the *LST* transform of the distribution function $G_g(y)$. The reward vector is $\underline{r}^{(2)} = [0, 1, 0]$, and the only nonzero entry of the branching probability matrix is $\Delta^{(2)}_{21} = 1$. Let us introduce the following notation:

$$H_g(s) = s + \lambda - \lambda G^{\sim}_g(s) \tag{23}$$

The non-zero entries of the 2nd row of $\mathbf{F}^{\sim *}(s,w)$ and $\mathbf{P}^{\sim *}(s,w)$ matrices are obtained by applying Equations (14) and (17):

$$\begin{aligned}
F^{\sim *}_{22}(s,w) &= \frac{1}{s+w+\lambda-\lambda G^{\sim}_g(s)} = \frac{1}{w+H_g(s)} \\
P^{\sim *}_{22}(s,w) &= \frac{s/w}{s+w+\lambda-\lambda G^{\sim}_g(s)} = \frac{s/w}{w+H_g(s)} \\
P^{\sim *}_{23}(s,w) &= \frac{\lambda(1-G^{\sim}_g(s))/w}{s+w+\lambda-\lambda G^{\sim}_g(s)} = \frac{\lambda(1-G^{\sim}_g(s))/w}{w+H_g(s)}
\end{aligned} \tag{24}$$

Uniformly distributed service time Let $G_U(y)$ indicate a uniform distribution defined between $\alpha(\geq 0)$ and $\beta(> \alpha)$. The non preemptive M/G/1/2/2 queue with uniformly distributed service time has been studied by Choi at al. in [8]. The extension to the *prs* service policy is developed in the following.

The LST transform of $G_U(y)$ is given by:

$$G^{\sim}_U(s) = \frac{1}{s}\frac{1}{\beta-\alpha}(e^{-\alpha s} - e^{-\beta s})$$

and substituting the actual value of $G^{\sim}_U(s) = G^{\sim}_g(s)$ in (23), we get:

$$H_U(s) = s + \lambda - \lambda G^{\sim}_U(s) = s + \lambda - \frac{\lambda}{s}\frac{1}{\beta-\alpha}(e^{-\alpha s} - e^{-\beta s})$$

According to the steps mentioned in Section 3.3, the symbolic inversion of Equations (24) is performed with respect to the transform variable w, followed by an integration with respect to the distribution of the service time $G_U(y)$. The inverse transformation with respect to w provides:

$$\begin{aligned}
F^{\sim}_{22}(s,y) &= e^{-yH_U(s)} \\
P^{\sim}_{22}(s,y) &= \frac{s}{H_U(s)}(1-e^{-yH_U(s)}) \\
P^{\sim}_{23}(s,y) &= \frac{\lambda(1-G^{\sim}_U(s))}{H_U(s)}(1-e^{-yH_U(s)})
\end{aligned}$$

Applying the integration step expressed by (10), the *LST* matrix functions $\mathbf{K}^{\sim}(s)$ and $\mathbf{E}^{\sim}(s)$ become:

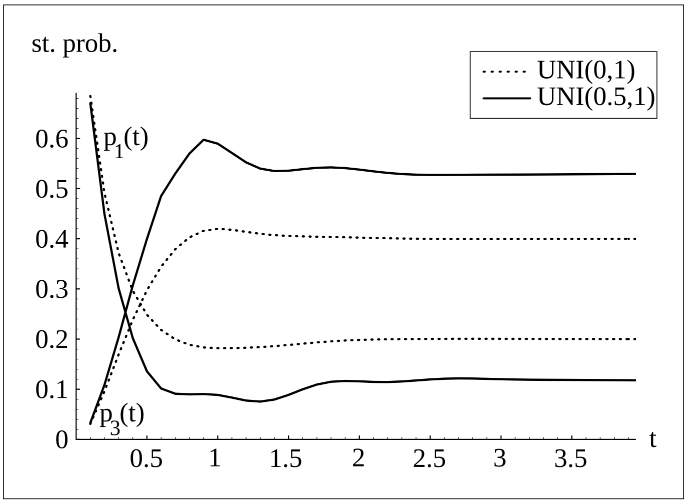

Figure 3 - Transient behavior of the state probabilities with uniformly distributed service time.

$$\mathbf{K}^{\sim}(s) = \begin{vmatrix} 0 & \dfrac{2\lambda}{s + 2\lambda} & 0 \\ \dfrac{1}{H_U(s)} \dfrac{1}{\beta - \alpha} (e^{-\alpha H_U(s)} - e^{-\beta H_U(s)}) & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} \tag{25}$$

and

$$\mathbf{E}^{\sim}(s) = \begin{vmatrix} \dfrac{s}{s + 2\lambda} & 0 & 0 \\ 0 & \dfrac{s}{H_U(s)}(1 - K^{\sim}_{21}(s)) & \dfrac{\lambda(1 - G^{\sim}_U(s))}{H_U(s)}(1 - K^{\sim}_{21}(s)) \\ 0 & 0 & 0 \end{vmatrix} \tag{26}$$

The LST of the transition probability matrix $\mathbf{V}^{\sim}(s)$ is obtained by solving (7). Finally, the time domain probabilities are calculated by numerically inverting (7) by resorting to the Jagerman method [14]. The plot of the state probabilities versus time for states s_1 and s_3 is depicted in Figure 3, for a submitting rate $\lambda = 2$, and for two different set of values ($\alpha = 0$, $\beta = 1$) and ($\alpha = 0.5$, $\beta = 1$). Figure 3 emphasizes the effect of the coefficient of variation of the service time on the state probabilities; a reduced coefficient of variation results in a more pronounced alternating behavior of the state probabilities.

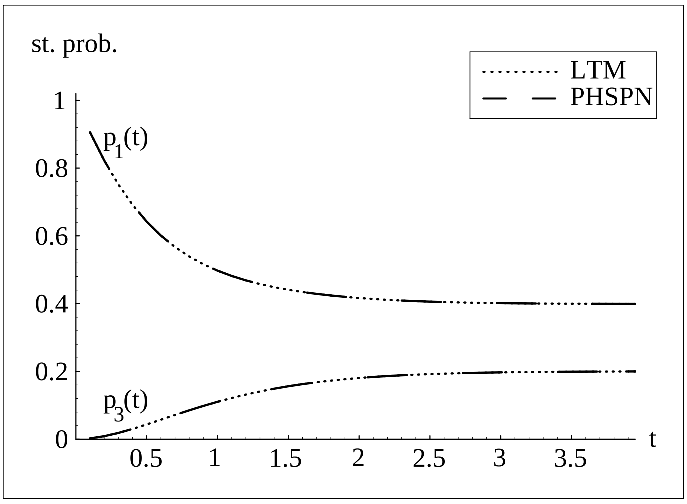

Figure 4 - Transient behavior of the state probabilities with *Erl_2* distributed service time.

EP distributed service time Let us consider the same preemptive M/G/1/2/2 in which the service time has an Erlang distribution of order 2 *(Erl_2)*. The *LST* $G_E^{\sim}(s)$ of the *Erl_2* with parameter τ is:

$$G_E^{\sim}(s) = \left(\frac{\tau}{s+\tau}\right)^2 \tag{27}$$

By substituting (27) into (23), we get:

$$H_E(s) = s + \lambda - \lambda G_E^{\sim}(s) = s + \lambda - \lambda \left(\frac{\tau}{s+\tau}\right)^2$$

The 1st and the 3rd row of the $\mathbf{K}^{\sim}(s)$ and $\mathbf{E}^{\sim}(s)$ matrices do not depend on the particular GEN distribution and remain unchanged. The nonzero entries of the second row can be obtained, as before, by a symbolic inverse transformation with respect to w followed by an the integration with respect to $G_g(y)$. Alternatively, since $G_g(y) = G_E(y)$ is *EP*, we can apply Theorem 1 to Equations (24).

$$K_{21}^{\sim}(s) = (-1)\tau^2 \frac{d\,F_{22}^{\sim *}(s,w)}{d\,w}\bigg|_{w=\tau} = \frac{\tau^2}{(\tau + H_E(s))^2} \tag{28}$$

$$E_{22}^{\sim}(s) = (-1)\tau^2 \frac{d\,P_{22}^{\sim *}(s,w)}{d\,w}\bigg|_{w=\tau} = \frac{s(2\tau + H_E(s))}{(\tau + H_E(s))^2} \tag{29}$$

$$E_{23}^{\sim}(s) = (-1)\tau^2 \frac{d\,P_{23}^{\sim *}(s,w)}{d\,w}\bigg|_{w=\tau} = \frac{\lambda(1 - G_E^{\sim}(s))(2\tau + H_E(s))}{(\tau + H_E(s))^2} \tag{30}$$

In this example, only EXP and *PH* firing times are considered. Hence, the transient probabilities can also be obtained by the well known method of the state space expansion [12]. However, if t_2 has a *PH* firing time but t_4 is non-*PH*, then only the above equations can be applied.

Similarly to the former case, the *LST* of the state probabilities are obtained by solving (7). The time domain probabilities are calculated by numerically inverting (7) by resorting to the Jagerman method [14]. The plot of the state probabilities versus time for states s_1 and s_3 (with $\tau = 2$, corresponding to a mean service time $2/\tau = 1$, and $\lambda = 0.5$.) are depicted in Figure 4 (dotted line). For the sake of comparison the results obtained by applying the method of the expanded *CTMC* [12] are reported in dashed line.

5 Conclusion

We have defined a new class of *MRSPN* called *AgeMRSPN*, which allow the inclusion of GEN transitions with associated age memory policy. This extension was motivated by the need of modeling systems in which the execution of tasks may follow a preemptive resume policy.

We have shown that the marking process subordinated to two consecutive regeneration time points can be, in general, a reward semi-Markov process. A binary reward variable is introduced to distinguish the states in which the execution of the service is interrupted and the states in which the execution is resumed with no loss of prior work. The transient analysis of a reward semi-Markov process has been derived in detail, in order to show how to obtain a double *LT-LST* closed form expression for the transient state probabilities of the general process.

An M/G/1/2/2 queuing system, considered as a case study example in previous literature [2, 7, 8, 5, 6], has been examined for the first time by introducing service policies of *prs* type and GEN firing distributions.

References

1. M. Ajmone Marsan, G. Balbo, A. Bobbio, G. Chiola, G. Conte, and A. Cumani. The effect of execution policies on the semantics and analysis of stochastic Petri nets. *IEEE Transactions on Software Engineering*, SE-15:832–846, 1989.
2. M. Ajmone Marsan and G. Chiola. On Petri nets with deterministic and exponentially distributed firing times. In *Lecture Notes in Computer Science*, volume 266, pages 132–145. Springer Verlag, 1987.
3. A. Bobbio. Stochastic reward models in performance/reliability analysis. *Journal on Communications*, XLIII:27–35, January 1992.
4. A. Bobbio and M. Telek. Task completion time. In *Proceedings 2nd International Workshop on Performability Modelling of Computer and Communication Systems (PMCCS2)*, 1993.
5. A. Bobbio and M. Telek. Computational restrictions for SPN with generally distributed transition times. In D. Hammer K. Echtle and D. Powell, editors, *First European Dependable Computing Conference (EDCC-1), Lecture Notes in Computer Science*, volume 852, pages 131–148, 1994.

6. A. Bobbio and M. Telek. Markov regenerative SPN with non-overlapping activity cycles. In *International Computer Performance and Dependability Symposium - IPDS95*, April 1995.
7. Hoon Choi, V.G. Kulkarni, and K. Trivedi. Transient analysis of deterministic and stochastic Petri nets. In *Proceedings of the 14-th International Conference on Application and Theory of Petri Nets*, Chicago, June 1993.
8. Hoon Choi, V.G. Kulkarni, and K. Trivedi. Markov regenerative stochastic Petri nets. *Performance Evaluation*, 20:337–357, 1994.
9. G. Ciardo, R. German, and C. Lindemann. A characterization of the stochastic process underlying a stochastic Petri net. *IEEE Transactions on Software Engineering*, 20:506–515, 1994.
10. G. Ciardo and C. Lindemann. Analysis of deterministic and stochastic Petri nets. In *Proceedings International Workshop on Petri Nets and Performance Models - PNPM93*, pages 160–169. IEEE Computer Society, 1993.
11. E. Cinlar. *Introduction to Stochastic Processes.* Prentice-Hall, Englewood Cliffs, 1975.
12. A. Cumani. Esp - A package for the evaluation of stochastic Petri nets with phase-type distributed transition times. In *Proceedings International Workshop Timed Petri Nets*, pages 144–151, Torino (Italy), 1985. IEEE Computer Society Press no. 674.
13. R. German and C. Lindemann. Analysis of stochastic Petri nets by the method of supplementary variables. *Performance Evaluation*, 20:317–335, 1994.
14. D.L. Jagerman. An inversion technique for the Laplace transform. *The Bell System Technical Journal*, 61:1995–2002, October 1982.
15. V.G. Kulkarni, V.F. Nicola, and K. Trivedi. On modeling the performance and reliability of multi-mode computer systems. *The Journal of Systems and Software*, 6:175–183, 1986.
16. C. Lindemann. An improved numerical algorithm for calculating steady-state solutions of deterministic and stochastic Petri net models. *Performance Evaluation*, 18:75–95, 1993.
17. A. Reibman, R. Smith, and K.S. Trivedi. Markov and Markov reward model transient analysis: an overview of numerical approaches. *European Journal of Operational Research*, 40:257–267, 1989.