Analysis of Second-Order Markov Reward Models*

G. Horváth, S. Rácz, M. Telek
Technical University of Budapest, 1521 Budapest, Hungary
e-mail: {hgabor,raczs,telek}@webspn.hit.bme.hu

Abstract

This paper considers the analysis of second-order Markov reward models. In these systems the reward accumulation during state sojourns is not deterministic, but follows a Brownian motion with a state dependent drift and variance parameter. We give the differential equations that describe the density function and the moments of the accumulated reward, and show the similarities compared to the first-order (ordinary) case. A randomization based numerical method is also presented which is numerically stable, has an error bound to control the precision, and allows the efficient analysis of large models. The computational cost of the proposed procedure is practically the same as the one of the analysis of first-order reward models, while the modeling power of second-order models is clearly larger.

1. Introduction

Reward models (i.e., discrete state stochastic processes extended with a continuous variable) are useful and efficient modeling tools for computing performability measures in a great many of practical systems. In typical applications a discrete state stochastic processes, referred to as structure state process, describes the behaviour of the considered system and a continuous variable, referred to as reward variable, represents the performance of the system. The extension of the structure state process with a reward variable allows a flexible definition and analysis of a large number of performance measures, which would be far more complex or impossible based on the structure state process itself.

In the past decades considerable effort has been made to analytically describe and numerically evaluate more and more general reward models. Some of these efforts aimed the generalization of the structure state process. Thanks to their numerical tractability, reward models associated with *continuous time Markov chains,* referred to as Markov reward models (MRMs), are very popular. There are solutions for reward models having *semi Markov* [1] or *Markov regenerative* structure state process [2], however their numerical solution is far less tractable.

An other direction in generalization is the introduction of different reward accumulation policies. *Rate rewards* assigned to the states determine the reward accumulation rate during the sojourn in the state. *Impulse rewards* might be assigned with the state transitions of the structure state process. Impulse reward increases the amount of accumulated reward instantaneously at state transitions [3].

The case when the accumulated reward is monotone increasing, i.e., there is no reward loss, is referred to *preemptive resume*. In contrast, there are cases when state transitions can result in loss of accumulated reward. The case when the accumulated reward is reset to zero (complete reward loss) is referred to *preemptive repeat*. Between the preemptive resume and the preemptive repeat cases various cases of partial reward loss models were defined [4, 5].

An other interesting generalization is to relax the homogenity assumption of MRMs. In [6] inhomogeneous MRMs were considered, where both the generator of the structure state process and the reward rates might be functions of the time and the reward level. It has been shown that in several cases the inhomogeneous behaviour does not increase the computational complexity of the analysis, while the modeling power is increased.

In this paper we increase the modeling power of MRMs in a different manner. The randomness of traditional MRMs is due to the stochastic nature of the structure state process, and given the trajectory of the structure state process the accumulated reward is a deterministic function. With second-order reward accumulation process it is possible to model another kind of randomness as well. In second-order reward models the accumulated reward is a random function of the trajectory. During a sojourn in a state the reward increment is a random variable whose mean and variance depend on the visited state and the time spent in that state. We consider the preemptive resume policy (i.e., no reward loss at state transitions) only, and do not consider impulse reward accumulation. However, the introduced solution method al-

* This work is partially supported by OTKA grant n. T-34972.

lows to relax these restrictions.

The idea to investigate second-order discrete-continuous combined processes is not new. Second-order fluid models are considered for example in [7, 8], with absorbing or reflecting behaviour at the lower bound.

The behaviour of second-order reward models is very similar to the behaviour of second-order fluid models with one essential difference. In second-order reward models the reward accumulation process is not bounded. In practical applications it is common to assume that the accumulation process has a positive drift (i.e., the mean accumulated reward is increasing), but we do not apply any restriction in the mathematical description of these models. The lack of bounds of the accumulation process makes the analysis of second-order reward models simpler than the one of second-order fluid models. One of the main contributions of this paper is the effective numerical procedure which is based on this relative simplicity of second-order reward models and allows the analysis of fairly large models. The resulting formulas are comparable with the ones of ordinary MRMs and the ones of second-order fluid models. During the discussion below we indicate these relations.

The paper is organized as follows. First, in section 2, we briefly introduce Brownian motions focusing on the properties that are used in the rest of the paper. We define second order Markov reward models on this bases in section 3. In section 4 and 5 we derive the differential equations that describe the density function and the moments of the accumulated reward. We present these measures both, in time and in Laplace domain. Section 6 describes a numerical method to compute the moments of the accumulated reward, and gives an error bound to control the precision. Section 7 provides a numerical example to demonstrate the application of second-order Markov reward models and section 8 concludes the paper.

2. Some properties of Brownian motion

The definition of Brownian motions[1] used in this paper is based on the definition provided in [9] page 490 and it is extended with non-zero drift:

Definition 1 *A (real-valued) stochastic process, $\mathcal{X}(t)$, is a Brownian motion with drift r and variance σ^2 if:*

- *$\mathcal{X}(t)$ has independent increments,*
- *the increment $\mathcal{X}(s+t)$ - $\mathcal{X}(s)$ for all $s,t \geq 0$ is normal distributed with mean rt and variance $\sigma^2 t$, where r is a real valued and σ is a positive constant.*
- *the sample paths of $\mathcal{X}(t)$ are almost surely continuous.*

1 This process is also referred to as Wiener process (e.g., in [9])

The density function of $\mathcal{X}(t)$ at time t is:

$$f(t,y) = \lim_{\Delta \to 0} \frac{1}{\Delta} \, Pr(y < \mathcal{X}(t) < y + \Delta)$$

and its double-sided Laplace transform[2] with respect to y is:

$$f^*(t,v) \; = \; \int_{-\infty}^{\infty} f(t,y) \, e^{-vy} \, dy.$$

Assuming $\mathcal{X}(0) = 0$ and using the second part of Definition 1 $f(t,y)$ and its Laplace transform can be expressed as:

$$\begin{aligned} f(t,y) &= \frac{1}{\sqrt{2\pi t \sigma^2}} \, e^{-\frac{(y-rt)^2}{2t\sigma^2}} \\ f^*(t,v) &= e^{-vrt+\frac{v^2}{2}\sigma^2 t}. \end{aligned}$$

Due to the quick decay of the normal distribution both, at large negative and at large positive values the double-sided Laplace transform exists for any finite v.

The first terms of the Taylor series of $f^*(t,v)$ around 0 are:

$$f^*(\Delta,v) \; = \; 1 \; - \; [v \, r - \frac{v^2}{2} \, \sigma^2]\Delta + o(\Delta). \qquad (1)$$

From the results of Wiener and Lévy we know that Brownian motion is the only higher order continuous time stochastic process with (almost surely) continuous trajectory [10].

3. Second-order Markov reward models

Definition 2 *A second-order Markov reward model is composed by a continuous time Markov chain and a Brownian motion with drift, where the drift and the variance parameter of the brownian motion is determined by the state of the Markov chain.*

Let the structure state process $\{\mathcal{Z}(t), \; t \geq 0\}$ be a finite state, continuous time Markov chain on state space $S = \{1, \ldots N\}$ with generator $\boldsymbol{Q} = \{q_{ij}\}$ and initial probability vector $\underline{\pi}$. Its transient probability vector $\underline{p}(t)$ satisfies:

$$\frac{d}{dt}\underline{p}(t) \; = \; \underline{p}(t) \, \boldsymbol{Q} \quad \text{with initial condition: } \underline{p}(0) = \underline{\pi}.$$

While the structure state process stays in state i ($i \in S$), reward is accumulated according to a Brownian motion with

2 Due to the nature of the considered problem we apply double-sided Laplace transform throughout this paper without mentioning it any more. The only difference of the single-sided and the double-sided Laplace transform, which is used in the paper, is the lack of initial value in the transform of a derivative: $\mathcal{L}_S\{f'(t)\} = sf^*(s) - f(0)$ while $\mathcal{L}_D\{f'(t)\} = sf^*(s)$

drift $-\infty < r_i < \infty$ and variance $0 \le \sigma_i^2 < \infty$, i.e., the accumulated reward process, $\mathcal{B}(t)$, is a Brownian motion with drift r_i and variance σ_i^2.

Figure 1 shows a sample realization of a second-order Markov reward model. On this figure state 2 has the largest assigned drift and variance ($r_2 = 3, \sigma_2^2 = 2$). By such a large variation it happens with a not negligible probability (for example between $t \approx 1.7$ and $t \approx 1.85$) that the amount of reward is less when the state is left than the one when the state is entered, even if the drift is large.

state	*rate*(r_i)	*variance*(σ_i^2)
0	0.1	0
1	2	0.1
2	3	2

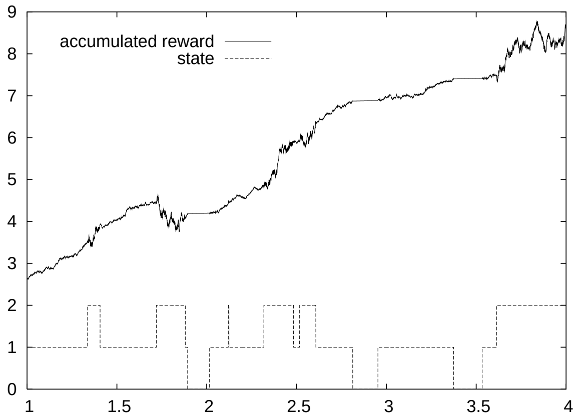

Figure 1. A sample realization of a second order reward model

In the sequel we use the following properties of second-order Markov reward models. If $\mathcal{Z}(t) = i$ and there is no state transition in the $(t, t+\Delta)$ interval the reward increment over this interval, $\mathcal{B}(t+\Delta) - \mathcal{B}(t)$, is normal distributed with mean $r_i\Delta$ and variance $\sigma_i^2\Delta$. If $\mathcal{Z}(t) = i$ and there is a state transition in the $(t, t+\Delta)$ interval to state k at time $t+\gamma$ ($\gamma < \Delta$) the distribution of the reward increment over the $(t, t+\Delta)$ interval is normal distributed with mean $r_i\gamma + r_k(\Delta-\gamma)$ and variance $\sigma_i^2\gamma + \sigma_k^2(\Delta-\gamma)$, since the sum of two normal distributed random variable is normal distributed as well and their first two cumulants (mean and variance) sum up.

For later use we define the diagonal matrices $\boldsymbol{R}$ and $\boldsymbol{S}$, constructed from r_i and σ_i^2:

$$\boldsymbol{R} = \begin{bmatrix} r_1 & & & \\ & r_2 & & \\ & & \ddots & \\ & & & r_N \end{bmatrix}, \quad \boldsymbol{S} = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_N^2 \end{bmatrix}.$$

Matrices $\boldsymbol{Q}$, $\boldsymbol{R}$ and $\boldsymbol{S}$ together with vector $\underline{\pi}$ define a second-order Markov reward model.

According to the mentioned continuity property of Brownian motion second-order Markov reward models are the most general Markov modulated reward models with (almost surely) continuous reward accumulation. Higher order models represent non-continuous reward accumulation and there is no other second order model.

As it is known for normal distributions the accumulated reward might become negative with commonly negligible, but positive probability and as it is emphasized in the above example the accumulated reward function is not monotone (as long as $\sigma > 0$). This model property has to be considered in practical application of second-order Markov reward models.

4. Analysis of accumulated reward

In this section we provide the differential equations that describe the density function and the moments of the accumulated reward. The applied analysis approach is the same as the one used for first order reward models, e.g. in [6]. First we introduce some definitions and notations.

The density function of the accumulated reward is[3]:

$$b_i(t,x) = \lim_{\Delta\to 0} \frac{1}{\Delta} Pr(x < \mathcal{B}(t) < x+\Delta | \mathcal{B}(0) = 0, \mathcal{Z}(0) = i),$$

and its Laplace transform with respect to x is:

$$b_i^*(t,v) \; = \; \int_{-\infty}^{\infty} b_i(t,x)\, e^{-vx}\, dx.$$

Furthermore, we introduce the following column vectors:

$$\underline{b}(t,x) \; = \; \{b_i(t,x)\} \text{ and } \underline{b}^*(t,v) \; = \; \{b_i^*(t,v)\}.$$

The following theorem provides the distribution of the accumulated reward of a second-order Markov reward model.

3 We use generalized density function when the distribution of $\mathcal{B}(t)$ is not continuous and the analysis remains the same.

Theorem 1 $\underline{b}^*(t,v)$ *satisfies the following differential equation:*

$$\frac{\partial}{\partial t}\,\underline{b}^*(t,v) + v\,\boldsymbol{R}\,\underline{b}^*(t,v) - \frac{v^2}{2}\,\boldsymbol{S}\,\underline{b}^*(t,v) = \boldsymbol{Q}\,\underline{b}^*(t,v), \tag{2}$$

with initial condition

$$\underline{b}^*(0,v) = \underline{h},$$

where $\underline{h}$ *is the column vector of ones.*

Proof of Theorem 1 The density of the reward at time $t+\Delta$ can be expressed by conditioning on the possible state transition in $(0,\Delta)$, the following way:

$$\begin{aligned} &b_i(t+\Delta,x) = \\ &(1+q_{ii}\Delta)\int_{-\infty}^{\infty} b_i(t,x-y)\; f_i(\Delta,y)\; dy \\ &+\sum_{k\in S,k\neq i} q_{ik}\Delta\int_{-\infty}^{\infty} b_k(t,x-y)\; f_{C_{i,k}}(\Delta,y)\; dy \\ &+o(\Delta). \end{aligned} \tag{3}$$

The first term represents the case when there is no state transition in $(0,\Delta)$, so the reward is accumulated according to a Brownian motion with drift r_i and variance σ_i^2. The density function of the accumulated reward at time Δ is $f_i(\Delta,y)$. To accumulate x amount of reward at time $t+\Delta$, $x-y$ amount of reward has to be accumulated in the remaining t long interval. The probability of no state transition in $(0,\Delta)$ is $1+q_{ii}\Delta+o(\Delta)$.
The second term corresponds to the case when a state transition takes place from state i to state k during $(0,\Delta)$. The probability of this case is $q_{ik}\Delta+o(\Delta)$. The accumulated reward at time Δ is normal distributed (see the end of section 3), but the parameters depends on the time of the state transition. We denoted the density function of the accumulated reward at time Δ by $f_{C_{i,k}}(\Delta,y)$, whose mean and variance are r'_{ik} and ${\sigma'}^2_{ik}$.
The probability of having more than one state transition in $(0,\Delta)$ is $o(\Delta)$.

The Laplace transformation of (3) with respect to x is:

$$\begin{aligned} b_i^*(t+\Delta,v) &= (1+q_{ii}\Delta)b_i^*(t,v)f_i^*(\Delta,v) \\ &+\sum_{k\in S,k\neq i} q_{ik}\Delta b_k^*(t,v)f_{C_{i,k}}^*(\Delta,v)\; +\; o(\Delta). \end{aligned}$$

Subtracting $b_i^*(t,v)$ from both sides, applying eq. (1), and dividing by Δ results:

$$\begin{aligned} &\frac{b_i^*(t+\Delta,v)-b_i^*(t,v)}{\Delta} = \\ &\sum_{k\in S} q_{ik}b_k^*(t,v) - b_i^*(t,v)\Big(v\,r_i - \frac{v^2}{2}\sigma_i^2\Big) \\ &- q_{ii}b_i^*(t,v)\underbrace{\Big(v\,r_i - \frac{v^2}{2}\sigma_i^2\Big)\Delta}_{\text{from } f_i^*(\Delta,v)} \\ &- \sum_{k\in S,k\neq i} q_{ik}b_k^*(t,v)\underbrace{\Big(v\,r'_{ik} - \frac{v^2}{2}{\sigma'}^2_{ik}\Big)\Delta}_{\text{from } f_{C_{i,k}}^*(\Delta,v)} \\ &+\frac{o(\Delta)}{\Delta} \end{aligned}$$

Where $r'_{ik} \leq \max(r_i, r_k)$, and ${\sigma'}^2_{ik} \leq \max(\sigma_i^2,\sigma_k^2)$.
Taking the limit $\Delta \to 0$, and using matrix notation yields (2).
Based on the definition of $b_i(t,x)$ at time zero $\mathcal{B}(0) = 0$. The density function of the accumulated reward is the Dirac delta function: $b_i(0,x) = \delta(x),\ \forall i \in S$ and its Laplace transform is constant 1, which is the initial condition of the differential equation (2). □

With appropriate Inverse Laplace and Laplace transformation Equation (2) can be expressed in original and in double Laplace transform domain:

Corollary 1 $\underline{b}(t,x)$ *is the solution of the following differential equation:*

$$\frac{\partial}{\partial t}\,\underline{b}(t,x) + \boldsymbol{R}\,\frac{\partial}{\partial x}\,\underline{b}(t,x) - \frac{1}{2}\,\boldsymbol{S}\,\frac{\partial^2}{\partial x^2}\,\underline{b}(t,x) = \boldsymbol{Q}\,\underline{b}(t,x), \tag{4}$$

with initial condition:

$$\underline{b}(0,x) = \underline{\delta}(x).$$

Corollary 2 *The Laplace transform* $\underline{b}^{**}(s,v)$ *of the accumulated reward with respect to* t *and* x *is:*

$$\underline{b}^{**}(s,v) \;=\; \Big[s\boldsymbol{I} \;-\; \boldsymbol{Q} \;+\; v\boldsymbol{R} \;-\; \frac{v^2}{2}\boldsymbol{S}\Big]^{-1}\underline{h} \tag{5}$$

The obtained results indicate the formal relation of first and second order Markov reward models in time domain:

$$\frac{\partial}{\partial t}\underline{b}(t,x) + \boldsymbol{R}\frac{\partial}{\partial x}\underline{b}(t,x) - \tfrac{1}{2}\boldsymbol{S}\frac{\partial^2}{\partial x^2}\underline{b}(t,x) = \boldsymbol{Q}\underline{b}(t,x)$$
$$\Updownarrow$$
$$\frac{\partial}{\partial t}\underline{b}(t,x) + \boldsymbol{R}\,\frac{\partial}{\partial x}\underline{b}(t,x) = \boldsymbol{Q}\;\underline{b}(t,x)$$

and in double transform domain:

$$\underline{b}^{**}(s,v) = \Big[s\boldsymbol{I} - \boldsymbol{Q} + v\boldsymbol{R} - \tfrac{v^2}{2}\boldsymbol{S}\Big]^{-1}\underline{h}$$
$$\Updownarrow$$
$$\underline{b}^{**}(s,v) = \Big[s\boldsymbol{I} - \boldsymbol{Q} + v\boldsymbol{R}\Big]^{-1}\underline{h}$$

The distribution of the accumulated reward can be computed using a partial differential equation solver based on eq. (4) or an ordinary differential equation solver and a numerical inverse transformation based on eq. (3) or inverse Laplace transformation from double transform domain based on eq. (5) (e.g., [11]). Unfortunately non of these procedures is applicable for second-order Markov reward models with more than 100 states. Nevertheless, similar to first-order Markov reward models, the moments of the accumulated reward can be analyzed with a more effective numerical method.

In the light of Eq. (4), we also can interpret the relation of second-order fluid models and second-order Markov reward models. The same partial differential equation characterize the system distribution of both models inside the valid region, but due to the fact that the reward accumulation is unbounded in second-order Markov reward models, but the fluid level is bounded (from one side (level 0) or from two sides (level 0 and the buffer size)) in second-order fluid models different boundary conditions apply to the different models. The solutions of a partial differential equation with different boundary conditions might differ significantly, hence unfortunately, the relatively simple solution of second-order Markov reward models is not applicable for the solution of second-order fluid models.

5. Moments of the accumulated reward

The nth moment of the accumulated reward at time t is denoted by $V_i^{(n)}(t)$:

$$V_i^{(n)}(t) \;=\; E(\mathcal{B}^n(t) \mid \mathcal{B}(0)=0,\; \mathcal{Z}(0)=i),$$

and the column vector of the nth moments is $\underline{V}^{(n)}(t) = \{V_i^{(n)}(t)\}$.

The analytical description of the moments of accumulated reward is provided by the following theorem.

Theorem 2 *$\underline{V}^{(n)}(t)$ is the solution of the following ordinary differential equation:*

$$\frac{\partial}{\partial t}\underline{V}^{(n)}(t) \;-\; n\,\boldsymbol{R}\,\underline{V}^{(n-1)}(t) \;- \frac{1}{2}\,n\,(n-1)\,\boldsymbol{S}\,\underline{V}^{(n-2)}(t) \;=\; \boldsymbol{Q}\,\underline{V}^{(n)}(t), \tag{6}$$

with initial condition:

$$\underline{V}^{(0)}(0) \;=\; \underline{h}, \qquad \textit{and}\;\; \underline{V}^{(n)}(0) \;=\; \underline{0},\; \forall n \geq 1.$$

Proof of Theorem 2 We start from eq. (2) and apply the following relation of the Laplace transform and the moments of the accumulated reward:

$$\underline{V}^{(n)}(t) \;=\; (-1)^n\,\frac{\partial^n}{\partial v^n}\,\underline{b}^*(t,v)\,|_{v=0}.$$

First we show by induction that:

$$\frac{\partial^n}{\partial v^n} v\,\underline{b}^*(t,v) \;=\; v\,\frac{\partial^n}{\partial v^n}\underline{b}^*(t,v) \;+\; n\,\frac{\partial^{n-1}}{\partial v^{n-1}}\underline{b}^*(t,v). \tag{7}$$

This statement clearly holds for $n=1$. Assuming that it is true for $n-1$, the nth is expressed as:

$$\frac{\partial^n}{\partial v^n} v\,\underline{b}^*(t,v) = \frac{\partial}{\partial v}\left(v\,\frac{\partial^{n-1}}{\partial v^{n-1}}\underline{b}^*(t,v) \;+\; (n{-}1)\frac{\partial^{n-2}}{\partial v^{n-2}}\underline{b}^*(t,v)\right) = v\,\frac{\partial^n}{\partial v^n}\underline{b}^*(t,v) \;+\; n\,\frac{\partial^{n-1}}{\partial v^{n-1}}\underline{b}^*(t,v).$$

Next, we show that:

$$\frac{\partial^n}{\partial v^n}\frac{v^2}{2}\,\underline{b}^*(t,v) \;=\; \frac{v^2}{2}\,\frac{\partial^n}{\partial v^n}\underline{b}^*(t,v) \;+\; nv\,\frac{\partial^{n-1}}{\partial v^{n-1}}\underline{b}^*(t,v) \;+\; \frac{1}{2}n(n-1)\,\frac{\partial^{n-2}}{\partial v^{n-2}}\underline{b}^*(t,v). \tag{8}$$

The proof is induction based again. For $n=1$ and $n=2$ eq. (8) holds. Assuming that it holds for $1\ldots n-1$, for the nth derivative we have:

$$\frac{\partial^n}{\partial v^n}\frac{v^2}{2}\,\underline{b}^*(t,v) \;= \frac{\partial}{\partial v}\left(\frac{v^2}{2}\,\frac{\partial^{n-1}}{\partial v^{n-1}}\underline{b}^*(t,v) \;+\; v(n{-}1)\,\frac{\partial^{n-2}}{\partial v^{n-2}}\underline{b}^*(t,v) \;+ \frac{1}{2}(n{-}1)(n{-}2)\,\frac{\partial^{n-3}}{\partial v^{n-3}}\underline{b}^*(t,v)\right) \;= \frac{v^2}{2}\,\frac{\partial^n}{\partial v^n}\underline{b}^*(t,v) \;+\; nv\,\frac{\partial^{n-1}}{\partial v^{n-1}}\underline{b}^*(t,v) \;+ \frac{1}{2}n(n-1)\,\frac{\partial^{n-2}}{\partial v^{n-2}}\underline{b}^*(t,v).$$

The nth derivative of $\frac{\partial}{\partial t}\underline{b}^*(t,v)$ and $\underline{b}^*(t,v)Q$ with respect to v is straightforward and using eq. (7) and (8) the theorem is given. □

6. Numerical Method

In this section we provide a randomization based numerical method that efficiently computes the moments of the accumulated reward at time t. Instead of using matrices $\boldsymbol{Q}$, $\boldsymbol{R}$ and $\boldsymbol{S}$ we introduce the following non-negative and substochastic matrices (i.e., their row-sum ≤ 1):

$$\boldsymbol{Q'} = \frac{1}{q}\,\boldsymbol{Q} \;+\; \boldsymbol{I}, \qquad \boldsymbol{R'} = \frac{1}{qd}\,\boldsymbol{R}, \qquad \boldsymbol{S'} = \frac{1}{qd^2}\,\boldsymbol{S},$$

where $q = \max_{i\in S}|q_{ii}|$, $d = \max_{i\in S}\{r_i, \sigma_i\}/q$. If there are negative reward rates the following model transformation provides non-negative, substochastic matrices: $\check{r} =$

$\min_{i \in S} r_i$, $\check{\boldsymbol{R}} = \boldsymbol{R} - \check{r}\boldsymbol{I}$ and $Pr(\mathcal{B}(t) < c) = Pr(\check{\mathcal{B}}(t) < c - \check{r}t)$. There are several advantages of using substochastic matrices in numerical methods. Multiplying only substochastic matrices and non-negative vectors avoids subtraction, which is usual source of numerical errors using floating point number representation. Furthermore, the result of the multiplication of substochastic matrices and a bounded non-negative vector is bounded as well. Exploiting this bound one can set the numerical precision of complex procedures.

The following theorem provides the moments of the accumulated reward using the introduced substochastic matrices.

Theorem 3 *The nth moment of the accumulated reward, $\underline{V}^{(n)}(t)$ can be expressed as follows:*

$$\underline{V}^{(n)}(t) \;=\; n!\, d^n \sum_{k=0}^{\infty} e^{-qt} \frac{(qt)^k}{k!} \underline{U}^{(n)}(k), \qquad (9)$$

where the following recursive relation holds for the $\underline{U}^{(n)}(k)$ coefficients:

$$\underline{U}^{(n)}(k+1) \;=\; \boldsymbol{R'}\, \underline{U}^{(n-1)}(k) + \frac{1}{2}\, \boldsymbol{S'}\, \underline{U}^{(n-2)}(k) \;+\; \boldsymbol{Q'}\, \underline{U}^{(n)}(k), \qquad (10)$$

with initial values:

$$\underline{U}^{(0)}(0) \;=\; \underline{h}, \quad \textit{and } \underline{U}^{(n)}(0) \;=\; \underline{0}, \;\; \forall n > 0.$$

Proof of Theorem 3 Substituting (9) into eq. (6) gives:

$$\begin{aligned} & n!\, d^n \sum_{k=0}^{\infty} e^{-qt} \frac{(qt)^k}{k!} \left[q\underline{U}^{(n)}(k+1) - q\underline{U}^{(n)}(k) \right] = \\ & n\,(n-1)!\, d^{n-1}\, qd\boldsymbol{R'} \sum_{k=0}^{\infty} e^{-qt} \frac{(qt)^k}{k!}\, \underline{U}^{(n-1)}(k) + \\ & \frac{1}{2}\, n(n-1)(n-2)!\, d^n q\boldsymbol{S'} \sum_{k=0}^{\infty} e^{-qt} \frac{(qt)^k}{k!}\, \underline{U}^{(n-2)}(k) + \\ & n!\, d^n\, q(\boldsymbol{Q'} - \boldsymbol{I}) \sum_{k=0}^{\infty} e^{-qt} \frac{(qt)^k}{k!}\, \underline{U}^{(n)}(k) \end{aligned}$$

With the given recursive formula for $\underline{U}^{(n)}(k)$, the above equality holds for each term having the same Poisson coefficient. □

The following theorem provides an error bound to control the precision of the procedure given in Theorem 3

Theorem 4 *The nth moment of the accumulated reward can be calculated as a finite sum up to G and an error vector $\underline{\xi}(G)$, whose elements are less than ϵ:*

$$\underline{V}^{(n)}(t) \;=\; n!\, d^n \sum_{k=0}^{G} e^{-qt} \frac{(qt)^k}{k!} \underline{U}^{(n)}(k) + \underline{\xi}(G),$$

where the limit, G, is calculated as:

$$G \;=\; \min_{g} \left(2\, d^n\, n!\, (qt)^n \sum_{k=g+n+1}^{\infty} e^{-qt} \frac{(qt)^k}{k!} < \epsilon \right). \qquad (11)$$

The proof of the theorem is in Appendix A.

Theorem 3 and 4 present a description of the moments of accumulated reward which allows an effective computation of these measures. Appendix B introduces a pseudo code of the procedure based on these theorems.

The computational complexity of the procedure is characterized by G and the complexity of the matrix vector multiplications in the $\underline{U}^{(j)} := 0.5\, \boldsymbol{S'} \cdot \underline{U}^{(j-2)} + \boldsymbol{R'} \cdot \underline{U}^{(j-1)} + \boldsymbol{Q'} \cdot \underline{U}^{(j)}$ step. The execution of this step G times dominates the execution time. Commonly, G has the same order of magnitude as qt, and since $\boldsymbol{R'}$ and $\boldsymbol{S'}$ are diagonal matrices and $\boldsymbol{Q'}$ is often sparse the complexity of the iterative step is equivalent with some vector-vector multiplications. If the mean number of non-zero elements in a row of $\boldsymbol{Q'}$ is m the complexity of the iterative step is equivalent with $m+2$ vector-vector multiplications and 3 vector summations where the cardinality of the square matrices and the vectors is $|S|$. The memory requirement of the procedure is determined by the size of the input data and the result. The size of memory needed to store the input data (matrix $\boldsymbol{Q}$, diagonal matrices $\boldsymbol{R'}$ and $\boldsymbol{S'}$ and vector $\underline{\pi}$) is the size of $(m+3) \times |S|$ floating point numbers and to calculate and store result we need memory for further $n+1$ vectors of size $|S|$.

7. Example

We study the performance of a tentative telecommunication system. A communication channel with capacity C serves traffic of two traffic classes. The service policy is such that class 1 has higher priority than class 2. There are N class 1 traffic sources which follow ON-OFF behaviour with exponentially distributed on and off periods, whose parameters are α and β, respectively. During its ON period a class 1 traffic source transmits data with rate r and variance σ^2, i.e., in a t long ON period the amount of transmitted data is normal distributed with mean rt and variance $\sigma^2 t$. In contrast with the behaviour of real communication systems this quantity can be negative with a given probability. If $r >> \sigma$ the probability that negative data is transmitted is negligible. During its OFF period a class 1 traffic does not transmit data. The class 2 traffic utilizes the remaining bandwidth of the channel. The performance measure of interest is the channel capacity available for class 2 traffic in the $(0, t)$ interval supposing that each class 1 sources are in the OFF phase at time 0. We define a second-order Markov reward model which describes the capacity available for class 2 traffic.

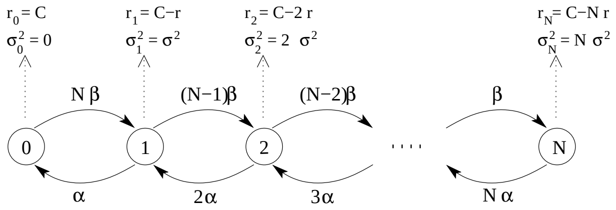

Figure 2. The structure of the background process

Figure 2 shows the behaviour of the "background" CTMC which characterize the instantaneous bandwidth available for class 2 traffic. The state numbers indicate the number of active (ON phase) class 1 customers. The figure also indicate the rate and variance associated with the states of the background process. During a visit in state i $r_i = C - i \cdot r$ and $\sigma_i^2 = i \cdot \sigma^2$.

Capacity of the channel:	$C = 32$
Number of CBR sources:	$N = 32$
Parameter of the ON period:	$\alpha = 4$
Parameter of the OFF period:	$\beta = 3$
Transmission rate of a CBR source:	$r = 1$
Variance of the transmission rate:	$\sigma^2 = 0, 1, 10$

Table 1. Parameters used in the example

Table 1 summarizes the considered values of model parameters. To demonstrate the difference between ordinary (first-order) and second order Markov reward models we evaluate the example with 3 different variance parameters. $\sigma^2 = 0$ corresponds to a first-order model, while $\sigma^2 = 1$ and $\sigma^2 = 10$ corresponds to second-order models.

The mean value of the accumulated reward with the 3 considered variance values are plotted in Figure 3. The figure verifies the assumption that the mean accumulated reward is independent of the variance parameter. To visualize the non-linear behaviour of the mean curve (only on this figure) we also plot the mean accumulated reward starting from steady state, which is a linear function of time.

Figure 4 depicts the second and third moment of the accumulated reward as a function of time for the 3 variance values. The figures demonstrate a (natural) general feature of second-order Markov reward models. The higher are the variance parameters of the states the bigger are the higher moments of the accumulated reward.

The numerical procedure presented in section 6 calculates only the moments of the accumulated reward. To compute the distribution of the accumulated reward one can solve the partial differential equation (4) numerically, which might be slow and inaccurate or can approximate the distribution based on its moments. For example, one can apply the method presented in [12] for calculating the bounds of the accumulated reward distribution based on its moments.

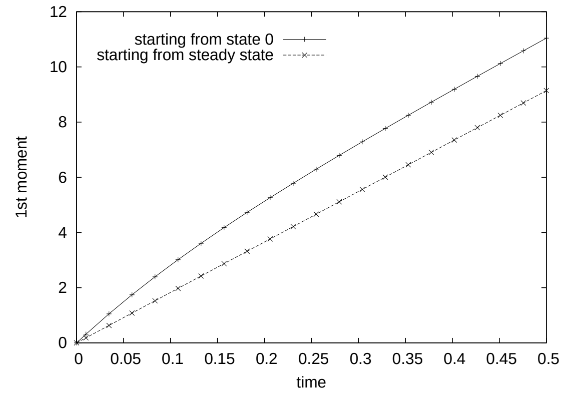

Figure 3. The mean of the accumulated reward

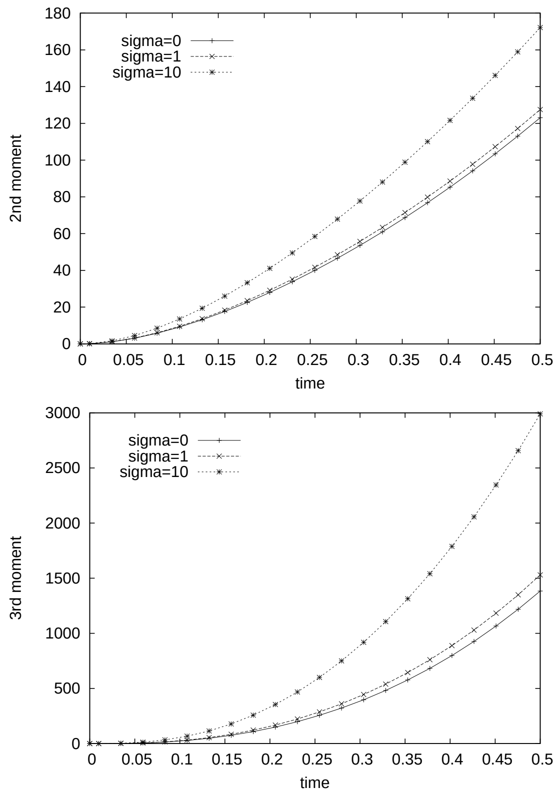

Figure 4. The 2nd and 3rd moment of the accumulated reward

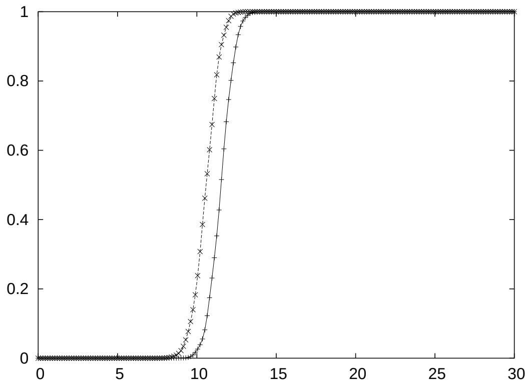

Figure 5. Bounds for the distribution of the accumulated reward with $\sigma^2 = 0$

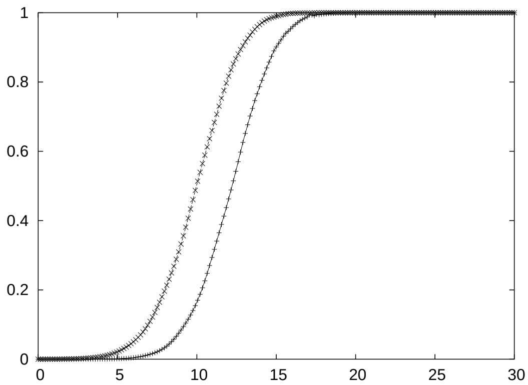

Figure 6. Bounds for the distribution of the accumulated reward with $\sigma^2 = 1$

Figures 5, 6 and 7 show the bounds of the accumulated reward distribution obtained with the different σ^2 parameters at time $t = 0.5$. Among other factors the tightness of the bounds depends on the number evaluated moments, which was 23 in these cases.

The results presented in this section were calculated with a C++ implementation of the randomization based numerical method presented in section 6. To be memory efficient our implementation uses sparse matrix representation.

The presented results have been compared to the results of a numerical ODE solver (working based on eq. 6 using trapezoid rule), and a second-order reward model simulation tool. The three solutions gave exactly the same results, however the randomization was far the fastest (the calculation of any figure presented above took less than one second on a PC running at 2.4 GHz).

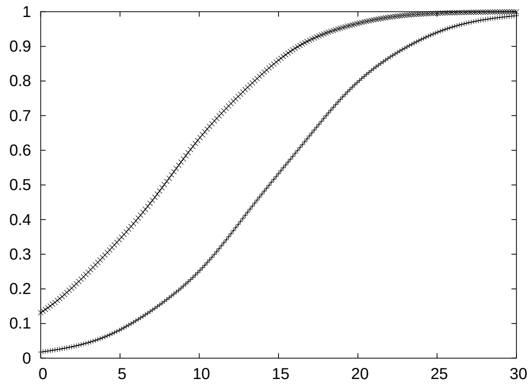

Figure 7. Bounds for the distribution of the accumulated reward with $\sigma^2 = 10$

Capacity of the channel:	$C = 200.000$
Number of CBR sources:	$N = 200.000$
Parameter of the ON period:	$\alpha = 4$
Parameter of the OFF period:	$\beta = 3$
Transmission rate of a CBR source:	$r = 1$
Variance of the transmission rate:	$\sigma^2 = 10$

Table 2. Parameters of the large model

To test the computational complexity of the numerical procedure we also evaluated the same example with a larger background process. Table 2 presents the parameters of the large model. Figure 8 depicts the first three moments of the accumulated reward at time $0.01, 0.02, 0.03, 0.04$ and 0.05. The overall computation time of the 5 evaluated time points was 3 hours on the same computer. At the final time point ($t = 0.05$) the number of required iterations was $G = 41,588$ with precision requirement $\epsilon = 10^{-9}$. In this case $q = 800,000$ and $qt = 40,000$. Due to the regular structure of $\boldsymbol{Q'}$ with 3 non-zero elements in each row (except the first and last ones) the number of floating point multiplications was

$$(\underbrace{3}_{\boldsymbol{Q'}} + \underbrace{1}_{\boldsymbol{R'}} + \underbrace{1}_{\boldsymbol{S'}}) \times \underbrace{200,001}_{|S|} \times \underbrace{4}_{\text{3 moments}}$$

in each of the G iteration steps, which took one hour.

8. Conclusion

This paper presents the analytical description of second-order Markov reward models. This description is similar to the one of second-order fluid models, but there are some dominant differences between them. The similarities and

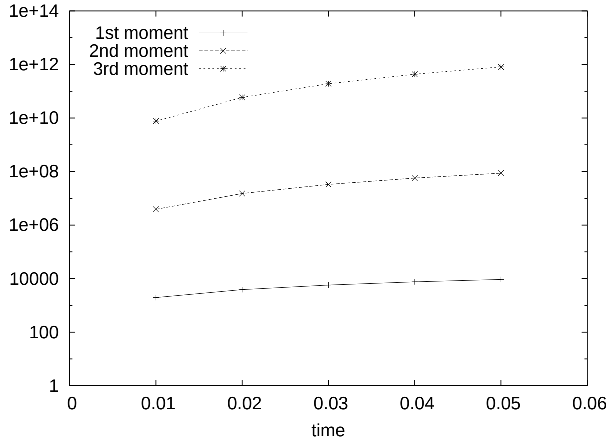

Figure 8. The moments of the accumulated reward of the large example

differences of second-order reward and second-order fluid models are emphasized as well.

Utilizing the special features of second-order Markov reward models an effective randomization based numerical procedure is presented which calculates the moments of accumulated reward.

The numerical analysis of a tentative communication system demonstrates the application and the analysis of second-order Markov reward models. Using the proposed analysis method the computational complexity and the memory requirement of the reward analysis of second-order models is practically identical with the ones of first-order models.

A. Proof of Theorem 4

Theorem 4 is based on the following lemmas.

Lemma 1 *For $n \le 2k$ the explicit solution of the recursion*

$$g_{n,k} \;=\; g_{n,k-1} \;+\; g_{n-1,k-1} \;+\; \frac{1}{2}\, g_{n-2,k-1} \qquad (12)$$

with initial conditions $g_{0,k} = 1,\ g_{1,k} = k,\ \forall k \ge 0;\ g_{n,0} = 0,\ \forall n > 0$ *is:*

$$g_{n,k} = \sum_{j=\max(0,n-k)}^{\lfloor n/2 \rfloor} 2^{-j} \frac{k!}{j!(k-n+j)!(n-2j)!} \qquad (13)$$

Proof of Lemma 1

Recursion (12) gives the nth coefficient of the $(1+x+\frac{1}{2}x^2)^k$ polynom, i.e., $(1+x+\frac{1}{2}x^2)^k = \sum_{n=0}^{2k} g_{n,k}\, x^n$. It can be checked looking at:

$$\Big(1+x+\frac{1}{2}x^2\Big)^k = \Big(1+x+\frac{1}{2}x^2\Big)\Big(1+x+\frac{1}{2}x^2\Big)^{k-1} = \Big(1+x+\frac{1}{2}x^2\Big)\sum_{n=0}^{2k-2} g_{n,k-1}\, x^n .$$

The power series expansion of $(1+x+\frac{1}{2}x^2)^k$ is:

$$\sum_{\ell=0}^{k} \binom{k}{\ell} x^\ell \sum_{j=0}^{\ell} \binom{\ell}{j} 2^{-j} x^j \;=\; \sum_{\ell=0}^{k} \sum_{j=0}^{\ell} \frac{k!}{(k-\ell)!(\ell-j)!j!}\, 2^{-j}\, x^{j+\ell} \;=\; \sum_{n=0}^{2k} x^n \underbrace{\sum_{j=\max(0,n-k)}^{\lfloor n/2 \rfloor} 2^{-j} \frac{k!}{j!(k-n+j)!(n-2j)!}}_{g_{n,k}},$$

where in the last step a new variable $n = \ell + j$ has been introduced[4]. □

Lemma 2 *Assuming $g_{n,k} = 0$ for all $n > 2k$ the elements of the $\underline{U}^{(n)}(k)$ vectors are upper-bounded by $g_{n,k}$ for all $n, k \ge 0$, i.e.*

$$\underline{U}^{(n)}(k) \;\le\; g_{n,k}\, \underline{h} \qquad (14)$$

Proof of Lemma 2 Due to the substochastic feature $\boldsymbol{R'}$, $\boldsymbol{S'}$ and $\boldsymbol{Q'}$ the following inequalities hold

$$\boldsymbol{R'}\ \underline{h} \le \underline{h},\quad \boldsymbol{S'}\ \underline{h} \le \underline{h} \text{ and } \boldsymbol{Q'}\ \underline{h} \le \underline{h}.$$

The proof is based on induction again. Based on the initial conditions of $g_{n,k} = 0$ and $\underline{U}^{(n)}(k)$ the lemma holds for $k = 0$. Assuming it also holds for k from eq. (10) we have

$$\begin{aligned}&\underline{U}^{(n)}(k+1) = \\ &\boldsymbol{R'}\ \underline{U}^{(n-1)}(k) \;+\; \tfrac{1}{2}\, \boldsymbol{S'}\ \underline{U}^{(n-2)}(k) \;+\; \boldsymbol{Q'}\ \underline{U}^{(n)}(k) \;\le \\ &\boldsymbol{R'}\ g_{n-1,k}\, \underline{h} \;+\; \tfrac{1}{2}\, \boldsymbol{S'}\ g_{n-2,k}\, \underline{h} \;+\; \boldsymbol{Q'}\ g_{n,k}\, \underline{h} \;\le \\ &g_{n-1,k}\, \underline{h} \;+\; \tfrac{1}{2}\, g_{n-2,k}\, \underline{h} \;+\; g_{n,k}\, \underline{h} \;= g_{n,k+1}\, \underline{h}\,. \qquad \Box\end{aligned}$$

Proof of Theorem 4 Starting from Lemma 1 and Lemma 2 $\underline{U}^{(n)}(k)$ can be further upper bounded:

$$\underline{U}^{(n)}(k) \;\le\; g_{n,k}\, \underline{h} \;\le\; \sum_{j=0}^{k} 2^{-j} \frac{k!}{(k-n)!}\, \underline{h} \;\le\; 2\frac{k!}{(k-n)!}\, \underline{h}.$$

4 We thank the correction of (13) to anonymous referee.

Then the error vector ($\underline{\xi}(G)$) is bounded by:

$$\underline{\xi}(G) \;=\; d^n\, n! \sum_{k=G+1}^{\infty} e^{-qt}\frac{(qt)^k}{k!}\,\underline{U}^{(n)}(k) \le$$
$$d^n\, n! \sum_{k=G+1}^{\infty} e^{-qt}\frac{(qt)^k}{k!}\cdot 2\frac{k!}{(k-n)!}\,\underline{h} \;=$$
$$2\, d^n\, n!\,(qt)^n \sum_{k=G+n+1}^{\infty} e^{-qt}\frac{(qt)^k}{k!}\,\underline{h},$$

from which eq. (11) follows. □

B. Implementation of the Numerical Algorithm

A pseudo code of an implementation of the numerical method based on Theorem 4 is described here.

Input	$\boldsymbol{Q}$	GENERATOR MATRIX OF THE CTMC
	$\boldsymbol{R}$	DIAGONAL MATRIX OF REWARD RATES
	$\boldsymbol{S}$	DIAGONAL MATRIX OF VARIANCES
	$\underline{P}$	INITIAL PROBABILITY VECTOR
	t	TIME OF ACCUMULATION
	n	ORDER OF MOMENT
	ϵ	COMPUTATION ACCURACY
Output	m	THE n-TH MOMENT OF ACC. REWARD

```
1 compute Q', R' and S'
  q := max_{i∈S} |q_ii|;
  d := max_{i∈S}{r_i, σ_i}/q;
  Q' := Q/q + I;
  R' := R/qd;
  S' := S/qd^2;
2 compute G
  G:=1;
  psum := 0;
  for  i := 1  to  G + n + 1
     psum := psum + Poisson(i; qt);
  while  2 · (qt)^n · d^n · n! · (1 − psum) > ε
   begin
     G := G + 1;
     psum := psum + Poisson(G + n + 1; qt);
   end
3 compute the n-th moment
  U^(0) = h;  U^(i) = 0 ,  i : 1 . . . n;
  for  i := 1  to  G  do
   begin
     for  j := n  downto  0  do
      U^(j) := 0.5 S'·U^(j−2) + R'·U^(j−1) + Q'·U^(j);
     m := m + U^(n)·Poisson(i; qt);
    end;
  m := P · m · n! · d^n
```

References

[1] G. Ciardo, R. Marie, B. Sericola, and K. Trivedi, "Performability analysis using semi-Markov reward processes," *IEEE Transactions on Computers*, vol. 39, pp. 1252–1264, 1990.

[2] M. Telek and A. Pfening, "Performance analysis of Markov Regenerative Reward Models," *Performance Evaluation*, vol. 27&28, pp. 1–18, 1996.

[3] M. A. Qureshi and W. H. Sanders, "Reward model solution methods with impulse and rate rewards: An algorithm and numerical results.," *Performance Evaluation*, pp. 413–436, 1994.

[4] A. Bobbio, V. G. Kulkarni, and M. Telek, "Partial loss in reward models," in *2nd Int. Conf. on Mathematical Methods in Reliability*, (Bordeaux, France), pp. 207–210, July 2000.

[5] V. Nicola, R. Martini, and P. Chimento, "The completion time of a job in a failure environment and partial loss of work," in *2nd Int. Conf. on Mathematical Methods in Reliability (MMR'2000)*, (Bordeaux, France), pp. 813–816, July 2000.

[6] M. Telek, A. Horváth, and G. Horváth, "Analysis of inhomogeneous markov reward models," in *International Conference on Numerical Solution of Markov Chains – NSMC 2003*, (Urbana, IL, USA), pp. 305–322, Sept 2003.

[7] S. Asmussen, "Stationary Distribution for Fluid Flow Models and Markov-Modulated Reflected Brownian Motion," Tech. Rep. TR R92-2015, Institute of Electronic Systems, Department of Mathematics and Computer Science, Aalborg University, Aalborg, Denmark, 1992.

[8] R. L. Karandikar and V. G. Kulkarni, "Second-order fluid flow models: Reflected brownian motion in a random environment," *Operations Research*, vol. 43, pp. 77–88, 1995.

[9] G. R. Grimmett and D. R. Stirzaker, *Probability and random processes*. Oxford University Press, 1992. second edition.

[10] W. Feller, *An Introduction to Probability Theory and its Applications - Vol. I and II*. New York: Wiley, 1968.

[11] G. L. Choudhury, D. M. Lucantoni, and W. Whitt, "Multi dimensional transform inversion with applications to the transient M/G/1 queue," *Annual Applied Probability*, vol. 4, pp. 719–740, 1994.

[12] S. Rácz, A. Tari, and M. Telek, "A distribution estimation method for bounding the reward measures of large MRMs," in *International Conference on Numerical Solution of Markov Chains – NSMC 2003*, (Urbana, IL, USA), pp. 341–342, Sept 2003. abstract.