A distribution estimation method for bounding the reward measures of large MRMs*

Sándor Rácz[1], Árpád Tari[2], Miklós Telek[2]
[1] Ericsson Research Hungary
Budapest, Hungary, Sandor.Racz@eth.ericsson.se
[2] Dept. of Telecommunications, Technical University of Budapest
1521 Budapest, Hungary, {arpi,telek}@webspn.hit.bme.hu

Abstract

This paper introduces an alternative approach for the numerical analysis of large Markov reward modes. Instead of the direct calculation of the distribution of reward measures, a two-step method is proposed. The first step is the analysis of the moments of required reward measures and the second step is the distribution estimation based on these moments. The focus of this paper is on the second step. We propose a numerical procedure and provide its detailed proof.

Numerical examples demonstrate the abilities of the proposed method. The examples verify the general feature of moment based distribution estimation, i.e., the bounds of the estimation are loose around the mean value and they are rather tight for extreme values. This property makes the proposed two-step method effective in bounding reward measures in the very unlikely region as it is the goal of the analysis of safety critical systems.

Keywords: large Markov reward models, distribution estimation, moment problem.

1 Introduction

The numerical analysis of discrete state systems is often limited by the size of the discrete state space. More and more effective numerical methods are needed to evaluate systems with increasing complexity. In this paper we propose an approach for the analysis of large Markov reward models (MRMs).

There are two main branches of numerical methods evaluating MRMs. The first branch of methods calculate the distribution of reward measures, e.g., [2, 3, 8, 4]. These methods are commonly based on randomization and their common feature is the evaluation of a two dimensional infinite summation. The other branch of methods evaluate only the moments of the reward measures [5, 16, 17]. Among these methods the recently published ones are also based on randomization and they evaluate a one dimensional infinite summation and a finite summation up to the number of required moments [17, 12, 13]. As a consequence, the computational complexity of the analysis of the first moment of reward measure is comparable with the transient analysis of the underlying Markov chain, because both of them require to evaluate a one dimensional infinite summation.

*This work was partially supported by Hungarian Scientific Research Fund (OTKA) under Grant No. T-34972.

Due to the mentioned general feature of the analysis methods the computational complexity of the methods calculating the moments is significantly less than the one of the methods calculating the distribution directly (hereafter referred to as direct methods). It also means that there are performance analysis problems which can not be evaluated using direct methods, but the moments of the reward measures can still be evaluated. According to our experiences the problems with state space of $10^3 - 10^6$ states fall in this class. To obtain approximate results on the distribution of reward measures in these cases we propose a two-step method. The first step is the analysis of moments and the second is the estimation of the distribution based on the moments. The first step is not discussed in this paper. We refer to [17, 12, 13] for details of the effective analysis of moments of MRMs.

Since the first effective analysis methods of moments were available we were looking for estimation methods of reward measure distribution based on its moments. We obtained a numerical procedure based on the properties of Hankel determinants of moments [10] which provided trustworthy results [11], but we could not prove the validity of this method still now. The book of Akhiezer [1] helped us a lot in understanding the basic rules of moments. On the base of this fundamental book we provide a detailed description of our numerical method to estimate the distribution of reward measures based on its moments and the proofs associated with the steps of the procedure.

Using this distribution estimation method one can evaluate a lower and an upper bounds of reward measure distribution also for those MRMs which can not be attacked by direct methods due to the size of the state space.

The rest of this paper is organized as follows. Section 2 collects the basic properties of moments applied in our distribution estimation method. Section 3 provides a high level description of the proposed method, while symbolic and numerical results are provided in Section 4 and 5, respectively. Section 6 concludes the paper.

2 Basic properties of moments

In this section we collect those properties of moments of real valued random variables which are utilized in the subsequent numerical analysis method.

2.1 Notations

Following [1] we introduce a set of notations.

$$\begin{aligned}
\mathfrak{M}_{2n} &= \text{the set of distributions with the same } 0,1,\ldots,2n \text{ moments.} && (1)\\
\sigma(\cdot) &= \text{a non-decreasing function} \quad (\sigma(x_1) \leq \sigma(x_2) \text{ if } x_1 \leq x_2) && (2)\\
\mu_i &= \int_{-\infty}^{\infty} x^i \, d\sigma(x) \quad (i = 0,1,2,\ldots) \quad \text{“the } i\text{th moment” of } \sigma(\cdot). && (3)\\
D_n &= \begin{vmatrix} \mu_0 & \mu_1 & \ldots & \mu_n \\ \mu_1 & \mu_2 & \ldots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \ldots & \mu_{2n} \end{vmatrix} \quad \text{the Hankel determinant of order } n. && (4)\\
P_0(x) &= 1 \quad \text{and} \quad P_n(x) = \frac{1}{\sqrt{D_{n-1}D_n}} \begin{vmatrix} \mu_0 & \mu_1 & \ldots & \mu_n \\ \mu_1 & \mu_2 & \ldots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \ldots & \mu_{2n-1} \\ 1 & x & \ldots & x^n \end{vmatrix} \quad (n = 1,2,\ldots) && (5)
\end{aligned}$$

orthonormal polinomials composed by the μ_i sequence.

$$\rho_n(x) = \frac{1}{\sum_{k=0}^{n} |P_k(x)|^2} \quad \text{the radius of the Hellinger circle.} \qquad (6)$$

The μ_i $(i = 0,1,\ldots,2n)$ sequence is said to be a positive sequence if the D_k $(k = 0,1,\ldots,n)$ determinants are positive.

2.2 The moment problem and its solvability

The moment problem plays an essential role in the theory of moments. It can be formulated as follows. Given a sequence of numbers μ_i $(i = 0,1,2,\ldots)$, under what conditions is it possible to find a positive bounded non-decreasing function $\sigma(\cdot)$ such that

$$\mu_i = \int_a^b x^i d\sigma(x) \quad , \qquad \text{for } \ i = 0,1,2,\ldots \quad .$$

Depending on the bounds a and b we distinguish three cases:

- Hamburger moment problem: $a = -\infty, \ b = \infty$,
- Stieltjes moment problem: $a = 0, \ b = \infty$,
- Hausdorff moment problem: $a = 0, \ b = 1$.

In this paper we focus on the first case. (I.e., we do not utilize the information on the bounds of the approximated distributions.)

Theorem 1 *[1] Let $\mu_0, \mu_1, \mu_2, \ldots$ be a sequence of real numbers. The Hamburger moment problem has a solution if and only if $D_n \geq 0, \ n = 0,1,\ldots$.*

Theorem 2 *[15] The solution of the Hamburger moment problem consists of infinite points of increase if and only if $D_n > 0, \qquad n = 0,1,\ldots$.*

Theorem 3 *[15] The solution of the Hamburger moment problem consists of exactly n distinct points of increase if and only if* $D_0 > 0, D_1 > 0 \ldots D_{n-1} > 0, D_n = D_{n+1} = \ldots = 0$. *The moment problem is determined in this case.*

An immediate consequence of Theorem 2 and 3 is that if μ_i are the moments of a distribution and $D_n = 0$ then all the higher Hankel determinants equal to 0 as well ($D_k = 0$ for all $k > n$).

2.3 Finite number of moments

Theorems 1 - 3 are about the infinite series μ_i and D_n, but in practice we have a finite number of moments to deal with. To bound a distribution based on its first $2n+1$ moments[1] we need to find the extreme members of the $\mathfrak{M}_{2n}$ class. At an arbitrary point C, the $\sigma(\cdot)$ distribution with positive sequence of moments $\mu_0, \ldots, \mu_{2n}$ is bounded by $\min_{\breve{\sigma}\in\mathfrak{M}_{2n}} \breve{\sigma}(C) \le \sigma(C) \le \max_{\hat{\sigma}\in\mathfrak{M}_{2n}} \hat{\sigma}(C)$. In the rest of this paper we investigate $\min_{\breve{\sigma}\in\mathfrak{M}_{2n}} \breve{\sigma}(C)$ and $\max_{\hat{\sigma}\in\mathfrak{M}_{2n}} \hat{\sigma}(C)$ in two steps. The first step is to determine the maximum mass the members of $\mathfrak{M}_{2n}$ can have at C, and the second step is to construct a distribution having this maximal mass at C. It will be shown that there is only one distribution composed by $n+1$ discrete points (including the one at C) with maximal mass in C and this distribution characterizes both the lower and the upper bound of the $\mathfrak{M}_{2n}$ class at C.

To simplify the discussion, we always study the bounds at point 0 with a proper transformation of moments. If the original point of interest is C then the moments of the distribution whose evaluated point is shifted to 0 are:

$$\mu_i' = \sum_{k=0}^{i} \binom{i}{k} (-C)^{i-k} \mu_k \tag{7}$$

Without loss of generality, from now on we assume that the point of interest is 0.

2.4 Maximum mass concentrated at 0

Theorem 4 *[1] If the sequence* $\mu_0, \mu_1, \ldots, \mu_{2n}$ *is positive and if* x *is an arbitrary real number then*

$$\max_{\sigma(\cdot)\in\mathfrak{M}_{2n}} \left(\sigma(x^+) - \sigma(x^-)\right) \le \rho_n(x) \ . \tag{8}$$

Theorem 4 gives the meaning of the introduction of $\rho_n(x)$. Indeed, $\rho_n(x)$ defines the maximal mass that can be located at point x given the first $2n+1$ moments. Following a completely different way of thinking than the one in [1], we obtained a different and computationally more effective way to determine maximal mass.

Theorem 5 *If* $\mu_0 = 1, \mu_1, \mu_2, \ldots, \mu_{2n}$ *is a positive sequence of moments of* $\sigma(\cdot)$ *the maximal*

[1] Throughout this paper the first k moments mean the $\mu_0, \mu_1, \ldots, \mu_{k-1}$ sequence.

mass of $\sigma(\cdot)$ *at* 0 *is*

$$p = \frac{\begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix}}{\begin{vmatrix} \mu_2 & \mu_3 & \cdots & \mu_{n+1} \\ \mu_3 & \mu_4 & \cdots & \mu_{n+2} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n+1} & \mu_{n+2} & \cdots & \mu_{2n} \end{vmatrix}}, \tag{9}$$

which means that

$$\rho_n(0) = p. \tag{10}$$

The proof of the theorem is provided in Appendix A.

Our way to obtain p was rather intuitive. A mass at point 0 does contribute to μ_0, but does not contribute to any $\mu_i, i > 0$. We locate a mass at 0 such that the Hankel determinant of the $\mu_0 - p, \mu_1, \ldots, \mu_{2n}$ sequence is just on the limit of positivity. Using $2n + 1$ moments the limit of positivity is reached at

$$\begin{vmatrix} \mu_0 - p & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix} = 0, \tag{11}$$

whose solution is given by Theorem 5.

Theorem 6 *If* $\mu_0, \mu_1, \mu_2, \ldots, \mu_{2n}$ *is a positive sequence of moments then the* $\mu_0 - p, \mu_1, \ldots, \mu_{2n}$ *sequence represents a determined moment problem.*

Proof: Since p is the solution of eq. (11) the order n Hankel determinant associated with the $\mu_0 - p, \mu_1, \mu_2, \ldots, \mu_{2n}$ sequence equals to zero and using Theorem 3 it implies Theorem 6. □

2.5 Maximum difference of the distribution bounds

The following theorem further increases the importance of $\rho_n(0)$ and p by giving an additional meaning to them.

Theorem 7 *[1] If* $\mu_0, \mu_1, \mu_2, \ldots, \mu_{2n}$ *is a positive sequence and* $\sigma_1(\cdot)$ *and* $\sigma_2(\cdot)$ *are members of* $\mathfrak{M}_{2n}$*, then for arbitrary real* x *we have:*

$$\left| \int_{-\infty}^{x^+} d\sigma_1(u) - \int_{-\infty}^{x^-} d\sigma_2(u) \right| \le \rho_n(x) \tag{12}$$

Theorem 7 provides the maximum difference of any two members of $\mathfrak{M}_{2n}$ at x. A direct consequence of Theorem 7 is that the difference of $\min_{\breve{\sigma} \in \mathfrak{M}_{2n}} \breve{\sigma}(x)$ and $\max_{\hat{\sigma} \in \mathfrak{M}_{2n}} \hat{\sigma}(x)$ cannot be larger than $\rho_n(x)$.

Having the difference between the lower and upper bounds it is enough to find one of them. The following theorem suggests a way to place the bounds.

Theorem 8 *If $\mu_0, \mu_1, \mu_2, \ldots, \mu_{2n}$ is a positive sequence, $\sigma(\cdot)$ and $\sigma^*(\cdot)$ are members of $\mathfrak{M}_{2n}$ and $\sigma^*(\cdot)$ is such that it has a mass of size $\rho_n(0) = p$ at 0 then*

$$\int_{-\infty}^{0^-} d\sigma(u) \geq \int_{-\infty}^{0^-} d\sigma^*(u) \tag{13}$$

$$\int_{-\infty}^{0^+} d\sigma(u) \leq \int_{-\infty}^{0^-} d\sigma^*(u) + p \tag{14}$$

The proof of the theorem is provided in Appendix B.

2.6 Construction of a reference distribution

According to Theorem 8 we have the bounds of the $\mathfrak{M}_{2n}$ class of distributions at 0 if we can obtain a reference distribution $\sigma^*(\cdot)$. $\sigma^*(\cdot)$ is such that it has a mass of size p at 0 and the rest of it is determined by the $\mu_0 - p, \mu_1, \ldots, \mu_{2n}$ sequence. Since the $\mu_0 - p, \mu_1, \ldots, \mu_{2n}$ sequence defines a determined moment problem $\sigma^*(\cdot)$ is unique and it has exactly n further points of increase (Theorem 3 and 6).

Let x_i and p_i $(i = 1, \ldots, n)$ denote the points and the associated value of increase of $\sigma^*(\cdot)$ excluding the one at 0, respectively. x_i and p_i are defined by the moments:

$$\mu_0 - p = \sum_{i=1}^{n} p_i \qquad \mu_k = \sum_{i=1}^{n} x_i^k \, p_i \qquad (k = 1, 2, \ldots, 2n-1) \tag{15}$$

These $2n$ equations can be solved in 2 steps.

Theorem 9 *[1] x_i, $(i = 1, \ldots, n)$ are the roots of the $P_n(x)$ polinomial defined by the $\mu_0 - p, \mu_1, \ldots, \mu_{2n}$ sequence.*

Having the $x_1, x_2, \ldots, x_n$ points the associated p_i values can be obtained from equation (15) with $k = 0, 1, \ldots, n-1$. In matrix form it is:

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_n \end{pmatrix} = \begin{pmatrix} \mu_0 - p \\ \mu_1 \\ \mu_2 \\ \vdots \\ \mu_{n-1} \end{pmatrix} \tag{16}$$

This, so-called, Vandermonde system can be solved efficiently using the algorithm provided in [9].

2.7 Special case with only negative or only positive roots

If all roots of $P_n(x)$ (i.e., $x_1, x_2, \ldots, x_n$) are negative or all of them are positive we can bound the $\mathfrak{M}_{2n}$ class without calculating the unknown x_is and p_is. This property can be checked without finding the roots of $P_n(x)$ by the *Liènard–Chipart criterion*:

Theorem 10 *[6] **(Liènard–Chipart)** Let $f(x)$ be a polynomial of order n:*

$$f(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n, \tag{17}$$

and T_i be the following series of determinants:

$$T_0 = a_0, \qquad T_1 = a_1, \qquad T_2 = \begin{vmatrix} a_1 & a_0 \\ a_3 & a_2 \end{vmatrix}, \qquad T_3 = \begin{vmatrix} a_1 & a_0 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{vmatrix}, \qquad \ldots$$

$$T_i = \begin{vmatrix} a_1 & a_0 & 0 & 0 & 0 & \ldots & 0 \\ a_3 & a_2 & a_1 & a_0 & 0 & \ldots & 0 \\ a_5 & a_4 & a_3 & a_2 & a_1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{2i-1} & a_{2i-2} & a_{2i-3} & a_{2i-4} & a_{2i-5} & \ldots & a_i \end{vmatrix} \tag{18}$$

If $a_0 > 0$ then the real part of the roots of $f(x)$ are all negative if and only if $T_0, T_1, \ldots, T_n$ are all positive.

We can also check if all the roots of $f(x)$ have positive real parts using Theorem 10. Let $\hat{f}(x) = f(-x)$. The coefficients of $\hat{f}(x)$ can be expressed by the ones of $f(x)$:

$$\hat{a}_i = (-1)^{n-i} a_i \qquad i = 0, 1, \ldots, n \tag{19}$$

If all roots of $\hat{f}(x)$ have negative real part then all roots of $f(x)$ have positive real part.

In the case when all roots of $P_n(x)$ $(x_1, x_2, \ldots, x_n)$ are negative the bounds are

$$\min_{\breve{\sigma} \in \mathfrak{M}_{2n}} \breve{\sigma}(x) = \mu_0 - p \ , \qquad \max_{\hat{\sigma} \in \mathfrak{M}_{2n}} \hat{\sigma}(x) = \mu_0 \ , \tag{20}$$

and when all roots of $P_n(x)$ $(x_1, x_2, \ldots, x_n)$ are positive the bounds are

$$\min_{\breve{\sigma} \in \mathfrak{M}_{2n}} \breve{\sigma}(x) = 0 \ , \qquad \max_{\hat{\sigma} \in \mathfrak{M}_{2n}} \hat{\sigma}(x) = p \ . \tag{21}$$

3 The algorithm

Based on the above general rules of moments we construct a numerical method in this section. The method provides an upper and a lower bound of a distribution at a given point C based on the first $2n+1$ moments of the distribution, $\mu_0 = 1, \mu_1, \mu_2, \ldots, \mu_{2n}$. The $\mu_0 = 1$ assumption comes from the probability application. The method can cope with any other positive μ_0 as well, but in this case $\rho_n(0)$ needs to be calculated based on its definition (6), because Theorem 5 can not be applied when $\mu_0 \neq 1$.

The main steps of the method are:

1. Checking the number of moments: we need an odd number of moments greater than 1 (including the 0th one).

2. Checking the D_k $(k = 0, 1, \ldots, n)$ sequence:

- If there is a k such that $D_k < 0$ then the μ_is can not be the moments of a non-decreasing distribution function.
- If there is a k such that $D_j = 0$ ($\forall j \geq k$) then the μ_is define a unique discrete distribution of k points and the discrete construction step of the procedure generate this distribution. (I.e., in this case the method provides the exact value of the distribution function.)
- If all D_k are positive (i.e., μ_i is a positive sequence) then there is a set of distributions having these first $2n + 1$ moments and we calculate the lower and upper bounds of this set at point C in the following steps of the algorithm.

3. Transforming the moments such that the point of interest is moved to 0 (eq. (7)).
4. Determining the maximum mass concentrated at 0 (Theorem 5, eq. (9)).
5. Checking if 0 is the leftmost or rightmost point of the reference discrete distribution via Theorem 10.
   - If 0 is the rightmost or leftmost point the bounds are given by eq. (20) and (21), respectively.
   - If $P_n(x)$ has both positive and negative roots then the procedure is continued with the following steps.
6. Determining the roots of $P_n(x)$ which are the points of the reference discrete distribution (Theorem 9).
7. Calculating the weights of the reference discrete distribution (eq. (16)).
8. Determining the lower and upper limits of the distribution at point 0 based on the sum of weights associated with negative roots and the maximum mass at point 0.

A block diagram of the method is presented in Figure 1. We implemented this computation method in Mathematica for getting symbolic results about the bounds and also in C for having fast and portable routine as well. The Mathematica implementation is provided in Appendix C.

4 Symbolic results

In this section we assume that the moments are transformed such that the point of interest is 0. Since Mathematica can perform symbolic calculations the results presented in this section can be obtained by calling the Mathematica routine `DiscreteD` (Appendix C) with moment vector `mom={1,`μ_1`, `μ_2`}` and `mom={1,`μ_1`, `μ_2`, `μ_3`, `μ_4`}`. Mathematica can evaluate the symbolic bounds also for 7 moments (in ~ 5 minutes), but it is too complex to be presented here. The symbolic bounds are not available for more than 9 moments because there is no general symbolic solution available for 5th and higher order equations.

4.1 Estimation based on 3 moments

Having the following sequence of moments: $\mu_0 = 1, \mu_1, \mu_2$ we distinguish two legal cases. When $D_1 = 0$ the moment problem is determined and the moments define a unique distribution.

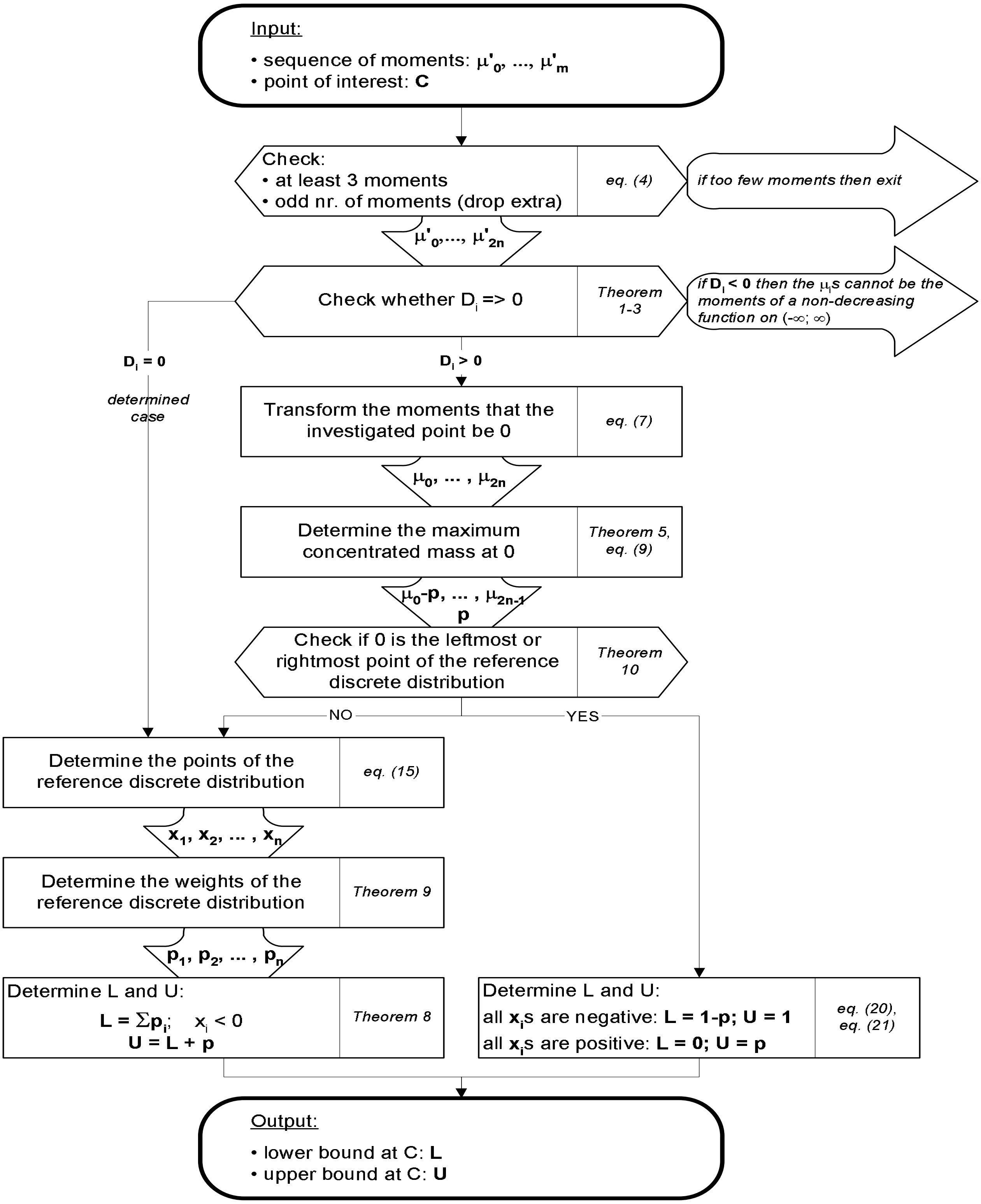


Figure 1: The block structure of the numerical procedure

When $D_1 > 0$ we can bound the limits of all distributions having the same first 2 moments. The case when $D_1 < 0$ can not be obtained by the moments of a real distribution.
The determined case: If $D_1 = \mu_2 - \mu_1^2 = 0$ then the moments determine a discrete distribution with only one point (indeed a deterministic distribution):

$$x_1 = \mu_1, \qquad p_1 = 1. \tag{22}$$

The undetermined case: If $D_1 = \mu_2 - \mu_1^2 > 0$ then we evaluate the bounds based on a discrete distribution with 2 points. One point is at 0 (where we need to bound the distribution) and the other one (x_1) is calculated based on eq. (15) together with the associated probability masses (p and p_1, respectively).

$$p = \frac{\mu_2 - \mu_1^2}{\mu_2} \tag{23}$$

$$x_1 = \frac{\mu_2}{\mu_1} \qquad p_1 = \frac{\mu_1^2}{\mu_2} \tag{24}$$

Note that $\mu_2 > 0$ because $\mu_2 - \mu_1^2 > 0$, thus the sign of x_1 is the same as the sign of μ_1. A degenerate case arises when $\mu_1 = 0$. In this case $p = 1$ and x_1 becomes irrelevant since the associated mass is $p_1 = 0$.

Finally the lower and upper bounds are:

$$\boxed{\begin{array}{lll} L = \frac{\mu_2-\mu_1^2}{\mu_2} & U = 1 & \text{if } \mu_1 < 0 \\ L = 0 & U = 1 & \text{if } \mu_1 = 0 \\ L = 0 & U = \frac{\mu_2-\mu_1^2}{\mu_2} & \text{if } \mu_1 > 0 \end{array}} \tag{25}$$

4.2 Estimation based on 5 moments

Having the moments: $\mu_0 = 1, \mu_1, \mu_2, \mu_3, \mu_4$, we discuss the two meaningful cases the determined one when $D_1 > 0$ and $D_2 = 0$ and the undetermined one when $D_1 > 0$ and $D_2 > 0$.
The determined case

$$D_1 = \begin{vmatrix} 1 & \mu_1 \\ \mu_1 & \mu_2 \end{vmatrix} > 0 \quad \text{and} \quad D_2 = \begin{vmatrix} 1 & \mu_1 & \mu_2 \\ \mu_1 & \mu_2 & \mu_3 \\ \mu_2 & \mu_3 & \mu_4 \end{vmatrix} = 0 \tag{26}$$

This determined moment problem defines a discrete distribution with 2 points by Theorem 3. The points and weights are calculated based on eq. (15). Let $r = \sqrt{-3\mu_1^2\mu_2^2 + 4\mu_2^3 + 4\mu_1^3\mu_3 - 6\mu_1\mu_2\mu_3 + \mu_3^2}$.

$$x_1 = \frac{\mu_1\mu_2 - \mu_3 + r}{2\mu_1^2 - 2\mu_2} \tag{27}$$

$$p_1 = \frac{1}{2} + \frac{2\mu_1^3 - 3\mu_1\mu_2 + \mu_3}{2r} \tag{28}$$

$$x_2 = \frac{\mu_1\mu_2 - \mu_3 - r}{2\mu_1^2 - 2\mu_2} \tag{29}$$

$$p_2 = \frac{1}{2} - \frac{2\mu_1^3 - 3\mu_1\mu_2 + \mu_3}{2r} \tag{30}$$

The undetermined case When $D_1 > 0$ and $D_2 > 0$

$$p = \frac{-\mu_2^3 + 2\mu_1\mu_2\mu_3 - \mu_3^2 - \mu_1^2\mu_4 + \mu_2\mu_4}{\mu_2\mu_4 - \mu_3^2}, \tag{31}$$

$$x_1 = \frac{\mu_2\mu_3 - \mu_1\mu_4 - q}{2(\mu_2^2 - \mu_1\mu_3)}, \tag{32}$$

$$p_1 = \frac{-{\mu_2}^4\mu_3 + 2{\mu_1}^2{\mu_3}^3 + 3\mu_1{\mu_2}^3\mu_4 - 5\mu_1^2\mu_2\mu_3\mu_4 + {\mu_1}^3{\mu_4}^2 - q(\mu_2^3 - 2\mu_1\mu_2\mu_3 + \mu_1^2\mu_4)}{2q(\mu_3^2 - \mu_2\mu_4)}, \tag{33}$$

$$x_2 = \frac{\mu_2\mu_3 - \mu_1\mu_4 + q}{2(\mu_2^2 - \mu_1\mu_3)} \tag{34}$$

$$p_2 = -\frac{\mu_2^2 - \mu_1\mu_3}{q}\left(-\mu_1 - \frac{(\mu_2^3 - 2\mu_1\mu_2\mu_3 + \mu_1^2\mu_4)(-\mu_2\mu_3 + \mu_1\mu_4 + q)}{2(\mu_2^2 - \mu_1\mu_3)(-\mu_3^2 + \mu_2\mu_4)}\right), \tag{35}$$

where $q = \sqrt{(-\mu_2\mu_3 + \mu_1\mu_4)^2 - 4(\mu_2^2 - \mu_1\mu_3)(\mu_3^2 - \mu_2\mu_4)}$. Note that $p + p_1 + p_2 = 1$.

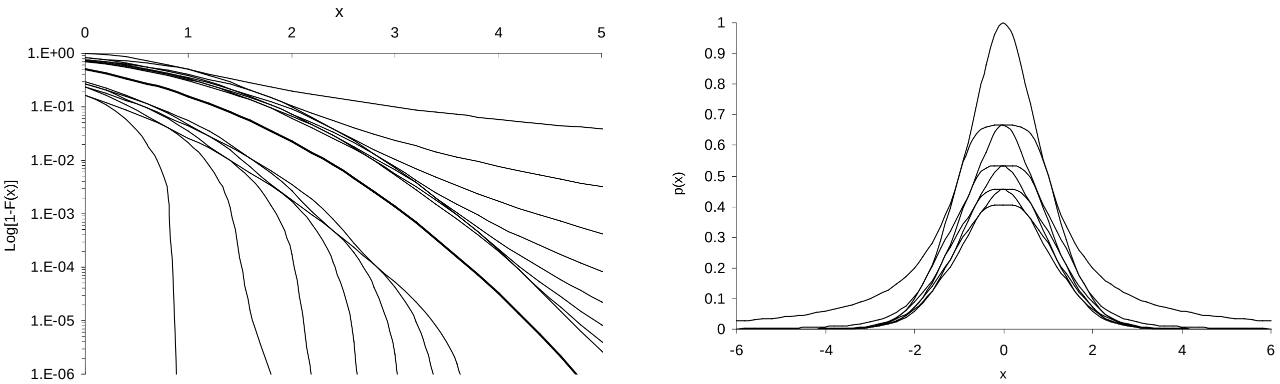
With 5 moments, it is much harder to formulate the bounds conditioned directly on the moments, but using the calculated point and weight sequences the bounds are:

$L = p1 + p2$	$U = 1$	if $x_1 < 0, x_2 < 0$
$L = p1$	$U = p_1 + p$	if $x_1 < 0, x_2 > 0$
$L = 0$	$U = p$	if $x_1 > 0, x_2 > 0$

(36)

5 Numerical results

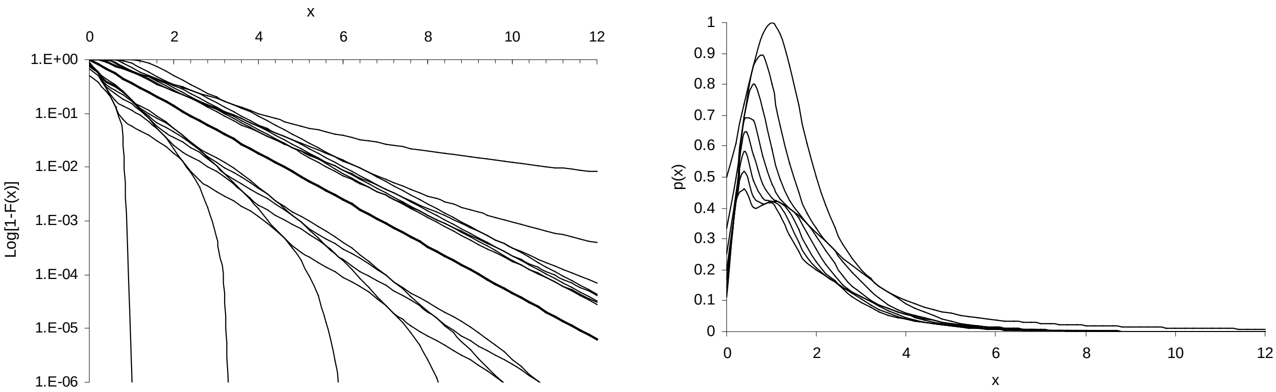
The application of the distribution estimation method for the analysis of large Markov reward models is demonstrated in [11]. In those examples we can not relate the results with any reference. In this section we apply the distribution estimation method for known distributions which allows us to relate the bounds with the exact values of the distribution function.



a) Bounds of the ccdf. b) Distance of upper and lower bounds (p).

Figure 2: Bounding the standard normal distribution based on $3, 5, \ldots, 17$ moments

We chose 4 distributions to evaluate based on their moments, the standard normal (mean= 0, var.= 1), the exponential (mean= 1), the Poisson (mean= 5) and the continuous uniform distribution between 0 and 1. We generated 17 moments of these distributions (the 0th moment is 1 in each cases) and estimate the distribution functions at several points based on these moments. In Figures 2a, 3a and 4a the ccdf and their bounds are depicted using logarithmic scale. While the upper and lower bounds should tend to the same limit (0, in case of ccdf) as $x \to \infty$, the figures indicate a "visually" non-decreasing region between the bounds. This misleading pictures

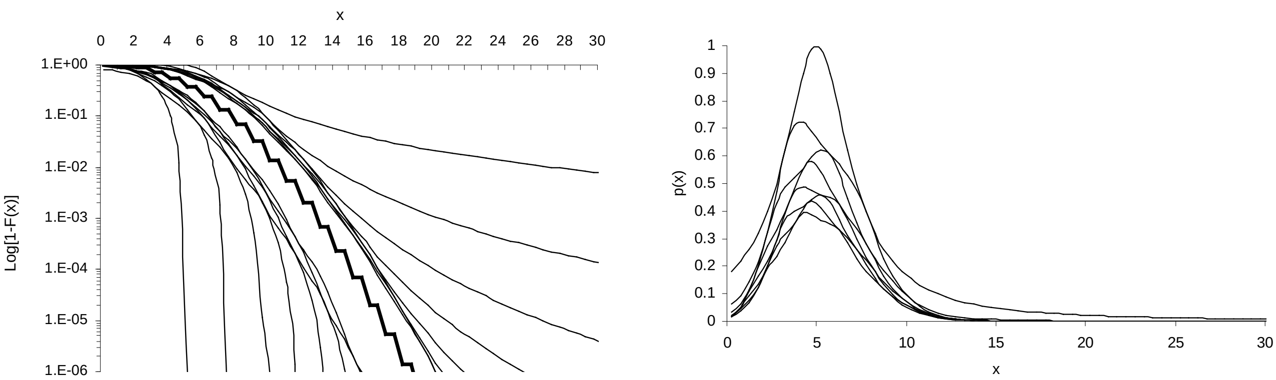


a) Bounds of the ccdf. b) Distance of upper and lower bounds (p).

Figure 3: Bounding the exponential distribution with mean 1 based on $3, 5, \dots, 17$ moments

mean that the difference of the bounds (the error of the estimation) decreases exponentially. In the figures, we applied logarithmic scale to emphasize another important property of the estimation. The relative error of the estimation remains more or less constant in a wide range, where the width of the range depends on the number of moments. I.e., at the extreme values of the distribution increasing the number of moments does not improve the bounds significantly, but extends the range where the bounds maintain the given level of relative error.

Figure 5a depicts the ccdf of the uniform distribution and its bounds with linear scale. This figure demonstrates the ability of our (Hamburger moment problem based) estimation in bounding distributions with finite support. The bounds based on more than 3 moments vanishes quickly at the limits of the distribution.

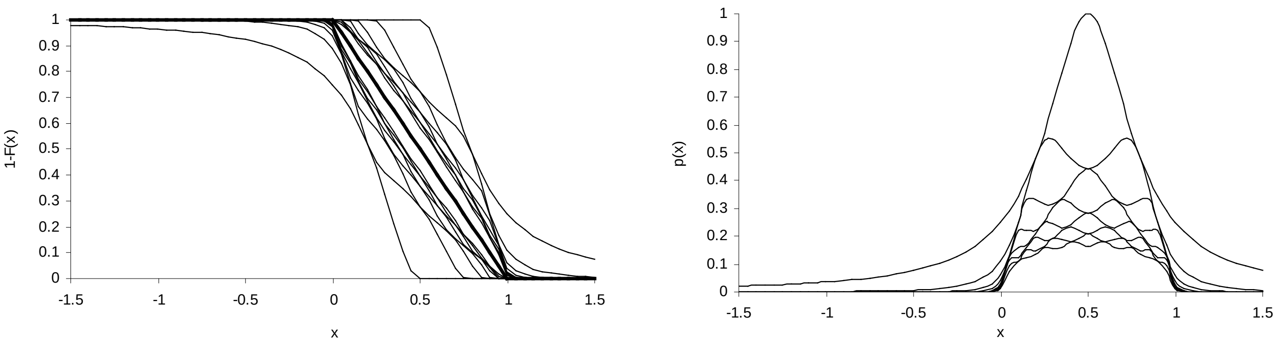


a) Bounds of the ccdf. b) Distance of upper and lower bounds (p).

Figure 4: Bounding the Poisson distribution with mean 5 based on $3, 5, \dots, 17$ moments

The lower bound of the ccdf reaches 0 at the point where all roots of (15) (with the moments of the shifted distribution) are negative. That is where the lower bounds break down in the figures with logarithmic scale. Beyond this limit the upper bound of the ccdf is p. As it is visible in the figures with logarithmic scale the upper bounds tend to an exponentially decreasing asymptotic limit which is a function of the number of moments. Indeed asymptotic limit is x^{-2p} as it is reported in [7].

Figures 2b, 3b, 4b and 5b depicts the difference of the bounds (i.e., the maximal mass p) as a function of x. All $p(x)$ functions has a maximum around the mean of the distribution and they decrease towards the extreme values in both directions. The figures obviously verify the middle line of (25) since the 3-moment curves reach 1 at the mean in all cases.



a) Bounds of the ccdf. b) Distance of upper and lower bounds (p).

Figure 5: Bounding the $(0,1)$ continuous uniform distribution based on $3, 5, \dots, 17$ moments

All figures indicate points where the bounds or the $p(x)$ curves calculated based on $2n-1$ and $2n+1$ moments coincide. Indeed the number of coinciding points is n. The reason of this property is associated with the layout of the reference discrete distribution with respect to the point of interest. When we bound extreme low (high) values of a distribution all roots of the reference distributions are on the right (left) of the point of interest. Between the extreme values there is a continuous transition of roots from one side to the other of the point of interest. The bounds based on $2n+1$ moments are calculated from a discrete distribution of n points (different from the point of interest). During the continuous transition of the points of the discrete distribution from one side to the other of the reference point there are n cases when a root coincide with the point of interest. In these cases the last two moments do not contain additional information for bounding the distribution and so the bound curves and the $p(x)$ curves coincide. Figures 2b, 3b, 4b and 5b also indicate that the maximal mass at the point of interest is a decreasing function of the number of moments and equality can occur only in the mentioned extreme points.

A crucial issue of the proposed distribution estimation method is the numerical instability. It is important to emphasize that all steps of the numerical procedure very much depends on the accuracy of the calculation. For example, the standard floating point precision of Mathematica provides negative Hankel determinant, D_{17}, for the $\mu_i = i!$ moment series (which is a positive sequence, it is the moment series of the exponential distribution) due to numerical errors. This example indicates that even the simplest step of the procedure can fail with "theoretically correct" moments series. Unfortunately, we do not always have "theoretically correct" moment sequence. If the moments are calculated by other complex computational methods like the ones for the moment analysis of large MRMs the resulted moments can accumulate the numerical errors of the preceding calculations. In these cases it is always recommended to check the positivity of the moment sequence in advance of the calculation.

The complexity as well as the accuracy of calculating p is practically identical with the calculation of the Hankel determinant. This fact suggests an easy practical approach to check the validity of p. We consider p to be valid as long as both, the numerator and the denominator of (9) are positive.

Finding the roots of the order n polinomial (Theorem 9) and the solution of the linear system (16) is more complex and more sensitive to numerical errors than calculating p. The majority of standard numerical packages can indicate ill conditioning or near-singularity which warn the presence of numerical problems. According to Theorem 3 the $P_n(x)$ polinomial has n real roots. The presence of complex roots also suggests numerical problems. According to our experience in the majority of the cases our numerical procedure provided trustworthy results (which follow the general trends of stable results) even in ill conditioned cases, but we can not evaluate the numerical error of these cases. We never experienced a case with incorrect bound in the sense that the calculated lower (upper) bound was higher (lower) than the exact value of the distribution at the given point. In case of serious numerical problems the number of moments used for the estimation has to be reduced by two.

The most stable distribution bounds can be obtained for extreme values, when the Liènard–Chipart criterion is fulfilled (Theorem 10). The complexity and the precision of calculating of the T_i determinants is similar to the one of calculating p and the search of the roots of $P_n(x)$ and the solution of the linear system is skipped in this case. Hence bounds provided by the numerical method are most trustworthy in these cases and the only possible check of precision is the positivity of the Hankel determinants.

6 Conclusion

This paper presents a moment-based distribution estimation procedure for bounding the distribution of reward measures of large MRMs. The procedure calculates a discrete reference distribution of $n+1$ points with maximal mass at the point of interest whose first $2n+1$ moments are identical with the sequence of known moments. The bounds of the unknown distribution are calculated from the left and right limits of this reference distribution at the point of interest. The paper presents the proof of this approach as well.

The numerical properties of the proposed method is investigated via estimations of known distributions. Interesting features of the distribution bounds are presented together with practical considerations on numerical stability.

A Proof of Theorem 5

To prove the theorem we need the following lemma:

Lemma 11 *[14]* ***(Jacobi)*** *Let A_{ij} denote the order $n-1$ minors of the $n \times n$ quadratic matrix A. Let $A\begin{pmatrix} i_1 \dots i_p \\ i_1 \dots i_p \end{pmatrix}$ denote the order $n-p$ minors of A, which can be obtained by deleting the rows and columns $i_1, i_2, \dots, i_p$. For these quantities the following equation holds:*

$$\begin{vmatrix} A_{i_1 i_1} & A_{i_1 i_2} \dots & A_{i_1 i_p} \\ A_{i_2 i_1} & A_{i_2 i_2} \dots & A_{i_1 i_p} \\ \vdots & \vdots \quad \ddots & \vdots \\ A_{i_n i_1} & A_{i_n i_2} \dots & A_{i_p i_p} \end{vmatrix} = |A|^{p-1} A \begin{pmatrix} i_1 \dots i_p \\ i_1 \dots i_p \end{pmatrix} \tag{37}$$

Proof: We prove the theorem by induction. Let us substitute the definitions of $\rho_n(z)$ and p into (10):

$$\frac{\begin{vmatrix} \mu_0 & \mu_1 & \dots & \mu_n \\ \mu_1 & \mu_2 & \dots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \dots & \mu_{2n} \end{vmatrix}}{\begin{vmatrix} \mu_2 & \mu_3 & \dots & \mu_{n+1} \\ \mu_3 & \mu_4 & \dots & \mu_{n+2} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n+1} & \mu_{n+2} & \dots & \mu_{2n} \end{vmatrix}} = \frac{1}{\sum_{k=0}^{n} |P_k(0)|^2} \tag{38}$$

1. For $n = 1$:
   The left hand side is:

$$\frac{\begin{vmatrix} 1 & \mu_1 \\ \mu_1 & \mu_2 \end{vmatrix}}{\mu_2} = \frac{\mu_2 - \mu_1^2}{\mu_2}\ , \tag{39}$$

   and the right hand side is:

$$\begin{aligned} \rho_n(0) &= \frac{1}{1 + |P_n(0)|^2} = \frac{1}{1 + \left| \frac{1}{\sqrt{\begin{vmatrix} 1 & \mu_1 \\ \mu_1 & \mu_2 \end{vmatrix} |1|}} \begin{vmatrix} 1 & \mu_1 \\ 1 & 0 \end{vmatrix} \right|^2} = \\ &= \frac{1}{1 + \left| \frac{1}{\sqrt{\mu_2 - \mu_1^2}} (-\mu_1) \right|^2} = \frac{1}{1 + \frac{\mu_1^2}{\mu_1 \mu_2 - \mu_1^2}} = \\ &= \frac{1}{\frac{\mu_2 - \mu_1^2 + \mu_1^2}{\mu_2 - \mu_1^2}} = \frac{\mu_2 - \mu_1^2}{\mu_2}\ . \end{aligned} \tag{40}$$

2. Let us assume that (38) is true for an arbitrary $n = k$ that is

$$\frac{\begin{vmatrix} \mu_0 & \mu_1 & \dots & \mu_k \\ \mu_1 & \mu_2 & \dots & \mu_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_k & \mu_{k+1} & \dots & \mu_{2k} \end{vmatrix}}{\begin{vmatrix} \mu_2 & \mu_3 & \dots & \mu_{k+1} \\ \mu_3 & \mu_4 & \dots & \mu_{k+2} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{k+1} & \mu_{k+2} & \dots & \mu_{2k} \end{vmatrix}} = \frac{1}{1 + (P_1(0))^2 + \ldots + (P_k(0))^2} \tag{41}$$

   We can write $(P_i(0))^2$ instead of $|P_i(0)|^2$ because it can be seen from (5) that $P_i(x)$ is real for every real x.

3. Now we show that (38) is true for $n = k+1$. The reciprocal of (38) is:

$$\frac{\begin{vmatrix} \mu_2 & \mu_3 & \dots & \mu_{k+2} \\ \mu_3 & \mu_4 & \dots & \mu_{k+3} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{k+2} & \mu_{k+3} & \dots & \mu_{2k+2} \end{vmatrix}}{\begin{vmatrix} \mu_0 & \mu_1 & \dots & \mu_{k+1} \\ \mu_1 & \mu_2 & \dots & \mu_{k+2} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{k+1} & \mu_{k+2} & \dots & \mu_{2k+2} \end{vmatrix}} = 1 + (P_1(0))^2 + \dots + (P_{k+1}(0))^2 \ . \tag{42}$$

Transforming the right hand side of the equation above we get:

$$\begin{aligned} 1 \quad & +(P_1(0))^2 + \dots + (P_k(0))^2 + (P_{k+1}(0))^2 = \\ = \quad & 1 + (P_1(0))^2 + \dots + (P_k(0))^2 + \frac{1}{D_{k+1}D_k} \begin{vmatrix} \mu_1 & \mu_2 & \dots & \mu_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_k & \mu_{k+1} & \dots & \mu_{2k+1} \\ 1 & 0 & \dots & 0 \end{vmatrix}^2 = \\ = \quad & 1 + \dots + (P_k(0))^2 + \frac{(-1)^{2k+2} \begin{vmatrix} \mu_1 & \dots & \mu_{k+1} \\ \vdots & \ddots & \vdots \\ \mu_{k+1} & \dots & \mu_{2k+1} \end{vmatrix}^2}{\begin{vmatrix} 1 & \dots & \mu_{k+1} \\ \vdots & \ddots & \vdots \\ \mu_{k+1} & \dots & \mu_{2k+2} \end{vmatrix} \begin{vmatrix} 1 & \dots & \mu_k \\ \vdots & \ddots & \vdots \\ \mu_k & \dots & \mu_{2k} \end{vmatrix}} \end{aligned} \tag{43}$$

We can write $\frac{1}{D_{k+1}D_k}$ instead of $\frac{1}{|D_{k+1}D_k|}$, because $\mu_0, \mu_1, \dots, \mu_{2n}$ is a positive sequence and so $D_i > 0$ for all $i = 1, \dots, n$. On the basis of (41) we can substitute $1 + \dots + (P_k(0))^2$ and this equals to the left hand side of (42):

$$\frac{\begin{vmatrix} \mu_2 & \dots & \mu_{k+1} \\ \vdots & \ddots & \vdots \\ \mu_{k+1} & \dots & \mu_{2k} \end{vmatrix}}{\begin{vmatrix} 1 & \dots & \mu_k \\ \vdots & \ddots & \vdots \\ \mu_k & \dots & \mu_{2k} \end{vmatrix}} + \frac{\begin{vmatrix} \mu_1 & \dots & \mu_{k+1} \\ \vdots & \ddots & \vdots \\ \mu_{k+1} & \dots & \mu_{2k+1} \end{vmatrix}^2}{\begin{vmatrix} 1 & \dots & \mu_{k+1} \\ \vdots & \ddots & \vdots \\ \mu_{k+1} & \dots & \mu_{2k+2} \end{vmatrix} \begin{vmatrix} 1 & \dots & \mu_k \\ \vdots & \ddots & \vdots \\ \mu_k & \dots & \mu_{2k} \end{vmatrix}} = \frac{\begin{vmatrix} \mu_2 & \dots & \mu_{k+2} \\ \vdots & \ddots & \vdots \\ \mu_{k+2} & \dots & \mu_{2k+2} \end{vmatrix}}{\begin{vmatrix} 1 & \dots & \mu_{k+1} \\ \vdots & \ddots & \vdots \\ \mu_{k+1} & \dots & \mu_{2k+2} \end{vmatrix}} \tag{44}$$

Hence we have to prove that:

$$\begin{vmatrix} \mu_2 & \dots & \mu_{k+1} \\ \vdots & \ddots & \vdots \\ \mu_{k+1} & \dots & \mu_{2k} \end{vmatrix} \begin{vmatrix} 1 & \dots & \mu_{k+1} \\ \vdots & \ddots & \vdots \\ \mu_{k+1} & \dots & \mu_{2k+2} \end{vmatrix} + \begin{vmatrix} \mu_1 & \dots & \mu_{k+1} \\ \vdots & \ddots & \vdots \\ \mu_{k+1} & \dots & \mu_{2k+1} \end{vmatrix}^2 = \begin{vmatrix} \mu_2 & \dots & \mu_{k+2} \\ \vdots & \ddots & \vdots \\ \mu_{k+2} & \dots & \mu_{2k+2} \end{vmatrix} \begin{vmatrix} 1 & \dots & \mu_k \\ \vdots & \ddots & \vdots \\ \mu_k & \dots & \mu_{2k} \end{vmatrix} \tag{45}$$

Let $p = 2$, $i_1 = 1$, $i_2 = k+1$ and $M = D_{k+1}$. By Lemma 11 it can be seen using (37) that:

$$\begin{vmatrix} M_{1,1} & M_{1,k+1} \\ M_{k+1,1} & M_{k+1,k+1} \end{vmatrix} = |M|\, M \begin{pmatrix} 1 & k+1 \\ 1 & k+1 \end{pmatrix} \tag{46}$$

Expanding the determinant on the left hand side:

$$M_{1,1}\, M_{k+1,k+1} - M_{k+1,1}\, M_{1,k+1} = |M|\; M \begin{pmatrix} 1 & k+1 \\ 1 & k+1 \end{pmatrix} \tag{47}$$

and $M_{1,k+1} = M_{k+1,1}$ because M is a symmetric matrix. Rearranging the equation we get:

$$|M|\; M \begin{pmatrix} 1 & k+1 \\ 1 & k+1 \end{pmatrix} + (M_{k+1,1})^2 = M_{1,1}\, M_{k+1,k+1} \tag{48}$$

Using the definitions of M_{ii} and $M \begin{pmatrix} 1 & k+1 \\ 1 & k+1 \end{pmatrix}$ equation (48) can be written as:

$$\begin{vmatrix} \mu_2 & \dots & \mu_{k+1} \\ \vdots & \ddots & \vdots \\ \mu_{k+1} & \dots & \mu_{2k} \end{vmatrix} \begin{vmatrix} 1 & \dots & \mu_{k+1} \\ \vdots & \ddots & \vdots \\ \mu_{k+1} & \dots & \mu_{2k+2} \end{vmatrix} + \begin{vmatrix} \mu_1 & \dots & \mu_{k+1} \\ \vdots & \ddots & \vdots \\ \mu_{k+1} & \dots & \mu_{2k+1} \end{vmatrix}^2 = \begin{vmatrix} \mu_2 & \dots & \mu_{k+2} \\ \vdots & \ddots & \vdots \\ \mu_{k+2} & \dots & \mu_{2k+2} \end{vmatrix} \begin{vmatrix} 1 & \dots & \mu_k \\ \vdots & \ddots & \vdots \\ \mu_k & \dots & \mu_{2k} \end{vmatrix}, \tag{49}$$

which is identical with (45) that had to be proven. □

B Proof of Theorem 8

Proof: Using the $\sigma_1(x) := \sigma^*(x)$ and the $\sigma_2(x) := \sigma(x)$ substitutions by Theorem 7 we have:

$$\left| \int_{-\infty}^{0^+} d\sigma^*(u) - \int_{-\infty}^{0^-} d\sigma(u) \right| \le p \,. \tag{50}$$

Resolving the absolute value sign it is:

$$\int_{-\infty}^{0^+} d\sigma^*(u) - p \le \int_{-\infty}^{0^-} d\sigma(u) \le \int_{-\infty}^{0^+} d\sigma^*(u) + p. \tag{51}$$

By the construction of $\sigma^*(x)$ it follows that

$$\int_{-\infty}^{0^+} d\sigma^*(u) = \int_{-\infty}^{0^-} d\sigma^*(u) + p \,. \tag{52}$$

Substituting it to the leftmost inequality we get:

$$\int_{-\infty}^{0^-} d\sigma^*(u) \le \int_{-\infty}^{0^-} d\sigma(u) \,, \tag{53}$$

which is (13). Now using the $\sigma_1(x) := \sigma(x)$ and the $\sigma_2(x) := \sigma^*(x)$ substitutions Theorem 7 gives

$$\left| \int_{-\infty}^{0^+} d\sigma(u) - \int_{-\infty}^{0^-} d\sigma^*(u) \right| \leq p \tag{54}$$

Resolving the absolute value sign we have:

$$\int_{-\infty}^{0^-} d\sigma^*(u) - p \leq \int_{-\infty}^{0^+} d\sigma(u) \leq \int_{-\infty}^{0^-} d\sigma^*(u) + p \tag{55}$$

whose rightmost inequality is (14). □

C Mathematica code of discrete distribution construction

```
DiscreteD[mom_] :=
(* Input: vector of moments (with point of interest = 0) *)
(* Output: Points and weights of the reference discrete distribution *)
Module[{n, k, j, p, mx, xv, root},

    (* number of given moments *)
    n = Dimensions[mom][[1]];
    k = (n - 1)/2;

    (* warning if wrong number of moments *)
    If[(n < 3) Or (Mod[n - 1, 2] != 0),
        Print["Few or even number of moments!!!"]];

    (* calculating the maximum mass at 0 *)
    p = Det[Hankelmx[mom]] / Det[ Hankelmx[Take[mom, -(n - 2)]] ];
    Print["p=", p];

    (* forming P_n(x) *)
    mx = Hankelmx[mom];
    mx[[1, 1]] = mx[[1, 1]] - p;
    Do[mx[[k + 1, j]] = x^(j - 1), {j, 1, k + 1}];

    (* points of the discrete distribution *)
    root = Solve[ Det[mx] == 0, x ];
    xv = Table[0, {k}];
    Do[xv[[j]] = root[[j]][[1]][[2]], {j, 1, k}];
    Print["roots=", xv];

    (* forming the Vandermonde-system *)
    mx = Table[0, {k}, {k}];
    Do[mx[[j]] = xv^(j - 1), {j, 1, k}];
    Do[xv[[j]] = mom[[j]], {j, 1, k}];
    xv[[1]] -= p;
```

```
    (* weights of the discrete distribution *)
    Print["weights=", pv = LinearSolve[mx, xv] ]
];

Hankelmx[mom_]:=Module[{i,j,n},
  n=(Dimensions[mom][[1]] - 1)/2;
  Table[mom[[i+j-1]],{i,n+1},{j,n+1}] ];
```

References

[1] N. I. Akhiezer. *The classical moment problem and some related questions in analysis.* Hafner publishing company, New York, 1965. (translation of Н. И. Ахиезер: Классическая Проблема Моментов и Некоторые Вопросы Анализа, published by Государственное Издательство Физико-Математической Литературы, Moscow, 1961).

[2] E. de Souza e Silva and H.R. Gail. Calculating cumulative operational time distributions of repairable computer systems. *IEEE Transactions on Computers*, C-35:322–332, 1986.

[3] E. de Souza e Silva and R. Gail. An algorithm to calculate transient distributions of cummulative rate and impulse based reward. *Commun. in Statist. – Stochastic Models*, 14(3):509–536, 1998.

[4] L. Donatiello and V. Grassi. On evaluating the cumulative performance distribution of fault-tolerant computer systems. *IEEE Transactions on Computers*, 1991.

[5] B.R. Iyer, L. Donatiello, and P. Heidelberger. Analysis of performability for stochastic models of fault-tolerant systems. *IEEE Transactions on Computers*, C-35:902–907, 1986.

[6] G. A. Korn and T. M. Korn. *Mathematical handbook for scientists and engineers.* Dover, 2000.

[7] B. G. Lindsay and P. Bassak. Moments determine the tail of a distribution (but not much else). *The American Statistician*, 54(4):248–251, 2000.

[8] H. Nabli and B. Sericola. Performability analysis: a new algorithm. *IEEE Transactions on Computers*, 45:491–494, 1996.

[9] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, 1993. http://lib-www.lanl.gov/numerical/bookcpdf.html.

[10] S. Rácz. *Numerical analysis of communication systems through Markov reward models.* PhD thesis, Technical University of Budapest, 2000.

[11] S. Rácz, Á. Tari, and M. Telek. MRMSolve: Distribution estimation of large Markov reward models. In *Tools 2002*, pages 72–81, London, England, April 2002. Springer, LNCS 2324.

[12] S. Rácz and M. Telek. Performability analysis of Markov reward models with rate and impulse reward. In M. Silva B. Plateau, W. Stewart, editor, *Int. Conf. on Numerical solution of Markov chains*, pages 169–187, Zaragoza, Spain, 1999.

[13] S. Rácz and M. Telek. Mission time analysis of large dependable systems. In *IPDS 2000*, pages 13–22, Chicago, IL, USA, March 2000. IEEE CS Press.

[14] P. Rózsa. *Lineális algebra és alkalmazásai.* Tankönyvkiadó, 1991. in Hungarian.

[15] Aldo Tagliani. Existence and stability of a discrete probability distribution by maximum entropy approach. *Applied Mathematics and Computation*, 110:105–114, 2000.

[16] M. Telek, A. Pfening, and G. Fodor. An effective numerical method to compute the moments of the completion time of Markov reward models. *Computers and mathematics with applications*, 36:8:59–65, 1998.

[17] M. Telek and S. Rácz. Numerical analysis of large Markovian reward models. *Performance Evaluation*, 36&37:95–114, Aug 1999.