# On the Properties of Moments of Matrix Exponential Distributions and Matrix Exponential Processes*

Levente Bodrog[2], András Horváth[1], Miklós Telek[2]

[1] Dipartimento di Informatica, Università di Torino, Italy

[2] Dept. of Telecommunications, Technical University of Budapest, Hungary

E-mail: horvath@di.unito.it, {bodrog,telek}@webspn.hit.bme.hu

**Abstract**

In this paper we provide properties of moments of matrix exponential distributions and joint moments of matrix exponential processes. Based on the provided properties, an algorithm is presented to compute any finite dimensional moments of these processes based on a set of required (low order) moments. This algorithm does not require the computation of any representation of the given process. We present some related examples to demonstrate the potential use of the properties of moments.

*Keywords:* Matrix exponential process, Markov arrival process, Matrix exponential distribution, phase type distribution, moment matching, inter-arrival time distribution, lag-correlation.

## 1 Introduction

Phase type (PH) distributions and Markovian arrival processes (MAP) are simple stochastic models that enjoy a simple stochastic interpretation based on Markov chains. They are widely used in traffic engineering because efficient numerical techniques are available for the solution of queueing models with PH distributions and MAPs [6, 10, 11]. Matrix exponential (ME) distributions and processes (MEP) are more general stochastic models than PH distributions and MAPs [2, 5]. They do not have a simple stochastic interpretation and most of the methods applied to Markovian models cannot be directly applied to them. Still, in recent years we have seen a growing interest in these models and several results for queueing systems have been presented [1, 3, 8, 4].

To apply PH, ME distributions, MAPs and MEPs in stochastic models, we need good understanding of their properties. In this paper we investigate issues regarding the moments of these distributions and processes. The paper is organised as follows. Section 2 presents results for PH and ME distributions, while Section 3 deals with MAPs and MEPs. We formulate the results for ME distributions and MEPs, but they are directly applicable to PH distributions and MAPs as well. We conclude with Section 4.

## 2 Matrix exponential and Phase type distributions

### 2.1 Basic definitions

Let $X$ be a continuous non-negative random variable with cumulative distribution function (cdf)

$$F(t) = Pr(X < t) = 1 - \alpha e^{At}\mathbb{1} \ , \tag{1}$$

where row vector $\alpha$ is referred to as the initial vector, square matrix $A$ as the generator and $\mathbb{1}$ as the closing vector. Without loss of generality (see [7]), we assume that the closing vector, $\mathbb{1}$, is a column vector of ones. When the cardinality of the vectors and the square matrix is $n$, $X$ is referred to as

*This work is partially supported by the Italian-Hungarian bilateral R&D programme, by OTKA grant n. T-34972, by MIUR through PRIN project Famous and by EEC project Crutial.



Dagstuhl Seminar Proceedings 07461
Numerical Methods for Structured Markov Chains
http://drops.dagstuhl.de/opus/volltexte/2008/1394

an order $n$ *matrix exponential distribution* (ME(n)). As $X$ is a continuous random variable, it has no probability mass at zero, i.e., $\alpha \mathbb{1} = 1$. The density, the Laplace transform and the moments of $X$ can be computed as

$$f(t) = \alpha e^{At}(-A)\mathbb{1} \ , \tag{2}$$

$$f^*(s) = E(e^{-sX}) = \alpha(sI - A)^{-1}(-A)\mathbb{1} \ , \tag{3}$$

$$\mu_n = E(X^n) = n!\alpha(-A)^{-n}\mathbb{1} \ . \tag{4}$$

In general, the elements of $\alpha$ and $A$ may be arbitrary real numbers. If $\alpha$ is a probability vector and $A$ is the generator matrix of a continuous-time Markov chain, then $X$ is an order n *phase type distributed* (PH(n)). $\alpha$ is a probability vector when $\alpha_i \geq 0$ $(\forall i = 1, \ldots, n)$ and $\alpha \mathbb{1} = 1$ and matrix $A$ is a transient Markovian generator when $A_{ii} < 0$ $(\forall i = 1, \ldots, n)$, $A_{ij} \geq 0$ $(\forall i, j = 1, \ldots, n, \ i \neq j)$ , $A\mathbb{1} \leq \mathbf{0}$, $A\mathbb{1} \neq \mathbf{0}$.

To ensure that $f(t)$ in (2) is a density function, $A$ has to fulfill the necessary but not sufficient condition that the real part of its eigenvalues are negative (consequently $A$ is non-singular).

The vector together with the square matrix, $(\alpha, A)$, is referred to as the representation of the ME (PH) distribution. In general, the $(\alpha, A)$ representation is not unique.

**Definition 1** *An $(\alpha, A)$ representation is* non-redundant *if the cardinality of vector $\alpha$ and square matrix $A$ is equal to the degree of the denominator of $f^*(s)$ (which is a rational function of $s$).*

Throughout the paper we assume that the representations of ME (PH) distributions are non-redundant.

## 2.2 Doubly infinite Hankel matrix

We define the doubly infinite Hankel matrix as

$$R = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \iddots \\ \cdots & r_{-4} & r_{-3} & r_{-2} & r_{-1} & r_0 & \cdots \\ \cdots & r_{-3} & r_{-2} & r_{-1} & r_0 & r_1 & \cdots \\ \cdots & r_{-2} & r_{-1} & r_0 & r_1 & r_2 & \cdots \\ \cdots & r_{-1} & r_0 & r_1 & r_2 & r_3 & \cdots \\ \cdots & r_0 & r_1 & r_2 & r_3 & r_4 & \cdots \\ \iddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

where $r_i = \alpha(-A)^{-i}\mathbb{1}$. For $i \geq 0$, $r_i$ is the $i$th reduced moment, $E(X^n)/n!$, while for $i < 0$, $r_i$ is related to the $i$th derivative of the cdf, since the $i$th derivative of the cdf of the ME distribution given by $\alpha$ and $A$ at $t = 0$ is

$$\left. \frac{d^i F(t)}{dt^i} \right|_{t=0} = -\alpha A^i \mathbb{1} = (-1)^{i+1}\alpha(-A)^i\mathbb{1} = (-1)^{i+1} \ r_{-i}.$$

**Theorem 1** *The rank of $R$ is $n$.*

**Proof 1** *$R$ can be expressed as*

$$R = \begin{bmatrix} \vdots \\ \alpha(-A)^2 \\ \alpha(-A)^1 \\ \alpha \\ \alpha(-A)^{-1} \\ \alpha(-A)^{-2} \\ \vdots \end{bmatrix} \begin{bmatrix} \cdots & (-A)^2\mathbb{1} & (-A)^1\mathbb{1} & \mathbb{1} & (-A)^{-1}\mathbb{1} & (-A)^{-2}\mathbb{1} & \cdots \end{bmatrix}.$$

*where the size of the first matrix is $\infty \times n$ and the size of the second matrix is $n \times \infty$. As a consequence the rank of $R$ is at most $n$.*

*The $n \times n$ sub-matrix of $R$ whose upper left element is $r_0$ is non-singular when $(\alpha, A)$ is non-redundant [13]. As a consequence the $\alpha(-A)^0, \alpha(-A)^{-1}, \ldots, \alpha(-A)^{-n+1}$ row vectors as well as the $(-A)^0 \mathbb{1}, (-A)^{-1} \mathbb{1}, \ldots, (-A)^{-n+1} \mathbb{1}$ column vectors are linearly independent.*

As a consequence of Theorem 1 the determinant of any $(n+1) \times (n+1)$ submatrix of R must be 0. Having the series $r_{i_1}, r_{i_2}, \ldots, r_{i_k}$, where $i_1 < i_2 < \ldots < i_k$ and $i_j \neq 0, \forall j \in 1, \ldots, k$

$$\det\left(R_{\{0,1,2,\ldots,n\},\{i,i+1,i+2,\ldots,i+n\}}\right) = 0, \quad i_1 \leq i \leq i_k - 2n \tag{5}$$

provides a set of $i_k - i_1 - 2n + 1$ equations with $i_k - i_1 - k$ unknowns. These equations are not necessarily linear in the unknowns.

## 2.3 Examples of application of (5)

**Generation of the $r_i$ series of ME distributions**

By $R_{c_1,c_2}$ where $c_1$ and $c_2$ are two sets of indices we denote the submatrix of $R$ which consists of rows according to $c_1$ and columns according to $c_2$. For example,

$$R_{\{1,3,4\},\{0,2\}} = \begin{bmatrix} r_1 & r_3 \\ r_3 & r_5 \\ r_4 & r_6 \end{bmatrix}.$$

Theorem 1 and (5) defines the relations in the $r_i$ series. Based on these relations we can compose an explicit algorithm to generate any element of the $r_i$ series of a ME(n) distribution based on the $r_0, r_1, \ldots, r_{2n-1}$ reduced moments without computing any representation of this distribution $(\alpha, A)$.

Given the first $2n$ reduced moments, $r_0, r_1, ..., r_{2n-1}$, $\det\left(R_{\{0,1,\ldots,n\},\{0,1,\ldots,n\}}\right) = 0$ and $\det\left(R_{\{0,1,\ldots,n\},\{-1,0,1,\ldots,n-1\}}\right) = 0$ give a single unknown linear equation for $r_{2n}$ and $r_{-1}$, respectively. Having determined $r_0, \ldots, r_{2n-1+i}$, $i > 0$, $(r_i, \ldots, r_{2n-1}$, $i < 0)$ $\det\left(R_{\{0,1,\ldots,n\},\{j,j+1,\ldots,j+n-1\}}\right) = 0$ with $j = i+1$ $(j = i-1)$ gives a single unknown linear equation for $r_{2n+i}$ $(r_{i-1})$.

**Non-unique PH(3) feasible solutions for $k = 2n - 1$**

Assuming that

$$r_0 = 1, r_{i_1} = r_{10} = 0.000295999, r_{i_2} = r_{20} = 6.9987\ 10^{-13}, r_{i_3} = r_{30} = 1.01758\ 10^{-23},$$
$$r_{i_4} = r_{40} = 5.24414\ 10^{-36}, r_{i_5} = r_{50} = 2.23015\ 10^{-49}$$

the set of equations in (5) is composed of $i_k - 2n + 1 = 45$ equations with $i_k - k = 45$ unknowns. There are several different solutions of this set of equations and there are more than one which are PH(3) feasible. E.g., the moments of

$$\alpha_1 = \begin{bmatrix} 0.1\ 0.2\ 0.7 \end{bmatrix} \qquad \alpha_2 = \begin{bmatrix} 0.586119\ 0.309469\ 0.104413] \end{bmatrix}$$

$$A_1 = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 1 & 0 & -3 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} -2.51342 & 0 & 0.796945 \\ 2.51342 & -2.51342 & 0 \\ 0 & 0.927169 & -0.927169 \end{bmatrix}$$

are solutions of (5) for this set of reduced moments. Table 1 lists the first reduced moments of the two PH(3). The differences vanish for large moments, but the numerical value of the relative differences in the last row of the table ($(r_i^{(1)} - r_i^{(2)})/r_i^{(1)}$) indicates that the $r_{10}$ reduced moments are identical while the $r_9$ and $r_{11}$ reduced moments are different. This periodicity of the moments remain valid forever, i.e, all tenth moments of the two PH(3) remains identical. It is interesting to note that the probability density function (pdf) of the two PH(3) exhibit a similar behaviour to that of the moments (see Figure 1). The two pdfs cross each others infinitely many times. Indeed, it implies that two PH(3) distributions whose pdfs are identical in infinitely many points are not necessarily identical.

| $r_i$ | 0. | 1. | 2. | 3. | 4. | 5. |
|---|---|---|---|---|---|---|
| $\{\alpha_1, A_1\}$ | 1. | 1.5 | 1.4875 | 1.03021 | 0.537956 | 0.224808 |
| $\{\alpha_2, A_2\}$ | 1. | 1.54536 | 1.50213 | 1.03248 | 0.538182 | 0.224824 |
| $r_i$ | 6. | 7. | 8. | 9. | 10. | 11. |
| $\{\alpha_1, A_1\}$ | 0.0782855 | 0.0233666 | 0.00610259 | 0.00141671 | 0.000295999 | 0.000056222 |
| $\{\alpha_2, A_2\}$ | 0.0782863 | 0.0233666 | 0.00610259 | 0.00141671 | 0.000295999 | 0.000056222 |
| rel. diff. | -1.10324 $10^{-5}$ | -1.51817 $10^{-6}$ | -1.78306 $10^{-7}$ | -1.51422 $10^{-8}$ | -3.66286 $10^{-16}$ | 3.94147 $10^{-10}$ |

Table 1: The reduced moments of PH(3) $\{\alpha_1, A_1\}$ and $\{\alpha_2, A_2\}$

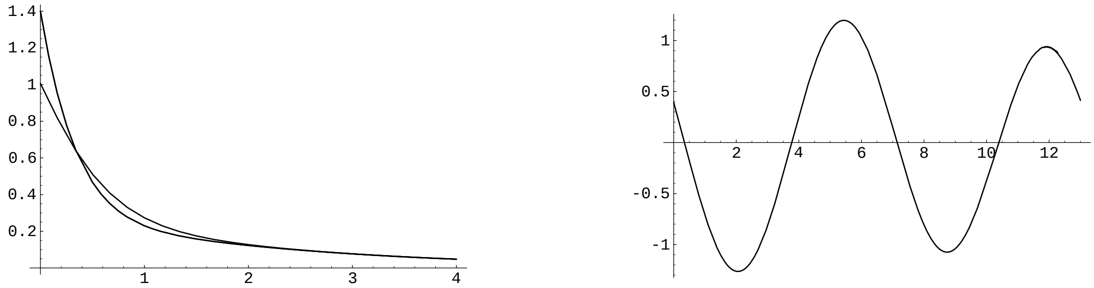

Figure 1: Density function of PH(3) $\{\alpha_1, A_1\}$ and $\{\alpha_2, A_2\}$ and their difference $(f_1(t) - f_2(t))e^{2.8t}$

**Fitting 3 reduced moments out of 5**

Assume that we are given the first five reduced moments of a random variable and our aim is to construct a PH distribution with two phases to approximate it. The five reduced moments are $r_1 = 1,\ r_2 = 1.25,\ r_3 = 4.1\dot{6},\ r_4 = 104.1\dot{6},\ r_5 = 8333.\dot{3}$. A PH distribution with two phases is determined by its first three reduced moments. However, we are not limited to use $r_1, r_2$ and $r_3$. For example, we can choose to fit reduced moments $r_1, r_3$ and $r_5$. In order to obtain PH distribution with two phases and reduced moments $r_1, r_3$ and $r_5$, we do the following. We assume that $r_1, r_3$ and $r_5$ are reduced moments of a PH distribution with two phases and we compute the first three reduced moments of this distribution based on (5). Then the fitting is performed based on these first three reduced moments.

In the following table we report fittings which are different in the choice of the utilised three reduced moments. In all the cases the mean is maintained and hence we have to choose two out of the four remaining reduced moments. The reduced moments that are set are indicated with bold characters. The reduced moments in the last row are chosen in such way that the sum of the relative errors in the reduced moments, $\sum_{i=2}^{5} |\hat{r}_i - r_i|/r_i$, are minimal. The distribution determined by this reduced moment set is referred to as *opt*.

| $r_2$ | $r_3$ | $r_4$ | $r_5$ |
|---|---|---|---|
| **1.25** | $\mathbf{4.1\dot{6}}$ | 35.5903 | 371.817 |
| **1.25** | 6.31469 | $\mathbf{104.1\dot{6}}$ | 1990.25 |
| **1.25** | 9.3729 | 265.489 | $\mathbf{8333.\dot{3}}$ |
| 1.09182 | $\mathbf{4.1\dot{6}}$ | $\mathbf{104.1\dot{6}}$ | 3353.48 |
| 1.05945 | $\mathbf{4.1\dot{6}}$ | 163.533 | $\mathbf{8333.\dot{3}}$ |
| 1.01543 | 2.27679 | $\mathbf{104.1\dot{6}}$ | $\mathbf{8333.\dot{3}}$ |
| 1.03826 | 3.22754 | 126.341 | 7047.53 |

In Figure 2 the pdf of the different fitted distribution are depicted. The legend indicates the reduced moments that are fitted. The figure depicts the fitting with three phases as well which matches all the five reduced moments. The logarithmic plot suggests that a better fitting of the lower moments results in a better body fitting, while the better fitting of the higher moments results better tail fitting.

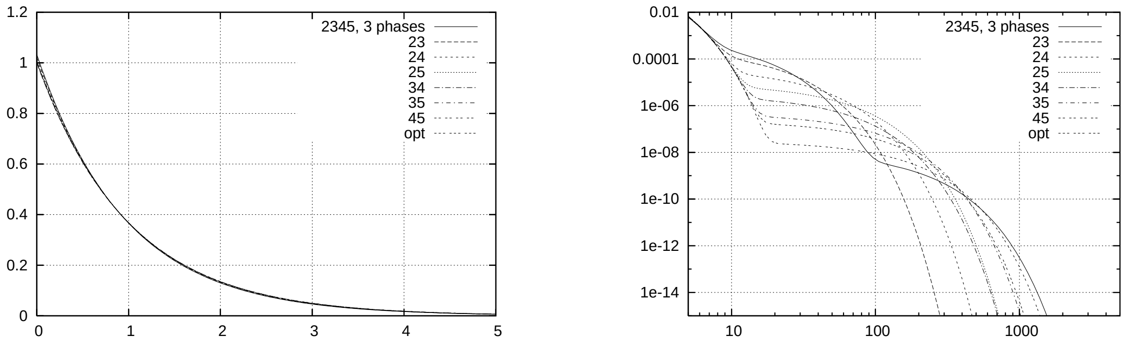


Figure 2: Body and tail of pdf of different fittings

**PH matching based on the derivatives of the cdf at zero**

The continuous PH(2) distributions are uniquely defined by their first 3 reduced moments and [12] proposes a way to compute an $\alpha$ probability vector and an $A$ generator matrix based on these reduced moments. We use this algorithm to create a "derivative matching" method for PH(2) distributions.

When the first three derivatives of the cdf of a PH(2) at $t = 0$ are $r_{-1} = 4,\ r_{-2} = 17,\ r_{-3} = 76$, we use

$$\det \begin{bmatrix} r_{-3} & r_{-2} & r_{-1} \\ r_{-2} & r_{-1} & r_0 \\ r_{-1} & r_0 & r_1 \end{bmatrix} = 0, \quad \det \begin{bmatrix} r_{-2} & r_{-1} & r_0 \\ r_{-1} & r_0 & r_1 \\ r_0 & r_1 & r_2 \end{bmatrix} = 0, \quad \det \begin{bmatrix} r_{-1} & r_0 & r_1 \\ r_0 & r_1 & r_2 \\ r_1 & r_2 & r_3 \end{bmatrix} = 0$$

and $r_0 = 1$ to obtain $r_1, r_2, r_3$. Solving the equations from left to right, the number of unknown is always 1, and the equations are linear in this unknown. The solution is

$$r_1 = \frac{4}{15}, \quad r_2 = \frac{17}{225}, \quad r_3 = \frac{76}{3375}.$$

From $r_1, r_2, r_3$ we obtain ([12]):

$$\alpha = [0.2 \ \ 0.8], \quad A = \begin{bmatrix} -3 & 3 \\ 0 & -5 \end{bmatrix}.$$

**An example of ME(3) feasibility check based on $r_{-2}, \ldots, r_3$**

One can perform a numerical ME(3) feasibility check based on the $r_{-j}, r_{-j+1}, \ldots, r_{5-j}$, $0 < j < 6$, series in three steps:

- compute the $r_0, r_1, \ldots, r_5$ series using the determinants of the $(n+1) \times (n+1)$ submatrices of $R$,
- calculate a vector $\alpha$ and a matrix $A$ based on $r_0, \ldots, r_{2n-1}$ such that $r_i = \alpha(-A)^{-i}\mathbb{1}$, $(i = 0, 1, \ldots, 2n-1)$ using the procedure of [13],
- decide the non-negativity of $f(t) = \alpha e^{At}(-A)\mathbb{1}$.

Starting from

$$r_{-3} = \frac{24}{5}, \quad r_{-2} = 0, \quad r_{-1} = \frac{6}{5}, \quad r_0 = 1, \quad r_1 = \frac{3}{5}, \quad r_2 = \frac{7}{20},$$

we can compute $r_3 = 5/24$ using $\det\left(R_{\{-3,-2,-1,0\},\{0,1,2,3\}}\right) = 0$, $r_4 = 1/8$ using $\det\left(R_{\{-2,-1,0,1\},\{0,1,2,3\}}\right) = 0$ and $r_5 = 3/40$ using $\det\left(R_{\{-1,0,1,2\},\{0,1,2,3\}}\right) = 0$.

Having $r_0, \ldots, r_5$ we compute $\alpha$ and $A$ which are

$$\alpha = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}, \quad A = \begin{bmatrix} -\frac{152}{9} & \frac{7702}{225} & -\frac{3962}{225} \\ -\frac{395}{81} & \frac{631}{81} & -\frac{341}{81} \\ \frac{305}{81} & -\frac{721}{81} & \frac{251}{81} \end{bmatrix}.$$

Then we plot $f(t) = \alpha e^{At}(-A)\mathbb{1}$ as shown in Figure 3. Assuming that one can decide the non-negativity of $f(t)$ based on Figure 3 we conclude that the given $\{r_{-2}, r_{-1}, r_0, r_1, r_2, r_3\}$ series is ME(3) feasible.

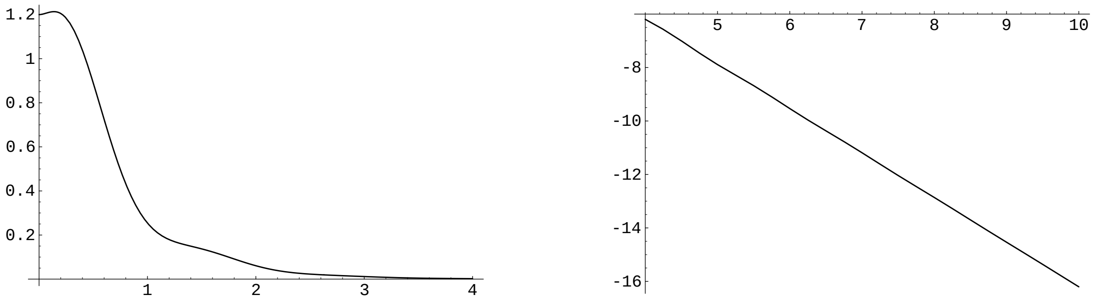


Figure 3: The density function, $f(t)$, and its tail on log-log scale constructed from the $\{r_{-2}, r_{-1}, r_0, r_1, r_2, r_3\}$ series

# 3 Matrix exponential and Markov arrival processes

## 3.1 Preliminaries

Let $X(t)$ be a stationary arrival process, defined by matrices $A_0$ and $A_1$, whose sequence of interarrival times is $X_0, X_1, \ldots$. The joint density of $X_0, X_1, \ldots, X_m$ is

$$f(x_0, x_1, \ldots, x_m) = \alpha e^{A_0 x_0} A_1 e^{A_0 x_1} A_1 \ldots e^{A_0 x_m} A_1 \mathbb{1}\ , \tag{6}$$

where $\alpha$ is the solution of $\alpha(-A_0)^{-1}A_1 = \alpha$ and $\alpha\mathbb{1} = 1$. The marginal densities of $f(x_0, x_1, \ldots, x_m)$ can be obtained using $\int_0^\infty e^{A_0 x} dx = (-A_0)^{-1}$. For example, $f(x_0, x_2, x_3, \ldots, x_m) = \alpha e^{A_0 x_0} A_1 (-A_0)^{-1} A_1 e^{A_0 x_2} A_1 \ldots e^{A_0 x_m} A_1 \mathbb{1}$.

The cardinality of the square matrices $A_0$ and $A_1$ is $n$. Similarly to the previous section, we consider the following cases:

- If $f(x_0, x_1, \ldots, x_m) \geq 0, \quad \forall\ m \geq 0$ and $\forall\ x_1, x_2, \ldots, x_m \geq 0$ and $\int_{x_1} \cdots \int_{x_m} f(x_0, x_1, \ldots, x_m) dx_1 \ldots dx_m = 1, \quad \forall\ m \geq 0$, then $X(t)$ is a matrix-exponential process of order $n$, MEP(n). MEP is identical to the rational arrival process defined in [2].

- If $A_0$ is a transient Markovian generator matrix and $A_1 \geq 0$, then $X(t)$ is a Markov arrival process of order $n$, MAP(n).

When $X(t)$ is a MEP(n), it has the following properties:

- The stationary inter-arrival time distribution is matrix-exponential with parameters $\alpha$ and $A_0$. Therefore, $A_0$ fulfills the conditions of ME distributions provided in the previous section.

- Starting from an arbitrary initial vector $(\alpha_0)$, the respective initial vectors at the consecutive inter-arrivals $(\alpha_1, \alpha_2, \ldots)$ satisfy $\alpha_i = \alpha_{i-1}G$, where $G = (-A_0)^{-1}A_1$. Matrix $G$ has the following properties:

- $\alpha G = \alpha$ and $G\mathbb{1} = \mathbb{1}$,
- $\mathbb{1} = G\mathbb{1} = (-A_0)^{-1} A_1 \mathbb{1}$ implies $-A_0\mathbb{1} = A_1\mathbb{1}$ and
- $\mathbb{1} = G\mathbb{1}$ implies that the respective initial vectors of the consecutive arrivals $(\alpha_1, \alpha_2, \ldots)$ satisfy $\alpha_i\mathbb{1} = 1$, if $\alpha_0\mathbb{1} = 1$.

- If $X(t)$ is a MEP(n), matrix $G$ has $n$ eigenvalues on the unit disk and one of them is 1 (otherwise the $\alpha_1, \alpha_2, \ldots$ series does not converge or $\lim_{i\to\infty} \alpha_i$ depends on $\alpha_0$).

When $X(t)$ is a MAP(n), it has the following additional properties:

- The phases of the system at arrival epochs form a DTMC, which means that matrix $G$ is a transition probability matrix, or stochastic matrix, i.e., the elements of $G$ are non-negative and not greater than 1.

- $\alpha$ is a probability vector.

The major differences of the MEP and the MAP cases are the following. In the case of MEPs the row sum and the diagonal element of $A_0$ can be positive, the elements of $\alpha$ and $G$ can be negative or greater than one and $A_1$ can contain negative elements. However, the row sums of $A_0 + A_1$ must be zero in both cases.

The pair of square matrices $(A_0, A_1)$ is referred to the representation of the MEP (MAP). In general, the $(A_0, A_1)$ representation is not unique.

**Definition 2** *An $(A_0, A_1)$ representation of cardinality n is said to be* non-redundant *if the inter-arrival time distribution defined by $(A_0, A_1)$ is a non redundant ME(n) distribution.*

Throughout the paper we assume that the representations of MEPs (MAPs) are non-redundant.

Since the inter-arrival times have a ME(n) distribution with generator $A_0$ and initial vector $\alpha$, the reduced moments of the inter-arrival times are (in accordance with (4)) $r_i = \alpha(-A_0)^{-i}\mathbb{1}$, and the results of the previous section are applicable to compute the relation of the elements of the $r_i$ series for $i = \{0, \pm 1, \pm 2, \ldots\}$.

The joint moments of the $a_0 = 0 < a_1 < a_2 < \ldots < a_{m-1}$-th inter-arrival times are

$$E(X_0^{i_1} X_{a_1}^{i_2} \ldots X_{a_{m-1}}^{i_m}) = \alpha\ i_1!(-A_0)^{-i_1} G^{a_1-a_0} i_2!(-A_0)^{-i_2} \ldots G^{a_{m-1}-a_{m-2}} i_m!(-A_0)^{-i_m}\mathbb{1}. \quad (7)$$

To shorten the notation we introduce $E = (-A_0)^{-1}$ and

$$\gamma_{k_2,\ldots,k_m}^{i_1,i_2,\ldots,i_m} = \frac{1}{\prod_{j=1}^m i_j!} E\left(X_0^{i_1} X_{k_2}^{i_2} \ldots X_{k_2+\ldots+k_m}^{i_m}\right) = \alpha E^{i_1} G^{k_2} E^{i_2} G^{k_3} E^{i_3} \ldots G^{k_m} E^{i_m}\mathbb{1}\ . \quad (8)$$

We refer to the $\gamma_{k_2,\ldots,k_m}^{i_1,i_2,\ldots,i_m}$ as reduced joint moments. Special cases of (7) include

- $m = 1$, $i_1 = i$:
$$E(X_0^i) = i!\alpha(-A_0)^{-i}\mathbb{1} = \gamma^i = r_i,$$
indicates that $\gamma_{k_2,\ldots,k_m}^{i_1,i_2,\ldots,i_m}$ is a generalisation of the reduced moment series. Since $\gamma^i = r_i$, we use the $\gamma_i$ notation to emphasise the relation with other $\gamma_{k_2,\ldots,k_m}^{i_1,i_2,\ldots,i_m}$ quantities and the $r_i$ notation to refer to the results of Section 2 in the remainder of the paper.

- $m = 2$, $a_1 = k$, $i_1 = i_2 = 1$: The joint mean,
$$E(X_0X_k) = \alpha(-A_0)^{-1}G^k(-A_0)^{-1}\mathbb{1} = \gamma_k^{1,1},$$
is the basic quantity to characterise the lag-k correlation of the process.

- $m = 2$, $a_1 = k$: The joint moments of the inter-arrival times $X_0$ and $X_k$,
$$E(X_0^{i_1} X_k^{i_2}) = i_1!i_2!\alpha(-A_0)^{-i_1} G^k (-A_0)^{-i_2}\mathbb{1} = i_1!i_2!\ \gamma_k^{i_1,i_2}\ ,$$
carries information about the joint distribution of the k-apart inter-arrival times. We refer to $E(X_0^iX_k^j)$ as lag-$k$ quantities.

The following subsections extend the concept of a Hankel matrix to find the relation between the various moments of the inter-arrival times of MEP(n)s. Similar to the ME distribution case, we present methods to compute the relations between all possible reduced moments of the inter-arrival times without computing any $(A_0, A_1)$ representation. We show that the first $2n-1$ reduced moments of the inter-arrival times, $E(X_0^i), 1 \leq i \leq 2n-1$, and the first $(n-1)^2$ lag-1 reduced moments, $E(X_0^i X_1^j), 1 \leq i, j \leq n-1$, uniquely determine all other reduced moments of a non-redundant MEP(n), and we present an algorithm for computing them.

## 3.2 Relation of reduced joint moments of a MEP(n) processes

Let

$$M_1(i_1, \ldots, i_m, k_2, \ldots, k_m) = \begin{bmatrix} \gamma_{k_2,\ldots,k_m}^{i_1,i_2,\ldots,i_m} & \gamma_{k_2,\ldots,k_{m-1}}^{i_1,i_2,\ldots,i_{m-1}} & \gamma_{k_2,\ldots,k_{m-1}}^{i_1,i_2,\ldots,i_{m-1}+1} & \ldots & \gamma_{k_2,\ldots,k_{m-1}}^{i_1,i_2,\ldots,i_{m-1}+n-1} \\ \gamma^{i_m} & r_0 & r_1 & \ldots & r_{n-1} \\ \gamma_{k_m}^{1,i_m} & r_1 & r_2 & \ldots & r_n \\ \gamma_{k_m}^{2,i_m} & r_2 & r_3 & \ldots & r_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma_{k_m}^{n-1,i_m} & r_{n-1} & r_n & \ldots & r_{2n-2} \end{bmatrix}$$

and

$$M_2(i_1, i_2, k) = \begin{bmatrix} \gamma_k^{i_1,i_2} & \gamma^{i_1} & \gamma_1^{i_1,1} & \ldots & \gamma_1^{i_1,n-1} \\ \gamma^{i_2} & r_0 & r_1 & \ldots & r_{n-1} \\ \gamma_{k-1}^{1,i_2} & r_1 & r_2 & \ldots & r_n \\ \gamma_{k-1}^{2,i_2} & r_2 & r_3 & \ldots & r_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma_{k-1}^{n-1,i_2} & r_{n-1} & r_n & \ldots & r_{2n-2} \end{bmatrix}.$$

**Theorem 2** *The rank of $M_1$ and $M_2$ is $n$.*

**Proof 2**

$$M_1(i_1, \ldots, i_m, k_2, \ldots, k_m) =$$

$$\begin{bmatrix} \alpha E^{i_1} G^{k_2} E^{i_2} \ldots G^{k_{m-1}} E^{i_{m-1}} \\ \alpha \\ \alpha E \\ \alpha E^2 \\ \vdots \\ \alpha E^{n-1} \end{bmatrix} \begin{bmatrix} G^{k_m} E^{i_m} 1\!\!1 & 1\!\!1 & E^1 1\!\!1 & E^2 1\!\!1 & \ldots & E^{n-1} 1\!\!1 \end{bmatrix},$$

$$M_2(i_1, i_2, k) = \begin{bmatrix} \alpha E^{i_1} G \\ \alpha \\ \alpha E \\ \alpha E^2 \\ \vdots \\ \alpha E^{n-1} \end{bmatrix} \begin{bmatrix} G^{k-1} E^{i_2} 1\!\!1 & 1\!\!1 & E^1 1\!\!1 & E^2 1\!\!1 & \ldots & E^{n-1} 1\!\!1 \end{bmatrix}$$

*and the rest of the proof is the same as for Theorem 1.*

From Theorem 2, the equation

$$\det(M_\ell(i_1,\ldots,i_m,k_2,\ldots,k_m)) = 0, \quad \ell = 1,2 \tag{9}$$

establishes the basic relation of the reduced joint moments of MEP(n)s. The upper left element of $M_1(i_1,\ldots,i_m,k_2,\ldots,k_m)$ and $M_2(i_1,i_2,k)$ are the ones with the highest order and (9) gives an explicit expression to compute these elements from the lower order reduced joint moments of the same process.

### 3.3 Computation of any reduced joint moment of MEP(n) processes

Assume that the parameters, $r_0 = 1$, $r_i = \gamma^i$ for $1 \le i \le 2n-1$ and $\gamma_1^{i,j}$ for $1 \le i,j \le n-1$ are known. Our aim is to compute $\gamma_{k_2,\ldots,k_m}^{i_1,i_2,\ldots,i_m}$ with $m > 1$. This can be done by the iterative application of (9).

Based on $M_1(i_1,\ldots,i_m,k_2,\ldots,k_m)$, it is possible to obtain an explicit expression for $\gamma_{k_2,\ldots,k_m}^{i_1,i_2,\ldots,i_m}$ in terms of $\gamma_{k_2,\ldots,k_{m-1}}^{i_1,i_2,\ldots,i_{m-1}+j}, 0 \le j \le n-1$, and $\gamma_{k_m}^{j,i_m}, 1 \le j \le n-1$, and $\gamma^j, 1 \le j$.

Then by constructing the matrices $M_1(i_1,\ldots,i_{m-1}+j,k_2,\ldots,k_{m-1}), 0 \le j \le n-1$ we can obtain explicit expressions for the quantities $\gamma_{k_2,\ldots,k_{m-1}}^{i_1,i_2,\ldots,i_{m-1}+j}, 0 \le j \le n-1$ in terms of $\gamma_{k_2,\ldots,k_{m-2}}^{i_1,i_2,\ldots,i_{m-2}+j}, 0 \le j \le n-1$, and $\gamma_{k_m}^{j_1,i_{m-1}+j_2}, 1 \le j_1 \le n-1, 0 \le j_2 \le n-1$, and $\gamma^j, 1 \le j$.

By repeating the above step finally we obtain an expression for $\gamma_{k_2,\ldots,k_m}^{i_1,i_2,\ldots,i_m}$ in terms of quantities such as $\gamma_k^{j,i}, 1 \le j \le n-1$, and $\gamma^j, 1 \le j$.

Quantities such as $\gamma^j = r_j, 1 \le j$ can be computed as described in Section 2. $M_2(i_1,i_2,k)$ will be used instead to deal with the quantities such as $\gamma_k^{j,i}, 1 \le j \le n-1$. Based on (9), it is possible the construct an explicit expression for $\gamma_k^{i_1,i_2}$ in terms of quantities such as $\gamma_1^{i_1,j}, 1 \le j \le n-1$, and $\gamma_{k-1}^{j,i_2}, 1 \le j \le n-1$, and $\gamma^j, 1 \le j$.

Since $1 \le i_1 \le n-1$, the quantities $\gamma_1^{i_1,j}, 1 \le j \le n-1$ are assumed to be known, while $\gamma^j, 1 \le j$ can be computed as described in Section 2. In order to deal with $\gamma_{k-1}^{j,i_2}, 0 \le j \le n-1$ we construct $M_2(j,i_2,k-1), 1 \le j \le n-1$ from which we can have expressions for these quantities in term of $\gamma_1^{j_1,j_2}, 1 \le j_1,j_2 \le n-1$, and $\gamma_{k-2}^{j,i_2}, 1 \le j \le n-1$, and $\gamma^j, 1 \le j$.

By successive application of the above step we obtain an expression for $\gamma_k^{j,i}, 1 \le j \le n-1$ in terms of $\gamma_1^{j,i}, 1 \le j \le n-1$, and $\gamma^j, 1 \le j$.

The quantities $\gamma_1^{j,i}, 1 \le j \le n-1$ can instead be computed by constructing the following special case of the $M_2(i_1,i_2,k)$ matrix.

$$M_2(i_1,i_2,1) = \tag{10}$$

$$\begin{bmatrix} \alpha E^{i_1} G \\ \alpha \\ \alpha E \\ \alpha E^2 \\ \vdots \\ \alpha E^{n-1} \end{bmatrix} \begin{bmatrix} E^{i_2}\mathbb{1} & \mathbb{1} & E^1\mathbb{1} & E^2\mathbb{1} & \ldots & E^{n-1}\mathbb{1} \end{bmatrix} = \begin{bmatrix} \gamma_1^{i_1,i_2} & \gamma^{i_1} & \gamma_1^{i_1,1} & \ldots & \gamma_1^{i_1,n-1} \\ \gamma^{i_2} & r_0 & r_1 & \ldots & r_{n-1} \\ \gamma^{i_2+1} & r_1 & r_2 & \ldots & r_n \\ \gamma^{i_2+2} & r_2 & r_3 & \ldots & r_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma^{i_2+n-1} & r_{n-1} & r_n & \ldots & r_{2n-2} \end{bmatrix}.$$

Based on (10), it is possible to construct an explicit expression for $\gamma_1^{i_1,i_2}$ in terms of quantities that are all known since $1 \le i_1 \le n-1$ or can be computed by the algorithm given in Section 2 and $\gamma_1^{i_1,i_2}$ is given for $1 \le i_1,i_2 \le n-1$.

**Application of the procedure for computing $E(X_0^2 X_2 X_3)$**

In the following we list the matrices that we need to compute $\gamma_{2,1}^{2,1,1}$ from which it is straightforward to determine $E(X_0^2 X_2 X_3)$ as a function of the known parameters. We assume that $n = 2$ and hence the matrices are of size $3 \times 3$ and we know $\gamma^i, 0 \le i \le 3$ and $\gamma_1^{1,1}$. All the matrices are used to obtain an expression for the (1,1) entry of the matrix based on the fact that the determinant of the matrix must be 0. In all the matrices bold characters are used to indicate the variables that we do not know yet.

$$M_1(2,1,1,2,1) = \begin{bmatrix} \gamma_{2,1}^{2,1,1} & \boldsymbol{\gamma_2^{2,1}} & \boldsymbol{\gamma_2^{2,2}} \\ \gamma^1 & \gamma^0 & \gamma^1 \\ \gamma_1^{1,1} & \gamma^1 & \gamma^2 \end{bmatrix}, M_1(2,1,2) = \begin{bmatrix} \gamma_2^{2,1} & \gamma^2 & \gamma^3 \\ \gamma^1 & \gamma^0 & \gamma^1 \\ \boldsymbol{\gamma_2^{1,1}} & \gamma^1 & \gamma^2 \end{bmatrix}, M_1(2,2,2) = \begin{bmatrix} \gamma_2^{2,2} & \gamma^2 & \gamma^3 \\ \gamma^2 & \gamma^0 & \gamma^1 \\ \boldsymbol{\gamma_2^{1,2}} & \gamma^1 & \gamma^2 \end{bmatrix},$$

$$M_2(1,1,2) = \begin{bmatrix} \gamma_2^{1,1} & \gamma^1 & \gamma_1^{1,1} \\ \gamma^1 & \gamma^0 & \gamma^1 \\ \gamma_1^{1,1} & \gamma^1 & \gamma^2 \end{bmatrix}, M_2(1,2,2) = \begin{bmatrix} \gamma_2^{1,2} & \gamma^1 & \gamma_1^{1,1} \\ \gamma^2 & \gamma^0 & \gamma^1 \\ \boldsymbol{\gamma_1^{1,2}} & \gamma^1 & \gamma^2 \end{bmatrix}, M_2(1,2,1) = \begin{bmatrix} \gamma_1^{1,2} & \gamma^1 & \gamma_1^{1,1} \\ \gamma^2 & \gamma^0 & \gamma^1 \\ \gamma^3 & \gamma^1 & \gamma^2 \end{bmatrix}$$

Based on the above matrices and the facts that $\gamma^0 = 1$ and $\gamma^i = r_i$ we have that

$$\begin{aligned} \gamma_{2,1}^{2,1,1} = \; & \frac{1}{((r_1)^2 - r_2)^4} \Big( r_2 \Big( \gamma_1^{1,1}(r_2)^4 - (r_1)^4 \gamma_1^{1,1} \Big( (\gamma_1^{1,1})^2 + 4\gamma_1^{1,1} r_2 - 8(r_2)^2 \Big) + \\ & (r_1)^2 \gamma_1^{1,1} r_2 \Big( 2(\gamma_1^{1,1})^2 - \gamma_1^{1,1} r_2 - 4(r_2)^2 \Big) + (r_1)^6 \Big( 2(\gamma_1^{1,1})^2 - 2\gamma_1^{1,1} r_2 - (r_2)^2 \Big) \Big) + \\ & r_1 \Big( (r_1)^2 - \gamma_1^{1,1} \Big)^2 \Big( r_2 \Big( -3\gamma_1^{1,1} + r_2 \Big) + (r_1)^2 \Big( \gamma_1^{1,1} + r_2 \Big) \Big) r_3 - \Big( (r_1)^2 - \gamma_1^{1,1} \Big)^3 (r_3)^2 \Big). \end{aligned}$$

We have seen that the $r_0 = 1$, $r_i$, $1 \le i \le 2n-1$ reduced moments and the $\gamma_1^{i,j}$, $1 \le i,j \le n-1$ reduced joint moments uniquely defines any reduced joint moment of a MEP(n). By this reason we define the basic moment set as follows.

**Definition 3** *The* basic moment set *of a MEP(n) is the set of* $r_0 = 1$, $r_i$, $1 \le i \le 2n-1$ *reduced moments together with the* $\gamma_1^{i,j}$, $1 \le i,j \le n-1$ *reduced joint moments. The basic moment set is composed of* $n^2$ *moments and an additive constraint,* $r_0 = 1$.

## 3.4 Dependent MEP(n) moments

Suppose we are given the $r_1, r_2, r_3, r_4, r_5$ reduced moments and the $\gamma_1^{1,1}, \gamma_1^{1,2}, \gamma_1^{1,3}, \gamma_1^{1,4}$ reduced joint moments of a MEP(3). Similar to the number of moments in the basic moment set, this is a set of $n^2$ moments and based on this fact one might expect to compute all other moments of the MEP(n) from this moment set.

We have that $\det(M_2(1,3,1)) = 0$ and $\det(M_2(1,4,1)) = 0$. These equations define relations for the given set of moments, that is,

$$M_2(1,3,1) = \begin{bmatrix} \gamma_1^{1,3} & \gamma^1 & \gamma_1^{1,1} & \gamma_1^{1,2} \\ \gamma^3 & \gamma^0 & \gamma^1 & \gamma^2 \\ \gamma^4 & \gamma^1 & \gamma^2 & \gamma^3 \\ \gamma^5 & \gamma^2 & \gamma^3 & \gamma^4 \end{bmatrix}, \quad M_2(1,4,1) = \begin{bmatrix} \gamma_1^{1,4} & \gamma^1 & \gamma_1^{1,1} & \gamma_1^{1,2} \\ \gamma^4 & \gamma^0 & \gamma^1 & \gamma^2 \\ \gamma^5 & \gamma^1 & \gamma^2 & \gamma^3 \\ \gamma^6 & \gamma^2 & \gamma^3 & \gamma^4 \end{bmatrix}.$$

It can be seen that $\det(M_2(1,3,1)) = 0$ and $\det(M_2(1,4,1)) = 0$ can be used to determine $\gamma_1^{1,3}, \gamma_1^{1,4}$, as a function of $r_1, r_2, r_3, r_4, r_5$ and $\gamma_1^{1,1}, \gamma_1^{1,2}$, using that $r_0, r_1, r_2, r_3, r_4, r_5$ also defines $r_6 = \gamma^6$.

The $\det(M_2(1,3,1)) = 0$ equation can also be used to show that the $r_1, r_2, r_3, r_4$ reduced moments and the $\gamma_1^{1,1}, \gamma_1^{1,2}, \gamma_1^{1,3}, \gamma_1^{2,1}, \gamma_1^{2,2}$ reduced joint moments uniquely determine the elements of the basic moment set.

From this example we draw the following conclusions. It is not obvious to see the dependencies of the various sets of moments. There is more than one set of $n^2$ moments that uniquely determine all moments of a MEP(n) process and, therefore, the choice of the basic moment set is not unique.

### 3.5 Two different MEPs with equal marginal distribution and lag-correlations

In [9] a procedure is presented for the construction of a MEP based on the first $2n-1$ moments of the marginal distribution and the first $2n-3$ lag-correlations. Here we show that there are different MEPs that realize the same marginal distribution and lag-correlations.

Assume that $n = 3$ and we are given the basic moment set of the MEP(3), in particular we have $\gamma_1^{1,2} = x_1$ and $\gamma_1^{2,1} = x_2$. Based on these parameters, as shown in Subsection 3.3, it is possible to compute the quantities that determine the lag-correlations. By following these computations, it can be seen that the given parameter set with $\gamma_1^{1,2} = x_1$ and $\gamma_1^{2,1} = x_2$ results in the same lag-correlations as the ones with $\gamma_1^{1,2} = x_2$ and $\gamma_1^{2,1} = x_1$.

As an example we consider the following two cases:

$$r_1 = \frac{559}{1350},\ r_2 = \frac{469081}{1215000},\ r_3 = \frac{4660039019}{12028500000},\ r_4 = \frac{4237895351171}{10825650000000},\ r_5 = \frac{42422816639765929}{107173935000000000},$$

$$E(X_0^1X_1^1) = \frac{1309691}{26730000},\ E(X_0^2X_1^2) = \frac{525968309171}{5412825000000},$$

$$E(X_0^1X_1^2) = \frac{703719119}{12028500000} \text{ or } \frac{1031769119}{12028500000},$$

$$E(X_0^2X_1^1) = \frac{1031769119}{12028500000} \text{ or } \frac{703719119}{12028500000}.$$

In both cases the lag-correlations are identical and the first lag-correlations are

$$E(X_0X_2) = \frac{4146491}{17820000},\ E(X_0X_3) = \frac{3011771}{21384000},\ E(X_0X_4) = \frac{39937801}{213840000},\ E(X_0X_5) = \frac{23351837}{142560000}.$$

In order to show that the two MEPs are different, in Figure 4 we depict the joint density of $X_0$ and $X_1$ for a given value of $X_0$.

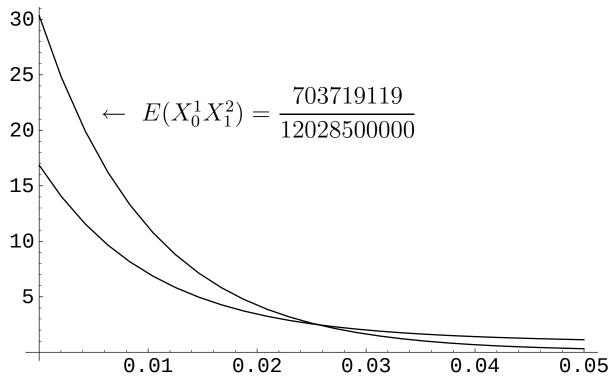


Figure 4: Joint density, $f_{X_0,X_1}(1,x)$, for the two MEPs with equal marginal moments and lag-correlation

# 4 Conclusion

This paper provides a methodology to investigate the relation of moments of ME distributions and MEPs. The presented results are also valid for PH distributions and MAPs as they are proper subsets of ME distributions and MEPs, respectively.

In our future work we intend to apply this compact moment representation of MEPs in moment matching and fitting.

## References


[1] S. Asmussen and M. Bladt. Renewal theory and queueing algorithms for matrix-exponential distributions. In *Proc. 1st Int. Conference on Matrix-Analytic Methods in Stochastic Models*, pages 313–341, 1996.

[2] S. Asmussen and M. Bladt. Point processes with finite-dimensional conditional probabilities. *Stochastic Processes and their Applications*, 82:127–142(16), July 1999.

[3] S. Asmussen and C. A. O'Cinneide. Matrix-exponential distributions – distributions with a rational Laplace transform. In S. Kotz and C. Read, editors, *Encyclopedia of Statistical Sciences*, pages 435–440, New York, 1997. John Wiley & Sons.

[4] M. Bladt and M. Neuts. Matrix-exponential distributions: Calculus and interpretations via flows. *Commun. Statist.-Stochastic Models*, 2003.

[5] M. Fackrell. Fitting with matrix-exponential distributions. *Stochastic Models*, 21:377–400, 2005.

[6] G. Latouche and V. Ramaswami. *Introduction to matrix analytic methods in stochastic modeling.* SIAM, 1999.

[7] L. Lipsky. *Queueing Theory: A linear algebraic approach.* MacMillan, New York, 1992.

[8] K. Mitchell, K. Sohraby, A. van de Liefvoort, and J. Place. Approximation models of wireless cellular networks using moment matching. In *Proc. Conf. on Computer Communications (IEEE Infocom)*, pages 189–197, 2000.

[9] K. Mitchell and A. van de Liefvoort. Approximation models of feed-forward G/G/1/N queueing networks with correlated arrivals. *Performance Evaluation*, 51:137–152, 2003.

[10] M. Neuts. A versatile markovian point process. *Journal of Applied Probability*, 16:764–779, 1979.

[11] M. Neuts. *Matrix-Geometric Solutions in Stochastic Models.* John Hopkins University Press, Baltimore, MD, USA, 1981.

[12] M. Telek and A. Heindl. Matching moments for acyclic discrete and continuous phase-type distributions of second order. *International Journal of Simulation Systems, Science & Technology*, 3(3-4):47–57, dec. 2002. Special Issue on: Analytical & Stochastic Modelling Techniques.

[13] A. van de Liefvoort. The moment problem for continuous distributions. Technical report, University of Missouri, WP-CM-1990-02, Kansas City, 1990.