



Multivariate matrix-exponential distributions

Bladt, Mogens; Nielsen, Bo Friis

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Bladt, M., & Nielsen, B. F. (2008). Multivariate matrix-exponential distributions. Technical University of Denmark, DTU Informatics, Building 321. (D T U Compute. Technical Report; No. 2008-02).

DTU Library

Technical Information Center of Denmark

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Multivariate matrix–exponential distributions

Mogens Bladt* and Bo Friis Nielsen †

April 28, 2008

Abstract

In this article we consider the distributions of non–negative random vectors with a rational Laplace transform. Hence the Laplace transforms are fractions between two multidimensional polynomials. These distributions are in the univariate case known as matrix–exponential distributions, since their densities can be written as linear combinations of the elements in the exponential of a matrix. For this reason we shall refer to multivariate distributions with rational Laplace transform as multivariate matrix–exponential distributions (MVME). The marginal distributions of an MVME are univariate matrix–exponential distributions.

We prove a characterization which states that a distribution is an MVME if, and only if, all positive linear combinations of the coordinates have a univariate matrix–exponential distribution. This theorem is analog to a well known characterization theorem for the multivariate normal distribution, however, the proof is different and involves theory for rational function based on continued fractions and Hankel determinants.

Keywords: Matrix–exponential; continued fractions; Hankel matrices; phase–type distribution;

AMS subject classification: 62H05; 60E10 .

*Institute for Applied Mathematics and Systems, National University of Mexico, Mexico City, bladt@stats.iimas.unam.mx

†Informatics and Mathematical Modelling, Technical University of Denmark, Kgs. Lyngby, bfn@imm.dtu.dk

1 Introduction

In one dimension, matrix–exponential distributions are defined as distributions on \mathbb{R}_+ with a rational Laplace transform, which in turn is equivalent to its density being a weighted sum of the elements of a matrix–exponential, thereby their name.

In this paper we define a class of distributions which we shall refer to as multivariate matrix–exponential distributions (MVME). They are defined in a natural way, inspired by the definition of univariate matrix–exponential distributions, as the distributions on \mathbb{R}_+^n having a rational (multidimensional) Laplace transform. A multidimensional rational function is the fraction between two multidimensional polynomials. The marginal distributions are hence univariate matrix–exponential distributions. The corresponding random variables are in general dependent.

The main purpose of this article is to characterize the MVME distributions in terms of one–dimensional matrix–exponential distributions. The main result states that a multivariate distribution is an MVME if and only if any non–negative, non–null linear combination of the coordinates are again matrix–exponentially distributed. This theorem is stated and proven in section 4.

Much of the analysis depends on the theories of one–dimensional matrix–exponential distributions and that of rational functions. We provide the necessary background on univariate matrix–exponential distributions in Section 2 as well as a historical review of two subclasses of MVME which were defined previously by Assaf et al. (1984) and Kulkarni (1989) respectively. We also re–formulate Kulkarni’s definition in terms of the structure of certain projections which is more in line with our characterization theorem. In Section 3 we discuss some theory of rational functions in terms of continued fractions, the theory of which turns out to provide a rich and powerful methodology for the analysis of multidimensional rational functions, in which Hankel matrices of certain reduced moments will play a particularly important role. The article is concluded in Section 5.

2 Preliminaries

In this section we provide some necessary background from the theory of one–dimensional matrix–exponential distributions. Thereafter we review two

classes of multivariate phase-type distributions, which are special cases of the MVME distributions.

The class proposed by Kulkarni (1989) is of particular interest, and we provide a characterization for this subclass in terms of the structure of the intensity matrices of certain projections.

For ease of exposition we shall only consider absolutely continuous distributions.

2.1 Univariate matrix-exponential distributions

Definition 2.1 *A non-negative random variable X is said to have a matrix-exponential distribution if the Laplace transform $L(s) = \mathbb{E}[\exp(-sX)]$ is a rational function in s .*

We explicitly state the parameterization of the Laplace transform as

$$L(s) = \mathbb{E}[\exp(-sX)] = \frac{f_1 s^{m-1} + f_2 s^{m-2} + \dots + f_m}{s^m + g_1 s^{m-1} + \dots + g_m},$$

where $f_m = g_m$. The following result is standard (see e.g. Asmussen & Bladt (1997) page 315 for a proof).

Lemma 2.1 *A random variable is matrix-exponentially distributed if and only if there exists a triple $(\boldsymbol{\beta}, D, \mathbf{d})$ such that the density $f(x)$ of X can be expressed as*

$$f(x) = \boldsymbol{\beta} e^{Dx} \mathbf{d} .$$

Here, $\boldsymbol{\beta}$ is a row vector of dimension m , \mathbf{d} is a column vector of the same dimension, and D is an $m \times m$ matrix, possibly with complex elements.

The triple $(\boldsymbol{\beta}, D, \mathbf{d})$ is called a representation of the matrix exponential distribution. The Laplace transform of X can be determined from a representation $(\boldsymbol{\beta}, D, \mathbf{d})$ as

$$L(s) = \boldsymbol{\beta} (sI - D)^{-1} \mathbf{d}, \tag{1}$$

where I is the identity matrix of dimension m . Any matrix-exponential distribution has infinitely many representations. The dimension of D is called the *order* of the representation. If the rational function $L(s)$ of Equation (1) cannot be reduced, we say that m is the *minimal order* or *degree* of the distribution (Asmussen & Bladt (1997)).

The non-centralized moments of a matrix-exponentially distributed random variable are easily derived by successive differentiation in the Laplace transform

$$M_i = \mathbb{E}(X^i) = i! \boldsymbol{\beta} (-D)^{-(i+1)} \mathbf{d} \quad i = 0, 1, \dots .$$

We shall also need the so-called reduced moments,

$$\mu_i = \frac{M_i}{i!}, \quad i = 0, 1, \dots . \quad (2)$$

Define recursively,

$$\psi_0 = \frac{1}{g_m}, \quad \psi_i = \sum_{j=0}^{i-1} (-1)^j \frac{\psi_{i-1-j} g_{m-1-j}}{g_m}.$$

Then the first $m - 1$ moments are given by

$$\mu_i = \sum_{j=0}^i (-1)^j f_{m-j} \psi_{i-j} .$$

From the Cayley-Hamilton theorem we have

$$\sum_{i=0}^m g_i D^{m-i} = 0 ,$$

where D can be taken from any representation. Since D is invertible, multiplying the equation with $D^{-(m+j+1)}$ we obtain that

$$\sum_{i=0}^m g_{m-i} (-1)^{m+j-i+1} (-D)^{-(m+j-i+1)} = \sum_{i=0}^m g_i (-1)^{j+i} (-D)^{-(i+j+1)} = \mathbf{0}.$$

Pre-multiplying with $\boldsymbol{\beta}$ and post-multiplying with \mathbf{d} we get

$$\sum_{i=0}^m g_i (-1)^{j+i} \frac{M_{i+j}}{(i+j)!} = 0.$$

We state this result in a slightly modified form in the following lemma.

Lemma 2.2 *The reduced moments of a matrix-exponential distribution satisfy*

$$\mu_{m+j} = \sum_{i=0}^{m-1} \frac{g_i}{g_m} (-1)^{m+i+1} \mu_{i+j} \quad \text{for } j \geq 0.$$

We notice that the rational distribution is hence characterized by $m - 1$ reduced moments together with the m coefficients g_0, \dots, g_{m-1} . In turn, this is equivalent to the distribution being characterized by $2m - 1$ reduced moments.

2.2 Two classes of multivariate phase-type distributions

There exist a vast amount of definitions concerning multivariate distributions of either exponential or gamma type in the literature (see e.g. Kotz, Balakrishnan & Johnson (2000)). Such distributions either have exponentially, or gamma distributed marginals. This has resulted in a rather extensive amount of distributions many of which are related or only differ from each other vaguely. A number of these distributions have a rational multidimensional Laplace transform. Also, the class of phase-type distributions, which generalize certain gamma type distributions, has been extended to a multivariate setting, first by Assaf et al. (1984) and later Kulkarni (1989). The latter class, which contains the former as a special case, provides an elegant construction of multivariate phase-type distributions in terms of a single underlying Markov jump process. Furthermore, it has a natural generalization to multivariate matrix-exponential distributions.

Assaf et al. (1984) introduced a class of multivariate phase-type distributions, denoted by MPH in the following, by considering the hitting times to different (possibly overlapping) subsets of the state-space. More specifically, we consider a phase-type generator (sub-intensity matrix) T of dimension m and let $\{J_t\}_{t \geq 0}$ denote the underlying Markov jump process. Let Γ_i , $i = 1, 2, \dots, n$ denote absorbing subsets of the state space. Let X_i denote the first hitting time of $\{J_t\}_{t \geq 0}$ to Γ_i . Then the n -dimensional vector $\mathbf{X} = (X_1, \dots, X_n)$ is said to have a phase-type distribution in the class MPH.

A rephrasing of the definition of MPH says that reward for X_k is accumulated with rate 1 in states belonging to Γ_k^c , where Γ_k^c is the complement of Γ_k . Based on this interpretation, Kulkarni (1989) introduced the class MPH*, which is a generalization of the MPH class. In the class MPH* reward for X_i is accumulated in state j with rate K_{ij} . There is no restriction on the phase-type generator T . If the total sojourn time in state j before absorption is denoted by Y_j we define an n dimensional random vector \mathbf{X} element-wise by $X_i = \sum_{j=1}^m K_{ij} Y_j$.

The following theorem gives an alternative characterization in terms of all non-negative projections and is not given explicitly in Kulkarni (1989).

Theorem 2.1 *A distribution in MPH* can be characterized by $\langle \mathbf{a}, \mathbf{X} \rangle$ being phase-type distributed with representation $(\boldsymbol{\alpha}, T(\mathbf{a}))$ with $T(\mathbf{a}) = \Delta(K\mathbf{a})^{-1}T$, where $\Delta(\mathbf{b})$ is the diagonal matrix with \mathbf{b} in the diagonal.*

Proof: Let \mathbf{X} denote a random variable with a distribution in MPH*. Then the i th component X_i of \mathbf{X} can be written as

$$X_i = \sum_{j=1}^m \sum_{k=1}^{N_k} K_{ij} Z_{jk} .$$

Here N_k denotes the number of visits to the transient state k in the $m + 1$ -dimensional continuous time Markov chain with m transient states and 1 absorbing state. Thus this Markov chain defines a continuous time phase-type distribution. We denote the transient part of the generator matrix by T . The random variables Z_{jk} are the k 'th sojourn in state j , while K_{ij} are non-negative real constants. Considering the distributions of the family of projections given by $\langle \mathbf{a}, \mathbf{X} \rangle$ we obtain

$$\langle \mathbf{a}, \mathbf{X} \rangle = \sum_{i=1}^n a_i \sum_{j=1}^m \sum_{k=1}^{N_k} K_{ij} Z_{jk} = \sum_{j=1}^m \left(\sum_{i=1}^n K_{ij} a_i \right) Z_{j.} ,$$

with $Z_{j.} = \sum_{k=1}^{N_k} Z_{jk}$. Before we proceed we have to introduce a technical condition. We will assume in the following that $K\mathbf{a} > 0$. If this condition is not true one can still proceed albeit some care is needed to get a proper phase-type representation. See Kulkarni (1989), Section 2, for how to handle this situation. Under the condition $K\mathbf{a} > 0$ we see that $\langle \mathbf{a}, \mathbf{X} \rangle$ is phase-type distributed with generator matrix $(\Delta(K\mathbf{a}))^{-1}T$. The argument is easily seen to apply in the reverse order as well.

□

3 Continued fractions

By a continued fraction we understand an expression on the form

$$d_0 + \frac{c_1}{d_1 + \frac{c_2}{d_2 + \frac{c_3}{d_3 + \dots}}} .$$

It is convenient to use the more compact notation

$$d_0 + \frac{c_1|}{|d_1|} + \frac{c_2|}{|d_2|} + \frac{c_3|}{|d_3|} + \dots .$$

A continued fraction is said to be finite if the sum above contains a finite number of terms. Of particular interest for our analysis are the C–continued fractions, which are expressions on the form

$$1 + \frac{c_1 s^{r_1}|}{|1|} + \frac{c_2 s^{r_2}|}{|1|} + \frac{c_3 s^{r_3}|}{|1|} + \dots .$$

If we consider the moment generating function $M(s)$ of a matrix–exponentially distributed random variable, then it has a power series expansion

$$M(s) = 1 + \mu_1 s + \mu_2 s^2 + \dots,$$

where μ_i 's are defined in (2).

Please note that $\mu_i > 0$ for all i . According to Perron (1957) Satz 3.5, any power series (Taylor series) with constant term 1 corresponds uniquely to a C–continued fraction. If furthermore the series is a power series expansion of a rational function, then the corresponding continued fraction is finite (Perron (1957), Satz 3.7). Particularly tractable are the regular C–continued fraction, where $r_i = 1$ for all i .

Lemma 3.1 *Let $M(s)$ be the moment generating function of a matrix–exponentially distributed random variable. Then the power series expansion of $M(s) = 1 + \mu_1 s + \mu_2 s^2 + \dots$ corresponds uniquely to a regular C–continued fraction.*

Proof: We already know from Perron (1957) that there is a unique and finite C-continued fraction, so all we need to prove is that $r_i = 1$ for all i . To this end we write

$$B_0(s) = 1 + \mu_1 s + \mu_2 s^2 + \dots = 1 + \frac{c_1 s^{r_1}}{B_1(s)},$$

where

$$B_1(s) = 1 + \frac{c_2 s^{r_2}}{|1|} + \frac{c_3 s^{r_2}}{|1|} + \dots + \frac{c_\nu s^{r_\nu}}{|1|}.$$

The first non-vanishing term of the power expansion $B_0(s) - 1$ is $\mu_1 s$. Hence $r_1 = 1$. Thus

$$B_1(s) = \frac{\mu_1 s}{B_0(s) - 1} = \frac{1}{1 + \mu'_1 s + \mu'_2 s^2 + \dots},$$

where $\mu'_i = \mu_i / \mu_1 \neq 0$ for all i . Then the first non-vanishing term in the power series expansion of $B_1(s) - 1$ is again of first order. Continuing this way until $B_\nu(s) = 1$ proves the result. □

For any continued fraction we may approximate it by a lower order n , where $n < \nu$ for the finite case,

$$\frac{C_n}{D_n} = d_0 + \frac{c_1}{|d_1|} + \frac{c_2}{|d_2|} + \frac{c_3}{|d_3|} + \dots + \frac{c_n}{|d_n|}.$$

The finite fraction C_n/D_n is called an approximant to the continued fraction. The following recursion scheme holds for the approximants of different orders

$$\begin{aligned} C_n &= d_n C_{n-1} + c_n C_{n-2} \\ D_n &= d_n D_{n-1} + c_n D_{n-2}, \end{aligned}$$

with the boundary conditions $C_{-1} = 1, C_0 = d_0$ and $D_{-1} = 0, D_0 = 1$. If we apply these recursive equation to regular C-continued fractions we get that the polynomials $C_n(s)$ and $D_n(s)$ (in the variable s) satisfy

$$\begin{aligned} C_n(s) &= 1 + \sum_{j=1}^n s^j \left(\sum_{0 \leq i_1 < i_2 < \dots < i_j < n-j} c_{i_1+1} c_{i_2+2} \dots c_{i_j+j} \right) \\ D_n(s) &= 1 + \sum_{j=1}^n s^j \left(\sum_{1 \leq i_1 < i_2 < \dots < i_j < n-j} c_{i_1+1} c_{i_2+2} \dots c_{i_j+j} \right), \end{aligned}$$

where we assume the sums over empty sets are zero. Thus we may write

$$\begin{aligned}
C_{2n}(s) &= 1 + \alpha_{n,1}s + \alpha_{n,2}s^2 + \cdots + \alpha_{n,n}s^n \\
D_{2n}(s) &= 1 + \beta_{n,1}s + \beta_{n,2}s^2 + \cdots + \beta_{n,n}s^n \\
C_{2n-1}(s) &= 1 + \gamma_{n,1}s + \gamma_{n,2}s^2 + \cdots + \gamma_{n,n}s^n \\
D_{2n-1}(s) &= 1 + \delta_{n,1}s + \delta_{n,2}s^2 + \cdots + \delta_{n,n-1}s^{n-1}.
\end{aligned}$$

In particular, we have that $\beta_{n,n} = c_2c_4 \cdots c_{2n}$ and $\gamma_{n,n} = c_1c_3 \cdots c_{2n-1}$ which shall turn out to be useful expressions.

We now consider the moment generating function $M(s)$ and its power series expansion in terms of the reduced moments. The power expansion of the k 'th order approximant is $C_k(s)/D_k(s)$. Then the first k terms of both power expansions coincide (see Perron (1957), Satz 3.2). Hence we may write

$$\frac{C_k(s)}{D_k(s)} = 1 + \mu_1s + \mu_2s^2 + \dots + \mu_k s^k + \tilde{\mu}_{k+1}s^{k+1} + \dots,$$

where $\tilde{\mu}_i$ are some constants. Now inserting the above expression for $C_k(s)$ and $D_k(s)$ with $k = 2n$ and $k = 2n - 1$ we get that

$$\begin{aligned}
\frac{1 + \alpha_{n,1}s + \cdots + \alpha_{n,n}s^n}{1 + \beta_{n,1}s + \cdots + \beta_{n,n}s^n} &= 1 + \mu_1s + \cdots + \mu_{2n}s^{2n} + \tilde{\mu}_{2n+1}s^{2n+1} + \dots \quad (3) \\
\frac{1 + \gamma_{n,1}s + \cdots + \gamma_{n,n}s^n}{1 + \delta_{n,1}s + \cdots + \delta_{n,n-1}s^{n-1}} &= 1 + \mu_1s + \cdots + \mu_{2n-1}s^{2n-1} + \tilde{\mu}'_{2n}s^{2n} + \dots
\end{aligned}$$

We can now solve for the constants $\alpha_{n,i}, \beta_{n,i}, \gamma_{n,i}$ and $\delta_{n,j}$, $i = 1, \dots, n, j = 1, \dots, n - 1$ by multiplying the numerators of the fractions onto their right hand sides and equating the coefficient to the terms s^i , $i = n + 1, n + 2, \dots, 2n$.

We get the following system of equations

$$\begin{aligned}
0 &= \mu_{n+1} + \mu_1\beta_{n,n} + \mu_2\beta_{n,n-1} + \cdots + \mu_n\beta_{n,1} \\
0 &= \mu_{n+2} + \mu_2\beta_{n,n} + \mu_3\beta_{n,n-1} + \cdots + \mu_{n+1}\beta_{n,1} \\
\dots &= \dots \\
0 &= \mu_{2n} + \mu_n\beta_{n,n} + \mu_{n+1}\beta_{n,n-1} + \cdots + \mu_{2n-1}\beta_{n,1}.
\end{aligned}$$

This is the same as

$$\begin{aligned}
-\mu_{n+1} &= \mu_1\beta_{n,n} + \mu_2\beta_{n,n-1} + \cdots + \mu_n\beta_{n,1} \\
-\mu_{n+2} &= \mu_2\beta_{n,n} + \mu_3\beta_{n,n-1} + \cdots + \mu_{n+1}\beta_{n,1} \\
\dots &= \dots \\
-\mu_{2n} &= \mu_n\beta_{n,n} + \mu_{n+1}\beta_{n,n-1} + \cdots + \mu_{2n-1}\beta_{n,1}.
\end{aligned}$$

By Cramér's rule,

$$\beta_{n,i} = \frac{\begin{vmatrix} \mu_1 & \dots & -\mu_{n+1} & \dots & \mu_n \\ \mu_2 & \dots & -\mu_{n+2} & \dots & \mu_{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ \mu_n & \dots & -\mu_{2n} & \dots & \mu_{2n-1} \end{vmatrix}}{\begin{vmatrix} \mu_1 & \mu_2 & \dots & \mu_n \\ \mu_2 & \mu_3 & \dots & \mu_{n+1} \\ \dots & \dots & \dots & \dots \\ \mu_n & \mu_{n+1} & \dots & \mu_{2n-1} \end{vmatrix}}.$$

In particular, for $\beta_{n,n}$ we get that

$$\beta_{n,n} = (-1)^n \frac{\psi_{n+1}}{\phi_n},$$

where ψ_n and ϕ_n are the Hankel determinants defined by

$$\phi_n = \begin{vmatrix} \mu_1 & \mu_2 & \dots & \mu_n \\ \mu_2 & \mu_3 & \dots & \mu_{n+1} \\ \dots & \dots & \dots & \dots \\ \mu_n & \mu_{n+1} & \dots & \mu_{2n-1} \end{vmatrix} \quad \text{and} \quad \psi_n = \begin{vmatrix} \mu_2 & \mu_3 & \dots & \mu_n \\ \mu_3 & \mu_4 & \dots & \mu_{n+1} \\ \dots & \dots & \dots & \dots \\ \mu_n & \mu_{n+1} & \dots & \mu_{2n-2} \end{vmatrix} \quad (4)$$

for all $n = 1, 2, 3, \dots$ for ϕ_n and for $n = 2, 3, \dots$ for ψ_n . Thus

$$\beta_{n,n} \phi_n = (-1)^{n-1} \psi_{n+1}.$$

Similarly, equating for $\frac{C_{2n-1}(s)}{D_{2n-1}(s)}$ we get that

$$\gamma_{n,n} \psi_n = (-1)^{n-1} \phi_n, \quad n = 1, 2, \dots$$

with $\psi_1 = 1$.

Inserting the expression $\beta_{n,n} = c_2 c_4 \cdots c_{2n}$ and $\gamma_{n,n} = c_1 c_3 \cdots c_{2n-1}$ we get that

$$\begin{aligned} c_2 c_4 \cdots c_{2n} \phi_n &= (-1)^n \psi_{n+1} \\ c_1 c_3 \cdots c_{2n-1} \psi_n &= (-1)^{n-1} \phi_n. \end{aligned}$$

We can extract the following information from these equations. If the moment generating function is an n 'th order rational function, then the corresponding continued fraction has at most $2n$ non-zero terms $c_0 = 1, c_1, c_2, \dots, c_{2n}$

and $c_\nu = 0$ for $\nu > 2n$. Then $\phi_{n+1} = 0$ by the second equation and consequently $\psi_{n+2} = 0$ by the first. Hence all higher order Hankel determinants $\phi_\nu = 0$ for $\nu > n$ and $\psi_\nu = 0$ for $\nu > n + 1$. Furthermore, we can retrieve the constants c_1, c_2, \dots , in terms of the Hankel matrices by defining $\phi_0 = 1$ and

$$c_1 = \phi_1, \quad c_{2n} = -\frac{\psi_{n+1}\phi_{n-1}}{\psi_n\phi_n}, \quad c_{2n+1} = -\frac{\phi_{n+1}\psi_n}{\psi_{n+1}\phi_n}$$

whenever the coefficients are non-zero.

The term $\alpha_{n,n}$ can similarly be calculated by considering the same equations as earlier but equating the terms of s^i , $i = n, n + 2, \dots, 2n - 1$ instead of s^i , $i = n + 1, n + 2, \dots, 2n$. Thus we get that

$$\alpha_{n,n} = \frac{\begin{vmatrix} 1 & \mu_1 & \mu_2 & \dots & \mu_n \\ \mu_1 & \mu_2 & \mu_3 & \dots & \mu_{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ \mu_n & \mu_{n+1} & \mu_{n+2} & \dots & \mu_{2n-1} \end{vmatrix}}{\begin{vmatrix} \mu_2 & \mu_3 & \dots & \mu_n \\ \mu_3 & \mu_4 & \dots & \mu_{n+1} \\ \dots & \dots & \dots & \dots \\ \mu_n & \mu_{n+1} & \dots & \mu_{2n-1} \end{vmatrix}}.$$

If the moment generating function is a rational function of order n , then $\alpha_{n,n}$ is the coefficient to s^n in the numerator and must be zero. Hence we have also proved that the Hankel determinant

$$H_n = \begin{vmatrix} 1 & \mu_1 & \mu_2 & \dots & \mu_n \\ \mu_1 & \mu_2 & \mu_3 & \dots & \mu_{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ \mu_n & \mu_{n+1} & \mu_{n+2} & \dots & \mu_{2n-1} \end{vmatrix}$$

is 0 when n is the order of the rational function. The rank of the matrix corresponding to H_n is $n - 1$ since the lower right sub-determinant is different from zero according to the analysis above. Hence $H_{n-1} \neq 0$ and by a similar argument we conclude that all $H_m \neq 0$ when $m < n$. Thus the order of the matrix-exponential distribution can be checked through the verification of the determinants to be the first time they are different from zero.

We collect these results in the following theorem, which shall turn out to play an important role in the proof for our main characterization theorem.

Theorem 3.1 Consider a matrix-exponential distributed random variable X with reduced moments $\mu_i = \mathbb{E}(X^i)/i!$. Then the rational moment generating function of X can be written as a finite and regular C -continued fraction

$$1 + \frac{c_1 s}{|1} + \frac{c_2 s}{|1} + \dots + \frac{c_{2n}}{|1}.$$

The coefficients c_i can be calculated in terms of the Hankel determinants (4) as follows:

$$c_1 = \phi_1, \quad c_{2n} = -\frac{\psi_{n+1}\phi_{n-1}}{\psi_n\phi_n}, \quad c_{2n+1} = -\frac{\phi_{n+1}\psi_n}{\psi_{n+1}\phi_n},$$

where $\phi_0 = 1$. The Hankel determinants $\phi_m = 0$ for $m > n$ and $\psi_m = 0$ for $m > n + 1$.

Example 3.1 Consider a distribution of a random variable X having a rational moment generating function $M(s)$ given by

$$M(s) = \frac{f_1 s + g_2}{s^2 + g_1 s + g_2}.$$

The moment generating function has a power series expansion

$$M(s) = 1 + \sum_{i=1}^{\infty} \mu_i s^i.$$

We now consider the continued fraction corresponding to the power expansion. Since it is a power expansion of a rational function, the corresponding continued fraction is finite. In the case of a regular C -continued fraction, where all non-trivial terms are of the form $a_i s$, we need $i = 4$ terms in order to express the given rational function in terms of a continued fraction.

By the remark above, we obtain the coefficients $a_i, i = 1, \dots, 4$ as follows:

$$\begin{aligned} a_1 &= \phi_1 = \mu_1 \\ a_2 &= -\frac{\psi_2 \phi_0}{\psi_1 \phi_1} = -\frac{\mu_2}{\mu_1} \\ a_3 &= -\frac{\phi_2 \psi_1}{\psi_2 \phi_1} = -\frac{\mu_1 \mu_3 - \mu_2^2}{\mu_1 \mu_2} \\ a_4 &= -\frac{\psi_3 \phi_1}{\phi_2 \psi_2} = -\frac{\mu_1 (\mu_2 \mu_4 - \mu_3^2)}{(\mu_1 \mu_3 - \mu_2^2) \mu_2} \end{aligned}$$

Then

$$M(s) = 1 + \frac{a_1 s}{|1|} + \frac{a_2 s}{|1|} + \frac{a_3 s}{|1|} + \frac{a_4 s}{|1|} = 1 + \frac{a_1 s}{1 + \frac{a_2 s}{1 + \frac{a_3 s}{1 + a_4 s}}}.$$

Substituting the expressions for μ_i into the a_i we get after a little algebra, that

$$M(s) = \frac{f_0^* s^2 + f_1^* s + g_2^*}{g_0^* s^2 + g_1^* s + g_2^*},$$

where

$$\begin{aligned} f_0^* &= \mu_2 \mu_4 - \mu_3^2 - \mu_1^2 \mu_4 + 2\mu_1 \mu_2 \mu_3 - \mu_2^3 \\ f_1^* &= -\mu_1 \mu_4 + \mu_3 \mu_2 + \mu_1^2 \mu_3 - \mu_1 \mu_2^2 \\ g_2^* &= \mu_1 \mu_3 - \mu_2^2 \\ g_0^* &= \mu_2 \mu_4 - \mu_3^2 \\ g_1^* &= -\mu_1 \mu_4 + \mu_3 \mu_2. \end{aligned}$$

Here we recognize in special

$$f_0^* = \begin{vmatrix} 1 & \mu_1 & \mu_2 \\ \mu_1 & \mu_2 & \mu_3 \\ \mu_2 & \mu_3 & \mu_4 \end{vmatrix}, \quad g_2^* = \begin{vmatrix} \mu_1 & \mu_2 \\ \mu_2 & \mu_3 \end{vmatrix}, \quad g_0^* = \begin{vmatrix} \mu_2 & \mu_3 \\ \mu_3 & \mu_4 \end{vmatrix}.$$

Thus, if X has two-dimensional rational function as moment generating function (or equivalently as Laplace transform), then $f_0^* = 0$.

4 Multivariate matrix-exponential distributions

We define multivariate matrix-exponential distributions as a natural extension of the univariate case.

Definition 4.1 *A non-negative random vector $\mathbf{X} = (X_1, \dots, X_n)$ of dimension n is said to have multivariate matrix-exponential distribution if the joint Laplace transform $L(\mathbf{s}) = \mathbb{E}[\exp(-\langle \mathbf{s}, \mathbf{X} \rangle)]$ is a multi-dimensional rational function, that is, a fraction between two multi-dimensional polynomials. Here $\langle \cdot, \cdot \rangle$ denotes the inner product in \mathbb{R}^n and $\mathbf{s} = (s_1, \dots, s_n)$. This class of distributions is denoted MVME.*

Our main theorem characterizes the class of MVME.

Theorem 4.1 *A vector $\mathbf{X} = (X_1, \dots, X_n)$ follows a multivariate matrix-exponential distribution if and only if $\langle \mathbf{a}, \mathbf{X} \rangle = \sum_{i=1}^n a_i X_i$ has a univariate matrix-exponential distribution for all non-negative vectors $\mathbf{a} \neq \mathbf{0}$.*

Proof: Suppose that \mathbf{X} has multivariate matrix-exponential distribution. Then using

$$\mathbb{E}[\exp(-s \langle \mathbf{a}, \mathbf{X} \rangle)] = \mathbb{E}[\exp(-\langle s\mathbf{a}, \mathbf{X} \rangle)]$$

we conclude that the left hand side is a rational function of s since the right hand side is, by definition, rational in $s\mathbf{a}$.

Now suppose that $\langle \mathbf{a}, \mathbf{X} \rangle$ has a rational Laplace transform, and hence also moment generating function, for all non-negative $\mathbf{a} \neq \mathbf{0}$. The dimensions of the representations for $\langle \mathbf{a}, \mathbf{X} \rangle$ are bounded by some m . Assume the contrary. The dimension of any distribution will be unaffected by normalizing \mathbf{a} by $\mathbf{a}\mathbf{e}$. Since \mathbf{a} is non-negative we can restrict the attention to the compact set (simplex) $\{\mathbf{a} \geq 0 : \mathbf{a}\mathbf{e} = 1\}$. If the dimension is unbounded, then there exists a sequence $\mathbf{a}_n \rightarrow \mathbf{a}_0$ in the simplex such that the corresponding dimension goes to infinity, contradicting the assumption that $\langle \mathbf{a}_0, \mathbf{X} \rangle$ has a rational Laplace transform. In what follows we shall assume that the dimensions of the representations are always of order m , though they may not be minimal.

Let $\mu_i(\mathbf{a}) = \mathbb{E}(\langle \mathbf{a}, \mathbf{X} \rangle^i)/i!$ denote the reduced moments of $\langle \mathbf{a}, \mathbf{X} \rangle$ as a function of \mathbf{a} . Then $\mu_i(\mathbf{a})$ is a sum of i -dimensional (multidimensional) monomials in \mathbf{a} . From Theorem 3.1 we get that the moment generating function of $\langle \mathbf{a}, \mathbf{X} \rangle$ can be written as a finite regular C -continued fraction of order at most $2m$. The Hankel determinants are again sums of monomials in \mathbf{a} , so the coefficients in the continued fraction are rational functions in \mathbf{a} . Hence, collapsing the continued fraction to a rational function, we conclude that the moment generating function, and hence its Laplace transform, is indeed a rational function in \mathbf{a} . The Hankel determinant $\phi_n(\mathbf{a})$ may vanish, but at most on a set of measure zero. The continuity of the multidimensional Laplace transform ensures that the coefficients of the univariate Laplace transform on the null set is obtained as a limit of the coefficients of the univariate Laplace transform outside this set.

It follows, that the functions $f_i(\mathbf{a})$ and $g_i(\mathbf{a})$ are also rational functions in \mathbf{a} .

□

Corollary 4.1 *Let $\mathbf{X} = (X_1, \dots, X_n)$ have a MVME distribution and let \mathbf{A} be a non-negative $m \times n$ matrix. Then $\mathbf{Y} = \mathbf{A}\mathbf{X}$ has a MVME distribution. In particular, all marginal distributions are again matrix-exponentially distributed.*

Proof: According to Theorem 4.1, \mathbf{Y} is MVME if and only if $\langle \mathbf{b}, \mathbf{Y} \rangle$ has a matrix-exponential distribution for all non-negative $\mathbf{b} \neq \mathbf{0}$. Now $\langle \mathbf{b}, \mathbf{A}\mathbf{X} \rangle = \langle \mathbf{b}\mathbf{A}, \mathbf{X} \rangle$ and hence has a matrix-exponential distribution.

□

Theorem 4.2 *Let $\mathbf{X} = (X_1, \dots, X_n)$ follow a multivariate matrix-exponential distribution. Then we may write its moment generating function for $\langle \mathbf{a}, \mathbf{X} \rangle$ as*

$$\frac{\tilde{f}_1(\mathbf{a})s^{m-1} + \tilde{f}_2(\mathbf{a})s^{m-2} + \dots + \tilde{f}_{m-1}(\mathbf{a})s + 1}{\tilde{g}_0(\mathbf{a})s^m + \tilde{g}_1(\mathbf{a})s^{m-1} + \dots + \tilde{g}_{m-1}(\mathbf{a})s + 1},$$

where the terms $\tilde{f}_i(\mathbf{a})$ and $\tilde{g}_i(\mathbf{a})$ are sums of monomials in \mathbf{a} of order $m - i$.

Proof: From Theorem 4.1 we know that the moment generating function of $\langle \mathbf{a}, \mathbf{X} \rangle$ can be written as

$$\frac{f_1^*(\mathbf{a})s^{m-1} + f_2^*(\mathbf{a})s^{m-2} + \dots + f_{m-1}^*(\mathbf{a})s + f_m^*(\mathbf{a})}{g_0^*(\mathbf{a})s^m + g_1^*(\mathbf{a})s^{m-1} + \dots + g_{m-1}^*(\mathbf{a})s + g_m^*(\mathbf{a})},$$

where $f_i^*(\mathbf{a})$ and $g_i^*(\mathbf{a})$ are sums of monomials of order $m(m+1) - i$. To see this, we notice that the order of the monomials in $g_0^*(\mathbf{a})$ is given by the sum of the indices of the diagonal elements of ψ_{m+1} in (4) which amounts to $m(m+1)$. Here $f_m^*(\mathbf{a}) = g_m^*(\mathbf{a})$ are of order $m(m+1) - m = m^2$. The assertion of the theorem is equivalent to divisibility of all coefficients $f_i^*(\mathbf{a})$ and $g_i^*(\mathbf{a})$ by $f_m^*(\mathbf{a})$.

In order to prove the divisibility we proceed as follows. Dividing through the numerator and denominator by $f_m^*(\mathbf{a})$ we obtain an expression

$$\frac{\tilde{f}_1(\mathbf{a})s^{m-1} + \tilde{f}_2(\mathbf{a})s^{m-2} + \dots + \tilde{f}_{m-1}(\mathbf{a})s + 1}{\tilde{g}_0(\mathbf{a})s^m + \tilde{g}_1(\mathbf{a})s^{m-1} + \dots + \tilde{g}_{m-1}(\mathbf{a})s + 1} = 1 + \mu_1(\mathbf{a})s + \mu_2(\mathbf{a})s^2 + \dots, \quad (5)$$

where $\tilde{f}_i(\mathbf{a})$ and $\tilde{g}_i(\mathbf{a})$ are now rational functions in \mathbf{a} . This equation is similar to (3) with $\alpha_{m,m} = 0$, $\alpha_{m,i} = \tilde{f}_{m-i}(\mathbf{a})$ and $\beta_{m,i} = \tilde{g}_{m-i}(\mathbf{a})$ which is solved by considering (4).¹

Write $\tilde{g}_i(\mathbf{a}) = P_{m-i}(\mathbf{a}) + E_{m-i}(\mathbf{a})$ where $P_{m-i}(\mathbf{a})$ is a sum of all, if any, $m-i$ 'th order monomials appearing in the expression for $\tilde{g}_i(\mathbf{a})$ while $E_{m-i}(\mathbf{a}) = \tilde{g}_i(\mathbf{a}) - P_{m-i}(\mathbf{a})$. Let $\boldsymbol{\mu}_m(\mathbf{a}) = (\mu_{m+1}(\mathbf{a}), \dots, \mu_{2m}(\mathbf{a}))'$, $\phi_n = \phi_n(\mathbf{a})$ the Hankel matrix (4) now depending of \mathbf{a} , $\mathbf{P}_m(\mathbf{a}) = (P_m(\mathbf{a}), \dots, P_1(\mathbf{a}))$ and $\mathbf{E}_m(\mathbf{a}) = (E_m(\mathbf{a}), \dots, E_1(\mathbf{a}))$. Then

$$-\boldsymbol{\mu}_m(\mathbf{a}) = \phi_m(\mathbf{a})\mathbf{P}_m(\mathbf{a}) + \phi_m(\mathbf{a})\mathbf{E}_m(\mathbf{a}).$$

Consider the j 'th equation. Here $\mu_{m+j}(\mathbf{a})$ is a sum of monomials of order $m+j$ as are the corresponding terms of $\phi_m(\mathbf{a})\mathbf{P}_m(\mathbf{a})$. Since $\phi_m(\mathbf{a})\mathbf{E}_m(\mathbf{a})$ is rational in \mathbf{a} and do not contain monomial terms of order $m+j$, we conclude by coefficient matching that

$$\phi_m(\mathbf{a})\mathbf{E}_m(\mathbf{a}) = \mathbf{0}.$$

Since $\phi_m(\mathbf{a})$ is non-singular we get that $\mathbf{E}_m(\mathbf{a}) = \mathbf{0}$ and hence all $\tilde{g}_i(\mathbf{a})$ are sums of monomials. From (5) we also see that the $\tilde{f}_i(\mathbf{a})$ are sums of monomials by multiplying both sides of the equation with the numerator and matching coefficients.

□

Corollary 4.2 *The number of free (reduced) moments for $m, n = 2$ is at most 7 (out of 9 potential).*

Proof: Define reduced cross-moments, $\kappa_{i,j} = \mathbb{E}(X_1^i X_2^j) / (i!j!)$. In particular, $\kappa_{i,0}$ and $\kappa_{0,i}$ are the usual i 'th order reduced moments of X_1 and X_2 respectively. Then

$$\mu_i = \sum_{j=0}^i a_1^j a_2^{i-j} \kappa_{j,i-j}.$$

From Theorem 4.2 we know that $\mu_2 - \mu_1^2$ divides $\mu_3 - \mu_1\mu_2$ and $\mu_3\mu_1 - \mu_2^2$ respectively. Thus there are constants $c_{i,j}$, such that

$$\mu_3 - \mu_1\mu_2 = (c_{1,0}a_1 + c_{0,1}a_2) (\mu_2 - \mu_1^2) \quad (6)$$

$$\mu_3\mu_1 - \mu_2^2 = (c_{2,0}a_1^2 + c_{1,1}a_1a_2 + c_{0,2}a_2^2) (\mu_2 - \mu_1^2). \quad (7)$$

¹Bo:check

By coefficient matching in (6) we get

$$\begin{aligned}
\kappa_{3,0} - \kappa_{1,0}\kappa_{2,0} &= c_{1,0} (\kappa_{2,0} - \kappa_{1,0}^2) \\
\kappa_{2,1} - \kappa_{1,0}\kappa_{1,1} - \kappa_{0,1}\kappa_{2,0} &= c_{0,1} (\kappa_{2,0} - \kappa_{1,0}^2) + c_{1,0} (\kappa_{1,1} - \kappa_{1,0}\kappa_{0,1}) \\
\kappa_{0,3} - \kappa_{0,1}\kappa_{0,2} &= c_{0,1} (\kappa_{0,2} - \kappa_{0,1}^2) \\
\kappa_{1,2} - \kappa_{0,1}\kappa_{1,1} - \kappa_{1,0}\kappa_{0,2} &= c_{1,0} (\kappa_{0,2} - \kappa_{0,1}^2) + c_{0,1} (\kappa_{1,1} - \kappa_{0,1}\kappa_{1,0}).
\end{aligned}$$

When $\kappa_{2,0} \neq \kappa_{1,0}^2$ and $\kappa_{0,2} \neq \kappa_{0,1}^2$ we see that $\kappa_{1,2}$ and $\kappa_{2,1}$ are uniquely given in terms of the other κ 's. Equation (7) establishes a connection between $\kappa_{1,2}$ and $\kappa_{2,1}$, which is compatible with the restrictions of (6).

Now, $\kappa_{2,0} = \kappa_{1,0}^2 \Rightarrow \kappa_{3,0} = \kappa_{1,0}^3$ and $\kappa_{0,2} = \kappa_{0,1}^2 \Rightarrow \kappa_{0,3} = \kappa_{0,1}^3$, which completes the proof. □

Inspired by Theorem 4.1 we propose the following definition of a multivariate phase-type distribution.

Definition 4.2 *A vector $\mathbf{X} = (X_1, \dots, X_n)$ has a multivariate phase-type distribution (MVPH) if $\langle \mathbf{a}, \mathbf{X} \rangle$ has a (univariate) phase-type distribution for all non-negative $\mathbf{a} \neq \mathbf{0}$.*

The following definition is a natural extension of the MPH* structure to matrix-exponential distributions.

Definition 4.3 *Let MME^* be the subclass of MVME, where $\langle \mathbf{a}, \mathbf{X} \rangle$ has representation $(\boldsymbol{\gamma}, (\Delta(K\mathbf{a}))^{-1}T, \mathbf{t})$, where $\boldsymbol{\gamma}$, K and T are constant vector and matrices, and $\mathbf{t} = -(\Delta(K\mathbf{a}))^{-1}T\mathbf{e}$. We say that the triple $(\boldsymbol{\gamma}, T, K)$ is a MME^* representation of the multivariate distribution.*

It is an open problem whether MME^* is a strict subset or equals the class of MVME. However,

Theorem 4.3 *There exists MVME distributions where the MVME order is strictly less than the MME^* order.*

Proof: The proof is based on the non-existence of a three dimensional MME^* representation of Krishnamoorthy and Parthasarathy's Multivariate Exponential for $n = 3$. For a discussion of this distribution see section 48.3.3

in Kotz, Balakrishnan & Johnson (2000) (with $\alpha = 1$). The distribution is defined through it's joint Laplace transform

$$|I + R\Delta(\mathbf{s})|^{-1} \quad ,$$

where R is a correlation matrix. To find a representation for $m = 3$ in $\text{MME}^*(3)$ we first parameterize

$$R = \begin{bmatrix} 1 & \rho & \tau \\ \rho & 1 & \eta \\ \tau & \eta & 1 \end{bmatrix} \quad .$$

Then

$$\begin{aligned} \tilde{g}_0 &= a_1 a_2 a_3 (1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2) \\ \tilde{g}_1 &= (a_1 a_2 (1 - \rho^2) + a_1 a_3 (1 - \tau^2) + a_2 a_3 (1 - \eta^2)) \\ \tilde{g}_2 &= (a_1 + a_2 + a_3) \\ \tilde{f}_1 &= 0 \\ \tilde{f}_2 &= 0 \end{aligned}$$

and the Laplace transform of $\langle \mathbf{X}, \mathbf{a} \rangle$ is given by

$$\frac{1}{s^3 \tilde{g}_0 + s^2 \tilde{g}_1 + s \tilde{g}_2 + 1} \quad .$$

Suppose now that we have a MME^* representation (γ, T, K) for this distribution in MME^* . It is immediately clear that we must have $K = I$. From equality of the Laplace transforms we must have

$$\begin{aligned} |T| &= \frac{-1}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2} \\ \begin{vmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{vmatrix} &= \frac{1}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2} \\ \begin{vmatrix} T_{11} & T_{13} \\ T_{31} & T_{33} \end{vmatrix} &= \frac{1}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2} \\ \begin{vmatrix} T_{22} & T_{23} \\ T_{32} & T_{33} \end{vmatrix} &= \frac{1}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2} \end{aligned}$$

$$T_{11} = -\frac{1 - \eta^2}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2}$$

$$T_{22} = -\frac{1 - \tau^2}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2}$$

$$T_{33} = -\frac{1 - \rho^2}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2}$$

$$(\gamma_1 + \gamma_2) \begin{vmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{vmatrix} - \gamma_1 \begin{vmatrix} T_{12} & T_{13} \\ T_{22} & T_{23} \end{vmatrix} + \gamma_2 \begin{vmatrix} T_{11} & T_{13} \\ T_{21} & T_{23} \end{vmatrix} = 0 \quad (8)$$

$$(\gamma_1 + \gamma_3) \begin{vmatrix} T_{11} & T_{13} \\ T_{31} & T_{33} \end{vmatrix} + \gamma_1 \begin{vmatrix} T_{12} & T_{13} \\ T_{32} & T_{33} \end{vmatrix} + \gamma_3 \begin{vmatrix} T_{11} & T_{12} \\ T_{31} & T_{32} \end{vmatrix} = 0$$

$$(\gamma_2 + \gamma_3) \begin{vmatrix} T_{22} & T_{23} \\ T_{32} & T_{33} \end{vmatrix} + \gamma_2 \begin{vmatrix} T_{21} & T_{23} \\ T_{31} & T_{33} \end{vmatrix} - \gamma_3 \begin{vmatrix} T_{21} & T_{22} \\ T_{31} & T_{32} \end{vmatrix} = 0$$

$$\gamma_1(T_{11} + T_{12} + T_{13}) = 0$$

$$\gamma_2(T_{21} + T_{22} + T_{23}) = 0$$

$$\gamma_3(T_{31} + T_{32} + T_{33}) = 0 .$$

We must have at least one $\gamma_i \neq 0$. Due to the symmetry we can without loss of generality take $\gamma_1 \neq 0$. We then have $T_{11} + T_{12} + T_{13} = 0$. Now denote the coefficients of γ_j in the i 'th equation of (8) by C_{ij} . We get

$$C_{11} = T_{11}T_{22} - T_{12}T_{21} - T_{12}T_{23} + T_{13}T_{22}$$

$$C_{12} = T_{11}T_{22} - T_{12}T_{21} + T_{11}T_{23} - T_{13}T_{21}$$

$$C_{21} = T_{11}T_{33} - T_{13}T_{31} + T_{12}T_{33} - T_{13}T_{32}$$

$$C_{23} = T_{11}T_{33} - T_{13}T_{31} + T_{11}T_{32} - T_{12}T_{31}$$

$$C_{32} = T_{22}T_{33} - T_{23}T_{32} + T_{21}T_{33} - T_{23}T_{31}$$

$$C_{33} = T_{22}T_{33} - T_{23}T_{32} - T_{21}T_{32} + T_{22}T_{31} .$$

By insertion of $T_{11} = -T_{12} - T_{13}$ in the first four equations we get

$$C_{11} = -T_{12}(T_{21} + T_{22} + T_{23})$$

$$C_{12} = -(T_{12} + T_{13})(T_{21} + T_{22} + T_{23})$$

$$C_{21} = -T_{13}(T_{31} + T_{23} + T_{33})$$

$$C_{23} = -(T_{12} + T_{13})(T_{31} + T_{32} + T_{33})$$

As $\gamma_2 C_{12} = \gamma_3 C_{23} = 0$ we must have

$$\begin{aligned}\gamma_1 C_{11} &= \gamma_1 T_{12}(T_{21} + T_{22} + T_{23}) = 0 \\ \gamma_1 C_{21} &= \gamma_1 T_{13}(T_{31} + T_{32} + T_{33}) = 0 \quad .\end{aligned}$$

We cannot have both T_{12} and T_{13} equal to zero as we cannot have both $T_{21} + T_{22} + T_{23} = 0$ and $T_{31} + T_{32} + T_{33} = 0$. Now again due to the symmetry we can assume without loss of generality that $T_{13} = 0$ and $T_{21} + T_{22} + T_{23} = 0$ while $T_{12} \neq 0$ and $T_{31} + T_{32} + T_{33} \neq 0$ to give

$$\begin{aligned}C_{32} &= -T_{23}(T_{31} + T_{32} + T_{33}) \\ C_{33} &= -(T_{21} + T_{23})(T_{31} + T_{32} + T_{33}) \quad .\end{aligned}$$

Since $T_{21} + T_{22} + T_{23} = 0$ we have $\gamma_3 = 0$, and as we cannot have both T_{13} and T_{23} equal to zero we conclude that $\gamma_2 = 0$. Now finally we must have

$$T_{11}T_{33} = \frac{T_{11}}{1-\eta^2} = \frac{T_{33}}{1-\rho^2} = \frac{1}{1+2\rho\tau\eta-\rho^2-\tau^2-\eta^2} \quad ,$$

which is only fulfilled when $\tau = \rho\eta$. Examination of this case reveals that for this special parameter setting we can find a representation in MPH* with

$$T = \begin{bmatrix} -\frac{1}{1-\rho^2} & \frac{1}{1-\rho^2} & 0 \\ \frac{\rho^2}{1-\rho^2} & -\frac{1-\eta^2\rho^2}{(1-\rho^2)(1-\eta^2)} & \frac{1}{1-\eta^2} \\ 0 & \frac{\eta^2}{1-\eta^2} & -\frac{1}{1-\eta^2} \end{bmatrix} \quad .$$

It should be clear that it is possible to get a representation in MME* whenever one of the three equations $\tau = \rho\eta$, $\rho = \tau\eta$, or $\eta = \rho\tau$ is true as the previous choice of $\gamma_1 = 1$ over $\gamma_2 = 1$ or $\gamma_3 = 1$ was arbitrary.

□

Example 4.1 (Marshall and Olkin's Bivariate Exponential)

The bivariate Marshall Olkin distribution Marshall & Olkin (1967) and Kotz, Balakrishnan & Johnson (2000) pp.362-369 is already in Assaf et al. (1984) as Example 5.1 p.699. The joint density for $x_1 \neq x_2$ is given by

$$f(x_1, x_2) = \begin{cases} \lambda_2(\lambda_1 + \lambda_{12})e^{-(\lambda_1 + \lambda_{12})x_1 - \lambda_2 x_2} & \text{for } 0 \leq x_2 < x_1 \\ \lambda_1(\lambda_2 + \lambda_{12})e^{-\lambda_1 x_1 - (\lambda_2 + \lambda_{12})x_2} & \text{for } 0 \leq x_1 < x_2 \end{cases} \quad .$$

There is a singularity on the line $x_1 = x_2$

$$f(x, x) = \lambda_{12}e^{-\lambda_{12}x} .$$

The joint density has Laplace transform

$$\frac{(\lambda_1 + \lambda_2 + \lambda_{12} + s_1 + s_2)(\lambda_1 + \lambda_{12})(\lambda_2 + \lambda_{12}) + s_1 s_2 \lambda_{12}}{(\lambda_1 + \lambda_2 + \lambda_{12} + s_1 + s_2)(\lambda_1 + \lambda_{12} + s_1)(\lambda_2 + \lambda_{12} + s_2)} ,$$

with MPH*(3) representation

$$\left((1, 0, 0), \begin{bmatrix} -(\lambda_1 + \lambda_2 + \lambda_{12}) & \lambda_2 & \lambda_1 \\ 0 & -(\lambda_1 + \lambda_{12}) & 0 \\ 0 & 0 & -(\lambda_2 + \lambda_{12}) \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \right) .$$

□

Example 4.2 (Kibbles's Bivariate Exponential)

This distribution is originally due to Kibble (1941) and is described in Kotz, Balakrishnan & Johnson (2000) p.350 and pp.350-377 as Moran and Downton's Bivariate Exponential. It was used in O'Kinneide (1990) to demonstrate that the order statistics of the components of an MPH* distributed vector \mathbf{X} do not necessarily have a rational Laplace transform. The distribution belongs to a more general system of mixtures, where the components are ME or MVME distributed and the mixing is due to a discrete multivariate distribution of ME type. The density is given by

$$f(x_1, x_2) = \lambda_1 \lambda_2 (1 - p) e^{-(\lambda_1 x_1 + \lambda_2 x_2)} \sum_{i=1}^{\infty} \frac{(\lambda_1 x_1 \lambda_2 p x_2)^{i-1}}{((i-1)!)^2} .$$

We can express $f(x_1, x_2)$ more compactly as $f(x_1, x_2) = \lambda_1 \lambda_2 I_0(2\sqrt{\lambda_1 x_1 \lambda_2 p x_2})$ by using the modified Bessel function of the first kind $I_0(z) = \sum_{j=0}^{\infty} \left(\frac{z}{2j!}\right)^{2j}$. The parameterization in Kotz, Balakrishnan & Johnson (2000) page 371 is obtained with $\rho = p$, $\theta_1 = \lambda_1 p$, and $\theta_2 = \lambda_2 p$.

The Laplace transform of $\langle \mathbf{a}, \mathbf{X} \rangle$ is

$$\frac{p\lambda_1\lambda_2}{a_1 a_2 s^2 + (a_2 \lambda_1 + a_1 \lambda_2) s + p\lambda_1 \lambda_2} .$$

The MME* representation of this distribution is

$$\left((1, 0) \begin{bmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2(1-p) & -\lambda_2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) .$$

□

Finally we show how the multivariate matrix–exponential distribution may naturally arise in applications such as renewal theory.

Lemma 4.1 *Let A_t and R_t be the age and residual life time respectively in a stationary renewal process with inter-arrival time density $f(x)$. The joint distribution of A_t and R_t is given by $\frac{f(x+y)}{\int_0^\infty xf(x)dx}$.*

Proof: The joint distribution of (A_t, R_t) in a renewal process is given by

$$P(R_t \geq x, A_t < y) = \int_{t-y}^t (1 - F(t+x-u))dU(u) \quad ,$$

see e.g. Yanushkevichius (1995). We obtain the result, inserting $1/\mathbb{E}X$ as the stationary renewal density and differentiating. □

Theorem 4.4 *In a stationary renewal process with matrix–exponentially distributed inter–arrival times with representation $(\boldsymbol{\alpha}, C)$, the joint distribution of age and residual life is a MME* with representation*

$$\left(\left(\frac{\boldsymbol{\alpha}(-C)^{-1}}{\mu_1}, \mathbf{0} \right), \begin{bmatrix} C & -C \\ 0 & C \end{bmatrix}, \begin{bmatrix} \mathbf{e} & \mathbf{0} \\ \mathbf{0} & \mathbf{e} \end{bmatrix} \right) \quad , \quad (9)$$

where m is the dimension of C .

Proof: We denote the density of the ME distribution with representation $(\boldsymbol{\alpha}, C)$ by $f(x)$. The corresponding mean is denoted by μ_1 . Using Theorem 4.1 we see that for $a_1 > 0$ and $a_2 > 0$, the random variable $Z = \langle \mathbf{a}, \mathbf{X} \rangle$, has density $g(z)$ given by

$$g(z) = \frac{1}{\mu_1} \int_0^{\frac{z}{a_1}} \boldsymbol{\alpha} \exp \left(C \left(x_1 + \frac{z - a_1 x_1}{a_2} \right) \right) \mathbf{c} dx_1 \frac{1}{a_2} \quad .$$

For $a_1 \neq a_2$ we get

$$\begin{aligned} g(z) &= \frac{1}{\mu_1} \boldsymbol{\alpha} C^{-1} \frac{1}{1 - \frac{a_1}{a_2}} \left[\exp \left(C \frac{z}{a_1} \right) \exp \left(-C \frac{z}{a_2} \right) - I \right] \exp \left(C \frac{z}{a_2} \right) \frac{\mathbf{c}}{a_2} \\ &= \frac{1}{\mu_1} \frac{1}{a_2 - a_1} \left(F \left(\frac{z}{a_1} \right) - F \left(\frac{z}{a_2} \right) \right) \quad , \end{aligned}$$

which is also the density of the ME distribution with the representation given by (9). This can be seen by direct evaluation of the latter. For $a_1 = a_2$ we get $g(z) = \frac{zf(z/a)}{a\mu_1}$. Again, by direct verification, we also obtain this density by evaluating (9) for $a_1 = a_2$.

□

5 Conclusions

In this article we have analyzed the general class of multivariate matrix–exponential distributions (MVME) defined as distributions with a rational multi–dimensional Laplace transform. This class of distributions generalizes several classes of multivariate exponential and Gamma distributions, which we shall treat in more detail in a forthcoming paper.

A main characterization theorem proves that distributions of MVME type are those which projections in any direction are univariate matrix–exponential distributions. An intimate connection to the theory of Hankel matrices, continued fractions, and the moment problem is used in the proof of this main theorem. Based on the theory of continued fractions we prove an important result concerning the order of matrix–exponential distributions and the vanishing of the Hankel determinant of the reduced moments. This result has previously been pointed out by van de Liefvoort (1990), however, we have not been able to pinpoint a proof in the literature.

Multivariate distributions may appear in a variety of situations, and since particularly phase–type distributions and matrix–exponential distributions, have turned out to be useful in the modeling of complex stochastic models, it is natural to consider such distributions in a broader generality. A trivial example where an MVME appears naturally is the joint distribution of the age and residual lifetime in a stationary renewal process.

Acknowledgments

The authors are grateful for the support by *Otto Mønstedts foundation, Danish research council for technology and production sciences grant no. 247-07-0090* and *Sistema Nacional de Investigadores, Mexico*.

References

- Asmussen, S. & Bladt, M. (1997). “Renewal Theory and Queueing Algorithms for Matrix-Exponential Distributions”. In Alfa, A. & Chakravathy, S. R., editors, *Matrix-analytic methods in stochastic models*, pages 313–341.
- Assaf, D.; Langberg, N. A.; Savits, T. H. & Shaked, M. (1984). “Multivariate Phase-Type Distributions”. *Operations Research*, 32(3):688–702.
- Kibble, W. F. (1941). “A two-variate gamma type distribution”. *Sankhy-a*, 5:137–150.
- Kotz, S.; Balakrishnan, N. & Johnson, N. L. (2000). *Continuous Multivariate Distributions*. John Wiley and Sons.
- Kulkarni, V. G. (1989). “A new class of multivariate phase type distributions”. *Operations Research*, 37(1):151–158.
- Marshall, A. W. & Olkin, I. (1967). “A Multivariate Exponential Distribution”. *Journal of the American Statistical Association*, 62(317):30–44.
- O’Cinneide, C. A. (1990). “On the Limitations of Multivariate Phsae-type Families”. *Operations Research*, 38(3):519–526.
- Perron, O. (1957). *Die lehre von den Kettenbrüchen*. B. G. Teubner Verlagsgesellschaft, Stuttgart.
- van de Liefvoort, A. (1990). “The moment problem for continuous distributions”. Technical Report WP-CM-1990-02, University of Missouri, Kansas City.
- Yanushkevichius, R. (1995). “Stability of one characterization by the properties of the renewal proces s”. *Lithuanian Mathematical Journal*, 35(2):158–162.