

# Explicit Identification of ME(3) Distributions

István Kolossváry

**Supervisor:** Miklós Telek  
Professor  
BME,  
Department of Telecommunications

**BME  
2011**

# Contents

<b>Acknowledgements</b>	<b>2</b>
<b>1 Introduction</b>	<b>3</b>
<b>2 Approximation of queuing systems</b>	<b>4</b>
2.1 PH and ME distributions . . . . .	5
2.1.1 Quasi-Birth-Death (QBD) processes . . . . .	8
2.2 Moment matching . . . . .	10
<b>3 General principles of the method</b>	<b>12</b>
3.1 The general approach . . . . .	12
<b>4 ME(3) distributions</b>	<b>13</b>
4.1 Two different eigenvalues . . . . .	14
4.2 One real and a complex conjugate pair of eigenvalues . . . . .	17
4.3 Visualization of ME(3) . . . . .	17
<b>5 Higher order ME distributions</b>	<b>18</b>
5.1 An explicit case ME(4) distribution . . . . .	19
5.2 Two real eigenvalues and a complex conjugate pair . . . . .	22
<b>6 Example</b>	<b>24</b>
6.1 Approximation of Weibull distributions . . . . .	25
<b>7 Conclusions and future work</b>	<b>28</b>
<b>References</b>	<b>29</b>

## Acknowledgements

I would like to thank Miklos Telek for drawing my attention to the field of queuing systems. His insight and useful remarks made the writing of this paper much easier.

# 1 Introduction

The analysis of queuing systems with generally distributed inter-arrival and/or service time is difficult in many cases. In the absence of efficient methods for performance analysis, simpler models are used that try to capture the behavior of the original system. In practice different distributions are used for approximation.

One such class of distributions are the phase type (PH) distributions. These distributions have a nice Markovian stochastic interpretation and the generator matrix of the overall Markov chain describing the behavior of the queuing system has a special structure. Due to this special structure efficient numerical methods, referred to as matrix geometric methods, are available to determine the stationary and transient behavior of the approximating system.

The computational complexity of these methods depend on the size of the PH distributions. The goal of the approximation is to reach greater accuracy without increasing the dimensions of the PH distributions too much. From this perspective the class of PH distributions is too small. A possible solution is to use a wider class of distributions for the fitting that can still be analyzed with the same efficient numerical methods. Matrix exponential (ME) distributions [1] satisfy these criteria.

ME distributions can be considered as a generalization of PH distributions since the generator of an ME distribution doesn't have to meet the restrictions of a PH distribution. This gives greater flexibility in the fitting so a better fitting accuracy can be achieved without increasing the size of the involved matrices. Furthermore, due to the same matrix exponential form of the density function, the same efficient numerical methods are available to analyze queuing systems with ME distributions. This is why ME distributions are gaining increasing attention in various application fields.

The drawback of the use of ME distributions is that it is difficult to decide whether the result of a fitting algorithm determines a valid ME distribution or not. In [2] a method is proposed based on the Laplace representation of ME distributions and [4] introduces a time domain counterpart that checks the non-negativity of the density function. However, even for order 3 ME distributions (ME(3)) neither of them give explicit conditions for all the cases.

The main aim of this paper is to show that for order 3 ME distributions it is possible to check the non-negative property of the density function in an explicit way in all possible cases. To this end an order reduction approach is introduced with which all the cases can be handled in a similar fashion. We obtain theorems that explicitly identify ME(3) distributions. These theorems

were implemented in *Mathematica*, to help decide whether the result of a fitting algorithm is a valid distribution. As an example, this implementation is used to illustrate the difference of PH and ME fitting on the Weibull distribution. Furthermore, we also study the effectiveness of the proposed approach on higher order ME distributions.

Accordingly, the outline of the subsequent sections is the following:

- We define phase type (PH) and matrix exponential (ME) distributions in Section 2. We discuss the advantages of using these distributions for the approximation of general arrival and/or service times in queuing systems. A matching procedure based on the moments of the general distribution is discussed and we point out its main drawback. Namely, we don't know if the result is a valid density function or not.
- To deal with this problem Section 3 introduces a new general approach to decide the non-negativity of matrix exponential functions, with which we will be able to handle all the different cases of ME(3) distributions in the same manner.
- In Section 4 ME(3) distributions are studied in detail. With the order reduction procedure of Section 3 we deduce necessary and sufficient conditions for ME(3) membership. In all of the cases the conditions can be explicitly checked without having to use any numerical methods.
- The efficiency of the approach is investigated in Section 5 on higher order ME distributions. We will see that there are just a few cases where the approach is able to produce further explicit conditions. The reason why it fails in all other cases is pointed out as well.
- We devote Section 6 to present an example where the result of the fitting is indeed better on the class of ME distributions than just restricting to PH distributions.

## 2 Approximation of queuing systems

In this section matrix exponential (ME), phase type (PH) and a special class of PH, acyclic phase type (APH) distributions are defined. After a brief discussion of their properties we turn our attention to Quasi-Birth-Death processes. They form a general class of queuing systems that can be efficiently analyzed by means of matrix geometric methods which we outline. We show that these methods can be applied when the inter-arrival and/or service time

has PH or ME distribution. We conclude the section by showing transformations between different representations of PH and ME distributions which is key in order to review a moment matching algorithm.

## 2.1 PH and ME distributions

We begin with the formal definition of phase type (PH) distributions.

**Definition 2.1.** Let  $\alpha$  and  $\mathbf{A}$  be a  $1 \times n$  row vector and an  $n \times n$  matrix respectively. We say that a random variable  $X$  is phase type distributed, denoted  $X \sim \text{PH}(\alpha, \mathbf{A})$ , if its cumulative distribution function (CDF) is

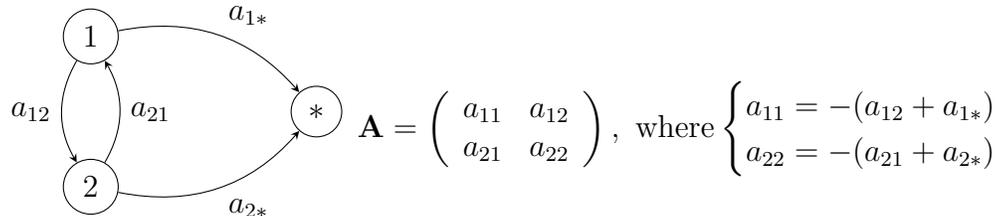
$$F(t) = \mathbf{P}(X < t) = 1 - \alpha e^{\mathbf{A}t} \mathbf{1}, \quad t \geq 0, \quad (1)$$

where  $\alpha$  and  $\mathbf{A}$  satisfy the following conditions:

- $\alpha_i \geq 0$ ,  $\alpha \cdot \mathbf{1} = 1$ , where  $\mathbf{1} = (1, 1, \dots, 1)^T$  is the closing vector,
- $\mathbf{A}_{ii} < 0$ ,  $\mathbf{A}_{ij} \geq 0$  ( $i \neq j$ ),  $\mathbf{A}\mathbf{1} \leq 0$ ,
- $\mathbf{A}$  is non-singular, i.e.  $\exists \mathbf{A}^{-1}$ .

We refer to  $\alpha$  as the initial probability vector and  $\mathbf{A}$  as the generator.

The constraints for  $\alpha$  and  $\mathbf{A}$  give us a very natural stochastic interpretation of the formal definition. Think of an exponential distribution with parameter  $\lambda$  as a Markov chain with two states 1 and \*. The distribution gives us the time needed for the transition from 1 to the absorbing state \* with rate  $\lambda$ . Now let us have  $n$  different transient states  $\{1, 2, \dots, n\}$ . Then a  $\text{PH}(\alpha, \mathbf{A})$  distribution gives the distribution of the elapsed time of the transition from the transient states to the absorbing state \*, assuming that the process started from state  $i$  with probability  $\alpha_i$  and  $\mathbf{A}$  gives the rates of transitions between the different transient states. This interpretation guarantees that a vector matrix pair  $(\alpha, \mathbf{A})$  satisfying the conditions in Definition 2.1 define a valid CDF. Illustration of the stochastic interpretation of a general PH distribution of order two:



Now let us neglect these constraints and this stochastic interpretation. Then  $(\alpha, \mathbf{A})$  might or might not define a valid CDF of the form (1). This way we get a wider class of distributions.

**Definition 2.2.** Let the row vector  $\alpha$  and the non-singular matrix  $\mathbf{A}$  be arbitrary. We say that the pair  $(\alpha, \mathbf{A})$  defines a matrix exponential distribution, if they determine a valid CDF of the form defined in (1).

In two dimensions the class of PH and ME distributions actually coincide. However, in any higher dimension PH distributions form a proper subset of ME distributions of the same order. This means that we can expect to get better results from fitting algorithms when we optimize on the class of ME distributions than if we just restrict to PH distributions. Another important advantage of ME distributions is that those analytic methods which do not make use of the constraints imposed on  $\alpha$  and  $\mathbf{A}$  can also be used to analyze matrix exponential distributions. The major drawback is that there is no guarantee that an arbitrary  $(\alpha, \mathbf{A})$  define a valid CDF.

Instead of the CDF we usually work with the probability density function (PDF). The PDF, Laplace transform and moments of a random variable  $X$  with PH or ME distribution all have explicit forms:

$$f(t) = \frac{d}{dt}F(t) = \alpha e^{\mathbf{A}t}(-\mathbf{A})\mathbf{1}, \quad (2)$$

$$f^*(s) = \mathbf{E}(e^{-sX}) = \alpha(s\mathbf{I} - \mathbf{A})^{-1}(-\mathbf{A})\mathbf{1} = \sum_{i=0}^{\infty} s^i (-1)^i \alpha(-\mathbf{A})^{-i}\mathbf{1}, \quad (3)$$

$$\mu_n = \mathbf{E}(X^n) = n! \alpha(-\mathbf{A})^{-n}\mathbf{1}, \quad (4)$$

with which different representations of PH and ME distributions can be given:

**Vector-matrix representation** We defined PH and ME distributions with this representation in Definitions 2.1 and 2.2 given by the parameters  $(\alpha, \mathbf{A})$ . One of the problems is that seemingly different  $(\alpha, \mathbf{A})$  and  $(\tilde{\alpha}, \tilde{\mathbf{A}})$  pairs can determine the same distribution. There is a similarity transformation which determines equivalence classes among these pairs. Namely,  $(\alpha, \mathbf{A})$  and  $(\tilde{\alpha}, \tilde{\mathbf{A}})$  determine the same distribution if and only if  $\exists \mathbf{B}$  such that it is non-singular and  $\tilde{\alpha} = \alpha\mathbf{B}$ ,  $\tilde{\mathbf{A}} = \mathbf{B}^{-1}\mathbf{A}\mathbf{B}$  and  $\mathbf{B}\mathbf{1} = \mathbf{1}$  ([8, Theorem 1]). The other problem is that it has too many parameters which is why this is not the representation used for optimization.

**PDF representation** If all the eigenvalues  $(\lambda_i)_{i=1}^n$  of  $\mathbf{A}$  are real, then the form of the PDF in (2) is  $\sum_{i=1}^n a_i t^{k_i} e^{\lambda_i t}$ , where  $k_i$  is a non-negative integer determined by the multiplicity of  $\lambda_i$  ( $i = 1, \dots, n$ ). The parameters  $(a_i, \lambda_i)_{i=1}^n$  uniquely define the distribution. They have to satisfy the constraint  $\int_{-\infty}^{\infty} f(t)dt = 1$ , so in fact we have  $2n - 1$  different parameters. This is a minimal representation, i.e. fewer parameters can't

determine the distribution. Matrices with complex eigenvalues also have unique PDF representations.

**Moments representation** The first  $2n$  moments  $(\mu_0, \dots, \mu_{2n-1})$  define a non-redundant PH distribution of order  $n$  (see [9]). In fact, if  $\alpha \mathbf{1} = 1$  then  $\mu_0 = 1$ , so we again have  $2n - 1$  parameters. We refer to  $(\mu_1, \dots, \mu_{2n-1})$  as the moments representation of the distribution. The moments representation is minimal and unique.

**Laplace representation** The Laplace transform in (3) is an order  $n$  rational function. This can be most easily seen from the first form in (3). The determinant of  $s\mathbf{I} - \mathbf{A}$  is an order  $n$  polynomial of  $s$  and the minors are of order  $n - 1$ . So the inverse of  $s\mathbf{I} - \mathbf{A}$  is indeed a rational function. The Laplace transform can be normalized such that

$$f^*(s) = \frac{a_{n-1}s^{n-1} + \dots + a_1s + a_0}{s^n + b_{n-1}s^{n-1} + \dots + b_0}.$$

Since  $\lim_{s \rightarrow 0} f^*(s) = 0$  the coefficients  $a_0$  and  $b_0$  are equal. The  $2n - 1$  coefficients of the properly normalized Laplace transform is referred to as the Laplace representation. Thus this representation is also minimal and unique.

Other representations also exist. There are techniques to transform one representation to another, which will be briefly discussed in Subsection 2.2. The advantage of the PDF, Moments and Laplace representations is that they are all minimal and unique which makes them favorable for use in finding the optimal fit for a general distribution. However, there are no explicit conditions that can guarantee that as a result we get a valid distribution.

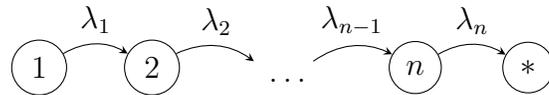
Previously, in [2] necessary and sufficient conditions were given in the Laplace transform domain for being a member of ME(3). These transform domain constraints still require the solution of transcendental equations. A time domain counterpart for ME(3) characterization was developed in [4] that in some cases still needs to use numerical methods. This paper presents a new approach that for the PDF representation gives explicit conditions for ME(3) membership. This procedure is discussed in detail in the following sections.

Here we also mention a special class of PH distributions. An acyclic phase type (APH) distribution can be interpreted as the absorption time of a special continuous-time Markov chain, where any given transient state is visited at most once (hence the name acyclic). This means that the states can be labeled in such a way that the generator  $\mathbf{A}$  of the transient states is

a triangular matrix. This implies that all the eigenvalues of  $\mathbf{A}$  are real, so APH is a proper subset of PH. APH distributions are easy to work with, since the special form of  $\mathbf{A}$  yields efficient methods for analysis, which differ from the ones used to analyze ME or PH distributions. Furthermore, all APH distributions can be transformed to a canonical form (CF(1)). Denoting the eigenvalues of  $\mathbf{A}$  by  $-\lambda_n \leq \dots \leq -\lambda_1 < 0$ , the form of the generator in the CF(1) form is

$$\mathbf{A} = \begin{pmatrix} -\lambda_1 & \lambda_1 & 0 & \dots & 0 \\ 0 & -\lambda_2 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & & -\lambda_{n-1} & \lambda_{n-1} \\ 0 & \dots & & 0 & -\lambda_n \end{pmatrix}.$$

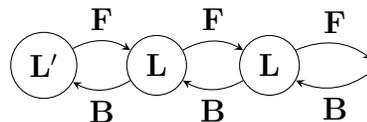
This corresponds to the Markov chain:



The absorbing state can only be reached from state  $n$  and the intensities of the transitions increase as the chain gets closer to the absorbing state. We will deal with APH distributions again in Section 6.

### 2.1.1 Quasi-Birth-Death (QBD) processes

Consider queuing systems where at any given time only one demand can arrive or leave the system. The inter-arrival and/or service times can have arbitrary distributions, but assume the system is regular, i.e. the distributions don't depend on how many demands are in the system. Assume that the behavior between arrivals is regulated by a background continuous-time Markov chain (CTMC) on a finite state space and another CTMC regulates the behavior between services (the two CTMCs can have different state spaces). Such systems can be modeled by a CTMC of the form



where  $\mathbf{F}$ ,  $\mathbf{B}$  and  $\mathbf{L}$  describe the phase transition with arrival, with service and the local behavior without arrival or service respectively. Note that the well

known birth-death processes are a very special case. Namely, if the inter-arrival time is exponentially distributed with parameter  $\lambda$  and the service time with parameter  $\mu$ , then  $\mathbf{F} = \lambda$ ,  $\mathbf{B} = \mu$ ,  $\mathbf{L} = -\lambda - \mu$  and there is only a 1 state background MC involved. The generator of a CTMC that represents a QBD process has a special block matrix structure:

$$\mathbf{Q} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \end{array} \begin{array}{ccccc} & 0 & 1 & 2 & 3 & \dots \\ \left( \begin{array}{cccc} \mathbf{L}' & \mathbf{F} & & & \\ \mathbf{B} & \mathbf{L} & \mathbf{F} & & \\ & \mathbf{B} & \mathbf{L} & \mathbf{F} & \\ & & \ddots & \ddots & \ddots \end{array} \right) \end{array}$$

where the numbers on the border indicate the number of demands in the system. If we assume that the inter-arrival and service times follow ME distributions with parameters  $(\alpha, \mathbf{A})$  and  $(\tilde{\alpha}, \tilde{\mathbf{A}})$  respectively, then for each block we obtain that

$$\begin{aligned} \mathbf{L} &= \mathbf{A} \oplus \tilde{\mathbf{A}}, & \mathbf{L}' &= \mathbf{A} \otimes \mathbf{I}, \\ \mathbf{F} &= (-\mathbf{A}\mathbb{1})\alpha \otimes \mathbf{I}, & \mathbf{B} &= \mathbf{I} \otimes (-\tilde{\mathbf{A}}\mathbb{1})\tilde{\alpha}, \end{aligned} \tag{5}$$

where the operators  $\oplus$  and  $\otimes$  denote the Kronecker sum and product of matrices. For example, to determine the local behavior  $\mathbf{L} = \mathbf{A} \oplus \tilde{\mathbf{A}} = \mathbf{A} \otimes \mathbf{I} + \mathbf{I} \otimes \tilde{\mathbf{A}}$  means that in the arrival process there is a phase transition without arrival while the service process stays put (hence the term  $\mathbf{A} \otimes \mathbf{I}$ ) or vice versa, the arrival process stays put and in the service process there is a phase transition without an actual service taking place (hence the term  $\mathbf{I} \otimes \tilde{\mathbf{A}}$ ). Similarly to determine the forward behavior  $\mathbf{F} = (-\mathbf{A}\mathbb{1})\alpha \otimes \mathbf{I}$  means that an arrival takes place and there is no change in the service process.

Being able to write  $\mathbf{L}'$ ,  $\mathbf{L}$ ,  $\mathbf{F}$ ,  $\mathbf{B}$  in the forms given in (5) means that the methods available for the analysis of QBD processes can also be applied to special QBD processes, where the inter-arrival and service times follow an ME distribution. As a result, when we approximate a general distribution with an ME distribution then at the same time we can give approximations, for example, for the stationary distribution of the original queuing system with the analysis of a special QBD process.

The methods used for the analysis of QBD processes are referred to as matrix geometric methods [6]. We give a brief outline on how to determine the stationary distribution of a QBD process, i.e. we want to determine the probability vector  $p = (p_0, p_1, p_2, \dots)$  that solves the system of equations

$$p\mathbf{Q} = 0 \quad \text{and} \quad p\mathbb{1} = 1,$$

where  $p_i$  are  $1 \times n$  vectors which elements  $p_{ij}$  give the probabilities that the system has  $i$  demands and is in state  $j$  in the stationary distribution. Note that the state space of the CTMC is the direct product of the state spaces of the two background CTMCs. We basically mimic the procedure of determining the stationary distribution for a simple birth-death process. Assume that we look for the solution in the form  $p_i = p_0 \mathbf{R}^i$ , hence the name matrix exponential distribution. Substituting this into  $p\mathbf{Q} = 0$ , we find that for every  $i \geq 1$

$$p_{i-1}\mathbf{F} + p_{i-1}\mathbf{R}\mathbf{L} + p_{i-1}\mathbf{R}^2\mathbf{B} = 0,$$

which holds if and only if  $\mathbf{R}$  satisfies the quadratic matrix equation

$$\mathbf{F} + \mathbf{R}\mathbf{L} + \mathbf{R}^2\mathbf{B} = 0. \quad (6)$$

There are efficient numerical methods to find the right solution of (6). Assuming we found  $\mathbf{R}$ , we still have to find the  $p_0$  that satisfies

$$p_0(\mathbf{L}' + \mathbf{R}\mathbf{B}) = 0 \quad \text{and} \quad 1 = p\mathbb{1} = p_0(\mathbf{I} - \mathbf{R})^{-1}\mathbb{1}.$$

Note that we have reduced the original problem to solving a finite linear system of equations, which can be solved in many different ways. So we determined the stationary distribution of the QBD process. Thus we can determine many other attributes of the queuing system, such as the expected number of demands in the system.

## 2.2 Moment matching

In this subsection we show transformations between different representations of PH or ME distributions that were introduced in Subsection 2.1. The main aim is to show how we can obtain a vector-matrix representation from a sequence of moments. A procedure that is able to do this, we refer to as moment matching. So to be able to approximate a general distribution, all we have to know are its first  $2n$  moments and then with a moment matching procedure we can generate a ME distribution that replicates those exact same moments that we started out with.

We already saw in equations (2), (3) and (4) that it is straightforward to obtain the PDF, Laplace transform or the moments of an ME distribution from its vector-matrix representation. The other direction is not so trivial.

Mark Fackrell showed a simple formula in his PHD dissertation (see [3, Theorem 5.1]) to obtain a vector-matrix representation from the Laplace transform. Assume that we are given the Laplace transform of an ME distribution in the form

$$f^*(s) = \frac{a_{n-1}s^{n-1} + \dots + a_1s + a_0}{s^n + b_{n-1}s^{n-1} + \dots + b_0}.$$

Then the ME distribution has a  $(\beta, \mathbf{T}, t)$  representation, where

$$\begin{aligned}\beta &= (a_0, a_1, \dots, a_{n-1}) \\ \mathbf{T} &= \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots \\ 0 & \dots & & 0 & 1 \\ -b_0 & -b_1 & \dots & b_{n-2} & b_{n-1} \end{pmatrix} \\ t &= (0, 0, \dots, 0, 1)^T,\end{aligned}$$

so that the distribution function of the ME distribution can be written in the form

$$F(u) = 1 - \beta e^{\mathbf{T}u} \mathbf{T}^{-1} t.$$

It is easy to calculate the inverse of  $\mathbf{T}$ :

$$\mathbf{T}^{-1} = \begin{pmatrix} \frac{-b_1}{b_0} & \frac{-b_2}{b_0} & \frac{-b_3}{b_0} & \dots & \frac{-b_{n-1}}{b_0} & \frac{-1}{b_0} \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \\ 0 & \dots & & & 0 & 0 \\ 0 & \dots & & & 1 & 0 \end{pmatrix}$$

Note that this is not exactly the representation we introduced as the definition, but it is not hard to find the similarity transformation that transforms  $(\beta, \mathbf{T}, t)$  into an  $(\alpha, \mathbf{A}, \mathbf{1})$  vector-matrix representation.

It is well known that the moments can be calculated from the Laplace transform with the formula

$$\mu_i = (-1)^i \frac{d^i}{ds^i} f^*(s) \Big|_{s=0}.$$

It is also possible to determine the Laplace transform from the moments (see [7, Section 2]). This way we can obtain an ME function from the given moments.

There is also another method using the procedure of Liefvoort (see [9]) and then using a similarity transformation to obtain a closing vector  $\mathbf{1}$ . For details we refer the Reader to [8, p. 11-12].

However in either case we don't know automatically whether the resulting vector-matrix representation actually defines a valid ME distribution. So it is vital to be able to tell if the resulting representation of a moment matching algorithm defines a valid distribution or not. It turns out that answering this question is difficult. The subsequent sections explore this question.

### 3 General principles of the method

In the end of Section 2 we gave the main motivation to study ME functions. In this Section we introduce the general idea of a new approach to obtain conditions for order  $n$  ME membership by studying the PDF representation (defined in Subsection 2.1). As a special case, we will see in Section 4 that the approach gives necessary and sufficient conditions for ME(3) membership that can be explicitly checked. Furthermore, the approach is applied to ME(4) distributions in Section 5.

Our goal is to explicitly determine whether a vector-square matrix pair,  $(\alpha, \mathbf{A})$ , determines a valid matrix exponential distribution with density

$$f(t) = \alpha e^{\mathbf{A}t} (-\mathbf{A}) \mathbb{1}$$

or not. We assume that the necessary condition  $\lim_{t \rightarrow \infty} f(t) = 0$  ( $\Leftrightarrow$  the real parts of the eigenvalues of  $\mathbf{A}$  are negative) holds and focus only on the non-negativity of  $f(t)$  in  $(0, \infty)$ . Instead of working with  $(\alpha, \mathbf{A})$  directly, we use the PDF representation which of course is uniquely determined from  $(\alpha, \mathbf{A})$  by the formula for the density function.

#### 3.1 The general approach

Let us consider a matrix exponential function of order  $n$  with distinct real eigenvalues, i.e.  $f(t) = \sum_{i=1}^n a_i e^{\lambda_i t}$ , where  $\lambda_n < \lambda_{n-1} < \dots < \lambda_1 < 0$  are the eigenvalues and  $a_i \neq 0$  are real constants. The idea is to divide the inequality by one of the  $e^{\lambda_i t}$  terms:

$$f(t) \geq 0 \quad \rightsquigarrow \quad \tilde{f}(t) = \sum_{i=1}^{n-1} a_i e^{(\lambda_i - \lambda_n)t} \geq -a_n,$$

which is a modified problem of one dimension less. As a result of this order reduction step, the complexity of the problem decreases. This gives the motivation to study the following problems simultaneously:

$$\tilde{f}(t) \geq 0, \quad \forall t \geq 0, \tag{7}$$

$$\tilde{f}(t) \geq b, \quad \forall t \geq 0, \tag{8}$$

$$\tilde{f}(t) = 0, \tag{9}$$

$$\tilde{f}(t) = b. \tag{10}$$

For example, equation (9) is needed to determine the extreme points of the matrix exponential function. Our approach will be to first solve (8), (9),

(10) for  $n = 2$  and then to trace back the order 3 problem of (7) to an order 2 problem of (8). The key is that these will be simple enough to solve explicitly, while for larger dimensions we get equations that can only be solved numerically.

When the eigenvalues are all distinct, it is practical to do the trace back so that all the exponents remain negative. This means that we divide by  $\exp(\lambda_1 t)$ . We note that  $a_1$ , the coefficient of  $\lambda_1$  must be positive.

In general, when the eigenvalues are not all distinct it is useful to divide by  $\exp(\lambda_i t)$ , where  $\lambda_i$  has multiplicity of one. This way we only have to compare the reduced function to a constant, otherwise we have to compare it to a different function which is a much tougher task. We show how the order reduction works in the following sections.

## 4 ME(3) distributions

In this section we apply the order reduction procedure of Section 3 to obtain explicit necessary and sufficient conditions for ME(3) membership. For completeness, we quote the results for the cases where explicit conditions have already been shown. In detail we deal with the remaining cases in separate subsections and prove our main results. Implementing the results into a Mathematica program, we illustrate the "shape" of the ME class if we fix the eigenvalues of  $\mathbf{A}$ .

In the case of ME(3) distributions we can distinguish four different cases according to the eigenvalue structure of  $\mathbf{A}$ :

- 1) three different negative real eigenvalues,
- 2) two different negative real eigenvalues,
- 3) one negative real eigenvalue,
- 4) one negative real and a complex conjugate pair.

In [4, Theorems 1 and 4] explicit formulas were given to decide ME(3) membership only in the cases of 1) and 3). We begin by quoting these results.

In case 1), where  $\mathbf{A}$  has three different real eigenvalues, the general form of the matrix exponential function is

$$f(t) = a_1 e^{\lambda_1 t} + a_2 e^{\lambda_2 t} + a_3 e^{\lambda_3 t}, \quad (11)$$

assuming without loss of generality that  $\lambda_3 < \lambda_2 < \lambda_1 < 0$ .

**Theorem 4.1.** *The matrix exponential function  $f(t)$  of the form (11) is non-negative for all  $t \geq 0$  if and only if*

- $a_1 + a_2 + a_3 \geq 0$  and
- $a_1 > 0$  and
- if  $a_2 < -a_1 \frac{\lambda_1 - \lambda_3}{\lambda_2 - \lambda_3}$ , then  $a_3 \geq a_1 \frac{\lambda_1 - \lambda_2}{\lambda_2 - \lambda_3} \left( -\frac{a_2 \lambda_2 - \lambda_3}{a_1 \lambda_1 - \lambda_3} \right)^{\frac{\lambda_1 - \lambda_3}{\lambda_1 - \lambda_2}}$ .

In case 3), where the single eigenvalue  $\lambda$  has multiplicity of three, the general form of the density function is

$$f(t) = (a_0 + a_1 t + a_2 t^2) e^{\lambda t}. \quad (12)$$

**Theorem 4.2.** *The matrix exponential function  $f(t)$  of the form (12) is non-negative for all  $t \geq 0$  if and only if*

$$a_0 > 0, \quad a_2 > 0 \quad \text{and} \quad a_1 \geq -2\sqrt{a_0 a_2}.$$

We now move forward by providing explicit formulas for cases 2) and 4) with the help of the order reduction approach introduced in Section 3.

## 4.1 Two different eigenvalues

We have to consider two cases. Assume that the eigenvalues are  $\lambda_2 < \lambda_1 < 0$  ( $\lambda_1$  is referred to as the dominant eigenvalue). In the case when the multiplicity of  $\lambda_1$  is one, the general form of the matrix exponential function is

$$f_1(t) = a_1 e^{\lambda_1 t} + (a_2 + a_{21} t) e^{\lambda_2 t}, \quad \text{where } a_1, a_{21} \neq 0. \quad (13)$$

In the other case when the multiplicity of the dominant eigenvalue is two, we can write

$$f_2(t) = (a_1 + a_{11} t) e^{\lambda_1 t} + a_2 e^{\lambda_2 t}, \quad \text{where } a_2, a_{11} \neq 0. \quad (14)$$

Dividing (13) or (14) by the exponential term of the single eigenvalue gives the following problem of type (8):

$$\hat{f}(t) = (g_1 + g_2 t) e^{\gamma t} \geq b \quad \forall t \geq 0, \quad \text{where } b, g_2 \neq 0. \quad (15)$$

Elementary calculations gives us the root of the function  $\hat{f}(t)$ , its extreme point (the root of  $d\hat{f}(t)/dt$ ) and its extreme value

$$t^* = \frac{-g_1}{g_2},$$

$$t_{opt} = -\frac{g_2 + g_1\gamma}{g_2\gamma} = \frac{-1}{\gamma} + t^*,$$

$$f_{opt} = \hat{f}(t_{opt}) = -\frac{g_2}{\gamma} \cdot \exp\left(-\frac{1}{\gamma} - \frac{g_1}{g_2}\right).$$

$t^*$  and  $t_{opt}$  coincide iff  $g_2 = 0$ , thus  $t^* \neq t_{opt}$ . Depending on the sign of  $\gamma$  and  $g_2$  there are four cases to consider. One of them is illustrated in Figure 1.

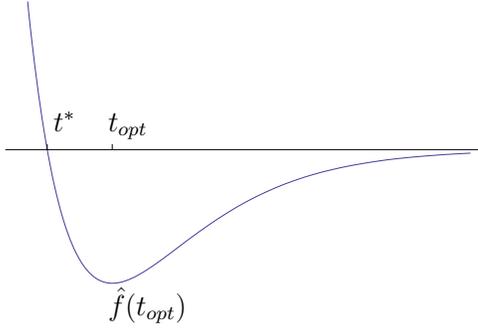


Figure 1: Structure of  $\hat{f}(t)$  when  $\gamma < 0$  and  $g_2 < 0$

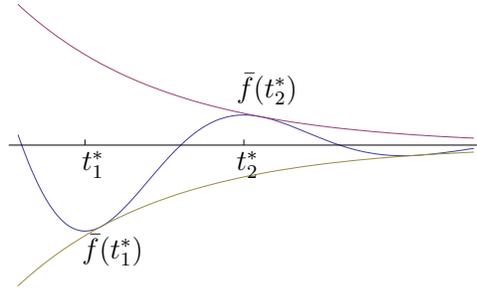


Figure 2: Structure of  $\bar{f}(t)$  when  $\bar{f}(t_1^*) < 0$

- $\gamma < 0$ ,  $g_2 < 0$ . The possible values of  $b$  depend on the sign of  $t_{opt}$ . If  $t_{opt} \leq 0$ , i.e.  $g_1 \leq -g_2/\gamma$  then  $b \leq \hat{f}(0) = g_1$ . Otherwise  $b \leq f_{opt}$  (see Figure 1). This gives us two possible necessary and sufficient conditions for  $\hat{f}(t) \geq b$ :

$$\gamma < 0, g_2 < 0, b \leq g_1 \leq \frac{-g_2}{\gamma} \quad (16)$$

$$\gamma, g_2 < 0, g_1 > \frac{-g_2}{\gamma}, b \leq -\frac{g_2}{\gamma} \cdot e^{-\frac{1}{\gamma} - \frac{g_1}{g_2}} \quad (17)$$

- $\gamma < 0$ ,  $g_2 > 0$ . This time the possible values of  $b$  depend on the sign of  $t^*$ . If  $t^* \leq 0$ , i.e.  $g_1 \geq 0$  then  $b \leq 0$ . Otherwise  $b \leq \hat{f}(0) = g_1$ . We thus gain two more conditions:

$$\gamma < 0, b \leq g_1 < 0 < g_2 \quad (18)$$

$$\gamma < 0, g_2 > 0, g_1 \geq 0, b \leq 0 \quad (19)$$

- $\gamma > 0$ ,  $g_2 < 0$ .  $\hat{f}(t) \geq b$  can't hold for any  $b$  since  $\lim_{t \rightarrow \infty} \hat{f}(t) = -\infty$ .
- $\gamma > 0$ ,  $g_2 > 0$ . The possible values of  $b$  again depend on the sign of  $t_{opt}$ . If  $t_{opt} \leq 0$ , i.e.  $g_1 \geq -g_2/\gamma$  then  $b \leq \hat{f}(0) = g_1$ . Otherwise  $b \leq f_{opt}$ . So the last two conditions are:

$$\gamma > 0, g_2 > 0, g_1 \geq \frac{-g_2}{\gamma}, b \leq g_1 \quad (20)$$

$$\gamma, g_2 > 0, g_1 < \frac{-g_2}{\gamma}, b \leq -\frac{g_2}{\gamma} \cdot e^{-\frac{1}{\gamma} - \frac{g_1}{g_2}} \quad (21)$$

Conditions (16), (17), (18) and (19) will be used after the trace back of (13). Similarly (20) and (21) will be used for (14). After these preparations we can prove the following.

**Theorem 4.3.**  $f_1(t)$  as defined in (13) is non-negative for  $t \geq 0$  if and only if  $\lambda_2 < \lambda_1 < 0$  and one of the following hold

$$i) a_{21} < 0, a_2 \leq \frac{-a_{21}}{\lambda_2 - \lambda_1}, a_1 \geq -a_2;$$

$$ii) a_{21} < 0, a_2 > \frac{-a_{21}}{\lambda_2 - \lambda_1}, a_1 \geq \frac{a_{21}}{\lambda_2 - \lambda_1} e^{\left(-\frac{1}{\lambda_2 - \lambda_1} - \frac{a_2}{a_{21}}\right)};$$

$$iii) a_{21} > 0, a_2 < 0, a_1 \geq -a_2;$$

$$iv) a_{21} > 0, a_2 \geq 0, a_1 > 0.$$

*Proof.* Dividing  $f_1(t)$  by  $e^{\lambda_1 t}$  results in an inequality of type (15). Substituting its parameters into (16), (17), (18) and (19) we obtain the theorem.  $\square$

**Theorem 4.4.**  $f_2(t)$  as defined in (14) is non-negative for  $t \geq 0$  if and only if  $\lambda_2 < \lambda_1 < 0$  and one of the following hold

$$i) a_{11} > 0, a_1 \geq \frac{-a_{11}}{\lambda_1 - \lambda_2}, a_2 \geq -a_1;$$

$$ii) a_{11} > 0, a_1 < \frac{-a_{11}}{\lambda_1 - \lambda_2}, a_2 \geq \frac{a_{11}}{\lambda_1 - \lambda_2} e^{\left(-\frac{1}{\lambda_1 - \lambda_2} - \frac{a_1}{a_{11}}\right)}.$$

*Proof.* Dividing  $f_2(t)$  by  $e^{\lambda_2 t}$  results in an inequality of type (15). Substituting its parameters into (20) and (21) results the statement of the theorem.  $\square$

## 4.2 One real and a complex conjugate pair of eigenvalues

The general form of the density function in this case is

$$f_3(t) = a_1 e^{\lambda_1 t} + a_2 \cos(\omega t + \phi) e^{\lambda_2 t}, \quad (22)$$

where  $t \geq 0$ ,  $a_2 > 0$ ,  $-\pi < \phi < \pi$ ,  $\lambda_1, \lambda_2 < 0$ . We want  $f_3(t) \geq 0$  to hold so  $\lambda_2 \leq \lambda_1 < 0$  should hold. The result of the order reduction step is

$$f_3(t) \geq 0 \Leftrightarrow \bar{f}(t) \geq b \quad (\forall t \geq 0),$$

where  $\bar{f}(t) = \cos(\omega t + \phi) e^{\lambda t}$ ,  $\lambda = \lambda_2 - \lambda_1$ ,  $b = \frac{-a_1}{a_2}$  and  $a_1 > 0$ . The extreme points of  $\bar{f}(t)$  are obtained at  $\bar{f}(t)' = 0$  which are

$$\begin{aligned} \lambda \cos(\omega t + \phi) &= \omega \sin(\omega t + \phi) && \Leftrightarrow \\ t_k^* &= \frac{\tan^{-1}\left(\frac{\lambda}{\omega}\right) - \phi + k\pi}{\omega}, \end{aligned}$$

where  $k \in \mathbb{Z}$ . Note that  $\cos(\omega t + \phi)$  and  $\sin(\omega t + \phi)$  cannot be zero at the same time and  $\tan(\omega t + \phi)$  is  $\pi/\omega$  periodic. Since  $\cos(\omega t + \phi)$  is  $2\pi/\omega$  periodic and  $e^{\lambda t}$  is monotone decreasing it is enough to consider the  $t_k^*$  which fall into  $[0, 2\pi/\omega]$  (see Figure 2). If  $k^* = -\lfloor (\tan^{-1}(\frac{\lambda}{\omega}) - \phi)/\pi \rfloor$ , then the extreme points in  $[0, 2\pi/\omega]$  are  $t_i^* = (\tan^{-1}(\frac{\lambda}{\omega}) - \phi + (k^* + i - 1)\pi) / \omega$ ,  $i = 1, 2$ . It only remains to check if  $\bar{f}(t_i^*) \geq b$ ,  $i = 1, 2$ .

**Theorem 4.5.**  *$f_3(t)$  as defined in (22) is non-negative for  $t \geq 0$  if and only if one of the following hold*

- $\lambda_1 = \lambda_2$  and  $a_1 \geq a_2 > 0$ ,
- $\lambda_2 < \lambda_1 < 0$ ,  $a_1 > 0$  and  $\bar{f}(t_i^*) \geq b$ ,  $i = 1, 2$ .

*Proof.* If  $\lambda_1 = \lambda_2$  then  $f_3(t) \geq 0$  simplifies to  $\cos(\omega t + \phi) \geq \frac{-a_1}{a_2}$ . It follows that  $\frac{-a_1}{a_2} \leq -1$ , i.e.  $a_1 \geq a_2$ . If  $\lambda_2 < \lambda_1 < 0$  then we proceed according to the analysis of  $\bar{f}(t) \geq b$ . □

## 4.3 Visualization of ME(3)

We saw in Subsection 2.1 that, without the normalizing constraint, the  $2n$  parameters of the PDF representation uniquely determine the ME( $n$ ) distribution it represents. For order three distributions this means 6 parameters. To be able to visualize the class of ME(3) distributions we chose to fix the

eigenvalues of  $\mathbf{A}$ . This way we only have the freedom to vary the three coefficients of the ME function. As a result we get very different behavior according to the eigenvalue structure of  $\mathbf{A}$ . Figures 3-6 were generated with *Mathematica*.

In Figure 3 the axes are labeled by the parameters  $ar = a1$ ,  $ac = a2$ ,  $\phi$  of the density function defined in (22). While in Figures 4-6 the axes are labeled by the parameters of the density functions defined in (11), (13) and (14) respectively. The complex case shows no symmetric behavior with respect to  $\phi$ . In each of the real cases we can of course find the positive space, where all the parameters are non-negative, but otherwise very different behavior can be observed. The circles in the figures start out from the origin.

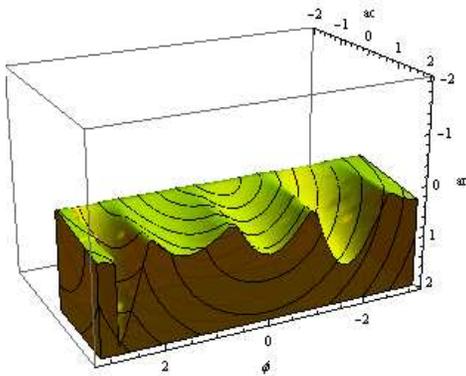


Figure 3: ME(3) complex case

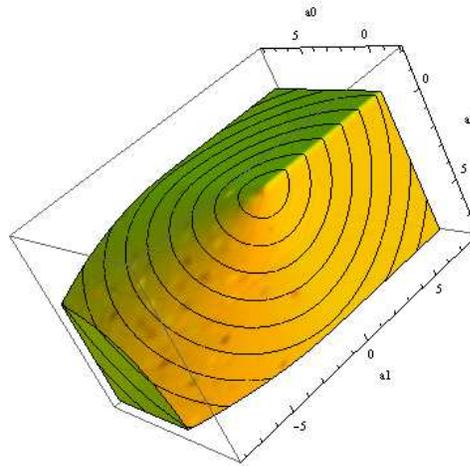


Figure 4: ME(3) real case:  
 $\lambda_1 = \lambda_2 = \lambda_3 = -5$

## 5 Higher order ME distributions

A logical related question is if the order reduction method leads us to similar results for higher order ME distributions. Unfortunately in almost all of the cases we can't expect to get explicit results, because we always end up needing to solve a transcendental equation that can't be solved explicitly anymore. If we allow ourselves to use numerical methods to solve such equations, then afterwards everything else remains explicit. Even without the help of numerical methods we can give different non-trivial necessary or sufficient conditions for higher order ME membership. We demonstrate this on two cases of ME(4) distributions.

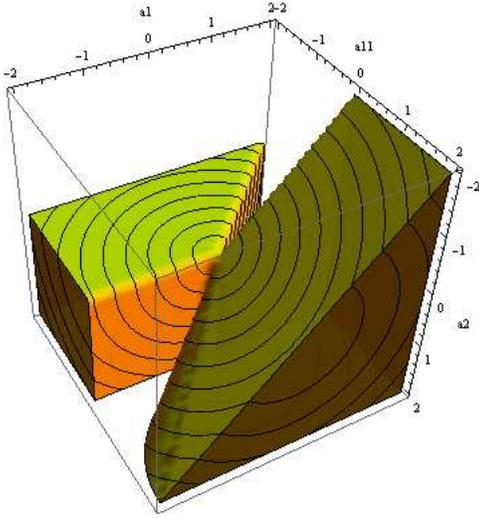


Figure 5: ME(3) real case:  
 $\lambda_1 = \lambda_2 = -3, \lambda_3 = -4$

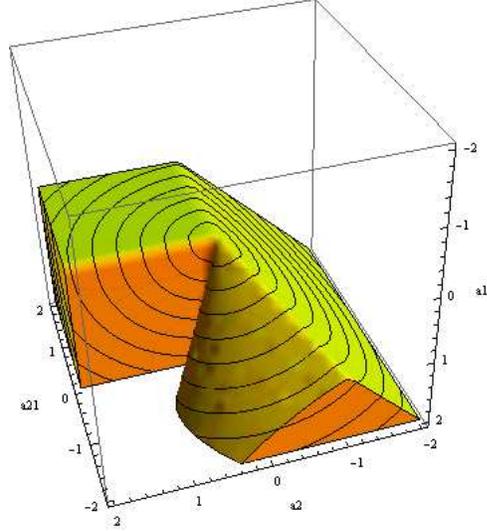


Figure 6: ME(3) real case:  
 $\lambda_1 = -3, \lambda_2 = \lambda_3 = -5$

In Subsection 5.1 we show one of the few cases where explicit necessary and sufficient conditions can be obtained with the approach. The proof is very similar to the ones already shown. Afterwards, in Subsection 5.2 we analyze an important case of ME(4) distributions. We identify the equation that can't be solved explicitly and with some simple reasoning we give a collection of non-trivial necessary or sufficient conditions to decide ME membership in this case. Note that attempts of characterizing the class of ME(4) distributions can't be found in the literature.

## 5.1 An explicit case ME(4) distribution

Let us consider  $4 \times 4$  matrices  $\mathbf{A}$ , such that have two different real eigenvalues, one of which has a multiplicity of three. Let  $\lambda_2 < \lambda_1 < 0$  denote the eigenvalues. In the case when the multiplicity of  $\lambda_1$  is one, the general form of the matrix exponential function is

$$f_1(t) = a_1 e^{\lambda_1 t} + (a_{20} + a_{21}t + a_{22}t^2)e^{\lambda_2 t}, \text{ where } a_1, a_{22} \neq 0. \quad (23)$$

In the other case when the multiplicity of the dominant eigenvalue is three, we can write

$$f_2(t) = (a_{10} + a_{11}t + a_{12}t^2)e^{\lambda_1 t} + a_2 e^{\lambda_2 t}, \text{ where } a_{12}, a_2 \neq 0. \quad (24)$$

In each case we apply the order reduction approach by dividing with the exponential term of the eigenvalue with single multiplicity. As a result we

obtain the inequality

$$\hat{f}(t) := (at^2 + bt + c)e^{\lambda t} \geq d \quad \forall t \geq 0, \text{ where } a, d \neq 0. \quad (25)$$

The parameters  $(a, b, c, d, \lambda)$  equal  $(a_{22}, a_{21}, a_{20}, -a_1, \lambda_2 - \lambda_1)$  in case (23) and  $(a_{12}, a_{11}, a_{10}, -a_2, \lambda_1 - \lambda_2)$  in case (24). We collect the conditions for case (23) in the following lemma.

**Lemma 5.1.** *For  $\lambda < 0$  and  $a, d \neq 0$ , the inequality*

$$\hat{f}(t) = (at^2 + bt + c)e^{\lambda t} \geq d,$$

*holds for every  $t \geq 0$  if and only if one of the following hold*

i)  $a < 0, t'_2 < 0, d \leq c$

ii)  $a < 0, t'_2 \geq 0, d \leq \min(c, \hat{f}(t'_2))$

iii)  $a > 0, b^2 \leq 4ac, d < 0$

iv)  $a > 0, b^2 > 4ac, t_2^* \leq 0, d < 0$

v)  $a > 0, b^2 > 4ac, t'_1 < 0 < t_2^*, d \leq c$

vi)  $a > 0, b^2 > 4ac, t'_1 \geq 0, d \leq \hat{f}(t'_1)$

*where  $t_2^*, t'_1$  and  $t'_2$  are defined in (26) and (27).*

*Proof.* Again we start by determining the roots  $t_{1,2}^*$  and the extreme points  $t'_{1,2}$  of  $\hat{f}(t)$ . The shape of  $\hat{f}(t)$  implies that  $t_1^* \leq t'_1 \leq t_2^* \leq t'_2$  (if  $t_{1,2}^* \in \mathbb{R}$ ). Elementary calculations give us the following

$$t_{1,2}^* = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \text{ and} \quad (26)$$

$$t'_{1,2} = \frac{-\tilde{b} \pm \sqrt{\tilde{b}^2 - 4\tilde{a}\tilde{c}}}{2\tilde{a}}, \text{ where } \begin{cases} \tilde{a} &= \lambda a \\ \tilde{b} &= \lambda b + 2a \\ \tilde{c} &= \lambda c + b \end{cases} \quad (27)$$

We differentiate three different cases according to the number of real roots of  $\hat{f}(t)$ .

- If  $t_1^*, t_2^* \notin \mathbb{R}$ , i.e.  $b^2 < 4ac$ , then

– if  $a > 0$ , then  $\hat{f}(t) > 0 \forall t$ . Thus arbitrary  $d < 0$  satisfies (25).

- if  $a < 0$ , then  $\hat{f}(t) < 0 \forall t$ . The possible values of  $d$  depend on the sign of  $t'_2$ . If  $t'_2 < 0$ , then  $d \leq \hat{f}(0) = c$ . If  $t'_2 \geq 0$ , then  $d \leq \min(c, \hat{f}(t'_2))$ .
- If  $t_1^* = t_2^* \in \mathbb{R}$ , i.e.  $b^2 = 4ac$ , then  $t_1^* = -b/2a$ ,  $t'_1 = -b/2a$  and  $t'_2 = -b/2a - 2/\lambda$ . The conditions for  $d$  are the same as in the previous case.
- If  $t_1^* \neq t_2^* \in \mathbb{R}$ , i.e.  $b^2 > 4ac$ , then
  - if  $a > 0$ , then the possible values of  $d$  depend on the sign of  $t'_1$  and  $t'_2$ . If  $t'_2 \leq 0$ , then arbitrary  $d < 0$  satisfies (25). If  $t'_1 < 0 < t'_2$ , then  $d \leq \hat{f}(0) = c$ . Finally, if  $t'_1 \geq 0$ , then  $d \leq \hat{f}(t'_1)$ .
  - if  $a < 0$ , then the possible values of  $d$  depend on the sign of  $t'_2$  the same way as in the first case.

By joining the obtained conditions, we get conditions i)-vi) of the lemma.  $\square$

From here we just have to substitute  $(a_{22}, a_{21}, a_{20}, -a_1)$  for  $(a, b, c, d)$  and  $\lambda_2 - \lambda_1$  for  $\lambda$  into the conditions of Lemma 5.1 and equations (26) and (27) to obtain the desired explicit necessary and sufficient conditions.

**Theorem 5.2.** *A matrix exponential function of the form (23) is non-negative for every  $t \geq 0$  if and only if one of the following hold*

- i)  $a_{22} < 0$ ,  $t'_2 < 0$ ,  $a_1 + a_{20} \geq 0$
- ii)  $a_{22} < 0$ ,  $t'_2 \geq 0$ ,  $-a_1 \leq \min(a_{20}, (a_{22}(t'_2)^2 + a_{21}t'_2 + a_{20})e^{(\lambda_2 - \lambda_1)t})$
- iii)  $a_{22} > 0$ ,  $a_{21}^2 \leq 4a_{22}a_{20}$ ,  $a_1 > 0$
- iv)  $a_{22} > 0$ ,  $a_{21}^2 > 4a_{22}a_{20}$ ,  $t_2^* \leq 0$ ,  $a_1 > 0$
- v)  $a_{22} > 0$ ,  $a_{21}^2 > 4a_{22}a_{20}$ ,  $t'_1 < 0 < t_2^*$ ,  $a_1 + a_{20} \geq 0$
- vi)  $a_{22} > 0$ ,  $a_{21}^2 > 4a_{22}a_{20}$ ,  $t'_1 \geq 0$ ,  $-a_1 \leq (a_{22}(t'_1)^2 + a_{21}t'_1 + a_{20})e^{(\lambda_2 - \lambda_1)t}$ ,

where

$$t_{1,2}^* = \frac{-a_{21} \pm \sqrt{a_{21}^2 - 4a_{22} \cdot a_{20}}}{2a_{22}}, \text{ and}$$

$$t'_{1,2} = \frac{-(\lambda_2 - \lambda_1)a_{21} \pm \sqrt{((\lambda_2 - \lambda_1)a_{21})^2 - 4(\lambda_2 - \lambda_1)a_{22}((\lambda_2 - \lambda_1)a_{20} + a_{21})}}{2(\lambda_2 - \lambda_1)a_{22}}$$

Following these lines, one can also obtain necessary and sufficient conditions when the matrix exponential function is of the form (24). We also expect to get explicit results in the case when  $\mathbf{A}$  has only one real eigenvalue with multiplicity of four. In this case the matrix exponential function is  $(a_0 + a_1t + a_2t^2 + a_3t^3)e^{\lambda t}$ . This is non-negative if and only if  $(a_0 + a_1t + a_2t^2 + a_3t^3)$  is non-negative, which can be decided since there are well known methods to determine the roots of cubic functions. From here we suspect to get explicit results for the ME(5) distribution with two different real eigenvalues, one of which has a multiplicity of four. In every other case we either need to solve a transcendental equation to obtain the extreme points or have to determine when a function  $f(t)$  dominates another function  $g(t)$  ( $\forall t > 0$ ) which can't be solved explicitly. We show this on an order 4 case in the next Subsection.

## 5.2 Two real eigenvalues and a complex conjugate pair

In this Subsection we consider  $4 \times 4$  matrices  $\mathbf{A}$ , such that have two different real eigenvalues and a complex conjugate pair. Perhaps this is the most important case from the point of applications, since a random  $4 \times 4$  matrix is most likely to have this type of eigenvalue structure. Note that  $4 \times 4$  matrices with two complex conjugate pairs of eigenvalues can't determine a valid distribution, since the PDF is guaranteed to go below zero because of the cosine terms.

Let  $\lambda_2 < \lambda_1 < 0$  denote the real eigenvalues and  $\lambda_c < 0$  be the real part of the complex eigenvalue. In this case the general form of the matrix exponential function is

$$f(t) = a_1e^{\lambda_1 t} + a_2e^{\lambda_2 t} + a_c \cos(\omega t + \phi)e^{\lambda_c t}, \quad (28)$$

where  $t \geq 0$ ,  $a_c > 0$ ,  $-\pi < \phi < \pi$ . We want  $f(t) \geq 0$  to hold so  $\lambda_c \leq \lambda_1$  should hold. This time we apply the order reduction step by dividing by  $a_c e^{\lambda_c t}$  to obtain

$$\hat{f}(t) := \frac{a_1}{a_c} e^{(\lambda_1 - \lambda_c)t} + \frac{a_2}{a_c} e^{(\lambda_2 - \lambda_c)t} \geq -\cos(\omega t + \phi). \quad (29)$$

Assume  $\lambda_1 = \lambda_c$ , then  $\hat{f}(t) \xrightarrow{t \rightarrow \infty} a_1/a_c$ , so

$$\text{If } \lambda_1 = \lambda_c, \text{ then } f(t) \geq 0 \forall t \geq 0 \Leftrightarrow a_1 \geq a_c. \quad (30)$$

From now on we assume that  $\lambda_1 > \lambda_c$ . Notice that  $a_1 > 0$  is necessary, since otherwise  $\hat{f}(t) \xrightarrow{t \rightarrow \infty} -\infty$ . The problem is that to go on with the explicit solution we need to solve the equation  $\hat{f}(t) = -\cos(\omega t + \phi)$ . Generally this

can't be done explicitly. Of course we could solve the equation numerically and proceed the usual way. Instead, we go on by giving some explicit necessary or sufficient conditions that collectively cover the vast majority of ME functions of the form (28).

Obviously  $\hat{f}(0) = a_1/a_c + a_2/a_c \geq -\cos\phi$  is a necessary condition. If  $a_2 > 0$  then in fact  $\hat{f}(0) \geq -\cos\phi$  is also sufficient, since in this case  $\hat{f}(t)$  is monotone increasing. So assume from now on that  $a_2 < 0$ . In this case, easy calculations give us the root of  $\hat{f}(t)$ :

$$t^* = \frac{\ln(-a_2/a_c) - \ln(a_1/a_c)}{\lambda_1 - \lambda_2}. \quad (31)$$

A less trivial necessary condition is that  $\cos(\omega t^* + \phi) > 0$ . We also give a sufficient condition. Namely, since the derivative of the cos function is bounded by one from above, if the derivative of  $\hat{f}(t)$  at  $t = 0$  is greater than one then the inequality  $\hat{f}(t) \geq -\cos(\omega t + \phi)$  holds.

Let us continue by giving extra conditions with the help of results from Section 4. We can bound  $\hat{f}(t)$  from below and above:

$$\begin{aligned} f_-(t) &:= a_1 e^{\lambda_1 t} + a_2 e^{\lambda_2 t} - a_c e^{\lambda_c t} \leq f(t) \\ f_+(t) &:= a_1 e^{\lambda_1 t} + a_2 e^{\lambda_2 t} + a_c e^{\lambda_c t} \geq f(t) \end{aligned}$$

If  $f_-(t)$  satisfies Theorem 4.1, then we can conclude that  $f(t)$  is a valid density function. If  $f_+(t)$  doesn't satisfy Theorem 4.1, then  $\exists t \geq 0 : f(t) < 0$ , so it is not a valid density function. It is still possible that neither of these are true.

It is easy to determine the points where  $f(t) = \hat{f}(t)$ . This holds if and only if  $\cos(\omega t + \phi) = -1$ , i.e.  $t_k^- = \frac{-\phi + (2k+1)\pi}{\omega}$ . For the only  $t_k^-$  that falls into  $[0, 2\pi/\omega]$  we have

$$f\left(\frac{-\phi + \pi}{\omega}\right) = f_-\left(\frac{-\phi + \pi}{\omega}\right).$$

So it is necessary that  $f_-\left(\frac{-\phi + \pi}{\omega}\right) \geq 0$ , otherwise  $f(t)$  is guaranteed to go below zero. We provide one final condition.

We can calculate the difference between  $f(t)$  and  $f_+(t)$ :

$$f_+(t) - f(t) = a_c (1 - \cos(\omega t + \phi)) e^{\lambda_c t}.$$

This is simple enough, so that we can determine the points where this difference has local maxima. We're not interested in the points where  $f(t) = f_+(t)$ , since we know that in these points  $f(t) \geq 0$ . Solving  $(f_+(t) - f(t))' = 0$  we

obtain

$$\begin{aligned} \lambda_c(1 - \cos(\omega t + \phi)) + \omega \sin(\omega t + \phi) &= 0 && \Leftrightarrow \\ \sin(\omega t + \phi) \left( \lambda_c \tan\left(\frac{\omega t + \phi}{2}\right) + \omega \right) &= 0 && \Leftrightarrow \\ t_k &= \frac{2 \tan^{-1}(\omega/|\lambda_c|) - \phi + 2k\pi}{\omega}. \end{aligned}$$

So if  $f(t) < 0$  for some  $t$ , then there is a "good chance" that  $f(t_k)$  is also negative. Hence it is worth checking whether  $f(t_k)$  is greater than zero or not. Of course, even the collection of all these mentioned conditions are not enough to claim that  $f(t) \geq 0$  if and only if these hold. It is not easy to show an example, however functions that satisfy all of these conditions but still go below zero exist.

We can conclude from Sections 3, 4 and 5 that the method introduced in Section 3 is able to explicitly handle all the cases of ME(3) distributions along the same train of thought. The approach can also be applied to higher order ME functions, usually resulting with the problem of solving a transcendental equation that requires numerical methods. We can either solve these equations numerically and proceed with the familiar steps of the method or we can try to give a collection of explicit necessary or sufficient conditions that cover the majority of cases. This is a big step forward in the analysis of ME functions compared to the previous results of [2] and [4]. We move forward by demonstrating the usefulness of these results by showing on a particular example that greater fitting accuracy can be achieved with the class of ME distributions.

## 6 Example

In this section we demonstrate the difference one can get as a result of approximating a distribution with the class of ME distributions or just with APH distributions, discussed in Subsection 2.1. It is difficult to compare the class of ME and PH distributions, since the same fitting algorithms are used and there is no efficient method to tell if a valid ME parameter pair  $(\alpha, \mathbf{A})$  does in fact define a PH distribution or not. We work with distributions only of order three.

To be able to determine the optimal approximating distribution, we need to define some sort of metric between distributions. Different metrics have been introduced, some of which give better results for the body of the distribution while others give better approximations for the tail of the distribution.

In our examinations we consider the relative entropy defined by

$$D(\hat{f}, f) := \int_{t=0}^{\infty} f(t) \log \left( \frac{f(t)}{\hat{f}(t)} \right) dt \geq 0,$$

where  $\hat{f}$  is the probability density function (PDF) of the original distribution while  $f$  is the PDF of the approximating distribution. The relative entropy gives a fairly good fit for the body and the tail at the same time. So the aim of the approximation is to determine the optimal parameters  $(\alpha, \mathbf{A})$  of the ME and PH distributions that minimizes  $D(\hat{f}, f)$ .

First, we find the optimal fit amongst all matrix exponential functions with the help of the moment matching procedure described in Subsection 2.2. Then we decide whether the result is a valid ME distribution by checking the conditions of the appropriate theorem of Section 4 implemented into a *Mathematica* function. This way we optimize over all ME(3) distributions.

Second, restricting the optimization to APH(3) distributions, we use *PhFit* (see [5]). With *PhFit* one can approximate continuous distributions or any empirical data. One can choose between different metrics and select the order of the APH distributions to use. As the output, it gives the initial vector  $\alpha$  and the intensities  $\lambda_1 \geq \lambda_2 \geq \lambda_3$  of the first canonical form of the optimal APH(3) fit. The matrix  $\mathbf{A}$  can be retrieved from the intensities:

$$\mathbf{A} = \begin{pmatrix} -\lambda_1 & 0 & 0 \\ \lambda_2 & -\lambda_2 & 0 \\ 0 & \lambda_3 & -\lambda_3 \end{pmatrix}.$$

We compare the two results by calculating the relative entropy in each case and illustrate the difference by plotting the PDFs. We studied Weibull distributions with different parameters.

## 6.1 Approximation of Weibull distributions

The Weibull distribution is used in many different fields, mainly due to its importance in extreme value theory. It is related to a number of other probability distributions. It has two parameters, the shape parameter  $k > 0$  and the scale parameter  $\lambda > 0$ . Its PDF is

$$\hat{f}_{k,\lambda}(x) = \begin{cases} \frac{k}{\lambda} \left( \frac{x}{\lambda} \right)^{k-1} e^{-(x/\lambda)^k}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0. \end{cases}$$

Depending on the value of the shape parameter, the PDF changes shape radically. To apply the moment matching procedure we need to calculate

the moments of a random variable  $X$  with Weibull distribution. The closed form of the  $n$ -th moment is

$$\mu_n = \mathbf{E}(X^n) = \lambda^n \Gamma\left(1 + \frac{n}{k}\right),$$

where  $\Gamma$  is the well-known gamma function, i.e.  $\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx$ . From here we see that the coefficient of variation (also called relative variance)

$$\frac{\mathbf{D}^2(X)}{\mathbf{E}^2 X} = \frac{\lambda^2 (\Gamma(1 + 2/k) - \Gamma^2(1 + 1/k))}{\lambda^2 \Gamma^2(1 + 1/k)}$$

does not depend on the scale parameter  $\lambda$ . Typically, experience shows that APH distributions give better approximations if the coefficient of variation is large. For this reason we show a case where the coefficient of variation is relatively small and also a case where it is large.

Consider the cases  $k_1 = 1.55$ ,  $\lambda_1 = 50$  and  $k_2 = 1.02$ ,  $\lambda_2 = 40$ . In the first case the coefficient of variation is 0.434 while in the second it is 0.961. With *Mathematica* we found that the moment matching gives the following results:

	$\alpha$	$\mathbf{A}$	$D(\hat{f}, f_{\alpha, \mathbf{A}})$
Case 1	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$\begin{pmatrix} -0.0539585 & 0.373227 & -0.286393 \\ -0.0082287 & -0.038395 & 0.025353 \\ -0.0065961 & 0.023570 & -0.040178 \end{pmatrix}$	0.00131286
Case 2	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$\begin{pmatrix} -0.0842461 & 0.286888 & -0.223354 \\ -0.0043999 & -0.051695 & 0.030653 \\ -0.0124034 & 0.034892 & -0.047457 \end{pmatrix}$	$1.3067 \times 10^{-5}$

In Case 1  $\mathbf{A}$  has a dominant real and a complex conjugate pair of eigenvalues, while in Case 2 all the eigenvalues are negative real numbers. We can already see that the relative entropy decreases if the relative variance grows, i.e. better approximations can be given for distributions with larger relative variance. Restricting to PH(3) distributions we found the following:

	$\alpha^T$	Intensities	$D(\hat{f}, f_{\alpha, \mathbf{A}})$
Case 1	$\begin{pmatrix} 0.73673 \\ 0.21406 \\ 0.04921 \end{pmatrix}$	(0.0776593, 0.0550558, 0.0536794)	0.00517755
Case 2	$\begin{pmatrix} 0.17460 \\ 0.18490 \\ 0.64049 \end{pmatrix}$	(0.1445179, 0.0970587, 0.0285208)	0.00202499

Again we can see that the accuracy of the approximation increased for the case where the relative variance was greater. By comparing the relative entropy for the same case with the different methods, we can see that in each case we got a better approximation with the moments matching method. In the second case the difference looks significant. To further compare the two methods we plotted the respective probability density functions side-by-side with the original Weibull distributions. Figures 7-10 show this.

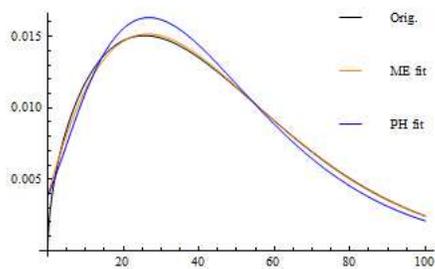


Figure 7: Approximation of Weibull distribution,  $k_1 = 1.55$ ,  $\lambda_1 = 50$

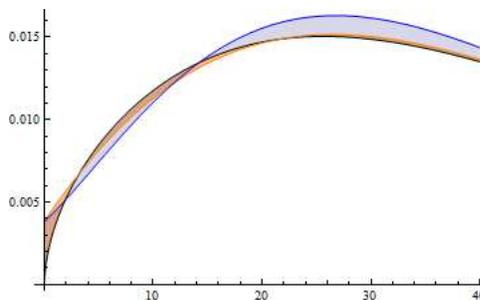


Figure 8: Inset of Fig. 7 , shading difference between orig. and approx.

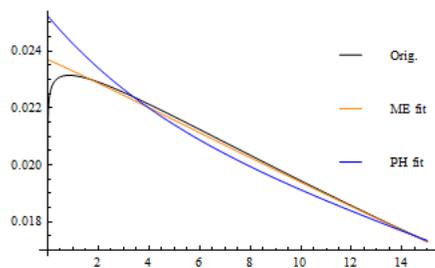


Figure 9: Approximation of Weibull distribution,  $k_2 = 1.02$ ,  $\lambda_2 = 40$

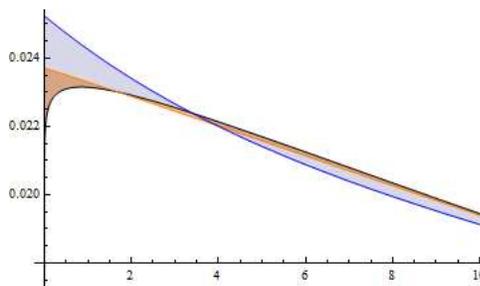


Figure 10: Inset of Fig. 9 , shading difference between orig. and approx.

All of these results lead us to the conclusion that it is indeed worth using ME distributions for approximating general distributions. It should also be mentioned that our experience was that unfortunately the moment matching algorithm very rarely gave a valid matrix exponential function as its result. It took some time even to find these parameters for the Weibull distribution. An attempt was also made with uniform distributions with no luck for any parameter pair.

## 7 Conclusions and future work

We saw that ME distributions are a useful tool in approximating general queuing systems. The most difficult part of checking the validity of ME distributions is the analysis of the non-negativity of ME density functions. We propose a general order reduction approach for the analysis of the non-negativity of order  $n$  ME functions. In case of order 3 ME functions this approach results in explicit expressions for all possible cases. The approach was also able to give useful results for order 4 ME distributions that have not been analyzed before. On a concrete example we demonstrated that it is indeed worth using ME distributions for approximations, since greater accuracy can be achieved.

Future work can take multiple directions. Research plans include finding non-trivial necessary or sufficient conditions for higher order ME membership. Typically, in practice higher order distributions are needed for adequate approximation, so an implementation of the method using numerical subroutines could be useful. Perhaps most valuable would be to develop a moment matching based optimization algorithm that stays within the class of ME distributions or at least gives valid results in the majority of cases. This would be useful, since we experienced in our tests that the applied moment matching procedure very rarely gave a valid distribution.

## References

- [1] ASMUSSEN, S. AND O’CINNEIDE, C. A. (1997). Matrix-exponential distributions – distributions with a rational Laplace transform. In *Encyclopedia of Statistical Sciences*, S. Kotz and C. Read, Eds. John Wiley & Sons, New York, 435–440.
- [2] BEAN, N. G., FACKRELL, M., AND TAYLOR, P. (2008). Characterization of matrix-exponential distributions. *Stochastic Models* 24, 339–363.
- [3] FACKRELL, M. W. (2003). Characterization of matrix-exponential distributions. Ph.D. thesis, The University of Adelaide.
- [4] HORVÁTH, A., RÁCZ, S., AND TELEK, M. (2009). Moments characterization of the class of order 3 matrix exponential distributions. In *16th Int.*

*Conf. on Analytical and Stochastic Modeling Techniques and Applications (ASMTA)*. LNCS, Vol. **5513**. Springer, 174 – 188.

- [5] HORVÁTH, A. AND TELEK, M. (2002). Phfit: A general purpose phase type fitting tool. In *Tools 2002*. Springer, LNCS 2324, London, England, 82–91.
- [6] LATOUCHE, G. AND RAMASWAMI, V. (1999). *Introduction to matrix analytic methods in stochastic modeling*. SIAM.
- [7] MITCHELL, K. AND VAN DE LIEFVOORT, A. (2003). Approximation models of feed-forward g/g/1/n queueing networks with correlated arrivals. *Performance Evaluation* 51, 137–152.
- [8] TELEK, M. AND HORVÁTH, G. (2007). A minimal representation of markov arrival processes and a moments matching method. *Performance Evaluation* 64, 9-12 (Aug.), 1153–1168.
- [9] VAN DE LIEFVOORT, A. (1990). The moment problem for continuous distributions. Tech. Rep. WP-CM-1990-02, University of Missouri, Kansas City.