Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 610907, 13 pages
http://dx.doi.org/10.1155/2014/610907

Research Article

Constructing Matrix Exponential Distributions by Moments and Behavior around Zero

Alessio Angius,[1] András Horváth,[1] Sami M. Halawani,[2] Omar Barukab,[2] Ab Rahman Ahmad,[2] and Gianfranco Balbo[1,2]

[1] *Dipartimento di Informatica, Università di Torino, Corso Svizzera 185, 10149 Torino, Italy*
[2] *Faculty of Computing and Information Technology in Rabigh, King Abdulaziz University, Rabigh, P.O. Box 344, Rabigh 21911, Saudi Arabia*

Correspondence should be addressed to András Horváth; horvath@di.unito.it

Received 1 September 2014; Accepted 5 December 2014; Published 24 December 2014

Academic Editor: Bo Shen

Copyright © 2014 Alessio Angius et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper deals with moment matching of matrix exponential (ME) distributions used to approximate general probability density functions (pdf). A simple and elegant approach to this problem is applying Padé approximation to the moment generating function of the ME distribution. This approach may, however, fail if the resulting ME function is not a proper probability density function; that is, it assumes negative values. As there is no known, numerically stable method to check the nonnegativity of general ME functions, the applicability of Padé approximation is limited to low-order ME distributions or special cases. In this paper, we show that the Padé approximation can be extended to capture the behavior of the original pdf around zero and this can help to avoid representations with negative values and to have a better approximation of the shape of the original pdf. We show that there exist cases when this extension leads to ME function whose nonnegativity can be verified, while the classical approach results in improper pdf. We apply the ME distributions resulting from the proposed approach in stochastic models and show that they can yield more accurate results.

1. Introduction

Probability distributions that can be expressed as the composition of exponential stages have gained widespread acceptance in recent years, specifically as ways of modeling nonexponential durations maintaining the Markov property of the underlying stochastic process. Distributions with these characteristics are represented by the family of phase type (PH) distributions which are given by the distribution of time to absorption in Markov chains (see, e.g., Chapter 2 in [1]). In the literature, mainly continuous-time PH distributions have been studied, but more recently some works that deal with the discrete-time version have also been proposed [2]. Phase type distributions are practically interesting since they allow studying with Markov models problems that are characterized by distributions with coefficient of variation either smaller or larger than one.

Matrix exponential (ME) distributions [3] are an extension of the PH class. They can provide more compact approximation than the PH class [4] and can capture cases in which the density functions that we want to represent may be zero on positive real numbers (i.e., they may exhibit multimodal shapes). ME distributions have the same algebraic form as PH distributions, but they do not enjoy a simple stochastic interpretation (a stochastic interpretation of ME distributions is given in [5], but this is much more complicated and less practical than that possessed by PH distributions).

The main reason to apply PH and ME distributions in stochastic modeling is that these distributions can be easily used as building blocks of more complex models. This fact is well-known for PH distributions because it is straightforward that if we are given a system in which all sojourn times are according to PH distributions and the next state distribution is Markov then the overall system behavior can be described by a Markov chain. Starting from [6], the construction of the Markov chain, often referred to as the expanded Markov chain, was proposed in the literature for several modeling formalisms, such as stochastic Petri nets (SPN), stochastic process algebras, or stochastic automata networks [7–9].

The downside of dealing with the expanded chain is that if the model has many states and/or it describes many activities performed in parallel, then the chain can have a huge number of states. To alleviate this problem several authors proposed techniques for the compact representation of the expanded chain. Such techniques are based on either Kronecker algebra (among the first works see, e.g., [10]) or, more recently, decision diagrams techniques (e.g., [11]).

The necessary theoretical background to use ME distributions as building blocks was developed instead only recently. In this case the overall model is not a Markov chain, but the transient system behavior can still be described by a set of ordinary differential equations. In the context of SPN, this was shown for a subclass of ME distributions in [12], while in [4] the result was extended to the whole family. In the context of quasi-birth-and-death processes the possibility of using ME distributions was investigated in [13].

In order to apply PH or ME distributions in stochastic modeling one needs methods to create distributions that capture real measured durations. Most of the existing techniques fall into two categories: those that are based on the maximum likelihood (ML) principle and those that aim to match moments.

One of the first works on ML estimation considered acyclic PH distributions [14] (i.e., PH distributions whose underlying Markov chain is acyclic), while an approach for the whole family, based on the expectation-maximization method, is proposed in [15]. Since these early papers, many new methods and improvements have been suggested for the whole PH family and for its subclasses (see, e.g., [16, 17]). Much less research tackled ML based fitting of ME distributions because of the lack of a practical stochastic interpretation. One such method, based on semi-infinite programming, is described however in Chapter 9 of [18] where the computational complexity of the problem is discussed and an algorithm is devised.

For what concerns moment matching methods the following results are available. For low-order (≤3) PH and ME distributions (the order of PH and ME distributions is the size of their generator matrices) either moment bounds and moment matching formulas are known in an explicit manner or there exist iterative numerical methods to check if given moments are possible to capture [19–21]. For higher orders there exist matching algorithms, but these often result in improper density functions (by proper pdf we mean a pdf which is nonnegative and normalized) and the validity check is a nontrivial problem [18, 22]. In [23] a simple method is provided that constructs a minimal order acyclic PH distribution given three moments. Characterization of moments of PH and ME distributions is discussed in [24]. Moreover, tool support is available for the construction of PH and ME distributions. Specifically, ML based fitting is implemented in PhFit [25] and a set of moment matching functions is provided in BuTools [26].

In this paper we consider moment matching of ME distributions by Padé approximation [18]. This method provides an order-n ME distribution given $2n-1$ moments. The problem is that the probability density function (pdf) associated with the distribution can be improper (i.e., it may assume negative values). We show that Padé approximation can be extended in such a way that it considers not only moments, but also the behavior of the original pdf around zero (called also zero-behavior in the sequel). We present examples where this extension leads to ME distributions with proper pdf for cases where the original approach results in improper pdf. Moreover, it can give a better approximation of the shape of the original pdf.

The paper is organized as follows. In Section 2, we provide the necessary theoretical background. The extended Padé approximation is introduced in Section 3. The relation between the zero-behavior of ME distribution and its moments is discussed in Section 4. Numerical illustration of the proposed approach is provided in Section 5. Application of ME distributions resulting from the extended Padé approximation in stochastic models is discussed in Section 6. The paper is concluded then in Section 7.

2. Background

The cumulative distribution function of a matrix exponential random variable, X, is of the form

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

where v is a row vector, referred to as the *initial vector*, A is a square matrix, referred to as the *generator*, and $\mathbb{1}$ is a column vector of ones. When the cardinality of the vector v is n, then the distribution is called *order-n matrix exponential distribution* (ME(n)). The vector-matrix pair (v, A) is called the representation of the distribution. Throughout the paper we assume that the distribution is without mass at time zero, that is, $v\mathbb{1} = 1$, but the extension to the case with mass at zero is straightforward. Note that the entries of v may be negative; that is, it is not necessarily a probability vector. If v is a probability vector and A is the infinitesimal generator of a transient Markov chain, then the distribution belongs to the class of PH distributions.

The pdf of ME distribution can be computed as

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

and the moment generator function can be expressed as

$$\begin{aligned} f^*(s) &= \mathscr{M}\{f(t)\} = E\left[e^{sX}\right] \\ &= \int_0^\infty e^{sx} f(x)\,dx = v(sI + A)^{-1} A\mathbb{1}, \end{aligned} \tag{3}$$

where I is the identity matrix. The last expression results in a rational function of the form

$$f^*(s) = \frac{a_0 + a_1 s + \cdots + a_{n-1}s^{n-1}}{b_0 + b_1 s + \cdots b_{n-1}s^{n-1} + s^n}, \tag{4}$$

where $a_0 = b_0$ if there is no probability mass at zero (i.e., if $v\mathbb{1} = 1$ as we have assumed before). If the ME distribution is *nonredundant*, then the number of poles of (4) is equal to the order of the matrix representation of the ME distribution [27]. In this paper, we consider only nonredundant distributions.

The moments can be obtained from the moment generating function and the ith moment is

$$m_i = E\left[X^i\right] = i!\nu(-A)^{-i}\mathbb{1}. \quad (5)$$

It can be seen from (4) that ME(n) distribution without probability mass at zero is determined by $2n-1$ parameters and, consequently, $2n-1$ moments can be matched by ME(n) if the moments are inside the moment bounds of the ME(n) family (i.e., if the $2n-1$ moments are such that they can be realized by ME(n) distribution).

There exist methods to construct ME distributions given moments. The applicability of these methods is limited by the facts that

(i) explicit moment bounds of the ME family are known only for low-order ($n \le 2$) ME distributions,

(ii) the validity check of the pdf for higher order ($n > 3$) ME distribution is an open research problem.

In the following we report those results that are relevant to our paper.

For order-2 distributions the PH and the ME families are equivalent and the moment bounds are provided in [19] in terms of the squared coefficient of variation defined as $c^2 = (m_2 - m_1^2)/m_1^2 = m_2/m_1^2 - 1$. The bounds are

$$\begin{gathered} 0 < m_1 < \infty, \\ \frac{1}{2} \le c^2 < \infty, \\ 3m_1^3\left(3c^2 - 1 + \sqrt{2}\left(1-c^2\right)^{3/2}\right) \le m_3 \le 6m_1^3c^2 \quad \text{if } \frac{1}{2} \le c^2 \le 1, \\ \frac{3}{2}m_1^3\left(1+c^2\right)^2 < m_3 < \infty \quad \text{if } 1 < c^2, \end{gathered} \quad (6)$$

meaning that if three moments satisfy the above conditions, then there exists a proper ME(2) distribution that realizes the moments.

For order-3 ME distributions the moment bounds are not known explicitly. In [21] an algorithm is proposed which, given ME(3) density, checks if it is proper. In some cases the algorithm requires the numerical solution of a transcendent equation. This means that, given 5 moments, one can create ME(3) distribution and then check its validity by the method provided in [21]. This check is implemented in the BuTools package [26].

3. Padé Approximation

Padé approximation for ME distributions exploits the fact that for any distribution the following relation holds between the moment generating function and the moments of the distribution:

$$g^*(s) = \sum_{i=0}^{\infty} \frac{m_i}{i!} s^i, \quad (7)$$

where we assume having $m_0 = 1$ which implies that the pdf is normalized. Accordingly, one may consider approximating the moment generating function by a rational function as

$$\begin{aligned} f^*(s) &= \frac{a_0 + a_1 s + \cdots + a_{n-1}s^{n-1}}{a_0 + b_1 s + \cdots b_{n-1}s^{n-1} + s^n} \\ &\approx \sum_{i=0}^{2n-1} \frac{m_i}{i!} s^i + \sum_{i=2n}^{\infty} \frac{m_i}{i!} s^i = g^*(s), \end{aligned} \quad (8)$$

where the $2n-1$ parameters of the moment generating function of the ME distribution, that is, a_i, $0 \le i \le n-1$ and b_i, $1 \le i \le n-1$, can be determined based on the first $2n-1$ moments of the distribution whose moments we aim to match. More precisely, we look for such parameters with which the first $2n$ coefficients of the polynomial

$$a_0 + a_1 s + \cdots + a_{n-1}s^{n-1} \quad (9)$$

are equal to the first $2n$ coefficients of the polynomial

$$\left(a_0 + b_1 s + \cdots b_{n-1}s^{n-1} + s^n\right) \sum_{i=0}^{2n-1} \frac{m_i}{i!} s^i. \quad (10)$$

The above corresponds to solving a linear system of $2n$ equations given as

$$a_i = \sum_{j=0}^{i} b_{i-j} \frac{m_j}{j!}, \quad (11)$$

with $a_i = 0$ if $i > n-1$, $b_0 = a_0$, $b_n = 1$, and $b_i = 0$ if $i > n$.

The above-described procedure results in a moment generating function that exactly matches the first $2n-1$ moments of the original distribution. It may, however, fail as there is no guarantee for the nonnegativity of the corresponding pdf.

The procedure can be extended to capture the behavior of the original pdf around zero. This is done by adding new equations to those provided by (11) and removing those referring to the highest order moments. For example, the pdf of ME(n) at zero assumes simply the value

$$f(0) = \lim_{s \to -\infty} -sf^*(s) = -a_{n-1}, \quad (12)$$

while the derivative at zero can be obtained as

$$f'(0) = \lim_{s \to -\infty} \left(s^2 f^*(s) + sf(0)\right) = a_{n-2} - a_{n-1}b_{n-1}. \quad (13)$$

Matching $f(0)$ means that the number of moments that can be captured is decreased to $2n-2$, while matching both $f(0)$ and $f'(0)$ means that we can capture $2n-3$ moments.

In general, the value of the kth derivative of the pdf at zero can be obtained by applying two properties of the moment generating functions. First, the moment generating function of the kth derivative of $f(t)$ is given by

$$\mathcal{M}\left\{f^{(k)}(t)\right\} = f^{(k)*}(s) = (-s)^k f^*(s) - \sum_{l=1}^{k} (-s)^{k-l} f^{(l-1)}(0). \quad (14)$$

Second, the value at zero can be obtained as

$$f^{(k)}(0) = \lim_{s \to -\infty} - s f^{(k)*}(s). \quad (15)$$

Equations (14) and (15) provide a recursive calculation procedure to calculate $f^{(k)}(0)$ as function of the parameters of the generating function. For example, in case of $n \geq 5$, we have

$$\begin{aligned} f^{(2)}(0) &= -a_{n-3} + a_{n-2}b_{n-1} + a_{n-1}\left(b_{n-2} - b_{n-1}^2\right), \\ f^{(3)}(0) &= a_{n-4} - a_{n-3}b_{n-1} + a_{n-2}\left(b_{n-1}^2 - b_{n-2}\right) \\ &\quad - a_{n-1}\left(b_{n-1}^3 - 2b_{n-2}b_{n-1} + b_{n-3}\right), \\ f^{(4)}(0) &= -a_{n-5} + a_{n-4}b_{n-1} + a_{n-3}\left(b_{n-2} - b_{n-1}^2\right) \\ &\quad + a_{n-2}\left(b_{n-1}^3 - 2b_{n-2}b_{n-1} + b_{n-3}\right) \\ &\quad + a_{n-1}\left(-b_{n-1}^4 + 3b_{n-2}b_{n-1}^2 - 2b_{n-3}b_{n-1} \right. \\ &\qquad \left. - b_{n-2}^2 + b_{n-4}\right). \end{aligned} \quad (16)$$

It can be seen from (12) that matching $f(0)$ directly determines a_{n-1}. Having determined a_{n-1}, (13) provides a linear equation that must be satisfied if one aims to capture $f'(0)$. In general, capturing $f^{(k)}(0)$ for $0 \leq k \leq m$ adds simple linear equations to those that are used to match the moments.

4. Relation of Derivatives at Zero and Moments by Hankel Matrices

In this section, we discuss the relation between the zero-behavior of ME distribution and its moments. Let us consider an order n ME distribution with representation (v, A) and introduce the $k \times k$ Hankel matrix

$$H_{i,k} = \begin{pmatrix} r_i & r_{i+1} & \cdots & r_{i+k-1} \\ r_{i+1} & r_{i+2} & \cdots & r_{i+k} \\ \vdots & \vdots & & \vdots \\ r_{i+k-1} & r_{i+k} & \cdots & r_{i+2(k-1)} \end{pmatrix}, \quad (17)$$

where $r_i = v(-A)^{-i}\mathbb{1}$. The entries of $H_{i,k}$ have the following interpretation:

(i) for $i > 0$ we have that r_i is related to the ith moment of the distribution through $r_i = m_i/i!$,

(ii) $r_0 = 1$,

(iii) for $i < 0$ we have that r_i is related to the derivative at zero of the cdf of the ME distribution because

$$\left.\frac{d^i F(t)}{dt^i}\right|_{t=0} = -vA^i\mathbb{1} = (-1)^{i+1} v(-A)^i \mathbb{1} = (-1)^{i+1} r_{-i}. \quad (18)$$

The rank of the $H_{i,k}$ equals $\min(n, k)$ [27] which implies that $\det(H_{i,k})$ is nonzero for $k \leq n$ and zero for $k > n$. As a consequence, given the quantities $r_i, r_{i+1}, \ldots, r_{i+2n-1}$, one can determine r_{i+2n} based on the equation

$$\det\left(H_{i,n+1}\right) = 0. \quad (19)$$

For example, given 3 moments of ME(2) distribution, the fourth moment can be calculated by considering

$$\det\left(H_{0,3}\right) = \det\begin{pmatrix} 1 & m_1 & \frac{m_2}{2} \\ m_1 & \frac{m_2}{2} & \frac{m_3}{6} \\ \frac{m_2}{2} & \frac{m_3}{6} & \frac{m_4}{24} \end{pmatrix} = 0, \quad (20)$$

or, given $f(0)$ and 2 moments, the third moment can be determined based on

$$\det\left(H_{-1,3}\right) = \det\begin{pmatrix} f(0) & 1 & m_1 \\ 1 & m_1 & \frac{m_2}{2} \\ m_1 & \frac{m_2}{2} & \frac{m_3}{6} \end{pmatrix} = 0. \quad (21)$$

The above relation of the derivatives and the moments can be exploited to transform moment matching procedures into zero-behavior matching procedures or into procedures matching part of the zero-behavior and some moments. Moreover, existing moment bounds can be transformed into bounds regarding zero-behavior.

For example, the set of inequalities provided in (6) can be transformed into inequalities regarding $f(0), m_1$, and m_2 by solving (21) for m_3 and using the result in (6). This leads to

$$\begin{aligned} & 0 < m_1 < \infty, \\ & \frac{1}{2} \leq c^2 < \infty, \\ & 3m_1^3\left(3c^2 - 1 + \sqrt{2}\left(1 - c^2\right)^{3/2}\right) \\ & \quad \leq \frac{3\left(4m_1^3 - 4m_1m_2 + f(0)m_2^2\right)}{2f(0)m_1 - 2} \leq 6m_1^3c^2 \\ & \qquad \text{if } \frac{1}{2} \leq c^2 \leq 1, \\ & \frac{3}{2}m_1^3\left(1 + c^2\right)^2 < \frac{3\left(4m_1^3 - 4m_1m_2 + f(0)m_2^2\right)}{2f(0)m_1 - 2} < \infty \\ & \qquad \text{if } 1 < c^2, \end{aligned} \quad (22)$$

meaning that if $f(0), m_1$, and m_2 satisfy the above conditions, then there exists a proper ME(2) distribution that realizes them.

5. Numerical Examples

In this section we provide numerical examples of considering the zero-behavior in constructing ME distributions. In Section 5.1, we show that there are cases in which considering the moments of the distribution leads to ME distribution with improper pdf, while considering also the behavior around zero results in a proper distribution. In Section 5.2, we provide an example for which moment matching gives a proper pdf, but the shape of the pdf can be improved by considering the value of the original pdf at zero instead of considering the highest order moment. In Section 5.3, we show that even if the original characteristics around zero are not possible to capture, it is possible to set the characteristics of the ME distribution around zero in such a way that the fitting of the shape of the pdf improves considerably.

5.1. Proper Distribution by Behavior around Zero. Consider the lognormal distribution whose pdf is given by

$$f_{\ln}(t) = \frac{e^{-(\log(t)-\alpha)^2/2\beta^2}}{\sqrt{2\pi}t\beta}, \tag{23}$$

with parameters

$$\alpha = 0, \qquad \beta = \frac{1}{2}. \tag{24}$$

Using our proposed approach, we derived two order-3 ME distributions to match the characteristics of the above pdf. The first one matches 5 moments and has an improper pdf as it assumes negative values around 0. Its representation is given by the vector-matrix pair

$$\begin{gathered}\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right),\\ \begin{pmatrix}-2.67775 & 20.4761 & -14.6936\\ -0.351307 & -1.18954 & 0.753936\\ -0.237024 & 0.601207 & -1.34227\end{pmatrix}.\end{gathered} \tag{25}$$

The second one matches three moments and also $f_{\ln}(0)$ and $f_{\ln}'(0)$ and results in a proper distribution. Its representation is

$$\begin{gathered}\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right),\\ \begin{pmatrix}-1.87683 & 21.0592 & -17.4174\\ -0.395566 & -1.22176 & 0.90445\\ -0.192765 & 0.633427 & -1.49278\end{pmatrix}.\end{gathered} \tag{26}$$

The validity of the second ME distribution was checked by using BuTools [26]. In Figure 1, we depicted the original pdf and the two matching ME(3) distributions. In Table 1, we provide the first five moments of the original and the second matching distribution (the moments of the first one are identical to those of the original distribution). For the second ME distribution, even if the 4th and 5th moments are not matched, they are close to those of the lognormal distribution and also the shape of the pdf is similar to that of the lognormal distribution.

5.2. Improving the Shape of the ME Pdf by Matching Exactly the Behavior around Zero. We consider now the three-parameter Weibull distribution commonly used to model time to failure. The pdf is of the form

$$f_{3w}(t) = e^{-((x+\alpha)/\beta)^{\gamma}}(x+\alpha)^{\gamma-1}\beta^{-\gamma}\gamma, \tag{27}$$

Table 1: Moments of the lognormal distribution and moments of the second matching ME distribution.

	1st	2nd	3rd	4th	5th
Lognormal	1.13315	1.64872	3.08022	7.38906	22.7599
Order-3 ME	1.13315	1.64872	3.08022	7.58974	24.5879

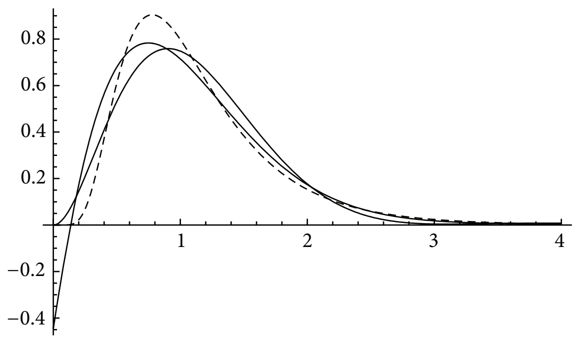

Figure 1: Pdf of the lognormal distribution with $\alpha = 0$ and $\beta = 1/2$ (dashed line); moment matching based on 5 moments (solid line curve with negative values around zero); moment matching based on 3 moments and the value of the original pdf at zero and the derivative of the original pdf at zero (solid line curve with no negative values around zero).

where γ is the so-called shape parameter, β is the scale parameter, and α determines the shift with respect to the two-parameter Weibull distribution. We assume that the parameters are

$$\gamma = \frac{1}{2}, \qquad \beta = 2, \qquad \alpha = \frac{1}{2}. \tag{28}$$

We constructed two order-2 ME distributions to match the three-parameter Weibull distribution. The first one matches the first three moments of the distribution and has a proper pdf, but the shape of the pdf is rather different from that of the original one. The second ME distribution matches instead two moments and the value of the original pdf at zero $f_{3w}(0)$. The validity of the matching ME distributions is guaranteed by the fact that the moments are in the area defined by (6). In Figure 2, we show the original pdf and the pdf of the two matching distributions. The representations of the two ME(2) distributions are

$$\begin{gathered}\left(\frac{1}{2},\frac{1}{2}\right), \qquad \begin{pmatrix}-0.228593 & -0.0909786\\ -0.0833333 & -0.0833333\end{pmatrix},\\ \left(\frac{1}{2},\frac{1}{2}\right), \qquad \begin{pmatrix}-0.716667 & -0.116667\\ -0.0833333 & -0.0833333\end{pmatrix}.\end{gathered} \tag{29}$$

In Figure 3, we show instead two order-3 ME distributions. The first one captures 5 moments of the shifted Weibull distribution, while the second one captures 3 moments as well as the value of the pdf and of the first derivative of the pdf at zero. It can be seen that not even 5 moments are enough to have ME distribution whose shape is similar to that of the original distribution. Instead, taking into account the

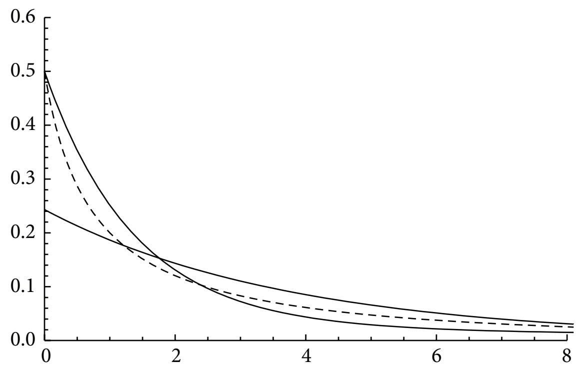

Figure 2: Pdf of a three-parameter Weibull distribution (dashed line); moment matching based on 3 moments (solid line curve far from the dashed line); moment matching based on 2 moments and the value of the original pdf at zero (solid line curve closer to the dashed line).

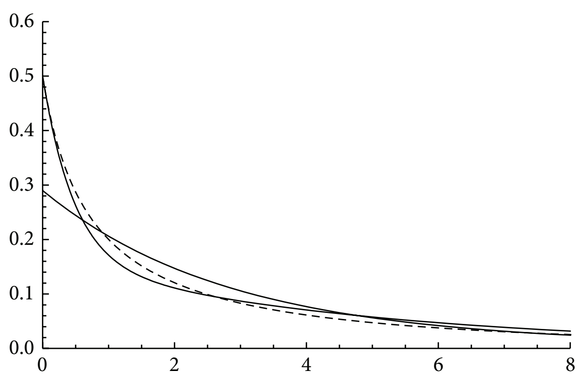

Figure 3: Pdf of a three-parameter Weibull distribution (dashed line); moment matching based on 5 moments (solid line curve far from the dashed line); moment matching based on 3 moments and the value and the first derivative of the original pdf at zero (solid line curve closer to the dashed line).

behavior around zero improves substantially the fitting of the shape. The representations of the two ME(3) distributions are

$$\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right),$$
$$\begin{pmatrix} -0.289888 & -0.0963572 & -0.15042 \\ 0.0163624 & -0.0565627 & -0.0129118 \\ -0.127473 & -0.0545485 & -0.0981994 \end{pmatrix},$$
$$\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right), \tag{30}$$
$$\begin{pmatrix} -0.688889 & 0.725397 & -1.20317 \\ 0.338519 & -0.720053 & 0.83709 \\ -0.44963 & 0.608942 & -0.948201 \end{pmatrix}.$$

In Table 2, we provide the first five moments of the original shifted Weibull distribution and those of the ME distributions depicted in Figures 2 and 3. The second ME(2) distribution gives a bad fit already of the third moment. It is interesting to note that the second ME(3) distribution gives

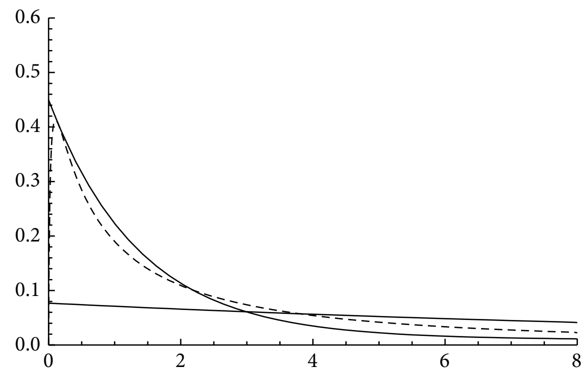

Figure 4: Pdf of the lognormal distribution with $\alpha = 1$ and $\beta = 1.8$ (dashed line); moment matching based on 5 moments (solid line curve far from the dashed line); moment matching based on 3 moments and the shape of the original pdf around zero (solid line curve closer to the dashed line).

a slightly better fit of the fourth and fifth moments than the first ME(2) distribution and this is thanks to the better fitting of the shape of the distribution.

5.3. Improving the Shape of the ME Pdf by Approximating the Behavior around Zero. We consider again a lognormal distribution, now with parameters

$$\alpha = 1, \qquad \beta = 1.8. \tag{31}$$

In this case, matching 5 moments results in ME(3) with pdf which is considerably far from the original one. Its representation is

$$\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right),$$
$$\begin{pmatrix} -0.0286601 & -0.028144 & -0.0281441 \\ -0.00912991 & -0.0109089 & -0.010908 \\ -0.0394053 & -0.0376263 & -0.0376272 \end{pmatrix}. \tag{32}$$

Matching $f(0)$ and four moments or $f'(0)$ and $f(0)$ and three moments results in ME(3)s that are not proper distributions. By observing the original pdf (Figure 4), one can see that it starts from 0, but then it immediately has a steep peak with a gradual monotonic decrease. Setting the characteristics of the ME(3) distribution around 0 as

$$f(0) = 0.45, \qquad f'(0) = -0.32 \tag{33}$$

and matching 3 moments result in a much more satisfactory fit of the original one (see Figure 4). The representation of this ME(3) is

$$\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right),$$
$$\begin{pmatrix} -0.433521 & -0.383529 & -0.387345 \\ 29.9426 & 26.2806 & 26.5629 \\ -29.9911 & -26.3291 & -26.6114 \end{pmatrix}. \tag{34}$$

TABLE 2: Moments of the three-parameter Weibull distribution and moments of the matching ME distributions.

	1st	2nd	3rd	4th	5th
3-parameter Weibull	6	152	9264	1.044×10^6	1.888×10^8
1st ME(2) (Figure 2)	6	152	9264	848870	9.929×10^7
2nd ME(2) (Figure 2)	6	152	6576	384384	2.81203×10^7
1st ME(3) (Figure 3)	6	152	9264	1.044×10^6	1.888×10^8
2nd ME(3) (Figure 3)	6	152	9264	878976	1.079×10^8

TABLE 3: Moments of the lognormal distribution with $\alpha = 1$ and $\beta = 1.8$ and moments of the second matching ME(3) distribution.

	1st	2nd	3rd	4th	5th
Lognormal	13.73	4817	4.314×10^7	9.882×10^{12}	5.745×10^{19}
ME(3)	13.73	4817	4.314×10^7	6.095×10^{11}	1.078×10^{16}

The choice given in (33) is motivated by the fact that the line $-0.32x + 0.45$ approximates well the initial decay of the lognormal pdf right after the peak. In general, numerical values for $f(0)$ and $f'(0)$ can be obtained by applying a linear fit to the original pdf.

In Table 3, we report the first five moments of the lognormal distribution and those of the second ME(3) distribution. The price we pay for fitting better the shape of the distribution is the distance of the 4th and 5th moments of the ME(3) from the original moments.

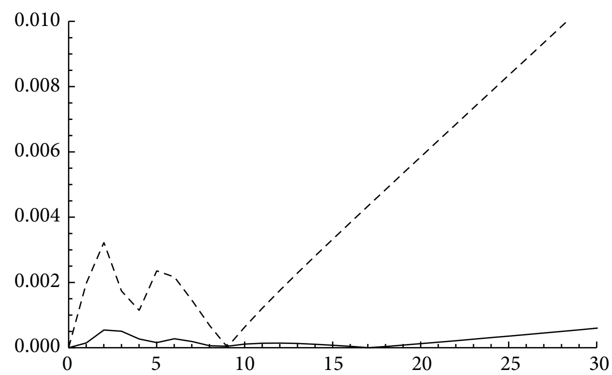

FIGURE 5: Relative error in the steady state queue length distribution of an M/G/1 queue applying the two ME distributions of Section 5.1; solid line: ME distribution with improper pdf capturing 5 moments; dashed line: ME distribution with proper pdf capturing 3 moments.

6. Application of ME Distributions in Stochastic Models

In this section, we show the impact of applying the ME distributions reported in Section 5 in stochastic modeling. In the following three subsections we illustrate three cases:

(i) the case when capturing the moments is more important than having a proper distribution or a good fit of the shape of the pdf,

(ii) the case when the use of an improper ME distribution leads to state probabilities that are either negative or larger than 1,

(iii) the case when capturing well the shape of distribution is more important than capturing the moments and leads to better approximations of the state probabilities.

6.1. M/G/1 Queue. In this section, we illustrate the goodness of the approximation applying the ME distributions in an M/G/1 queue and assuming that the steady state queue length distribution is the measure of interest.

As a first example we apply the two ME distributions presented in Section 5.1. We assume hence that the service time distribution of the queue follows a lognormal distribution with parameters $\alpha = 0$ and $\beta = 1/2$. The arrival intensity is such that the probability of the empty queue is 1/4. In Figure 5, we depicted the relative error in the steady state distribution applying the two ME distributions. Both ME distributions capture exactly the probability of the empty queue since it depends only on the mean service time. The other local minima appear in the relative error when the exact and the approximate queue length distributions cross each other. Furthermore, in an M/G/1 queue the first m moments of the steady state queue length distribution are determined by the first $m + 1$ moments of the service time distribution. Accordingly, the ME distribution with improper pdf, but capturing 5 moments, leads the queue length distribution in which 4 moments are exact. The ME distribution with proper pdf captures instead only 2 moments of the queue length exactly. As the stationary queue length distribution is of a regular shape, it is likely that the more moments are captured the better the approximation is. In Figure 5, it can be observed that indeed the ME distribution with improper pdf, but capturing more moments, outperforms the other ME distribution.

In our second example, we use as service time distribution the three-parameter Weibull distribution of Section 5.2 and apply as approximations the two ME distributions given in Figure 3. The relative error in the stationary queue length distribution is depicted in Figure 6. The same reasoning applies as in the preceding paragraph: both ME distributions capture the empty queue probability and one of them captures 4 moments of the queue length while the other one matches only 2. In this case as well, the ME distribution that captures more moments gives more precise steady state probabilities.

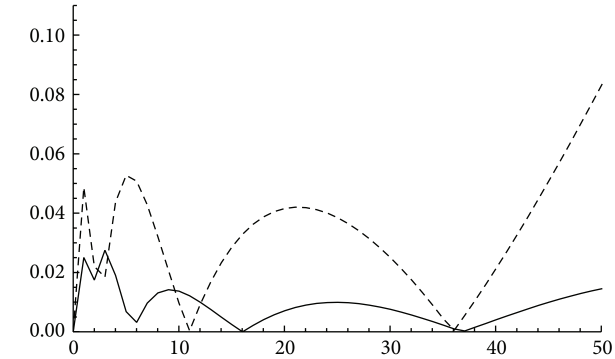

FIGURE 6: Relative error in the steady state queue length distribution of an M/G/1 queue applying the two ME distributions of Figure 3; solid line: ME distribution capturing 5 moments; dashed line: ME distribution capturing 3 moments and the shape.

6.2. Improper State Probabilities. We use in this section an extremely simple model to show that the negativity of the pdf of ME distribution, when applied in a stochastic model, can lead to negative or larger-than-one state probabilities. The model is depicted in Figure 7(a) and formulated as a stochastic Petri net [7]. It contains two concurrent activities modeled by transitions t_1 and t_2. There are four possible states, as the tokens can be distributed in the following four ways: P_1P_3, P_2P_3, P_1P_4, and P_2P_4. The state transition diagram is depicted in Figure 7(b). At the beginning firing times are chosen for both t_1 and t_2 according to their firing time distributions. Let us denote these firing times by F_1 and F_2. If $F_1 < F_2$, then after F_1 time units t_1 fires and the state becomes P_2P_3. Subsequently, after $F_2 - F_1$ time units, t_2 fires and the model arrives to its final state, that is, to state P_2P_4. If $F_2 < F_1$, then t_2 fires first, leading to state P_1P_4, and afterwards the firing of t_1 takes the model to state P_2P_4.

We assume that the firing time of transition t_1 is according to the ME(3) distribution with improper pdf reported in Section 5.1, and the firing time of t_2 is exponential with parameter λ. As described in [4], the overall behavior of the model can be described by a block matrix which algebraically plays the same role as the infinitesimal generator of a Markov chain (if the firing time distributions of t_1 and t_2 were of PH, then Q would be a proper infinitesimal generator). This block matrix is of the form

$$Q = \begin{pmatrix} A - \lambda I & -(A\mathbb{1}) & \lambda I & 0 \\ 0 & -\lambda & 0 & \lambda \\ 0 & 0 & A & -(A\mathbb{1}) \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (35)$$

where A is the generator of the ME(3) distribution and the 0s represent zero matrices of appropriate sizes. The block rows of the matrix correspond to the states in the order P_1P_3, P_2P_3, P_1P_4, and P_2P_4. For a detailed description of the construction of Q the reader is referred to [4]. Here we mention only that

(i) the matrix $A - \lambda I$ describes the parallel evolution of transitions t_1 and t_2,

(ii) the column vector $-(A\mathbb{1})$ describes the firing of transition t_1,

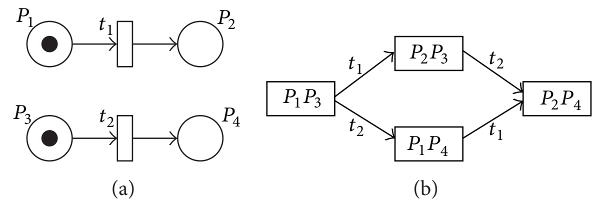

FIGURE 7: A simple Petri net with two concurrent activities (a) and the associated state transition diagram (b).

(iii) the matrix $-\lambda I$ describes the firing of t_2 maintaining the current state of t_1.

We assume that the initial state is P_1P_3 and, accordingly, the initial situation is described by the vector $\pi_0 = |\nu, 0, \ldots, 0|$, where ν is the initial vector of the ME(3) distribution. State probabilities at time t then can be obtained as π_0 and Q were the descriptors of a Markov chain; that is, we need to calculate the vector $\pi_0 e^{xQ}$ and sum up the entries according to the block structure [4]. As Q is small, namely, it is an 8 times 8 matrix, the calculation of $\pi_0 e^{xQ}$ can be done easily in many numerical computing environments.

In Figures 8 and 9 we depicted the transient state probabilities of the model for two different values of λ. In the first case, two states have negative probabilities in the time interval $[0, 0.3]$, while in the second case we find one state with larger-than-one probabilities in $[0, 0.2]$ and another state with negative probabilities in $[0, 0.3]$. This means that the negativity of the pdf of the applied ME(3) distribution can easily lead to negative and larger-than-one state probabilities in the overall distribution of the model. Moreover, the way this phenomenon presents itself depends on the remaining parameters of the model.

6.3. Importance of the Shape of the Pdf. In this section we evaluate a system composed of two machines and an intermediate finite buffer (depicted in Figure 10). This small system is often used as a building block in the approximate analysis of long production lines [28]. The first machine puts parts into the buffer and the time needed to produce a part follows the lognormal distribution considered in Section 5.2. The production of the current part is subject to a forced restart mechanism (the use of this mechanism in computing systems is studied in detail in [29]); that is, after a given amount of time the part under operation is discarded and the machine starts the production of a new one from scratch. Forced restart happens after an exponentially distributed amount of time with parameter λ. Since the lognormal distribution has a heavy tail, the forced restart mechanism decreases the mean production time of the machine. The second machine takes parts from the buffer and the time needed to treat a part follows a two-stage Erlang distribution with mean equal to 4. The first (second) machine is blocked when the buffer is full (empty). The buffer can accommodate at most 5 parts at a time.

We will evaluate the system both by simulation and by applying the two approximating ME distributions introduced

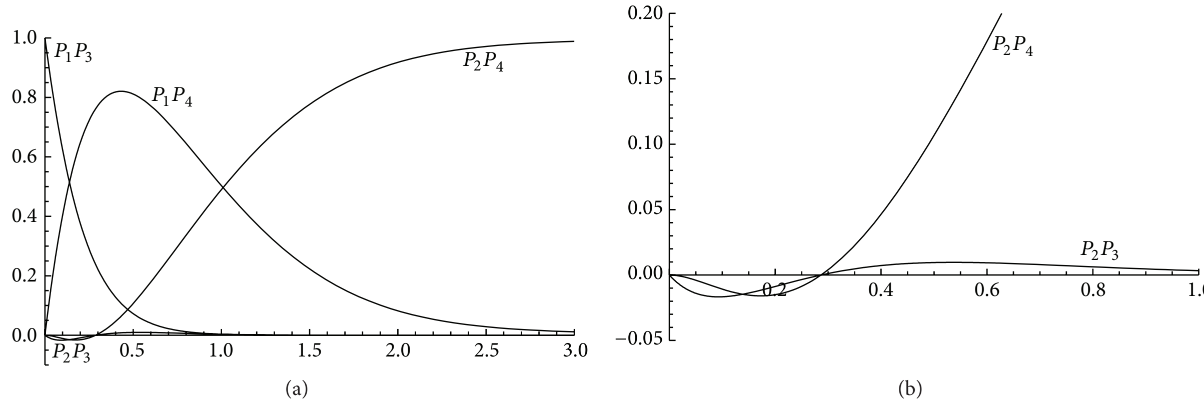

FIGURE 8: State probabilities in case of $\lambda = 5$ (a) and zoom on the negative probabilities (b).

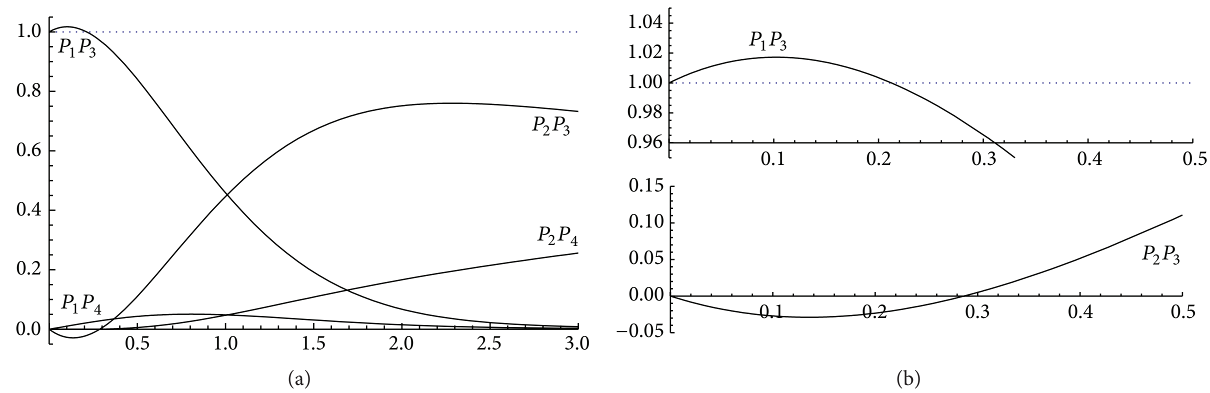

FIGURE 9: State probabilities in case of $\lambda = 0.1$ (a) and zoom on the larger-than-one and on the negative probabilities (b).

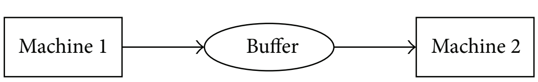

FIGURE 10: Machine-buffer-machine model.

in Section 5.2. Denoting by (v_1, A_1) the representation of the approximating ME distribution and by (v_2, A_2) the representation of the two-stage Erlang distribution, the behavior of the model is captured by the matrix

$$Q = \begin{pmatrix} A_1 - \lambda I + \lambda \mathbb{1} v_1 & -A_1 \mathbb{1} (v_1 \otimes v_2) & 0 & 0 & \cdots \\ I \otimes (-A_2 \mathbb{1}) & \bullet & (-A_1 \mathbb{1} v_1) \otimes I & 0 & \cdots \\ 0 & I \otimes (-A_2 \mathbb{1} v_2) & \bullet & (-A_1 \mathbb{1} v_1) \otimes I & \cdots \\ \cdots & \ddots & \ddots & \ddots & \cdots \\ \cdots & 0 & I \otimes (-A_2 \mathbb{1} v_2) & \bullet & (-A_1 \mathbb{1}) \otimes I \\ & \cdots & 0 & -A_2 \mathbb{1} (v_1 \otimes v_2) & A_2 \end{pmatrix}, \tag{36}$$

where the blocks in the diagonal represented by $\bullet$ are equal to

$$(A_1 \oplus A_2) - \lambda I + (\lambda \mathbb{1} v_1 \otimes I) \tag{37}$$

and $\otimes$ and $\oplus$ denote the Kronecker product and sum operators, respectively. The first block row corresponds to empty buffer (only the first machine works), while the last one refers to the full buffer (only the second machine works). As before, we do not detail the construction of Q but mention that

(i) the matrix $A_1 - \lambda I + \lambda \mathbb{1} v_1$ describes the evolution of the first machine together with possible forced restarts that reinitialize the corresponding ME distribution,

(ii) the term $A_1 \oplus A_2$ captures the parallel execution of the two machines which is then corrected by the term $-\lambda I + (\lambda \mathbb{1} v_1 \otimes I)$ in order to include the forced restart mechanism,

(iii) the matrix $(-A_1 \mathbb{1} v_1) \otimes I$ describes the completion of a part as well as the restart of the first machine and

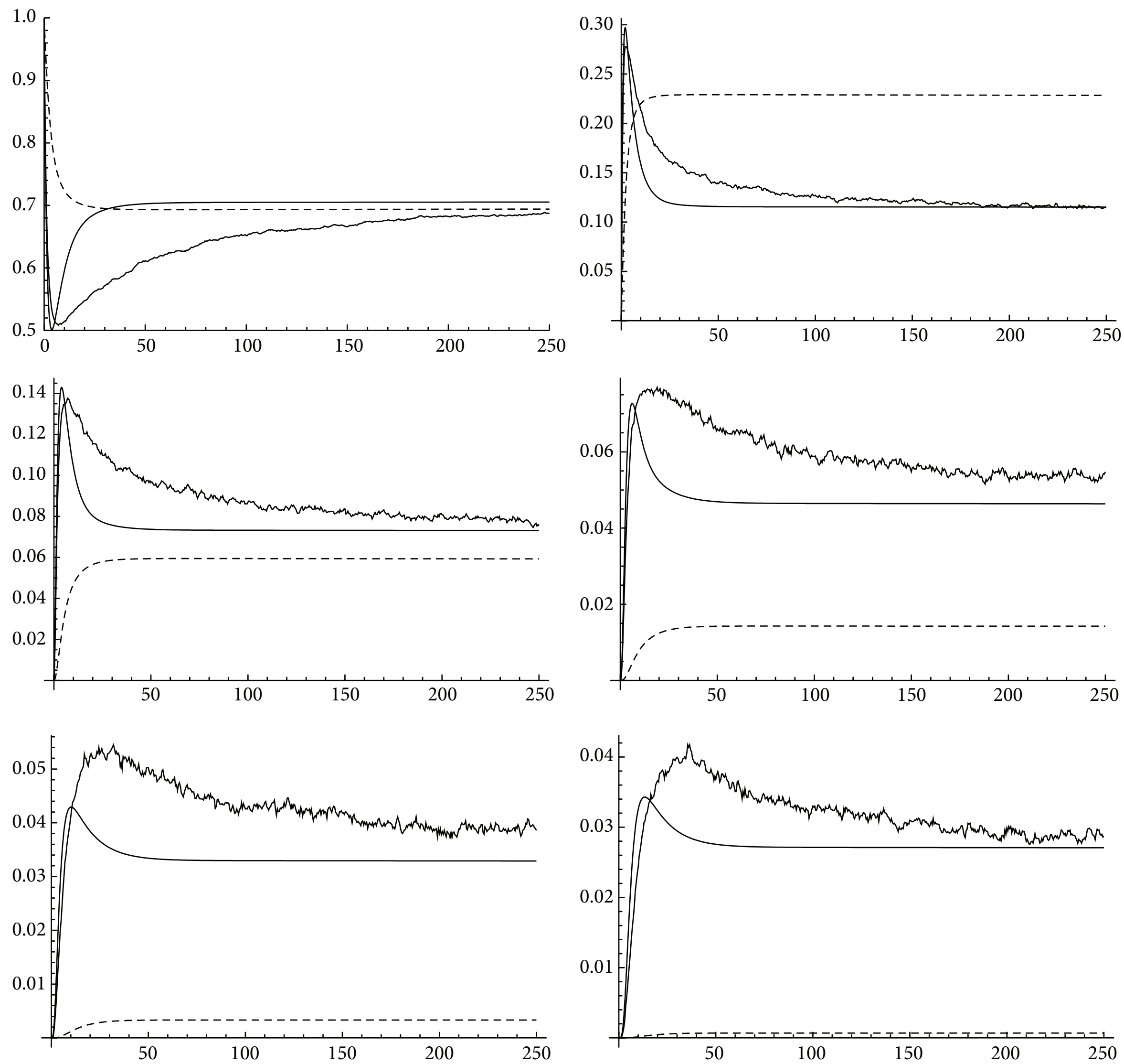

Figure 11: Transient probabilities of all possible buffer levels (from empty to full) with $\lambda = 0$; line with zigzag: simulation; dashed line: ME distribution matching 5 moments; solid line: ME distribution matching 3 moments and shape.

maintains the state of the second one (the term $I \otimes (-A_2 \mathbb{1} v_2)$ does the opposite),

(iv) the term $v_1 \otimes v_2$ describes the start of a new job for both machines (this term is used when the system leaves either the empty or the full states).

The matrix Q is 42 times 42. The first (last) three rows correspond to the empty (full) buffer situation and the remaining 4×9 rows describe the intermediate buffer levels. The transient analysis can be performed easily using a numerical computing environment.

We evaluated the model with $\lambda = 0, 0.01$ and 0.1. For what concerns simulation, the presented results are based on 50000 runs which are enough to obtain a satisfactory indication of the target distribution used to check the goodness of the approximation provided by the two approximating ME distributions. The analysis based on the ME distributions requires a few seconds on an ordinary laptop, while 50000 simulation runs require about one minute.

The results in case of $\lambda = 0$, that is, without forced restart mechanism, are depicted in Figure 11. It can be seen that the second ME distribution, which captures only three moments while taking into greater account the shape of the original pdf, outperforms the first ME distribution that captures 5 moments. It gives a more precise view on both the transient period and the long run probabilities.

In Figure 12 we provide the results for $\lambda = 0.01$. The state probabilities are less precise than in the previous case. This is due to the fact that the forced restart mechanism alters the moments of the time needed by the first machine to produce a part. This change in the moments is not captured well in the approximating model. For instance, in the original model the forced restart mechanism decreases the mean time to produce a part in the first machine from 13.7 to 9.6. In the

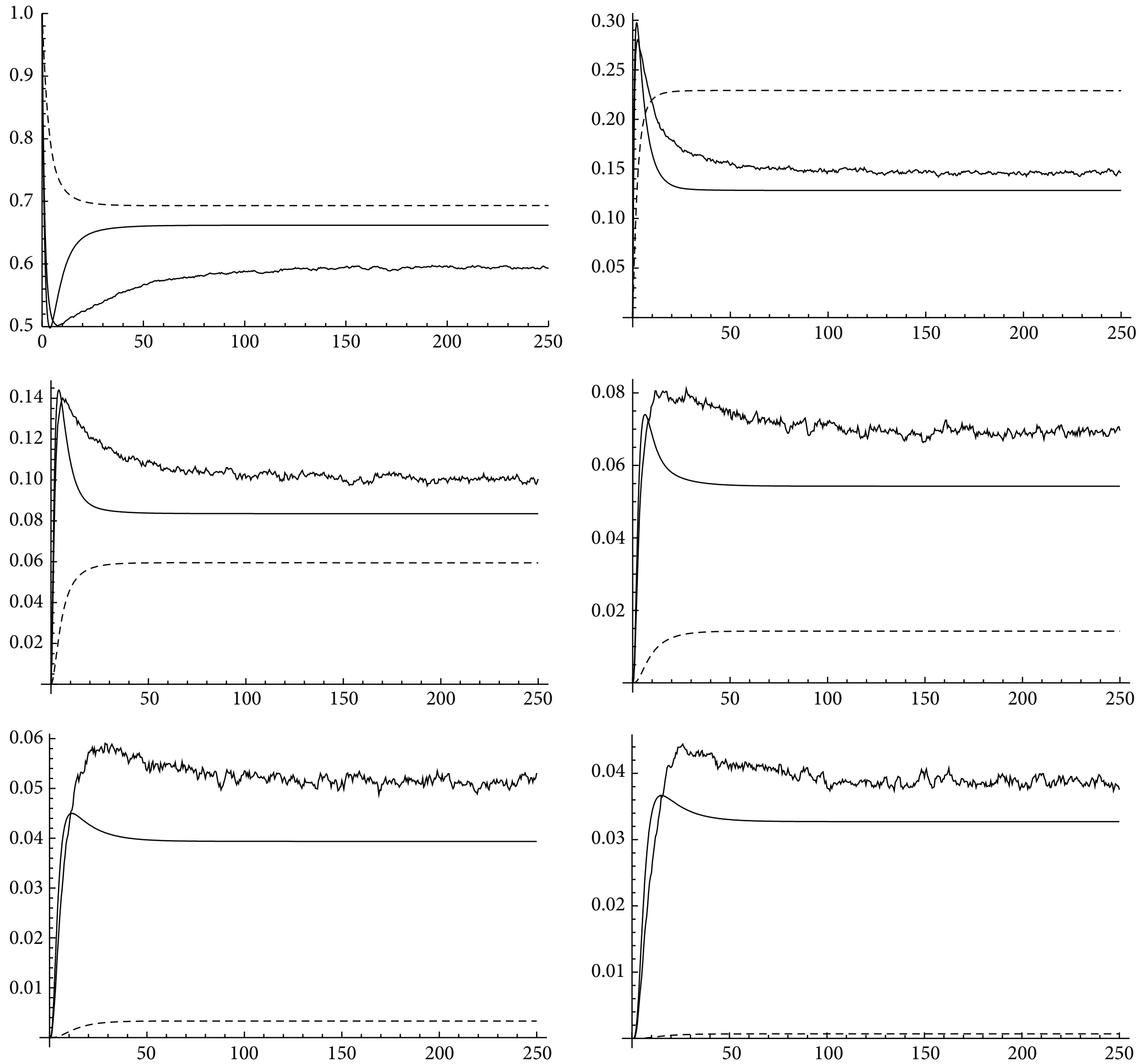

Figure 12: Transient probabilities of all possible buffer levels (from empty to full) with $\lambda = 0.01$; line with zigzag: simulation; dashed line: ME distribution matching 5 moments; solid line: ME distribution matching 3 moments and shape.

approximating model with the first ME distribution the same mean is about 13 and with the second ME distribution it is 11.4. In this case, the second ME distribution yields a better approximation of the state probabilities as well.

At last, we evaluated the model with $\lambda = 0.1$ (Figure 13). In this case the forced restart is more frequent and in the original model the mean time to produce a part for the first machine is 5.6. The same mean in the approximating model is about 13 with the first ME distribution and about 6 with the second ME distribution. The second ME distribution captures better the real mean because its pdf follows much more precisely the original lognormal distribution in the interval $[0, 10]$. Accordingly, the second ME distribution gives much better approximation of the state probabilities as well.

6.4. Discussion. There exist models, like the M/G/1 queue, in which the moments of the involved random variables determine directly moments of the measure of interest. In case of these models, if the pdf of the measure of interest is likely to have a regular shape, then it can be more convenient to capture more moments even if the shape of the approximating is a poor representation of the original one. We have seen (Section 6.1) that even a pdf with negative values can give more accurate results if it matches more moments than a proper pdf. There is no guarantee however that the negativity of the pdf does not appear in the distribution of the measure of interest, thus providing undesirable effects. Indeed, it is easy to encounter models (like the one in Section 6.2) in which the negativity of the pdf translates into negative transient probabilities. Hence it is always advisable to avoid the use of improper distribution functions.

When distributions are placed into a more complex context then the goodness of fitting the shape of the pdf can have a strong impact on the adequacy of the approximation of the overall system behavior. We have shown in Section 6.3 that

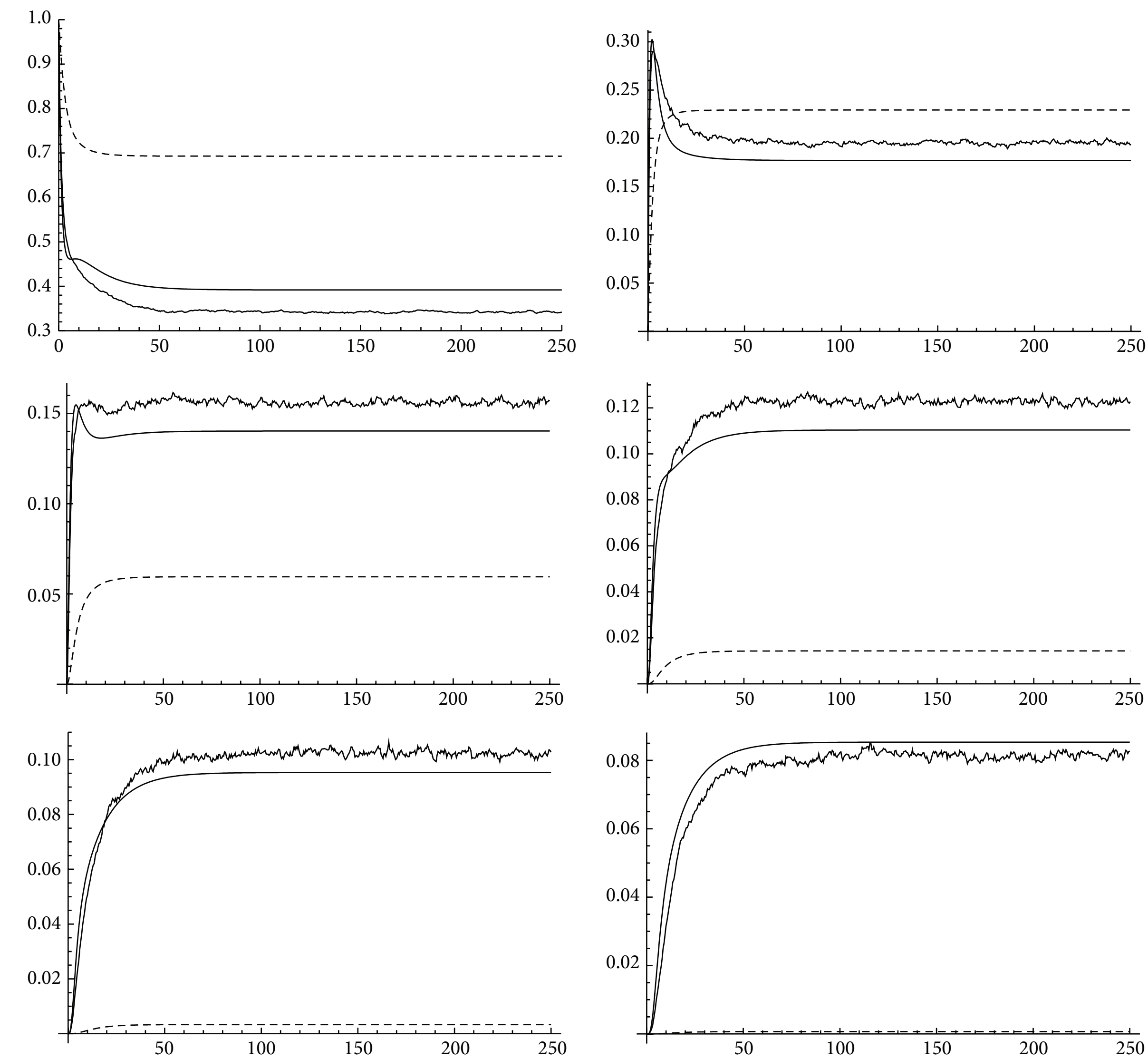

Figure 13: Transient probabilities of all possible buffer levels (from empty to full) with $\lambda = 0.1$; line with zigzag: simulation; dashed line: ME distribution matching 5 moments; solid line: ME distribution matching 3 moments and shape.

this happens when two or more activities are performed in parallel and it is important to capture precisely the probability that one is completed before the others.

7. Conclusions

In this paper, we proposed using the behavior of the pdf around zero in constructing matrix exponential distributions. We have shown that matching these characteristics can be incorporated easily into the well-known Padé approximation. We illustrated by numerical examples that matching the behavior around zero can be beneficial when matching only the moments results in improper density functions or in a density function whose behavior differs a lot from the original one when evaluated at points close to zero.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work has been supported in part by "Advanced Methodologies for the Analysis and Management of Future Internet" (AMALFI) project sponsored by Universitå di Torino and Compagnia di San Paolo and by Project Grant no. 10-15-1432/HICI from King Abdulaziz University, Saudi Arabia.

References

[1] M. Neuts, *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*, Dover, 1981.

[2] A. Bobbio, A. Horváth, M. Scarpa, and M. Telek, "Acyclic discrete phase type distributions: properties and a parameter estimation algorithm," *Performance Evaluation*, vol. 54, no. 1, pp. 1–32, 2003.

[3] S. Asmussen and M. Bladt, "Renewal theory and queueing algorithms for matrix-exponential distributions," in *Matrix-Analytic Methods in Stochastic Models*, vol. 183 of *Lecture Notes in Pure and Applied Mathematics*, pp. 313–341, Dekker, New York, NY, USA, 1996.

[4] P. Buchholz, A. Horváth, and M. Telek, "Stochastic Petri nets with Low variation matrix exponentially distributed firing time," *International Journal of Performability Engineering*, vol. 7, no. 5, pp. 441–454, 2011.

[5] M. Bladt and M. F. Neuts, "Matrix-exponential distributions: calculus and interpretations via flows," *Stochastic Models*, vol. 19, no. 1, pp. 113–124, 2003.

[6] A. Cumani, "ES—a package for the evaluation of stochastic Petri nets with Phase-type distributed transition times," in *Proceedings of the International Workshop on Timed Petri Nets*, pp. 144–151, Torino, Italy, 1985.

[7] M. Ajmone Marsan, G. Balbo, G. Conte, S. Donatelli, and G. Franceschinis, *Modelling with Generalized Stochastic Petri Nets*, John Wiley & Sons, 1995.

[8] A. El-Rayes, M. Kwiatkowska, and G. Norman, "Solving in finite stochastic process algebra models through matrix-geometric methods," in *Proceedings of the 7th Process Algebras and Performance Modelling Workshop (PAPM '99)*, J. Hillston and M. Silva, Eds., pp. 41–62, University of Zaragoza, 1999.

[9] I. Sbeity, L. Brenner, B. Plateau, and W. J. Stewart, "Phase-type distributions in stochastic automata networks," *European Journal of Operational Research*, vol. 186, no. 3, pp. 1008–1028, 2008.

[10] M. Scarpa and A. Bobbio, "Kronecker representation of stochastic Petri nets with discrete PH distributions," in *Proceedings of the IEEE International Computer Performance and Dependability Symposium (IPDS '98)*, pp. 52–62, Durham, NC, USA, September 1998.

[11] G. Ciardo and A. S. Miner, "A data structure for the efficient Kronecker solution of GSPNs," in *Proceedings of the 8th International Workshop on Petri Nets and Performance Models (PNPM '99)*, pp. 22–31, Zaragoza, Spain, 1999.

[12] P. Buchholz and M. Telek, "Stochastic Petri nets with matrix exponentially distributed firing times," *Performance Evaluation*, vol. 67, no. 12, pp. 1373–1385, 2010.

[13] N. G. Bean and B. F. Nielsen, "Quasi-birth-and-death processes with rational arrival process components," *Stochastic Models*, vol. 26, no. 3, pp. 309–334, 2010.

[14] A. Bobbio and A. Cumani, "ML estimation of the parameters of a PH distribution in triangular canonical form," in *Computer Performance Evaluation*, pp. 33–46, Elsevier, 1992.

[15] O. N. S. Asmussen and M. Olsson, "Fitting phase-type distributions via the EM algorithm," *Scandinavian Journal of Statistics*, vol. 23, no. 4, pp. 419–441, 1996.

[16] A. Thümmler, P. Buchholz, and M. Telek, "A novel approach for fi tting probability distributions to trace data with the EM algorithm," in *Proceedings of the International Conference on Dependable Systems and Networks (DSN '05)*, pp. 712–721, 2005.

[17] A. Panchenko and A. Thümmler, "Efficient phase-type fitting with aggregated traffic traces," *Performance Evaluation*, vol. 64, no. 7-8, pp. 629–645, 2007.

[18] M. Fackrell, *Characterization of matrix-exponential distributions [Ph.D. thesis]*, Faculty of Engineering, Computer and Mathematical Sciences, University of Adelaide, 2003.

[19] M. Telek and A. Heindl, "Matching moments for acyclic discrete and continuous phase-type distributions of second order," *International Journal of Simulation Systems, Science & Technology*, vol. 3, no. 3-4, 2003.

[20] G. Horváth and M. Telek, "A canonical representation of order 3 phase type distributions," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4748, pp. 48–62, 2007.

[21] A. Horváth, S. Rácz, and M. Telek, "Moments characterization of order 3 matrix exponential distributions," in *Proceedings of the 16th International Conference on Analytical and Stochastic Modeling Techniques and Applications (ASMTA '09)*, vol. 5513 of *Lecture Notes in Control and Information Sciences*, pp. 174–188, Madrid, Spain, 2009.

[22] M. Telek and G. Horváth, "A minimal representation of Markov arrival processes and a moments matching method," *Performance Evaluation*, vol. 64, no. 9–12, pp. 1153–1168, 2007.

[23] A. Bobbio, A. Horváth, and M. Telek, "Matching three moments with minimal acyclic phase type distributions," *Stochastic Models*, vol. 21, no. 2-3, pp. 303–326, 2005.

[24] L. Bodrog, A. Horváth, and M. Telek, "Moment characterization of matrix exponential and Markovian arrival processes," *Annals of Operations Research*, vol. 160, pp. 51–68, 2008.

[25] A. Horváth and M. Telek, "PhFit: a general phase-type fitting tool," in *Proceedings of the 12th International Conference on Computer Performance Evaluation, Modelling Techniques and Tools (TOOLS '02)*, vol. 2324 of *Lecture Notes in Computer Science*, Imperial College, London, UK, 2002.

[26] BuTools, http://webspn.hit.bme.hu/~butools.

[27] A. van de Liefvoort, "The moment problem for continuous distributions," Tech. Rep. WP-CM-1990-02, University of Missouri, Kansas City, Mo, USA, 1990.

[28] M. Colledani, M. Ekvall, T. Lundholm, P. Moriggi, A. Polato, and T. Tolio, "Analytical methods to support continuous improvements at Scania," *International Journal of Production Research*, vol. 48, no. 7, pp. 1913–1945, 2010.

[29] K. Wolter, *Stochastic Models for Fault Tolerance: Restart, Rejuvenation and Checkpointing*, Springer, Heidelberg, Germany, 2010.

Hindawi

Submit your manuscripts at
http://www.hindawi.com

Advances in Operations Research

Advances in Decision Sciences

Journal of Applied Mathematics

Algebra

Journal of Probability and Statistics

The Scientific World Journal

International Journal of Differential Equations

International Journal of Combinatorics

Advances in Mathematical Physics

Journal of Complex Analysis

Journal of Mathematics

Mathematical Problems in Engineering

Abstract and Applied Analysis

Discrete Dynamics in Nature and Society

International Journal of Mathematics and Mathematical Sciences

Journal of Discrete Mathematics

Journal of Function Spaces

International Journal of Stochastic Analysis

Journal of Optimization