A Closed-Form Solution for Mapping General Distributions to Minimal PH Distributions

Takayuki Osogami[1] and Mor Harchol-Balter[1]

Department of Computer Science, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213, USA
{osogami, harchol}@cs.cmu.edu

Abstract. Approximating general distributions by phase-type (PH) distributions is a popular technique in queueing analysis, since the Markovian property of PH distributions often allows analytical tractability. This paper proposes an algorithm for mapping a general distribution G to a PH distribution where the goal is to find a PH distribution which matches the first three moments of G. Since efficiency of the algorithm is of primary importance, we first define a particular subset of the PH distributions, which we refer to as EC distributions. The class of EC distributions has very few free parameters, which narrows down the search space, making the algorithm efficient – In fact we provide a closed-form solution for the parameters of the EC distribution. Our solution is general in that it applies to any distribution whose first three moments can be matched by a PH distribution. Also, our resulting EC distribution requires a nearly minimal number of phases, always within one of the minimal number of phases required by any acyclic PH distribution. Lastly, we discuss numerical stability of our solution.

1 Introduction

Motivation There is a very large body of literature on the topic of approximating general distributions by phase-type (PH) distributions, whose Markovian properties make them far more analytically tractable. Much of this research has focused on the specific problem of finding an algorithm which maps any general distribution, G, to a PH distribution, P, where P and G agree on the first three moments. Throughout this paper we say that G is *well-represented* by P if P and G agree on their first three moments. We choose to limit our discussion in this paper to three-moment matching, because matching the first three moments of an input distribution has been shown to be effective in predicting mean performance for variety of many computer system models [4, 5, 19, 23, 27]. Clearly, however, three moments might not always suffice for every problem, and we leave the problem of matching more moments to future work.

Moment-matching algorithms are evaluated along four different measures:

The number of moments matched – In general matching more moments is more desirable.

The computational efficiency of the algorithm – It is desirable that the algorithm have short running time. Ideally, one would like a closed-form solution for the parameters of the matching PH distribution.

The generality of the solution – Ideally the algorithm should work for as broad a class of distributions as possible.

The minimality of the number of phases – It is desirable that the matching PH distribution, P, have very few phases. Recall that the goal is to find P which can replace the input distribution G in some queueing model, allowing a Markov chain representation of the problem. Since it is desirable that the state space of this resulting Markov chain be kept small, we want to keep the number of phases in P low.

This paper proposes a moment-matching algorithm which performs very well along all four of these measures. Our solution matches three moments, provides a closed form representation of the parameters of the matching PH distribution, applies to all distributions which can be well-represented by a PH distribution, and is nearly minimal in the number of phases required.

The general approach in designing moment-matching algorithms in the literature is to start by defining a subset $\mathcal{S}$ of the PH distributions, and then match each input distribution G to a distribution in $\mathcal{S}$. The reason for limiting the solution to a distribution in $\mathcal{S}$ is that this narrows the search space and thus improves the computational efficiency of the algorithm. Observe that n-phase PH distributions have $\Theta(n^2)$ free parameters [16] (see Figure 1), while $\mathcal{S}$ can be defined to have far fewer free parameters. For all computationally efficient algorithms in the literature, $\mathcal{S}$ was chosen to be some subset of the acyclic PH distributions, where an acyclic PH distribution is a PH distribution whose underlying continuous time Markov chain has no transition from state i to state j for all $i > j$. One has to be careful in how one defines the subset $\mathcal{S}$, however. If $\mathcal{S}$ is too small it may limit the space of distributions which can be well-represented.[1] Also, if $\mathcal{S}$ is too small it may exclude solutions with minimal number of phases.

In this paper we define a subset of the PH distributions, which we call EC distributions. EC distributions have only six free parameters which allows us to derive a closed-form solution for these parameters in terms of the input distribution G. The set of EC distributions is general enough, however, that for all distributions G that can be well-represented by a PH distribution, there exists an EC distribution, E, such that G is well-represented by E. Furthermore, the class of EC distributions is broad enough such that for any distribution G, that is well-represented by an n-phase acyclic PH distribution, there exists an EC distribution E with at most $n+1$ phases, such that G is well-represented by E.[2]

---

[1] For example, let G be a distribution whose first three moments are 1, 2, and 12. The system of equations for matching G to a 2-phase Coxian$^+$ distribution (see Figure 2) with three parameters (λ_1, λ_2, p) results in either λ_1 or λ_2 being negative. As another example, it can be shown that the generalized Erlang distribution is not general enough to well-represent all the distributions with low variability (see [17]).

[2] Ideally, one would like to evaluate the number of phases with respect to the minimal (possibly-cyclic) PH distribution, i.e., the PH distribution is not restricted to be

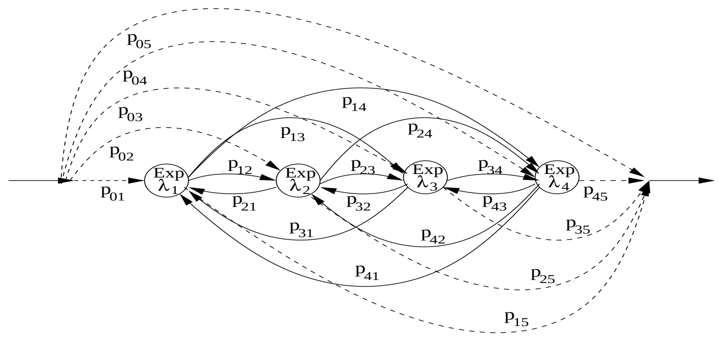


Fig. 1. A PH distribution is the distribution of the absorption time in finite state continuous time Markov chain. The figure shows a 4-phase PH distribution. There are $n = 4$ states, where the ith state has exponentially-distributed sojourn time with rate λ_i. With probability p_{0i} we start in the ith state, and the next state is state j with probability p_{ij}. Each state i has probability p_{i5} that the next state will be the absorbing state. The absorption time is the sum of the times spent in each of the states.

Preliminary Definitions Formally, we will use the following definitions:

Definition 1. *A distribution G is* **well-represented** *by a distribution F if F and G agree on their first three moments.*

The normalized moments, introduced in [18], help provide a simple representation and analysis of our closed-form solution. These are defined as follows:

Definition 2. *Let μ_k^F be the k-th moment of a distribution F for $k = 1, 2, 3$. The* **normalized k-th moment** *m_k^F of F for $k = 2, 3$ is defined to be $m_2^F = \frac{\mu_2^F}{(\mu_1^F)^2}$ and $m_3^F = \frac{\mu_3^F}{\mu_1^F \mu_2^F}$.*

Notice the correspondence to the coefficient of variability C_F and skewness γ_F of F: $m_2^F = C_F^2 + 1$ and $m_3^F = \nu_F \sqrt{m_2^F}$, where $\nu_F = \frac{\mu_3^F}{(\mu_2^F)^{3/2}}$. ($\nu_F$ and γ_F and closely related, since $\gamma_F = \frac{\bar{\mu}_3^F}{(\bar{\mu}_2^F)^{3/2}}$, where $\bar{\mu}_k^F$ is the centralized k-th moment of F for $k = 2, 3$.)

Definition 3. *$\mathcal{PH}_3$ refers to the set of distributions that are well-represented by a PH distribution.*

It is known that a distribution G is in $\mathcal{PH}_3$ iff its normalized moments satisfy $m_3^G > m_2^G > 1$ [10]. Since any nonnegative distribution G satisfies $m_3^G \geq m_2^G \geq 1$ [13], almost all the nonnegative distributions are in $\mathcal{PH}_3$.

---

acyclic. However, the necessary and sufficient number of phases required to well-represent a given distribution by a (possibly-cyclic) PH distribution is unknown.

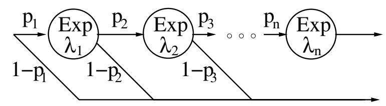


Fig. 2. An n-phase Coxian distribution is a particular n-phase PH distribution whose underlying Markov chain is of the form in the figure, where $0 \leq p_i \leq 1$ and $\lambda_i > 0$ for all $0 \leq i \leq n$. An n-phase Coxian$^+$ distribution is a particular n-phase Coxian distribution with $p_1 = 1$.

Definition 4. *$OPT(G)$ is defined to be the minimum number of necessary phases for a distribution G to be well-represented by an acyclic PH distribution.*[3]

Previous Work Prior work has contributed a very large number of moment matching algorithms. While all of these algorithms excel with respect to some of the four measures mentioned earlier (number of moments matched; generality of the solution; computational efficiency of the algorithm; and minimality of the number of phases), they all are deficient in at least one of these measures as explained below.

In cases where matching only two moments suffices, it is possible to achieve solutions which perform very well along all the other three measures. Sauer and Chandy [21] provide a closed-form solution for matching two moments of a general distribution in $\mathcal{PH}_3$. They use a two-branch hyper-exponential distribution for matching distributions with squared coefficient of variability $C^2 > 1$ and a generalized Erlang distribution for matching distributions with $C^2 < 1$. Marie [15] provides a closed-form solution for matching two moments of a general distribution in $\mathcal{PH}_3$. He uses a two-phase Coxian$^+$ distribution[4] for distributions with $C^2 > 1$ and a generalized Erlang distribution for distributions with $C^2 < 1$.

If one is willing to match only a subset of distributions, then again it is possible to achieve solutions which perform very well along the remaining three measures. Whitt [26] and Altiok [2] focus on the set of distributions with $C^2 > 1$ and sufficiently high third moment. They obtain a closed-form solution for matching three moments of any distribution in this set. Whitt matches to a two-branch hyper-exponential distribution and Altiok matches to a two-phase Coxian$^+$ distribution. Telek and Heindl [25] focus on the set of distributions with $C^2 \geq \frac{1}{2}$ and various constraints on the third moment. They obtain a closed-form solution for matching three moments of any distribution in this set, by using a two-phase Coxian$^+$ distribution.

Johnson and Taaffe [10, 9] come closest to achieving all four measures. They provide a closed-form solution for matching the first three moments of any distribution $G \in \mathcal{PH}_3$. They use a mixed Erlang distribution with common order.

[3] The number of necessary phases in general PH distributions is not known. As shown in the next section, all the previous work on computationally efficient algorithms for mapping general distributions concentrates on a subset of acyclic PH distributions.

[4] Coxian$^+$ and Coxian distributions are particular PH distributions shown in Figure 2.

Unfortunately, this mixed Erlang distribution does not produce a minimal solution. Their solution requires $2OPT(G) + 2$ phases in the worst case.

In complementary work, Johnson and Taaffe [12, 11] again look at the problem of matching the first three moments of any distribution $G \in \mathcal{PH}_3$, this time using three types of PH distributions: a mixture of two Erlang distributions, a Coxian$^+$ distribution, and a general PH distribution. Their solution is nearly minimal in that it requires at most $OPT(G) + 2$ phases. Unfortunately, their algorithm requires solving a nonlinear programing problem and hence is very computationally inefficient.

Above we have described the prior work focusing on moment-matching algorithms (three moments), which is the focus of this paper. There is also a large body of work focusing on fitting the *shape* of an input distribution using a PH distribution. Of particular recent interest has been work on fitting heavy-tailed distributions to PH distributions, see for example the work of [3, 6, 7, 14, 20, 24]. There is also work which combines the goals of moment matching with the goal of fitting the shape of the distribution, see for example the work of [8, 22]. The work above is clearly broader in its goals than simply matching three moments. Unfortunately there's a tradeoff: obtaining a more precise fit requires many more phases. Additionally it can sometimes be very computationally inefficient [8, 22].

The Idea Behind the EC Distribution In all the prior work on computationally efficient moment-matching algorithms, the approach was to match a general input distribution G to some subset $\mathcal{S}$ of the PH distributions. In this paper, we show that by using the set of EC distributions as our subset $\mathcal{S}$, we achieve a solution which excels in all four desirable measures mentioned earlier. We define the EC distributions as follows:

Definition 5. *An n-phase EC (Erlang-Coxian) distribution is a particular PH distribution whose underlying Markov chain is of the form in Figure 3.*

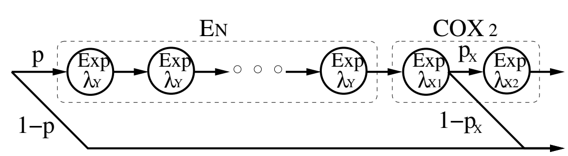


Fig. 3. The Markov chain underlying an EC distribution, where the first box above depicts the underlying continuous time Markov chain in an N-phase Erlang distribution, where $N = n - 2$, and the second box depicts the underlying continuous time Markov chain in a two-phase Coxian$^+$ distribution. Notice that the rates in the first box are the same for all states.

We now provide some intuition behind the creation of the EC distribution. Recall that a Coxian distribution is very good for approximating any distribution with high variability. In particular, a two-phase Coxian distribution is known to well-represent any distribution that has high second and third moments (any distribution G that satisfies $m_2^G > 2$ and $m_3^G > \frac{3}{2} m_2^G$) [18]. However a Coxian distribution requires many more phases for approximating distributions with lower second and third moments. (For example, a Coxian distribution requires at least n phases to well-represent a distribution G with $m_2^G \leq \frac{n+1}{n}$ for integers $n \geq 1$) [18]. The large number of phases needed implies that many free parameters must be determined which implies that any algorithm that tries to well-represent an arbitrary distribution using a minimal number of phases is likely to suffer from computational inefficiency.

By contrast, an n-phase Erlang distribution has only two free parameters and is also known to have the least normalized second moment among all the n-phase PH distributions [1]. However the Erlang distribution is obviously limited in the set of distributions which it can well-represent.

Our approach is therefore to combine the Erlang distribution with the two-phase Coxian distribution, allowing us to represent distributions with all ranges of variability, while using only a small number of phases. Furthermore the fact that the EC distribution has very few free parameters allows us to obtain closed-from expressions for the parameters (n, p, λ_Y, λ_{X1}, λ_{X2}, p_X) of the EC distribution that well-represents any given distribution in $\mathcal{PH}_3$.

Outline of Paper We begin in Section 2 by characterizing the EC distribution in terms of normalized moments. We find that for the purpose of moment matching it suffices to narrow down the set of EC distributions further from six free parameters to five free parameters, by optimally fixing one of the parameters.

We next present three variants for closed-form solutions for the remaining free parameters of the EC distribution, each of which achieves slightly different goals. The first closed-form solution provided, which we refer to as *the simple solution*, (see Section 3) has the advantage of simplicity and readability; however it does not work for all distributions in $\mathcal{PH}_3$ (although it works for almost all). This solution requires at most $OPT(G) + 2$ phases. The second closed-form solution provided, which we refer to as *the improved solution*, (see Section 4.1) is defined for all the input distributions in $\mathcal{PH}_3$ and uses at most $OPT(G) + 1$ phases. This solution is only lacking in numerical stability. The third closed-form solution provided, which we refer to as *the numerically stable solution*, (see Section 4.2) again is defined for all input distributions in $\mathcal{PH}_3$. It uses at most $OPT(G) + 2$ phases and is numerically stable in that the moments of the EC distribution are insensitive to a small perturbation in its parameters.

2 EC Distribution: Motivation and Properties

The purpose of this section is twofold: to provide a detailed characterization of the EC distribution, and to discuss a narrowed-down subset of the EC distributions with only five free parameters (λ_Y is fixed) which we will use in our moment-matching method. Both of these results are summarized in Theorem 1.

To motivate the theorem in this section, consider the following story. Suppose one is trying to match the first three moments of a given distribution G to a distribution P which consists of a generalized Erlang distribution (in a generalized Erlang distribution the rates of the exponential phases may differ) followed by a two-phase Coxian$^+$ distribution. If the distribution G has sufficiently high second and third moments, then a two-phase Coxian$^+$ distribution alone suffices and we need zero phases of the generalized Erlang distribution. If the variability of G is lower, however, we might try appending a single-phase generalized Erlang distribution to the two-phase Coxian$^+$ distribution. If that doesn't suffice, we might append a two-phase generalized Erlang distribution to the two-phase Coxian$^+$ distribution. If our distribution G has very low variability we might be forced to use many phases of the generalized Erlang distribution to get the variability of P to be low enough. Therefore, to minimize the number of phases in P, it seems desirable to choose the rates of the generalized Erlang distribution so that the overall variability of P is minimized.

Continuing with our story, one could express the appending of each additional phase of the generalized Erlang distribution as a "function" whose goal is to reduce the variability of P yet further. We call this "function ϕ."

Definition 6. *Let X be an arbitrary distribution.* **Function** *ϕ maps X to $\phi(X)$ such that $\phi(X) = Y * X$, where Y is an exponential distribution with rate λ_Y independent of X, $Y*X$ is the convolution of Y and X, and λ_Y is chosen so that the normalized second moment of $\phi(X)$ is minimized. Also, $\phi^l(X) = \phi(\phi^{l-1}(X))$ refers to the distribution obtained by applying function ϕ to $\phi^{l-1}(X)$ for integers $l \geq 1$, where $\phi^0(X) = X$.*

Observe that, when X is a k-phase PH distribution, $\phi(X)$ is a $(k+1)$-phase PH distribution whose underlying Markov chain can be obtained by appending a state with rate λ_Y to the Markov chain underlying X, where λ_Y is chosen so that $m_2^{\phi(X)}$ is minimized. In theory, function ϕ allows each successive exponential distribution which is appended to have a different first moment. The following theorem shows that if the exponential distribution Y being appended by function ϕ is chosen so as to minimize the normalized second moment of $\phi(X)$ (as specified by the definition), then the first moment of each successive Y is always *the same* and is defined by the simple formula shown in (1). The theorem below further characterizes the normalized moments of $\phi^l(X)$.

Theorem 1. *Let $\phi^l(X) = Y_l * \phi^{l-1}(X)$ and let $\lambda_{Y_l} = \frac{1}{\mu_1^{Y_l}}$ for $l = 1, ..., N$. Then,*

$$\lambda_{Y_l} = \frac{1}{(m_2^X - 1)\mu_1^X} \tag{1}$$

for $l = 1, ..., N$.

The normalized moments of $Z_N = \phi^N(X)$ are:

$$m_2^{Z_N} = \frac{(m_2^X - 1)(N+1) + 1}{(m_2^X - 1)N + 1}; \tag{2}$$

$$m_3^{Z_N} = \frac{m_2^X m_3^X}{((m_2^X - 1)(N+1) + 1)\left((m_2^X - 1)N + 1\right)^2} + \frac{(m_2^X - 1)N\left(3m_2^X + (m_2^X - 1)(m_2^X + 2)(N+1) + (m_2^X - 1)^2(N+1)^2\right)}{((m_2^X - 1)(N+1) + 1)\left((m_2^X - 1)N + 1\right)^2}. \tag{3}$$

Observe that, when X is a k-phase PH distribution, $\phi^N(X)$ is a $(k+N)$-phase PH distribution whose underlying Markov chain can be obtained by appending N states with rate λ_Y to the Markov chain underlying X, where λ_Y is chosen so that $m_2^{\phi(X)}$ is minimized. The remainder of this section will prove the above theorem and a corollary.

Proof (Theorem 1).

We first characterize $Z = \phi(X) = Y * X$, where X is an arbitrary distribution with a finite third moment and Y is an exponential distribution. The normalized second moment of Z is $m_2^Z = \frac{m_2^X+2y+2y^2}{(1+y)^2}$, where $y = \frac{\mu_1^Y}{\mu_1^X}$. Observe that m_2^Z is minimized when $y = m_2^X - 1$, namely,

$$\mu_1^Y = (m_2^X - 1)\mu_1^X. \tag{4}$$

Observe that when equation (4) is satisfied, the normalized second moment of Z satisfies:

$$m_2^Z = 2 - \frac{1}{m_2^X}, \tag{5}$$

and the normalized third moment of Z satisfies:

$$m_3^Z = \frac{1}{m_2^X(2m_2^X - 1)} m_3^X + \frac{3(m_2^X - 1)}{m_2^X}. \tag{6}$$

We next characterize $Z_l = \phi^l(X) = Y_l * \phi^{l-1}(X)$ for $2 \le l \le N$: By (5) and (6), (2) and (3) follow from solving the following recursive formulas (where we use b_l to denote $m_2^{\phi^l(X)}$ and B_l to denote $m_3^{\phi^l(X)}$):

$$b_{l+1} = 2 - \frac{1}{b_l}; \tag{7}$$

$$B_{l+1} = \frac{B_l}{b_l(2b_l - 1)} + \frac{3(b_l - 1)}{b_l}. \tag{8}$$

The solution for (7) is given by

$$b_l = \frac{(b_1 - 1)l + 1}{(b_1 - 1)(l - 1) + 1} \tag{9}$$

for all $l \ge 1$, and the solution for (8) is given by

$$B_l = \frac{b_1 B_1 + (b_1 - 1)(l - 1)\left(3b_1 + (b_1 - 1)(b_1 + 2)l + (b_1 - 1)^2 l^2\right)}{((b_1 - 1)l + 1)\left((b_1 - 1)(l - 1) + 1\right)^2} \tag{10}$$

for all $l \ge 1$. Equations (9) and (10) can be easily verified by substitution into (7) and (8), respectively. This completes the proof of (2) and (3).

The proof of (1) proceeds by induction. When $l = 1$, (1) follows from (4). Assume that (1) holds when $l = 1, ..., t$. Let $Z_t = \phi^t(X)$. By (2), which is proved above, $m_2^{Z_t} = \frac{(m_2^X-1)(t+1)+1}{(m_2^X-1)t+1}$. Thus, by (4)

$$\mu_1^{Y_{t+1}} = (m_2^{Z_t} - 1)\mu_1^{Z_t} = (m_2^X - 1)\mu_1^X.$$

$\square$

Corollary 1. *Let $Z_N = \phi^N(X)$. If $X \in \{F \mid \ 2 < m_2^F\}$, then* $Z_N \in \left\{F \middle| \frac{N+2}{N+1} < m_2^F < \frac{N+1}{N}\right\}$.

Corollary 1 suggests the number N of times that function ϕ must be applied to X to bring $m_2^{Z_N}$ into the desired range, given the value of m_2^X. Observe that any Coxian$^+$ distribution is in $\{F \mid 2 < m_2^F\}$

Proof (Corollary 1). By (2), $m_2^{Z_N}$ is a continuous and monotonically increasing function of m_2^X. Thus, the infimum and the supremum of $m_2^{Z_N}$ are given by evaluating $m_2^{Z_N}$ at the infimum and the supremum, respectively, of m_2^X. When $m_2^X \to 2$, $m_2^{Z_N} \to \frac{N+2}{N+1}$. When $m_2^X \to \infty$, $m_2^{Z_N} \to \frac{N+1}{N}$. □

3 A Simple Closed-Form Solution

Theorem 1 implies that the parameter λ_Y of the EC distribution can be fixed without excluding the distributions of lowest variability from the set of EC distributions. In the rest of the paper, we constrain λ_Y as follows:

$$\lambda_Y = \frac{1}{(m_2^X - 1)\mu_1^X}, \tag{11}$$

and derive closed-form representations of the remaining free parameters (n, p, λ_{X1}, λ_{X2}, p_X), where these free parameters will determine m_2^X and μ_1^X in (11). Obviously, at least three degrees of freedom are necessary to match three moments. As we will see, the additional degrees of freedom allow us to accept all input distributions in $\mathcal{PH}_3$, use a smaller number of phases, and achieve numerical stability.

We introduce the following sets of distributions to describe the closed-form solutions compactly:

Definition 7. *Let $\mathcal{U}_i$, $\mathcal{M}_i$, and $\mathcal{L}$ be the sets of distributions defined as follows:*

$$\begin{aligned}
\mathcal{U}_0 &= \left\{F \middle| m_2^F > 2 \text{ and } m_3^F > 2m_2^F - 1\right\}, \\
\mathcal{U}_i &= \left\{F \middle| \frac{i+2}{i+1} < m_2^F < \frac{i+1}{i} \text{ and } m_3^F > 2m_2^F - 1\right\}, \\
\mathcal{M}_0 &= \left\{F \middle| m_2^F > 2 \text{ and } m_3^F = 2m_2^F - 1\right\}, \\
\mathcal{M}_i &= \left\{F \middle| \frac{i+2}{i+1} < m_2^F < \frac{i+1}{i} \text{ and } m_3^F = 2m_2^F - 1\right\}, \\
\mathcal{L} &= \left\{F \middle| m_2^F > 1 \text{ and } m_2^F < m_3^F < 2m_2^F - 1\right\},
\end{aligned}$$

for nonnegative integers i. Also, let $\mathcal{U}^+ = \cup_{i=1}^{\infty}\mathcal{U}_i$, $\mathcal{M}^+ = \cup_{i=1}^{\infty}\mathcal{M}_i$, $\mathcal{U} = \mathcal{U}_0 \cup \mathcal{U}^+$, and $\mathcal{M} = \mathcal{M}_0 \cup \mathcal{M}^+$.

These sets are illustrated in Figure 4. The next theorem provides the intuition behind the sets $\mathcal{U}$, $\mathcal{M}$, and $\mathcal{L}$; namely, for all distributions X, the distributions X and $A(X)$ are in the same classification region (Figure 4).

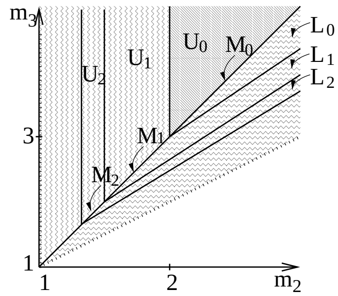


Fig. 4. A classification of distributions. The dotted lines delineate the set of all non-negative distributions G ($m_3^G \geq m_2^G \geq 1$).

Lemma 1. *Let $Z_N = A^N(X)$ for integers $N \geq 1$. If $X \in \mathcal{U}$ (respectively, $X \in \mathcal{M}$, $X \in \mathcal{L}$), then $Z_N \in \mathcal{U}$ (respectively, $Z_N \in \mathcal{M}$, $Z_N \in \mathcal{L}$) for all $N \geq 1$.*

Proof. We prove the case when $N = 1$. The theorem then follows by induction. Let $Z = A(X)$. By (2), $m_2^X = \frac{1}{2-m_2^Z}$, and

$$\begin{aligned} m_3^Z &= (\text{respectively}, <, \text{ and } >) \quad \frac{2m_2^X - 1}{m_2^X(2m_2^X - 1)} + 3\frac{m_2^X - 1}{m_2^X} \\ &= (\text{respectively}, <, \text{ and } >) \quad 2m_2^Z - 1, \end{aligned}$$

where the last equality follows from $m_2^X = \frac{1}{2-m_2^Z}$. $\square$

By Corollary 1 and Lemma 1, it follows that:

Corollary 2. *Let $Z_N = A^N(X)$ for $N \geq 0$. If $X \in \mathcal{U}_0$ (respectively, $X \in \mathcal{M}_0$), then $Z_N \in \mathcal{U}_N$ (respectively, $Z_N \in \mathcal{M}_N$).*

The corollary implies that for all $G \in \mathcal{U}_N \cup \mathcal{M}_N$, G can be well-represented by an $(N+2)$-phase EC distribution with no mass probability at zero ($p = 1$), since, for all $F \in \mathcal{U}_0 \cup \mathcal{M}_0$, F can be well-represented by two-phase Coxian$^+$ distribution, and $Z_N = A^N(X)$ can be well-represented by $(2+N)$-phase EC distribution. It can also be easily shown that for all $G \in \mathcal{L}_N$, G can be well-represented by an $(N+2)$-phase EC distribution with nonzero mass probability at zero ($p < 1$).

From these properties of $A^N(X)$, it is relatively easy to provide a closed-form solution for the parameters (n, p, λ_{X1}, λ_{X2}, p_X) of an EC distribution Z so that a given distribution G is well-represented by Z. Essentially, one just needs to find an appropriate N and solve $Z = A^N(X)$ for X in terms of normalized moments, which is immediate since N is given by Corollary 1 and the normalized moments of X can be obtained from Theorem 1. A little more effort is necessary to minimize the number of phases and to guarantee numerical stability.

In this section, we give a simple solution, which assumes the following condition on A: $A \in \mathcal{PH}_3^-$, where $\mathcal{PH}_3^- = \mathcal{U} \cup \mathcal{M} \cup \mathcal{L}$. Observe $\mathcal{PH}_3^-$ includes almost all distributions in $\mathcal{PH}_3$. Only the borders between the $\mathcal{U}_i$'s are not included. We also analyze the number of necessary phases and prove the following theorem:

Theorem 2. *Under the simple solution, the number of phases needed to well-represent any distribution G by an EC distribution is at most $OPT(G) + 2$.*

The Closed-Form Solution: The solution differs according to the classification of the input distribution G. When $G \in \mathcal{U}_0 \cup \mathcal{M}_0$, a two-phase Coxian$^+$ distribution suffices to match the first three moments. When $G \in \mathcal{U}^+ \cup \mathcal{M}^+$, G is well-represented by an EC distribution with $p = 1$. When $G \in \mathcal{L}$, G is well-represented by an EC distribution with $p < 1$. For all cases, the parameters $(n, p, \lambda_{X1}, \lambda_{X2}, p_X)$ are given by simple closed formulas.

(i) If $G \in \mathcal{U}_0 \cup \mathcal{M}_0$, then a two-phase Coxian$^+$ distribution suffices to match the first three moments, i.e., $p = 1$ and $n = 2$ ($N = 0$). The parameters $(\lambda_{X1}, \lambda_{X2}, p_X)$ of the two-phase Coxian$^+$ distribution are chosen as follows [25, 18]:

$$\lambda_{X1} = \frac{u + \sqrt{u^2 - 4v}}{2\mu_1^G}, \quad \lambda_{X2} = \frac{u - \sqrt{u^2 - 4v}}{2\mu_1^G}, \quad \text{and} \quad p_X = \frac{\lambda_{X2}\mu_1^G(\lambda_{X1}\mu_1^G - 1)}{\lambda_{X1}\mu_1^G},$$

where $u = \frac{6-2m_3^G}{3m_2^G - 2m_3^G}$ and $v = \frac{12-6m_2^G}{m_2^G(3m_2^G - 2m_3^G)}$.

(ii) If $G \in \mathcal{U}^+ \cup \mathcal{M}^+$, Corollary 1 specifies number, n, of phases needed:

$$n = \min\left\{ k \middle| m_2^G > \frac{k}{k-1} \right\} = \left\lfloor \frac{m_2^G}{m_2^G - 1} + 1 \right\rfloor, \tag{12}$$

($N = \left\lfloor \frac{m_2^G}{m_2^G - 1} - 1 \right\rfloor$). Next, we find the two-phase Coxian$^+$ distribution $X \in \mathcal{U}_0 \cup \mathcal{M}_0$ such that G is well-represented by Z, where $Z(\cdot) = Y^{(n-2)*}(\cdot) * X(\cdot)$ and Y is an exponential distribution satisfying (1), $Y^{(n-2)*}$ is the $(n-2)$-th convolution of Y, and $Y^{(n-2)*} * X$ is the convolution of $Y^{(n-2)*}$ and X. To shed light on this expression, consider i.i.d. random variables V_1, ... V_k whose distributions are Y and a random variable V_{k+1}. Then random variable $\sum_{t=1}^{k+1} V_t$ has distribution Z. By Theorem 1, this can be achieved by setting

$$m_2^X = \frac{(n-3)m_2^G - (n-2)}{(n-2)m_2^G - (n-1)}; \quad m_3^X = \frac{\beta m_3^G - \alpha}{m_2^X}; \quad \mu_1^X = \frac{\mu_1^G}{(n-2)m_2^X - (n-3)}, \tag{13}$$

where

$$\alpha = (n-2)(m_2^X - 1)\left(n(n-1)(m_2^X)^2 - n(2n-5)m_2^X + (n-1)(n-3)\right),$$
$$\beta = \left((n-1)m_2^X - (n-2)\right)\left((n-2)m_2^X - (n-3)\right)^2.$$

Thus, we set $p = 1$, and the parameters $(\lambda_{X1}, \lambda_{X2}, p_X)$ of X are given by case (i), using the first moment and the normalized moments of X specified by (13).

(iii) If $G \in \mathcal{L}$, then let

$$p = \frac{1}{2m_2^G - m_3^G}, \quad m_2^W = pm_2^G, \quad m_3^W = pm_3^G, \quad \text{and} \quad \mu_1^W = \frac{\mu_1^G}{p}. \tag{14}$$

G is then well-represented by distribution Z, where $Z(\cdot) = W(\cdot)p + 1 - p$. To shed light on this expression, consider a random variables V_1 whose distribution is W, where W is an EC distribution whose first moment and normalized moments are specified by (14). Then,

$$V_2 = \begin{cases} V_1 & \text{with probability } p \\ 0 & \text{with probability } 1-p. \end{cases}$$

has distribution Z, since $\Pr(V_2 < t) = p\Pr(V_1 < t) + (1-p)$.

Observe that p satisfies $0 \leq p < 1$ and W satisfies $W \in \mathcal{M}$. If $W \in \mathcal{M}_0$, the parameters of W are provided by case (i), using the normalized moments specified by (14). If $W \in \mathcal{M}^+$, the parameters of W are provided by case (ii), using the normalized moments specified by (14).

Figure 5 shows a graphical representation of the simple solution.

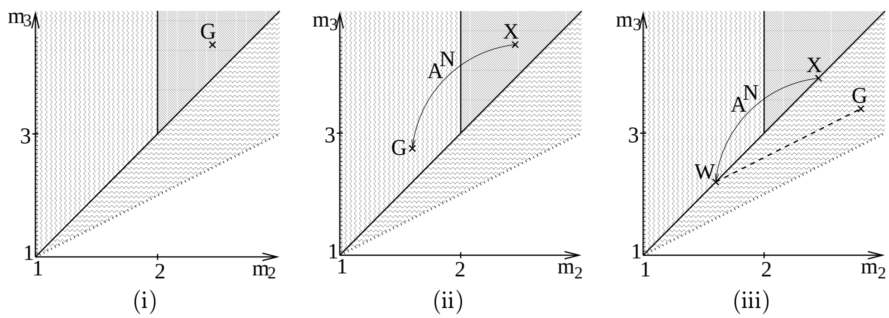


Fig. 5. A graphical representation of the simple solution. Let G be the input distribution. (i) If $G \in \mathcal{U}_0 \cup \mathcal{M}_0$, G is well-represented by a two-phase Coxian$^+$ distribution X. (ii) If $G \in \mathcal{U}^+ \cup \mathcal{M}^+$, G is well-represented by $A^N(X)$, where X is a two-phase Coxian$^+$ distribution. (iii) If $G \in \mathcal{L}$, G is well-represented by Z, where Z is $W = A^N(X)$ with probability p and 0 with probability $1-p$ and X is a two-phase Coxian$^+$ distribution.

Analyzing the Number of Phases Required The proof of Theorem 2 relies on the following theorem:

Theorem 3. *[18] Let $\mathcal{S}^{(n)}$ denote the set of distributions that are well-represented by an n-phase acyclic PH distribution. Let $\mathcal{S_V}^{(n)}$ and $\mathcal{E}^{(n)}$ be the sets defined by:*

$$\mathcal{S_V}^{(n)} = \left\{ F \middle| m_2^F > \frac{n+1}{n} \text{ and } m_3^F \geq \frac{n+3}{n+2} m_2^F \right\};$$
$$\mathcal{E}^{(n)} = \left\{ F \middle| m_2^F = \frac{n+1}{n} \text{ and } m_3^F = \frac{n+2}{n} \right\}$$

for integers $n \geq 2$. Then $\mathcal{S}^{(n)} \subset \mathcal{S_V}^{(n)} \cup \mathcal{E}^{(n)}$ for integers $n \geq 2$.

Proof (Theorem 2). We will show that (i) if a distribution G is in $\mathcal{S_V}^{(l)} \cap (\mathcal{U} \cup \mathcal{M})$, then at most $l+1$ phases are used, and (ii) if a distribution G is in $\mathcal{S_V}^{(l)} \cap \mathcal{L}$, then at most $l+2$ phases are used. Since $\mathcal{S}^{(l)} \subset \mathcal{S_V}^{(l)} \cup \mathcal{E}^{(l)}$ by Theorem 3, this completes the proof. Notice that the simple solution is not defined when $G \in \mathcal{E}^{(l)}$.

(i) Suppose $G \in \mathcal{U} \cup \mathcal{M}$. If $G \in \mathcal{S_V}^{(l)}$, then by (12) the EC distribution provided by the simple solution has at most $l+1$ phases. (ii) Suppose $G \in \mathcal{L}$. If $G \in \mathcal{S_V}^{(l)}$, then $m_2^W = \frac{m_2^G}{2m_2^G - m_3^G} > \frac{l+2}{l+1}$. By (12), the EC distribution provided by the simple solution has at most $l+2$ phases. $\square$

4 Variants of Closed-Form Solutions

In this section, we present two refinements of the simple solution (Section 3), which we refer to as the improved solution and the numerically stable solution.

4.1 An Improved Closed-Form Solution

We first describe the properties that the improved solution satisfies. We then describe the high level ideas behind the construction of the improved solution. Figure 6 is an implementation of the improved solution. See [17] for details on how the high level ideas described above are realized in the improved solution.

Properties of the Improved Solution This solution is defined for all the input distributions $G \in \mathcal{PH}_3$ and uses a smaller number of phases than the simple solution. Specifically, the number of phases required in the improved solution is characterized by the following theorem:

Theorem 4. *Under the improved solution, the number of phases needed to well-represent any distribution G by an EC distribution is at most $OPT(G)+1$.*

For a proof of the theorem, see [17].

High Level Ideas Consider an arbitrary distribution $G \in \mathcal{PH}_3$. Our approach consists of two steps, the first of which involves constructing a baseline EC distribution, and the second of which involves reducing the number of phases in this baseline solution. If $G \in \mathcal{PH}_3^-$, then the baseline solution used is simply given by the simple solution (Section 3). If $G \notin \mathcal{PH}_3^-$, then to obtain the baseline EC distributing we first find a distribution $W \in \mathcal{PH}_3^-$ such that $\frac{m_3^W}{m_2^W} = \frac{m_3^G}{m_2^G}$ and $m_2^W < m_2^G$ and then set p such that G is well-represented by distribution Z, where $Z(\cdot) = W(\cdot)p + 1 - p$. (See Section 3 for an explanation of Z). The parameters of the EC distribution that well-represents W are then obtained by the simple solution (Section 3).

Next, we describe an idea to reduce the number of phases used in the baseline EC distribution. The simple solution (Section 3) is based on the fact that a distribution X is well-represented by a two-phase Coxian distribution when $X \in \mathcal{U}_0 \cup \mathcal{M}_0$. In fact, a wider range of distributions are well-represented by the set of two-phase Coxian distributions. In particular, if

$$X \in \left\{ F \middle| \frac{3}{2} \leq m_2^X \leq 2 \text{ and } m_3^X = 2m_2^X - 1 \right\},$$

then X is well-represented by a two-phase Coxian distribution. In fact, the above solution can be improved upon yet further. However, for readability, we postpone this to [17].[5]

$(n,\, p,\, \lambda_Y,\, \lambda_{X1},\, \lambda_{X2},\, p_X) = \texttt{Improved}(\mu_1^G,\, \mu_2^G,\, \mu_3^G)$

<u>Input:</u> the first three moments of a distribution G: μ_1^G, μ_2^G, and μ_3^G.

<u>Output:</u> parameters of the EC distribution, $(n,\, p,\, \lambda_Y,\, \lambda_{X1},\, \lambda_{X2},\, p_X)$

1. $m_2^G = \frac{\mu_2^G}{(\mu_1^G)^2};\quad m_3^G = \frac{\mu_3^G}{\mu_1^G \mu_2^G}.$

2. $p = \begin{cases} \frac{(m_2^G)^2 + 2m_2^G - 1}{2(m_2^G)^2} & \text{if } m_3^G > 2m_2^G - 1, \text{ and } \frac{1}{m_2^G - 1} \text{ is an integer,} \\ \frac{1}{2m_2^G - m_3^G} & \text{if } m_3^G < 2m_2^G - 1, \\ 1 & \text{otherwise.} \end{cases}$

3. $\mu_1^W = \frac{\mu_1^G}{p};\quad m_2^W = pm_2^G;\quad m_3^W = pm_3^G.$

4. $n = \begin{cases} \left\lfloor \frac{m_2^W}{m_2^W - 1} \right\rfloor & \text{if } m_3^W = 2m_2^W - 1, \text{ and } m_2^W < 2 \\ \left\lfloor \frac{m_2^W}{m_2^W - 1} + 1 \right\rfloor & \text{otherwise.} \end{cases}$

5. $m_2^X = \frac{(n-3)m_2^W - (n-2)}{(n-2)m_2^W - (n-1)};\quad \mu_1^X = \frac{\mu_1^W}{(n-2)m_2^X - (n-3)}.$

6. $\alpha = (n-2)(m_2^X - 1)\left(n(n-1)(m_2^X)^2 - n(2n-5)m_2^X + (n-1)(n-3)\right).$

7. $\beta = \left((n-1)m_2^X - (n-2)\right)\left((n-2)m_2^X - (n-3)\right)^2.$

8. $m_3^X = \frac{\beta m_3^W - \alpha}{m_2^X}.$

9. $u = \begin{cases} 1 & \text{if } 3m_2^X = 2m_3^X \\ \frac{6 - 2m_3^X}{3m_2^X - 2m_3^X} & \text{otherwise} \end{cases};\quad v = \begin{cases} 0 & \text{if } 3m_2^X = 2m_3^X \\ \frac{12 - 6m_2^X}{m_2^X(3m_2^X - 2m_3^X)} & \text{otherwise} \end{cases}.$

10. $\lambda_{X1} = \frac{u + \sqrt{u^2 - 4v}}{2\mu_1^X};\ \lambda_{X2} = \frac{u - \sqrt{u^2 - 4v}}{2\mu_1^X};\ p_X = \frac{\lambda_{X2}\mu_1^X(\lambda_{X1}\mu_1^X - 1)}{\lambda_{X1}\mu_1^X};\ \lambda_Y = \frac{1}{(m_2^X - 1)\mu_1^X}.$

Fig. 6. An implementation of the improved closed-form solution.

4.2 A Numerically Stable Closed-Form Solution

The improved solution (Section 4.1) is not numerically stable when $G \in \mathcal{U}$ and m_2^G is close to $\frac{l+1}{l}$ for integers $l \geq 1$, i.e., on the borders between $\mathcal{U}_i$'s. In this section, we present a numerically stable solution. We first describe the properties that the numerically stable solution satisfies. We then describe the high level ideas behind the construction of the numerically stable solution. Figure 6 is an implementation of the numerically stable solution. See [17] for details on how the high level ideas described above are realized in the numerically stable solution.

Properties of the Numerically Stable Solution The numerically stable solution uses at most one more phase than the improved solution and is defined

[5] While this further improvement reduces the number of necessary phases by one for many distributions, it does not improve the worst case performance.

$(n,\, p,\, \lambda_Y,\, \lambda_{X1},\, \lambda_{X2},\, p_X) = \texttt{Stable}(\mu_1^G,\, \mu_2^G,\, \mu_3^G)$
If $m_3^G \le 2m_2^G - 1$, use `Improved`.
Otherwise, replace steps 2-4 of `Improved` as follows:
2. $n = \left\lceil \frac{3m_2^G - 2 + \sqrt{(m_2^G)^2 - 2m_2^G + 2}}{2(m_2^G - 1)} \right\rceil$.
3. $p = \frac{1}{2m_2^G}\left(\frac{n-1}{n-2} + \frac{n}{n-1}\right)$.
4. $\mu_1^W = \frac{\mu_1^G}{p}$; $\quad m_2^W = pm_2^G$; $\quad m_3^W = pm_3^G$.

Fig. 7. An implementation of the numerically stable closed-form solution.

for all the input distributions in $\mathcal{PH}_3$. Specifically, the number of phases required in the numerically stable solution is characterized by the following theorem:

Theorem 5. *Under the numerically stable solution, the number of phases needed to well-represent any distribution G by an EC distribution is at most $OPT(G)+2$.*

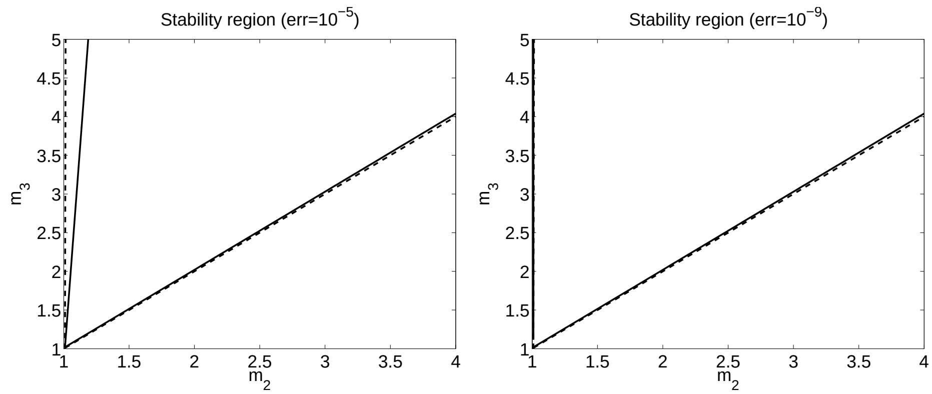
A proof of Theorem 5 is given in [17].

The EC distribution, Z, that is provided by the numerically stable solution is numerically stable in the following sense:

Proposition 1. *Let Z be the EC distribution provided by the numerically stable solution, where the input distribution G is well-represented by Z. Let $(n,\, p,\, \lambda_Y,\, \lambda_{X1},\, \lambda_{X2},\, p_X)$ be the parameters of Z. Suppose that each parameter p, λ_Y, λ_{X1}, λ_{X2}, and p_X has an error Δp, $\Delta\lambda_Y$, $\Delta\lambda_{X1}$, $\Delta\lambda_{X2}$, and Δp_X, respectively, in absolute value. Let $\Delta\mu_1^Z = |\mu_1^Z - \mu_1^G|$ be the error of the first moment of Z and let $\Delta m_i^Z = |m_i^Z - m_i^G|$ be the error of the i-th normalized moment of Z for $i = 2, 3$. If $\frac{\Delta p}{p}$, $\frac{\Delta\lambda_Y}{\lambda_Y}$, $\frac{\Delta\lambda_{X1}}{\lambda_{X1}}$, $\frac{\Delta\lambda_{X2}}{\lambda_{X2}}$, and $\frac{\Delta p_X}{p_X} < \epsilon = 10^{-5}$ (respectively, $\epsilon = 10^{-9}$), then $\frac{\Delta\mu_1^Z}{\mu_1^Z} < 0.01$ and $\frac{\Delta m_i^Z}{m_i^Z} < 0.01$ for $i = 2, 3$, provided that the normalized moments of G satisfies the condition in Figure 8 (a) (respectively, (b)).*

In Proposition 1, ϵ was chosen to be 10^{-5} and 10^{-9}, respectively. These correspond to the precisions of the `float` (six decimal digits) and `double` (ten decimal digits) data type in C, respectively. In Figure 8 (b), it is impossible to distinguish the set of all non-negative distributions from the set of distributions for which the stability guarantee of Proposition 1 holds. Closed form formulas for the curves in Figure 8 and a proof of Proposition 1 are given in [17].

High Level Ideas Achieving the numerical stability is based on the same idea as treating input distributions which are not in $\mathcal{PH}_3^-$. Namely, we first find an EC distribution W such that $\frac{m_3^W}{m_2^W} = \frac{m_3^G}{m_2^G}$ and $m_2^W < m_2^G$ so that the solution is numerically stable for W, and then set p such that G is well-represented by $Z(\cdot) = W(\cdot)p + 1 - p$. (See Section 3 for an explanation of Z).



(a) $\epsilon = 10^{-5}$ (b) $\epsilon = 10^{-9}$

Fig. 8. If the normalized moments of G lie between the two solid lines, then the normalized moments of the EC distribution Z, provided by the numerically stable solution, are insensitive to the small change ($\epsilon = 10^{-5}$ for (a) and $\epsilon = 10^{-9}$ for (b)) in the parameters of Z. The dotted lines delineate the set of all nonnegative distributions G ($m_3^G \geq m_2^G \geq 1$).

5 Conclusion

In this paper, we propose a closed-form solution for the parameters of a PH distribution, P, that well-represents a given distribution G. Our solution is the first that achieves all of the following goals: (i) the first three moments of G and P agree, (ii) any distribution G that is well-represented by a PH distribution (i.e., $G \in \mathcal{PH}_3$) can be well-represented by P, (iii) the number of phases used in P is at most $OPT(G) + c$, where c is a small constant, (iv) the solution is expressed in closed form. Also, the numerical stability of the solution is discussed.

The key idea is the definition and use of EC distributions, a subset of PH distributions. The set of EC distributions is defined so that it includes minimal PH distributions, in the sense that for any distribution, G, that is well-represented by n-phase acyclic PH distribution, there exists an EC distribution, E, with at most $n + 1$ phases such that G is well-represented by E. This property of the set of EC distributions is the key to achieving the above goals (i), (ii), and (iii). Also, the EC distribution is defined so that it has a small number (six) of free parameters. This property of the EC distribution is the key to achieving the above goal (iv). The same ideas are applied to further reduce the degrees of freedom of the EC distribution. That is, we constrain one of the six parameters of the EC distribution without excluding minimal PH distributions from the set of EC distributions.

We provide a complete characterization of the EC distribution with respect to the normalized moments; the characterization is enabled by the simple definition of the EC distribution. The analysis is an elegant induction based on the recursive definition of the EC distribution; the inductive analysis is enabled by a solution to a nontrivial recursive formula. Based on the characterization, we provide three variants of closed-form solutions for the parameters of the EC distribution that well-represents any input distribution, G, that can be well-represented by a PH distribution ($G \in \mathcal{PH}_3$).

One *take-home lesson* from this paper is that the moment-matching problem is better solved with respect to the above four goals by sewing together two or more types of distributions, so that one can gain the best properties of both. The EC distribution sews the two-phase Coxian distribution and the Erlang distribution. The point is that these two distributions provide several different and complementary desirable properties.

Future work includes assessing the minimality of our solution with respect to general (cyclic) PH distributions. If our solution is not close to minimal, then finding a minimal cyclic PH distribution that well-represents any given distribution G is also important. While acyclic PH distributions are well characterized in [18], the minimum number of phases required for a general (cyclic) PH distribution to well-represent a given distribution is not known.

Acknowledgement We would like to thank Miklos Telek for his help in improving the presentation and quality of this paper.

References

1. D. Aldous and L. Shepp. The least variable phase type distribution is Erlang. *Communications in Statistics - Stotchastic Models*, 3:467 – 473, 1987.
2. T. Altiok. On the phase-type approximations of general distributions. *IIE Transactions*, 17:110 – 116, 1985.
3. A. Feldmann and W. Whitt. Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. *Performance Evaluation*, 32:245 – 279, 1998.
4. H. Franke, J. Jann, J. Moreira, P. Pattnaik, and M. Jette. An evaluation of parallel job scheduling for ASCI blue-pacific. In *Proceedings of Supercomputing '99*, pages 679 – 691, November 1999.
5. M. Harchol-Balter, C. Li, T. Osogami, A. Scheller-Wolf, and M. S. Squillante. Analysis of task assignment with cycle stealing under central queue. In *Proceedings of ICDCS '03*, pages 628–637, May 2003.
6. A. Horváth and M. Telek. Approximating heavy tailed behavior with phase type distributions. In *Advances in Matrix-Analytic Methods for Stochastic Models*, pages 191 – 214. Notable Publications, July 2000.
7. A. Horváth and M. Telek. Phfit: A general phase-type fitting tool. In *Proceedings of Performance TOOLS 2002*, pages 82 – 91, April 2002.
8. M. A. Johnson. Selecting parameters of phase distributions: Combining nonlinear programming, heuristics, and Erlang distributions. *ORSA Journal on Computing*, 5:69 – 83, 1993.
9. M. A. Johnson and M. F. Taaffe. An investigation of phase-distribution moment-matching algorithms for use in queueing models. *Queueing Systems*, 8:129 – 147, 1991.

10. M. A. Johnson and M. R. Taaffe. Matching moments to phase distributions: Mixtures of Erlang distributions of common order. *Communications in Statistics — Stochastic Models*, 5:711 – 743, 1989.
11. M. A. Johnson and M. R. Taaffe. Matching moments to phase distributions: Density function shapes. *Communications in Statistics — Stochastic Models*, 6:283 – 306, 1990.
12. M. A. Johnson and M. R. Taaffe. Matching moments to phase distributions: Nonlinear programming approaches. *Communications in Statistics — Stochastic Models*, 6:259 – 281, 1990.
13. S. Karlin and W. Studden. *Tchebycheff Systems: With Applications in Analysis and Statistics*. John Wiley and Sons, 1966.
14. R. E. A. Khayari, R. Sadre, and B. Haverkort. Fitting world-wide web request traces with the EM-algorithm. *Performance Evalutation*, 52:175 – 191, 2003.
15. R. Marie. Calculating equilibrium probabilities for $\lambda(n)/c_k/1/n$ queues. In *Proceedings of* Performance 1980, pages 117 – 125, 1980.
16. M. F. Neuts. *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. The Johns Hopkins University Press, 1981.
17. T. Osogami and M. Harchol-Balter. A closed-form solution for mapping general distributions to minimal PH distributions. Technical Report CMU-CS-03-114, School of Computer Science, Carnegie Mellon University, 2003.
18. T. Osogami and M. Harchol-Balter. Necessary and sufficient conditions for representing general distributions by Coxians. In *Proceedings of TOOLS '03*, September 2003.
19. T. Osogami, M. Harchol-Balter, and A. Scheller-Wolf. Analysis of cycle stealing with switching cost. In *Proceedings of Sigmetrics '03*, pages 184–195, June 2003.
20. A. Riska, V. Diev, and E. Smirni. Efficient fitting of long-tailed data sets into PH distributions. *Performance Evaluation*, 2003 (to appear).
21. C. Sauer and K. Chandy. Approximate analysis of central server models. *IBM Journal of Research and Development*, 19:301 – 313, 1975.
22. L. Schmickler. Meda: Mixed Erlang distributions as phase-type representations of empirical distribution functions. *Communications in Statistics — Stochastic Models*, 8:131 – 156, 1992.
23. M. Squillante. Matrix-analytic methods in stochastic parallel-server scheduling models. In *Advances in Matrix-Analytic Methods for Stochastic Models*. Notable Publications, July 1998.
24. D. Starobinski and M. Sidi. Modeling and analysis of power-tail distributions via classical teletraffic methods. *Queueing Systems*, 36:243 – 267, 2000.
25. M. Telek and A. Heindl. Matching moments for acyclic discrete and continuous phase-type distributions of second order. *International Journal of Simulation*, 3:47 – 57, 2003.
26. W. Whitt. Approximating a point process by a renewal process: Two basic methods. *Operations Research*, 30:125 – 147, 1982.
27. Y. Zhang, H. Franke, J. Moreira, and A. Sivasubramaniam. An integrated approach to parallel scheduling using gang-scheduling, backfilling, and migration. *IEEE Transactions on Parallel and Distributed Systems*, 14:236 – 247, 2003.