A closed-form solution for mapping general distributions to minimal PH distributions

Takayuki Osogami[1]　　Mor Harchol-Balter[2]

February 2003

CMU-CS-03-114

School of Computer Science

Carnegie Mellon University

Pittsburgh, PA 15213

[1]Carnegie Mellon University, Computer Science Department. Email: osogami@cs.cmu.edu

[2]Carnegie Mellon University, Computer Science Department. Email: harchol@cs.cmu.edu. This work was supported by NSF Career Grant CCR-0133077, by NSF ITR Grant 99-167 ANI-0081396 and by Spinnaker Networks via Pittsburgh Digital Greenhouse Grant 01-1.

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.

Keywords: closed form, algorithm, moment matching, Coxian distribution, phase-type distribution, EC distribution, normalized moment, matrix analytic

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. Matching three moments is desirable because it has been found to lead to sufficient accuracy in modeling many computer systems [4, 16]. Matching only two moments often does not suffice since the performance of some queueing models has been shown to be heavily dependent on the third moment of distributions in the model [22, 6].

Moment-matching algorithms are evaluated along four different measures:

The number of moments matched – In general matching more moments is more desirable. Matching three moments often suffices and is a popular approach.

The 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 a $P \in PH$ 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 for designing moment-matching algorithms is to start by defining a subset S of the PH distributions, and then match each input distribution G to a distribution in S. The reason for

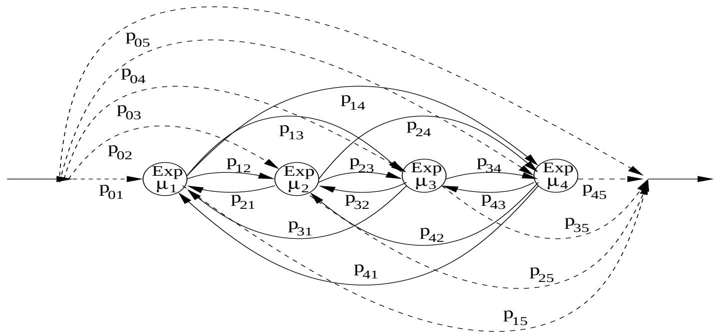


Figure 1: *A four-phase PH distribution. There are $n = 4$ states, where the ith state has exponentially-distributed service time with rate μ_i. With probability p_{0i} we start in the ith state, and the next state is state j with probability p_{ij}. Each state i has probability p_{i5} that it will be the last state. The value of the distribution is the sum of the times spent in each of the states.*

limiting the solution to a distribution in S is that this narrows the search space and thus improves the efficiency of the algorithm. Observe that PH distributions have $\Theta(n^2)$ free parameters, see Figure 1 [14], while S can be defined to have far fewer free parameters. For all efficient algorithms in the literature, S was chosen to be some subset of the acyclic PH distributions. One has to be careful in how one defines the subset S, however: If S is too small it may limit the space of distributions which can be well-represented.[1] Also, if 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 $G \in \mathcal{PH}$ 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.

Preliminary definitions

Formally, we will use the following definitions:

[1]For example, a two-phase Coxian distribution has three parameters (μ_1, μ_2, p). However the set of two-phase Coxian distributions is still not large enough to well-represent distributions with squared coefficient of variability $C^2 = 1$, unless the third moment is 6; for distributions G with $C^2 = 1$, the system of equations always gives the solution ($\mu_1 = 1/E[G]$, $\mu_2 = 0$, $p_2 = 0$), which is an exponential distribution. As another example, it can also be shown that the generalized Erlang distribution is not general enough to well-represent all the distributions with low variability (Lemma 9.1).

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, which were introduced in [15], help provide a simple representation and analysis of our closed-form solution. Formally, the normalized moments are defined as follows:

Definition 2 *Let $E[X^k]$ be the k-th moment of a distribution X for $k = 1, 2, 3$. The* **normalized k-th moment** m_k^X *of X for $k = 2, 3$ is defined to be*

$$m_2^X = \frac{E[X^2]}{(E[X])^2} \quad \text{and} \quad m_3^X = \frac{E[X^3]}{E[X]E[X^2]}.$$

Notice the correspondence to the coefficient of variability C and skewness γ: $m_2^X = C^2 + 1$ and $m_3^X = \gamma\sqrt{m_2}$.

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

It is known that a distribution G is in $\mathcal{PH}$ iff its normalized moments satisfy $m_3^G > m_2^G > 1$ [8]. Since any nonnegative distribution G satisfies $m_3^G \geq m_2^G \geq 1$ [11], almost all the nonnegative distributions are in $\mathcal{PH}$.

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, where an acyclic PH distribution is a PH distribution in which there is no transition from state i to state j for all $i > j$.*[2]

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; 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 [17] provide a closed-form solution for matching two moments of a general distribution in $\mathcal{PH}$. They use a two-branch hyper-exponential

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

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 [13] provides a closed-form solution for matching two moments of a general distribution in $\mathcal{PH}$. He uses a two-phase Coxian distribution 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 [21] and Altiok [2] focus on only 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 [20] focus on only 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 [8, 7] come closest to achieving all four measures. They provide a closed-form solution for matching the first three moments of any distribution $G \in \mathcal{PH}$. They use a mixed Erlang distribution with common order. 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 [10, 9] again look at the problem of matching the first three moments of any distribution $G \in \mathcal{PH}$, this time using three types of PH distributions: a mixture of two Erlang distributions, a Coxian distribution without mass probability at zero, 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 computationally very expensive.

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 hyperexponential distributions, see for example the work of [3, 19, 12]. There is also work which combines the goals of moment matching with the goal of fitting the shape of the distribution, see for example Schmickler [18] and Johnson [5]. 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 expensive [18, 5].

The idea behind the EC distribution

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

Definition 5 *A PH distribution is said to be an n-phase EC distribution if with probability $1-p$, the value is zero, and with probability p, the value is an Erlang-($n-2$) distribution followed by a two-phase Coxian distribution for integers $n \geq 2$ (see Figure 2).*

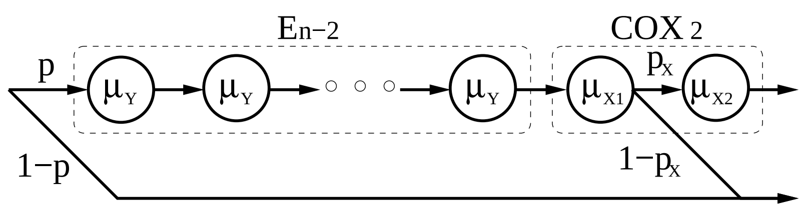


Figure 2: *An EC distribution.*

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 $m_3^G > \frac{3}{2}m_2^G$) [15]. However a Coxian distribution requires many more phases for approximating distributions with lower second and third moments (e.g. 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 $n \geq 1$) [15]. This in turn means that many free parameters are necessary in the solution of matching to a Coxian distribution, which results in the moment-matching algorithm being inefficient.

By contrast, an Erlang distribution has only two free parameter and is also known to have the least normalized second moment among all the PH distributions with a fixed number of phases [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-k 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 a closed-from expression of the parameters (n, p, μ_Y, μ_{X1}, μ_{X2}, p_X) of the EC distribution that well-represents any given distribution in $\mathcal{PH}$.

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 $\mathcal{PH}$ (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) is defined for all the input distributions in $\mathcal{PH}$ 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 5) again is defined for all input distributions in $\mathcal{PH}$. 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 Motivation for the EC distribution and important properties

The purpose of this section is two fold: 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) which we will use in our moment-matching method.

2.1 Motivating story

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 an "operation" whose goal is to reduce the variability of P yet further. We call this "Operation A."

Definition 6 *Let X be an arbitrary distribution.* **Operation A** *converts X to $A(X)$ such that*

$$A(X) = Y + X,$$

where Y is an exponential distribution independent of X, and the mean of Y is chosen so that the normalized second moment of $A(X)$ is minimized. Also, $A^m(X) = A(A^{m-1}(X))$ refers to the distribution obtained by applying operation A to $A^{m-1}(X)$ for integers $m \geq 1$, where $A^0(X) = X$.

Observe that operation A in theory allows each successive exponential distribution which is appended to have a different mean. The following theorem shows that if the exponential distribution Y being appended by operation A is chosen so as to minimize the normalized second moment of $A(X)$ (as specified by the definition), then the mean 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 $A^m(X)$.

Theorem 1 *Let $A^m(X) = Y_m + A^{m-1}(X)$ for $m = 1, ..., n$. Then,*

$$E[Y_m] = (m_2^X - 1)E[X] \tag{1}$$

for $m = 1, ..., n$. The normalized moments of $Z = A^n(X)$ are:

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

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

The remainder of this section will prove the above theorem and a corollary.

2.2 Proof of Theorem 1 and a corollary

Proof:[Theorem 1]

We first characterize $Z = A(X) = Y + X$, 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 $y = \frac{E[Y]}{E[X]}$. Since $\frac{\partial}{\partial y} m_2^Z = \frac{2\left(y - (m_2^X - 1)\right)}{(1+y)^3}$, m_2^Z is minimized when $y = m_2^X - 1$, namely,

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

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 $Z = A^n(X) = Y_n + A^{n-1}(X)$: By (5) and (6), (2) and (3) follow from solving the following recursive formulas (where we use b_n to denote $m_2^{A^n(X)}$ and B_n to denote $m_3^{A^n(X)}$):

$$b_{n+1} = 2 - \frac{1}{b_n}; \tag{7}$$

$$B_{n+1} = \frac{B_n}{b_n(2b_n - 1)} + \frac{3(b_n - 1)}{b_n}. \tag{8}$$

The solution for (7) is given by

$$b_n = \frac{(b_1 - 1)n + 1}{(b_1 - 1)(n - 1) + 1} \tag{9}$$

for all $n \geq 1$, and the solution for (8) is given by

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

for all $n \geq 1$. (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 $n = 1$, (1) follows from (4). Assume that (1) holds when $n = 1, ..., k$. Let $Z = A^k(X)$. By (2), which is proved above, $m_2^Z = \frac{(m_2^X-1)(k+1)+1}{(m_2^X-1)k+1}$. Thus, by (4)

$$E[Y_{k+1}] = (m_2^Z - 1)E[Z] = \left(\frac{(m_2^X - 1)(k+1)+1}{(m_2^X - 1)k - 1} - 1\right)(E[X] + k(m_2^X - 1)E[X]) = (m_2^X - 1)E[X].$$

■

Corollary 1 *Let $Z = A^n(X)$. If $X \in \left\{F \mid \ 2 < m_2^F\right\}$, then $Z \in \left\{F \mid \frac{n+2}{n+1} < m_2^F < \frac{n+1}{n}\right\}$.*

Corollary 1 suggests the number of times that operation A must be applied to X to bring m_2^Z into the desired range, given the value of m_2^X. Notice that by choosing X to be a two-phase Coxian distribution, m_2^X can take on any value greater than 2.

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

3 A simple closed-form solution

Theorem 1 implies that the parameter μ_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 μ_Y as follows:

$$\mu_Y = \frac{1}{(m_2^X - 1)E[X]}, \tag{11}$$

and derive closed-form representation of the remaining free parameters ($n, p, \mu_{X1}, \mu_{X2}, p_X$), where these free parameters will determine m_2^X and $E[X]$ 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 $\mathcal{PH}$, use a smaller number of phases, and achieve numerical stability.

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

Definition 7 *Let U_1, U_2, M_1, M_2, and L be the sets of distributions defined as follows:*

$$\begin{aligned} U_1 &= \left\{F \mid m_2^F > 2 \text{ and } m_3^F > 2m_2^F - 1\right\}, \quad U_2 = \left\{F \mid 1 < m_2^F \leq 2 \text{ and } m_3^F > 2m_2^F - 1\right\}, \\ M_1 &= \left\{F \mid m_2^F \geq 2 \text{ and } m_3^F = 2m_2^F - 1\right\}, \quad M_2 = \left\{F \mid 1 < m_2^F < 2 \text{ and } m_3^F = 2m_2^F - 1\right\}, \\ L &= \left\{F \mid m_2^F > 1 \text{ and } m_2^F < m_3^F < 2m_2^F - 1\right\}. \end{aligned}$$

Also, let $U = U_1 \cup U_2$ and $M = M_1 \cup M_2$.

These sets are illustrated in Figure 3 The next theorem provides the intuition behind the sets U, M, and

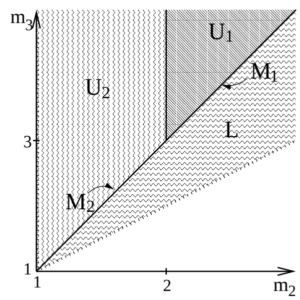


Figure 3: *A classification of distributions. The dotted lines delineate the set of all nonnegative distributions G ($m_3^G \geq m_2^G \geq 1$).*

L; namely, for all distributions X, the distributions X and $A(X)$ are in the same classification region (Figure 3).

Theorem 2 *Let $Z = A^n(X)$ for integers $n \geq 1$. If $X \in U$ (respectively, $X \in M$, $X \in L$), then $Z \in U$ (respectively, $Z \in M$, $Z \in L$) 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),

$$m_2^X = \frac{1}{2 - m_2^Z} \tag{12}$$

and

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

where the last equality follows from (12). ■

From the detailed characterization of $A^n(X)$ given in Section 2, it is relatively easy to provide a closed-form solution for the parameters $(n, p, \mu_{X1}, \mu_{X2}, p_X)$ 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 $G = A^n(X)$ 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 G: $G \in \mathcal{PH}^-$ where

$$\mathcal{PH}^- = \left(U \cap \left\{F \mid m_2^F \neq \frac{n+1}{n} \text{ for integers } n \geq 1\right\}\right) \bigcup \left((M \cup L) \cap \left\{F \mid m_2^F \neq \frac{n+1}{n+2} m_3^F \text{ for integers } n \geq 1\right\}\right).$$

Observe $\mathcal{PH}^-$ includes almost all distributions in $\mathcal{PH}$. We also analyze the number of necessary phases:

Theorem 3 *The number of phases of the EC distribution provided by the simple solution is at most* $OPT(G) + 2$.

The closed-form solution:

The solution differs according to the classification of the input distribution G. When $G \in U_1 \cup M_1$, a two-phase Coxian distribution suffices to match the first three moments. When $G \in U_2 \cup M_2$, G is well-represented by an EC distribution with $p = 1$. When $G \in L$, G is well-represented by an EC distribution with $p < 1$. For all cases, the parameters $(n, p, \mu_{X1}, \mu_{X2}, p_X)$ are given by simple closed formulas.

(i) If $G \in U_1 \cup M_1$, then a two-phase Coxian distribution suffices to match the first three moments, i.e., $p = 1$ and $n = 2$. The parameters $(\mu_{X1}, \mu_{X2}, p_X)$ of the two-phase Coxian distribution are chosen as follows [20, 15]:

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

where

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

(ii) If $G \in U_2 \cup M_2$, Corollary 1 specifies the necessary number of phases, n:

$$n = \min\left\{k \mid m_2^G > \frac{k}{k-1}\right\} = \left\lfloor \frac{m_2^G}{m_2^G - 1} + 1 \right\rfloor. \tag{13}$$

Next, we find the two-phase Coxian distribution $X \in U_1 \cup M_1$ such that G is well-represented by

$$Z = \sum_{k=1}^{n-2} Y_k + X,$$

where Y_k is an exponential distribution satisfying (1) for $k = 1, ..., n-2$. By Theorem 1, this can be achieved by setting

$$m_2^X = \frac{(n-3)m_2^G - (n-2)}{(n-2)m_2^G - (n-1)}, \quad m_3^X = \frac{\gamma m_3^G - \beta}{m_2^X}, \quad \text{and} \quad E[X] = \frac{E[G]}{(n-2)m_2^X - (n-3)}, \tag{14}$$

for $k = 1, ..., n-2$, where

$$\begin{aligned} \beta &= (n-2)(m_2^X - 1)\left(n(n-1)(m_2^X)^2 - n(2n-5)m_2^X + (n-1)(n-3)\right), \\ \gamma &= \left((n-1)m_2^X - (n-2)\right)\left((n-2)m_2^X - (n-3)\right)^2. \end{aligned}$$

Thus, we set $p = 1$, and the parameters $(\mu_{X1}, \mu_{X2}, p_X)$ of X are provided by case (i), using the mean and the normalized moments of X specified by (14).

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

$$p = \frac{1}{2m_2^G - m_3^G}, \quad m_2^W = pm_2^G, \quad m_3^W = pm_3^G, \quad \text{and} \quad E[W] = \frac{E[G]}{p}. \tag{15}$$

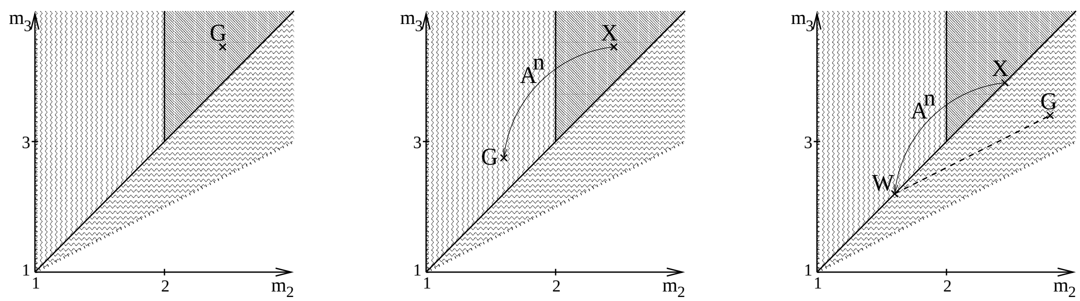
G is then well-represented by Z:

$$Z = \begin{cases} W & \text{with probability } p \\ 0 & \text{with probability } 1-p, \end{cases}$$

where W is an EC distribution with a mean and normalized moments specified by (15). Observe that p satisfies $0 \le p < 1$ and W satisfies $W \in M_1 \cup M_2$. If $W \in M_1$, the parameters of W are provided by case (i), using the normalized moments specified by (15). If $W \in M_2$, the parameters of W are provided by case (ii), using the normalized moments specified by (15).

A graphical representation

Figure 4 shows a graphical representation of the simple solution.



(i) (ii) (iii)

Figure 4: *A graphical representation of the simple solution: Let G be the input distribution. (i) If $G \in U_1 \cup M_1$, G is well-represented by a two-phase Coxian distribution X. (ii) If $G \in U_2 \cup M_2$, G is well-represented by $A^n(X)$, where X is a two-phase Coxian distribution. (iii) If $G \in L$, 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

We now prove Theorem 3, whose proof relies on the following theorem:

Theorem 4 *[15] Let $S^{(n)}$ denote the set of distributions that are well-represented by an n-phase acyclic PH distribution. Let $S_V^{(n)}$ and $E^{(n)}$ be the sets defined as follows:*

$$\begin{aligned} S_V^{(n)} &= \left\{ F \mid m_2^F > \frac{n+1}{n} \text{ and } m_3^F \geq \frac{n+3}{n+2} m_2^F \right\}; \\ E^{(n)} &= \left\{ F \mid m_2^F = \frac{n+1}{n} \text{ and } m_3^F = \frac{n+2}{n} \right\} \end{aligned}$$

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

Proof:[Theorem 3] We will show that (i) if a distribution G is in $S_V^{(n)} \cap (U \cup M)$, then at most n+1 phases are used, and (ii) if a distribution G is in $S_V^{(n)} \cap L$, then at most n+2 phases are used. Since $S^{(n)} \subset S_V^{(n)} \cup E^{(n)}$ by Theorem 4, this completes the proof. Notice that the simple solution is not defined when the input $G \in E^{(n)}$.

(i) Suppose $G \in U \cup M$. If $G \in S_V^{(n)}$, then by (13) the EC distribution provided by the simple solution has at most $n+1$ phases. (ii) Suppose $G \in L$. If $G \in S_V^{(n)}$, then $m_2^W = \frac{m_2^G}{2m_2^G - m_3^G} > \frac{n+2}{n+1}$. By (13), the

EC distribution provided by the simple solution has at most $n+2$ phases. ■

4 An improved closed-form solution

In this section, we present a refinement of the simple solution (Section 3), which we refer to as the improved solution. This solution is defined for all the input distributions $G \in \mathcal{PH}$ and uses a smaller number of phases than the simple solution.

Theorem 5 *The number of phases of the EC distribution provided by the improved solution is at most* $OPT(G)+1$.

Ideas behind the improvement For a distribution $G \notin \mathcal{PH}^-$, we first find a distribution $W \in \mathcal{PH}^-$ 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

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

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 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 $X \in U_1 \cup M_1$. In fact, a wider range of distributions are well-represented by the set of two-phase Coxian distributions. In particular, if

$$X \in \left\{ F \mid \frac{3}{2} \le m_2^X \le 2 \text{ and } m_3^X = 2m_2^X - 1 \right\}, \tag{16}$$

then X is well-represented by a two-phase Coxian distribution. By fully exploiting the power of two-phase Coxian distribution, one can construct a solution that uses a smaller number of phases. For readability, we use Condition (16). Further improvement is possible and discussed briefly in Appendix A.[3]

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

We first characterize the distribution $Z = A^n(X)$ as a function of the normalized moments of X and n when the normalized moments of X satisfy (16).

Theorem 6 *Let $Z = A^n(X)$. If $X \in \left\{F \mid \frac{3}{2} \leq m_2^F \leq 2 \text{ and } m_3^F = 2m_2^F - 1\right\}$, then*

$$Z \in \left\{F \mid \frac{n+1}{n} \leq m_2^F \leq \frac{n}{n-1} \text{ and } m_3^F = 2m_2^F - 1\right\}.$$

Proof: By Theorem 1, m_2^Z is a continuous and monotonically increasing function of m_2^X, Thus, $\frac{n+1}{n} \leq m_2^Z \leq \frac{n}{n-1}$ follows by simply evaluating m_2^Z at the lower and upper bound of m_2^X. $m_3^Z = 2m_2^Z - 1$ follows from Theorem 2. ■

The closed-form solution: (i) If $G \in U \cap \mathcal{PH}^-$, then the simple solution (Section 3) provides the parameters $(n, p, \mu_{X1}, \mu_{X2}, p_X)$.

(ii) If $G \in U \cap (\mathcal{PH}^-)^c$, then let

$$n = \frac{2m_2^G - 1}{m_2^G - 1}; \quad m_2^W = \frac{1}{2}\left(\frac{n-1}{n-2} + \frac{n}{n-1}\right); \quad m_3^W = \frac{m_3^G}{m_2^G} m_2^W; \quad p_W = \frac{m_2^W}{m_2^G}; \quad E[W] = \frac{E[G]}{p_W}. \tag{17}$$

G is then well-represented by Z:

$$Z = \begin{cases} W & \text{with probability } p_W \\ 0 & \text{with probability } 1 - p_W, \end{cases}$$

where W is an EC distribution with a mean and normalized moments specified by (17). The parameters $(n, \mu_{X1}, \mu_{X2}, p_X)$ of W are provided by the simple solution (Section 3), by using the normalized moments of W specified by (17). Also, set $p = p_W$, since W has not pass probability at zero.

(iii) If $G \in M \cup L$, then the simple solution (Section 3) provides the parameters $(n, p, \mu_{X1}, \mu_{X2}, p_X)$, except that if the number n of phases calculated by (13) is $n > 2$, then n is decremented by one. Theorem 6 guarantees that parameters obtained with the reduced n are still feasible.

Analyzing the number of phases required Now, we prove Theorem 5.

Proof:[Theorem 5] Recall the definition of $S_V^{(n)}$ and $E^{(n)}$ in Theorem 4. We will show that if a distribution $G \in S_V^{(n)} \cup E^{(n)}$, then at most $n+1$ phases are used. Since $S^{(n)} \subset S_V^{(n)} \cup E^{(n)}$ by Theorem 4, the proof is completed.

(i) Suppose $G \in U$. If $G \in S_V^{(n)}$, then the simple solution (Section 3) is used and at most $n+1$ phases are used. (ii) Suppose $G \in M$. If $G \in S_V^{(n)}$, then the number of phases used in the improved solution is one less than the simple solution. Therefore, at most n phases are used. If $G \in E^{(n)}$, then exactly n phases are used. (iii) Suppose $G \in L$. If $G \in S_V^{(n)}$, then the number of phases used in the improved solution is one less than the simple solution. Therefore, at most $n+1$ phases are used. ■

5 A numerically stable closed-form solution

The improved solution (Section 4) is not numerically stable when $G \in U$ and m_2^G is close to $\frac{n+1}{n}$ for integers $n \geq 1$. In this section, we present a numerically stable solution. The numerically stable solution uses at most one more phase than the improved solution and is defined for all the input distributions in $\mathcal{PH}$:

Theorem 7 *The number of phases of the EC distribution provided by the numerically stable solution is at most $OPT(G) + 2$.*

We also evaluate the numerical stability of the solution.

The closed-form solution: Achieving the numerical stability is based on the same idea as treating input distributions which are not in $\mathcal{PH}^-$. 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:

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

(i) <u>If $G \in M \cup L$</u>, then the improved solution (Section 4) provides the parameters (n, p, μ_{X1}, μ_{X2}, p_X).

(ii) If $G \in U$, first obtain the number of necessary phases n:

$$
\begin{aligned}
n &= k \text{ s.t. } \frac{1}{2}\left(\frac{k-1}{k-2}+\frac{k}{k-1}\right) \le m_2^G < \frac{1}{2}\left(\frac{k-1}{k-2}+\frac{k-2}{k-3}\right) \\
&= \left\lceil \frac{3m_2^G - 2 + \sqrt{(m_2^G)^2 - 2m_2^G + 2}}{2(m_2^G - 1)} \right\rceil . \qquad (18)
\end{aligned}
$$

Then, let

$$
m_2^W = \frac{1}{2}\left(\frac{n-1}{n-2}+\frac{n}{n-1}\right), \quad m_3^W = \frac{m_3^G}{m_2^G} m_2^W, \quad p_W = \frac{m_2^W}{m_2^G}, \quad \text{and} \quad E[W] = \frac{E[G]}{p}. \qquad (19)
$$

G is then well-represented by Z:

$$
Z = \begin{cases} W & \text{with probability } p \\ 0 & \text{with probability } 1-p, \end{cases}
$$

where W is an EC distribution with a mean and normalized moments specified by (19). The parameters $(n, \mu_{X1}, \mu_{X2}, p_X)$ of W are provided by the simple solution (Section 3), by using the normalized moments of W specified by (19), except that (18) is used as the number n of phases. Also, set $p = p_W$, since W has no mass probability at zero.

Analysis of the number of phases required and the numerical stability

From the construction of the solution, it is immediate that the numerically stable solution uses at most one more phase than the improved solution (Section 4). Thus, Theorem 7 follows from Theorem 5.

In the rest of this section, we evaluate the numerical stability of the EC distribution Z that is provided by the numerically stable solution. Formally, we show that Z is numerically stable in the following sense:

Theorem 8 *Let Z be the EC distribution provided by the numerically stable solution, where the input distribution G is well-represented by Z. Let $(n, p, \mu_Y, \mu_{X1}, \mu_{X2}, p_X)$ be the parameters of Z. Suppose that each parameter p, μ_Y, μ_{X1}, μ_{X2}, and p_X has an error Δp, $\Delta\mu_Y$, $\Delta\mu_{X1}$, $\Delta\mu_{X2}$, and Δp_X, respectively, in absolute value. Let $\Delta E[Z]$ be the error of the mean of Z and let Δm_i^Z be the error of the i-th normalized moment of Z for $i = 2, 3$; namely, $\Delta E[Z] = |E[Z] - E[G]|$ and $\Delta m_i^Z = |m_i^Z - m_i^G|$. If $\frac{\Delta p}{p}$, $\frac{\Delta \mu_Y}{\mu_Y}$, $\frac{\Delta \mu_{X1}}{\mu_{X1}}$, $\frac{\Delta \mu_{X2}}{\mu_{X2}}$, and $\frac{\Delta p_X}{p_X} < \epsilon = 10^{-5}$, then $\frac{\Delta E[Z]}{E[Z]} < 0.01$ 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 5.*

In Theorem 8, ϵ was chosen to be 10^{-5}. This corresponds to the precision of the `float` data type in

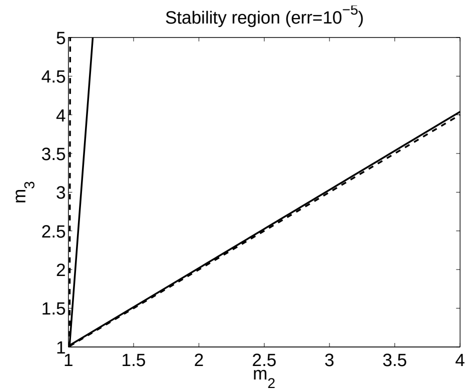


Figure 5: *If the normalized moments of G lie between the two solid lines, then the normalized moments of the EC distribution Z, provided by the numerically stable solution, are insensitive to the small change ($\epsilon = 10^{-5}$) in the parameters of Z. The dotted lines delineate the set of all nonnegative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

C, which is six decimal digits. The precision of the `double` data type in C is 10 decimal digits, which corresponds to an ϵ of 10^{-9}. Figure 6 in Appendix B is the corresponding figure to Figure 5, for the case of $\epsilon = 10^{-9}$. In Figure 6 (drawn to the same scale as Figure 5) it is impossible to distinguish the set of all non-negative distributions from the set of distributions for which the stability guarantee of Theorem 8 holds. The proof of Theorem 8 is given in Appendix B.

6 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., $G \in \mathcal{PH}$) 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 in $\mathcal{PH}$.

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 two-phase Coxian distribution and the Erlang distribution provide several different desirable properties, which are complementary to each other, and the combination of these two distributions provides all the desirable properties needed to achieve the four goals.

References

[1] D. Aldous and L. Shepp. The least variable phase type distribution is Erlang. *Communications in Statistics - Stochastic 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] 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 23rd International Conference on Distributed Computing Systems (ICDCS '03)*, Providence, RI, May 2003 (to appear).

[5] M. A. Johnson. Selecting parameters of phase distributions: Combining nonlinear programming, heuristics, and Erlang distributions. *ORSA Journal on Computing*, 5:69 – 83, 1993.

[6] M. A. Johnson and M. F. Taaffe. A graphical investigation of error bounds for moment-based queueing approximations. *Queueing Systems*, 8:295–312, 1991.

[7] 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.

[8] 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.

[9] M. A. Johnson and M. R. Taaffe. Matching moments to phase distributions: Density function shapes. *Communications in Statistics — Stochastic Models*, 6:283–306, 1990.

[10] M. A. Johnson and M. R. Taaffe. Matching moments to phase distributions: Nonlinear programming approaches. *Communications in Statistics — Stochastic Models*, 6:259–281, 1990.

[11] S. Karlin and W. Studden. *Tchebycheff Systems: With Applications in Analysis and Statistics*. John Wiley and Sons, 1966.

[12] R. E. A. Khayari, R. Sadre, and B. Haverkort. Fitting world-wide web request traces with the EM-algorithm. In *Proceedings of SPIE: Internet Performance and Control of Network Systems II*, volume 4523, pages 211–220, 2001.

[13] R. Marie. Calculating equilibrium probabilities for $\lambda(n)/c_k/1/n$ queues. In *Proceedings of* Performance 1980, pages 117 – 125, 1980.

[14] M. F. Neuts. *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. The Johns Hopkins University Press, 1981.

[15] T. Osogami and M. Harchol-Balter. Necessary and sufficient conditions for representing general distributions by Coxians. Technical Report CMU-CS-02-178, School of Computer Science, Carnegie Mellon University, September 2002.

[16] T. Osogami, M. Harchol-Balter, and A. Scheller-Wolf. Analysis of cycle stealing with switching cost. In *Proceedings of ACM Sigmetrics 2003 Conference on Measurement and Modeling of Computer Systems*, San Diego, CA, June 2003 (to appear).

[17] C. Sauer and K. Chandy. Approximate analysis of central server models. *IBM Journal of Research and Development*, 19:301 – 313, 1975.

[18] L. Schmickler. Meda: Mixed Erlang distributions as phase-type representations of empirical distribution functions. *Communications in Statistics — Stochastic Models*, 8:131 – 156, 1992.

[19] D. Starobinski and M. Sidi. Modeling and analysis of power-tail distributions via classical teletraffic methods. *Queueing Systems*, 36:243–267, 2000.

[20] M. Telek and A. Heindl. Moment bounds for acyclic discrete and continuous phase type distributions of second order. In *Proceedings of UK Performance Evaluation Workshop, UKPEW 2002*, 2002.

[21] W. Whitt. Approximating a point process by a renewal process: Two basic methods. *Operations Research*, 30:125 – 147, 1982.

[22] W. Whitt. On approximations for queues, iii: Mixtures of exponential distributions. *AT&T Bell Laboratories Technical Journal*, 63:163 – 175, 1984.

A Further improvement

Further improvement on the improved solution (Section 4) is possible by fully exploiting the power of the two-phase Coxian distribution: if

$$X \in \left\{ F \mid \frac{3}{2} \le m_2^F \le 2 \text{ and } \frac{4}{3} m_2^F \le m_3^F \le \frac{6(m_2^F - 1)}{m_2^F} \right\}, \tag{20}$$

then X is well-represented by a two-phase Coxian distribution. The following theorem provides information on the number of necessary phases:

Theorem 9 *Let* $Z = A^n(X)$. *If* $X \in \left\{ F \mid \frac{3}{2} \le m_2^F \le 2 \text{ and } \frac{4}{3} m_2^F \le m_3^F \le \frac{6(m_2^F - 1)}{m_2^F} \right\}$, *then* $Z \in \left\{ F \mid \frac{n+1}{n} \le m_2^F \le \frac{n}{n-1} \text{ and } l \le m_3^F \le u \right\}$, *where*

$$\begin{aligned} l &= \frac{-(n^3 - 5n^2 + 6n)(m_2^Z)^3 + (3n^3 - 12n^2 + 8n + 12)(m_2^Z)^2 - (3n^3 - 9n^2 + n + 10)m_2^Z - n^3 + 2n^2 + n - 2}{3m_2^Z}, \\ u &= \frac{n(m_2^Z - 1)\left((n-1)(n-2)(m_2^Z)^2 - (2n+1)(n-2)m_2^Z + (n-1)(n+1)\right)}{m_2^Z}. \end{aligned}$$

Proof: By Theorem 1, m_2^Z is a continuous and monotonically increasing function of m_2^X, Thus, $\frac{n+1}{n} \le m_2^Z \le \frac{n}{n-1}$ follows by simply evaluating m_2^Z at the lower and upper bound of m_2^X.

Since m_2^Z is a continuous and monotonically increasing function of m_2^X, m_2^Z and m_2^X have a one-to-one correspondence. Also, for fixed m_2^X, m_3^Z is an increasing function of m_3^X. Thus, $l \le m_3^Z \le u$ follows by simply evaluating m_3^Z, given by Theorem 1, at the lower and upper bound of m_3^X and substituting $m_2^X = \frac{(n-3)m_2^Z - (n-2)}{(n-2)m_2^Z - (n-1)}$. ∎

B Details of the sensitivity analysis

Proof:[Theorem 8] Let $\Delta p \le \epsilon p$, $\Delta \mu_Y \le \epsilon \mu_Y$, $\Delta \mu_{X1} \le \epsilon \mu_{X1}$, $\Delta \mu_{X2} \le \epsilon \mu_{X2}$, and $\Delta p_X \le \epsilon p_X$. We first obtain the sensitivity of the normalized moments of the two-phase Coxian distribution X to its parameters:

$$\begin{aligned} \Delta E[X] &\le \left|\frac{\partial E[X]}{\partial \mu_{X1}}\right| \Delta \mu_{X1} + \left|\frac{\partial E[X]}{\partial \mu_{X2}}\right| \Delta \mu_{X2} + \left|\frac{\partial E[X]}{\partial p_X}\right| \Delta p_X = \frac{1}{\mu_{X1}^2}\Delta \mu_{X1} + \frac{p_X}{\mu_{X2}^2}\Delta \mu_{X2} + \frac{1}{\mu_{X2}}\Delta p_X, \\ \Delta m_2^X &\le \left|\frac{\partial m_2^X}{\partial u}\right| \Delta u + \left|\frac{\partial m_2^X}{\partial v}\right| \Delta v = \frac{2}{v}\Delta u + \frac{2(u-1)}{v^2}\Delta v, \\ \Delta m_3^X &\le \left|\frac{\partial m_3^X}{\partial u}\right| \Delta u + \left|\frac{\partial m_3^X}{\partial v}\right| \Delta v = 3\left(\frac{1}{(u-1)^2} + \frac{1}{v}\right)\Delta u + \frac{3u}{v}\Delta v, \end{aligned}$$

where

$$\Delta u \le \Delta \mu_{X1} + \Delta \mu_{X2} \quad \text{and} \quad \Delta v \le \mu_{X2}\Delta \mu_{X1} + \mu_{X1}\Delta \mu_{X2}.$$

Next, we obtain the sensitivity of $W = X + E_{n-2} = X + \sum_{k=1}^{n-2} Y_k$:

$$\Delta E[W] \le \Delta E[X] + \Delta E[E_{n-2}]$$

$$\Delta m_2^W \leq \left|\frac{\partial m_2^W}{\partial m_2^X}\right| \Delta m_2^X + \left|\frac{\partial m_2^W}{\partial z}\right| \Delta z = \frac{1}{(1+z)^2}\Delta m_2^X + \left|\frac{2(z(m_2^{E_{n-2}}-1)-(m_2^X-1))}{(1+z)^3}\right| \Delta z$$

$$\Delta m_3^W \leq \left|\frac{\partial m_3^W}{\partial m_2^X}\right| \Delta m_2^X + \left|\frac{\partial m_3^W}{\partial m_3^X}\right| \Delta m_3^X + \left|\frac{\partial m_3^W}{\partial z}\right| \Delta z$$

where

$$E[E_{n-2}] = (n-2)E[Y], \quad m_2^{E_{n-2}} = \frac{n-1}{n-2}, \quad m_3^{E_{n-2}} = \frac{n}{n-2}, \quad z = \frac{E[E_{n-2}]}{E[X]},$$

$$\Delta E[E_{n-2}] \leq (n-2)\Delta E[Y] = \frac{(n-2)}{(\mu_Y)^2}\Delta\mu_Y, \quad \Delta z \leq \frac{1}{E[X]}\Delta E[E_{n-2}] + \frac{E[E_{n-2}]}{(E[X])^2}\Delta E[X],$$

$$\frac{\partial m_3^W}{\partial m_2^X} = \frac{z(m_3^X(2+m_2^{E_{n-2}}z)+z(6-m_2^{E_{n-2}}(3+(-3+m_3^{E_{n-2}})z)))}{(1+z)(m_2^X+2z+m_2^{E_{n-2}}z^2)^2}, \quad \frac{\partial m_3^W}{\partial m_3^X} = \frac{m_2^X}{(1+z)(m_2^X+2z+m_2^{E_{n-2}}z^2)},$$

and

$$\frac{\partial m_3^W}{\partial z} = \frac{-(m_2^X)^2(m_3^X-3)+m_2^{E_{n-2}}z^2(6-3m_2^{E_{n-2}}z^2+m_3^{E_{n-2}}z(4+(2+m_2^{E_{n-2}})z))}{(1+z)(m_2^X+2z+m_2^{E_{n-2}}z^2)^2}$$
$$+\frac{-m_2^X(m_3^X(2+2(2+m_2^{E_{n-2}})z+3m_2^{E_{n-2}}z^2)+z(6z-m_2^{E_{n-2}}(6-6z^2+m_3^{E_{n-2}}z(3+2z))))}{(1+z)(m_2^X+2z+m_2^{E_{n-2}}z^2)^2}.$$

Finally, we derive the sensitivity of Z:

$$\Delta E[Z] = p\Delta E[W] + E[W]\Delta p, \quad \Delta m_2^Z = \frac{1}{p}\Delta m_2^W + \frac{m_2^W}{p^2}\Delta p, \quad \text{and} \quad \Delta m_3^Z = \frac{1}{p}\Delta m_2^W + \frac{m_3^W}{p^2}\Delta p.$$

■

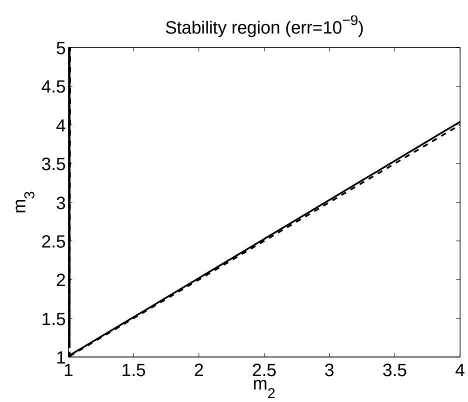


Figure 6: *If the normalized moments of G lie between the two solid lines, then the normalized moments of the EC distribution Z, provided by the numerically stable solution, are insensitive to the small change ($\epsilon = 10^{-9}$) in the parameters of Z. The dotted lines delineate the set of all nonnegative distributions G ($m_3^G \geq m_2^G \geq 1$).*

C The generalized Erlang distribution is not general enough

Lemma 9.1 *Let $Z = Y + X$, where Y is an exponential distribution. If $X \in \left\{F \mid m_2^F \leq 2 \text{ and } m_3^F \leq 3\right\}$, then $Z \in \left\{F \mid m_3^F < 3\right\}$.*

Proof: The normalized third moment of Z is

$$m_3^Z = \frac{m_2^X m_3^X + 3m_2^X y + 6y^2 + 6y^3}{(m_2^X + 2y + 2y^2)(1+y)},$$

where $y = \frac{E[Y]}{E[X]}$. m_3^Z is an increasing function of m_2^X and m_3^X, since

$$\frac{\partial m_3^Z}{\partial m_2^X} = \frac{2m_3^X y}{(m_2^X + 2y + 2y^2)^2} > 0 \quad \text{and} \quad \frac{\partial m_3^Z}{\partial m_3^X} = \frac{2m_2^X}{(m_2^X + 2y + 2y^2)(1+y)} > 0$$

Therefore, m_3^Z is maximized when $m_2^X = 2$ and m_3^X. Thus, $m_3^Z = \frac{3(1+y+y^2+y^3)}{(1+y+y^2)(1+y)}$. It is easy to see $\frac{1+y+y^2+y^3}{(1+y+y^2)(1+y)} < 1$ for all $0 < y < \infty$. Hence, $m_3^Z < 3$. ∎