Matching more than three moments with acyclic phase type distributions

András Horváth[1], Miklós Telek[2]

[1] Dip. di Informatica, Università di Torino

[2] Dept. of Telecommunications, Technical University of Budapest

E-mail: horvath@di.unito.it, telek@hit.bme.hu

May 17, 2006

Abstract

This paper presents a matching procedure for generating an acyclic phase type distribution of order N given the first $2N - 1$ moments, if they are feasible. The matching procedure uses an iterative approach and, theoretically, it can be applied to match an arbitrary number of moments. The first step of the iterative procedure contains the solution of an equation of order N and the order is decreased by one in each consecutive step. Apart of these equations the procedure makes use of explicit expressions. The practical applicability of the proposed procedure is limited by the numerical accuracy of the solution of these equations and the complexity of the involved expressions. We present examples for matching more than 10 moments with acyclic phase type distributions.

Keywords: acyclic phase type distribution, moments matching.

1 Introduction

The applicability of phase type distributions in stochastic models of real systems is determined by the availability of appropriate phase type fitting procedures[1]. There are two main categories of phase type fitting algorithms: numerical optimization [1, 5] and procedures that make use of explicit expressions. Up

---
[1]A survey of fitting algorithms can be found in [6].

to now explicit expressions for matching only the first three moments of acyclic phase type distributions (APH) have been known [9, 7, 8, 2]. This paper presents an iterative approach to match an arbitrary number of moments with acyclic phase type distributions.

The solution of matching an arbitrary number of moments with matrix exponential distributions, which is an even richer class of distributions, containing the APH class, has been known for more than a decade, see [10]. However, the applicability of this matching procedure is limited because the class of matrix exponential distributions and the properties of the matching method have the following disadvantages.

- The matrix exponential class contains the phase type and the acyclic phase type classes, but it also contains distributions which cannot be represented as the time to absorption in a Markov chain. These non-Markovian distributions might inhibit the use of the widely applied Markovian solution techniques.

- The matching procedure calculates a matrix for any set of moments without indicating if the matrix (together with the predefined initial and final vectors) corresponds to a matrix exponential distribution or not. When the procedure is called with the moments of a matrix exponential distribution of order N it results in a true matrix exponential distribution of order N, but when it is called with an invalid set of moments it results in a matrix which does not correspond to a real distribution. Unfortunately, it is hard to check if a given matrix corresponds to a real matrix exponential distribution [4].

The analytical description of the moment bounds of matrix exponential distributions would avoid the second disadvantage, but they are not known beyond $N = 2$. The procedure presented in this paper avoids the above mentioned disadvantages. APH matching results in a Markovian representation by definition and the feasibility of the solution is obvious (a solution with real, positive intensities and real probabilities between 0 and 1 is feasible).

With the procedure presented in this paper, matching the moments of an acyclic phase type distribution of order N (APH(N)) requires the solution of an equation of order N. As a consequence, the matching procedure has a symbolic solution up to $N = 4$ and numerical techniques are applicable for larger N. Here we present the symbolic solution for $N = 3$ (matching 5 moments). Because of the complexity of the symbolic solution for $N = 4$ we use a numerical technique for $N \geq 4$. The practical applicability of the proposed procedure is limited by the complexity of the involved expressions and by the numerical

accuracy of the solution of the equation of order N. We present numerical results for matching more than 10 moments.

2 Notations and theoretical background

2.1 Normalized moments

Based on the ordinary moments m_i we define the ith normalized moment as follows:

$$n_i = \frac{m_i}{m_{i-1}m_1} \ , \qquad i \geq 1. \tag{1}$$

Note that throughout this paper we consider only non-defective distributions whose 0th moment is assumed to be 1. As a consequence the first normalized moment of the considered distributions is 1 (i.e., $m_0 = 1$ and $n_1 = 1$).

We introduce the normalized moments for two reasons.

- They simplify the problem of matching. With the use of normalized moments one can match the first moment and the normalized moments, n_i, $2 \leq i \leq N-1$, separately.

- They simplify the obtained expressions. E.g., the ith normalized moment of an Erlang(N) distribution is $\dfrac{N+i-1}{N}$.

Based on (1), the ordinary moments can be calculated from the normalized moments using

$$m_i = n_i m_{i-1} m_1 = m_1^i \prod_{k=1}^{i} n_k \ . \tag{2}$$

2.2 Acyclic phase type distributions

Definition 1. *An acyclic phase type distribution of order N, denoted by APH(N), is the distribution of the time to absorption in an acyclic Markov chain with N transient states and an absorbing one.*

Theorem 2. *[3] Any APH(N) distribution can be represented in the canonical form (CF) depicted in Figure 1 such that $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N$.*

Definition 3. *The representation of an APH(N) in the structure of Figure 1 without any restrictions on the order of the intensities is called its series representation.*
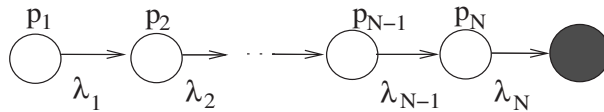
Figure 1: Canonical representation of APH(N) distributions

Both, the series and the canonical representations of an APH(N) distribution are defined by two vectors $\underline{p} = \{p_1, \ldots, p_N\}$ and $\underline{\lambda} = \{\lambda_1, \ldots, \lambda_N\}$.

Corollary 4. *The $\underline{\lambda}$ vectors of the different series representations of an APH(N) contains the same intensities. They differ only in the order of these intensities.*

Proof. The corollary is a consequence of Cumani's result [3], but it can be derived also from the fact that the eigenvalues of any triangular matrix are determined by the diagonal elements. □

2.3 The $2N$th and the $2N+1$th normalized moments of APH(N) distributions

As shown by Cumani in [3] an APH(N) distribution is characterized by $2N - 1$ parameters. According to [10], the first $2N - 1$ moments of an APH(N) characterize the distribution[2]. As a consequence, one can calculate the higher moments of an APH(N) in terms of the first $2N - 1$ moments based on [10] (see Appendix A). It is straightforward to transform this result into the relation of the normalized moments

$$n_{2N} = \mathcal{F}_{2N}(n_2, \ldots, n_{2N-1}), \quad n_{2N+1} = \mathcal{F}_{2N+1}(n_2, \ldots, n_{2N}). \tag{3}$$

For $N = 2$ the second and third normalized moments determine all higher normalized moments. In this case the fourth and fifth normalized moments are

$$n_4 = \frac{2n_2(9n_2 + 2(n_3 - 6)n_3)}{3(n_2 - 2)n_3} \ , \tag{4}$$

$$n_5 = \frac{5n_3(4n_2(2n_3 - 3n_4) + 3n_4^2)}{6(2n_3 - 3n_2)n_4} \ . \tag{5}$$

---

[2]Indeed [10] proves this statement for the whole matrix exponential class using reduced moments.

For $N = 3$, n_2, n_3, n_4 and n_5 determine all higher normalized moments. In this case

$$n_6 = \Bigg(n_3(25n_2(2n_2(16n_3^2 - 36n_3n_4 + 9n_4^2) - 3n_3(n_4 - 8)n_4^2)) +$$

$$n_3(120n_2(6n_2 + n_3(n_4 - 4) - 3n_4)n_4n_5 - 36(n_2 - 2)n_4^2n_5^2) \Bigg) \Big/ \qquad (6)$$

$$\Bigg(10n_4(18n_2^2 + 6n_3n_4 + n_2n_3(4n_3 - 3(8 + n_4)))n_5 \Bigg) ,$$

$$n_7 = \Bigg(7n_4(-8n_3n_5^2(3n_4(3n_5 - 5n_6) + 5n_6^2) + 5n_2(12n_5^2n_6^2 + 12n_3n_5(6n_4n_5 - 5n_4n_6 - 4n_5n_6) +$$

$$n_3^2(75n_4^2 - 180n_4n_5 + 48n_5^2 + 80n_5n_6))) \Bigg) \Big/ \Bigg(20n_5(5n_3(4n_2(2n_3 - 3n_4) + 3n_4^2) + \qquad (7)$$

$$6(3n_2 - 2n_3)n_4n_5)n_6 \Bigg) .$$

Because of their complexity, we do not provide the expressions for higher order moments. These expressions can be calculated up to $N = 8$ in about two minutes with a standard personal computer (1.5GHz processor and 524MB RAM). The computation for $N = 9$ has such high memory requirement that it cannot be performed with a computer with 524MB RAM.

3 The matching procedure

It is straightforward to compute the normalized moments, n_i, $2 \leq i \leq 2N - 1$, of an APH(N) distribution in terms of its $2N - 1$ parameters. The inverse, i.e. to determine the parameters given the normalized moments, is non-trivial. In this section we present an iterative approach which, at each step, determines two parameters of the APH(N) distribution.

The section starts with the building blocks and then describes the procedure itself. Section 3.1 investigates the effects of extending an APH($N - 1$) distribution with an additional phase. Based on properties of the extension, the problem of matching $2N - 2$ normalized moments with an APH(N) distribution can be reduced to the problem of matching $2N - 4$ normalized moments with an APH($N - 1$) distribution. This reduction step is described in Section 3.2. The complete algorithm is presented in Section 3.4.

3.1 The APH($N-1$)-EXP structure

The extension of an APH($N-1$) distribution with an additional phase, as depicted in Figure 2, is characterized by two parameters, p and λ. The following theorem and its corollary describe the effect of the extension on the moments and normalized moments, respectively.
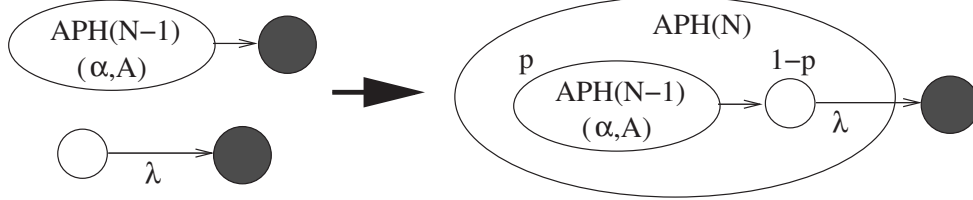

Figure 2: Realization of an APH(N) distribution by the composition of an APH($N-1$) distribution and an additional phase

Theorem 5. *The moments of the resulting APH($N-1$)-EXP distribution, m'_i, can be expressed in terms of the moments of the APH($N-1$) distribution, m_i, and the two parameters of the extension (p and λ), using*

$$m'_i = i!\lambda^{-i}\left(1 + p\sum_{j=1}^{i}\frac{\lambda^j\, m_j}{j!}\right).\tag{8}$$

Proof. Let $f_L(s)$ be the Laplace transform of the APH($N-1$) distribution. The Laplace transform of the APH($N-1$)-EXP distribution, $g_L(s)$, is

$$g_L(s) = p f_L(s)\frac{\lambda}{s+\lambda} + (1-p)\frac{\lambda}{s+\lambda} = \frac{\lambda}{s+\lambda}\left(p f_L(s) + 1 - p\right).\tag{9}$$

To obtain its moments we first take the ith derivative of (9)

$$\frac{d^i}{ds^i}g_L(s) = \sum_{j=0}^{i}\binom{i}{j}\frac{d^{i-j}}{ds^{i-j}}\left(\frac{\lambda}{s+\lambda}\right)\frac{d^j}{ds^j}\left(p f_L(s)+1-p\right) =$$

$$\sum_{j=1}^{i}\binom{i}{j}\lambda(i-j)!(-1)^{i-j}(s+\lambda)^{-i+j-1}\, p\frac{d^j}{ds^j}f_L(s) + \lambda i!(-1)^i(s+\lambda)^{-i-1}\left(p f_L(s)+1-p\right)$$

then by the limit $s \to 0$

$$(-1)^i m'_i \;=\; \sum_{j=1}^{i}\binom{i}{j}(i-j)!(-1)^{i-j}\lambda^{-i+j}\, p(-1)^j m_j + i!(-1)^i\lambda^{-i}\,.$$

6

Multiplying both sides by $(-1)^i$ gives (8). $\qquad\square$

To simplify the following expressions and to eliminate the dependence on the first moment, we replace the parameters, p and λ, by

$$a = m_1\lambda \quad \text{and} \quad b = ap. \tag{10}$$

Indeed $a = \frac{m_1}{1/\lambda}$ is the ratio of the means of the APH($N-1$) and the EXP parts of the APH($N-1$)-EXP distribution.

Corollary 6. *The normalized moments of the APH(N − 1)-EXP distribution, n_i', can be expressed in terms of the normalized moments of the APH(N − 1) distribution, n_i, using*

$$n_i' = \frac{i\left(1 + b\sum_{j=1}^{i} \frac{a^{j-1}}{j!} \prod_{k=1}^{j} n_k\right)}{(1+b)\left(1 + b\sum_{j=1}^{i-1} \frac{a^{j-1}}{j!} \prod_{k=1}^{j} n_k\right)} \tag{11}$$

Proof. From the definition of the normalized moments, using (8) and (2) we have

$$n_i' = \frac{m_i'}{m_{i-1}'m_1'} = \frac{i!\lambda^{-i}\left(1 + p\sum_{j=1}^{i} \frac{\lambda^j}{j!} \prod_{k=1}^{j} n_k m_1^j\right)}{\left((i-1)!\lambda^{-i+1}\left(1 + p\sum_{j=1}^{i-1} \frac{\lambda^j}{j!} \prod_{k=1}^{j} n_k m_1^j\right)\right)\left(pm_1 + \lambda^{-1}\right)}.$$

Introducing the a and b parameters results in (11). $\qquad\square$

3.2 Reduction to one phase less

Given $2N-2$ normalized moments, n_i', $2 \le i \le 2N-1$, we aim to find n_i, $2 \le i \le 2N-3$, and a and b such that the extension of an APH($N-1$) distribution whose normalized moments are n_i with an exponential phase according to Figure 2 with parameters a and b, results in an APH(N) distribution with normalized moments n_i'. To this end we derive the following two corollaries of Theorem 5.

Corollary 7. *The inverse function of (8), i.e., the moments of the APH(N − 1) distribution as a function*

of the moments of the APH(N − 1)-EXP distribution, have the following from

$$m_i = \frac{\lambda m_i' - i m_{i-1}'}{\lambda p}.$$ (12)

Proof. We prove the corollary by induction. (12) holds for $i = 1$. Assuming that (12) holds for $j \in \{1, 2, \ldots, i-1\}$ from (8) we have

$$\frac{\lambda^i}{i!} m_i' = 1 + p \sum_{j=1}^{i-1} \frac{\lambda^j}{j!} m_j + p \frac{\lambda^i}{i!} m_i = 1 + p \sum_{j=1}^{i-1} \frac{\lambda^j}{j!} \frac{\lambda m_j' - j m_{j-1}'}{\lambda p} + p \frac{\lambda^i}{i!} m_i .$$ (13)

That is

$$\frac{\lambda^i}{i!} m_i' = 1 + \underbrace{\sum_{j=1}^{i-1} \left(\frac{\lambda^j m_j'}{j!} - \frac{\lambda^{j-1} m_{j-1}'}{(j-1)!} \right)}_{\frac{\lambda^{i-1} m_{i-1}'}{(i-1)!}} + p \frac{\lambda^i}{i!} m_i ,$$ (14)

which gives (12) after some manipulation. □

Corollary 8. *The normalized moments of the APH(N − 1) distribution as a function of the normalized moments of the APH(N − 1)-EXP distribution, have the following from*

$$n_i = \frac{n_{i-1}'(1 + b)\left(n_i'(1 + b) - i \right)}{a\left(n_{i-1}'(1 + b) - (i - 1) \right)} , \qquad i > 1 .$$ (15)

Proof. From the definition of the normalized moments and (12) we have

$$
\begin{aligned}
n_i &= \frac{m_i}{m_{i-1} \, m_1} = \frac{\dfrac{\lambda m_i' - i m_{i-1}'}{\lambda p}}{\dfrac{\lambda m_{i-1}' - (i - 1) m_{i-2}'}{\lambda p} \dfrac{a}{\lambda}} \\
&= \frac{(\lambda m_i' - i m_{i-1}')\lambda}{(\lambda m_{i-1}' - (i - 1) m_{i-2}')a} = \frac{\left(\lambda \dfrac{m_i'}{m_{i-1}'} - i \right) \dfrac{m_{i-1}'}{m_{i-2}'} \lambda}{\left(\lambda \dfrac{m_{i-1}'}{m_{i-2}'} - (i - 1) \right) a} \\
&= \frac{\left(\lambda n_i' \dfrac{1 + b}{\lambda} - i \right) n_{i-1}' \dfrac{1 + b}{\lambda} \lambda}{\left(\lambda n_{i-1}' \dfrac{1 + b}{\lambda} - (i - 1) \right) a} ,
\end{aligned}
$$ (16)

8

where we applied $m'_1 = (1 + b)/\lambda$. $\hfill \square$

Based on (15), the n_i, $2 \leq i \leq 2N - 3$, normalized moments of the APH$(N - 1)$ distribution are explicitly known as a function of n'_i, $2 \leq i \leq 2N - 3$, and a and b. The two other unknowns, a and b, can be determined based on the remaining two normalized moments n'_{2N-2} and n'_{2N-1}. An equation to determine a is constructed by the following steps:

Step 1 take (11) for $i = 2N - 2$,

Step 2 substitute n_{2N-2} by $\mathcal{F}_{2N-2}(n_2, ..., n_{2N-3})$ as provided by [10],

Step 3 substitute n_i, $2 \leq i \leq 2N - 3$ applying (15).

The resulting equation has two unknowns, a and b, and enjoys the following property.

Theorem 9. *The equation to determine a is linear in a and hence it can be rearranged to have an explicit expression for a as a function of n'_i, $2 \leq i \leq 2N - 2$, and b.*

In a similar manner, an equation to determine b is constructed by the following steps:

Step 1 take (11) for $i = 2N - 1$,

Step 2 substitute n_{2N-1} by $\mathcal{F}_{2N-1}(n_2, ..., n_{2N-2})$ as provided by [10],

Step 3 substitute n_i, $2 \leq i \leq 2N - 2$ applying (15).

The resulting equation has a single unknown variable, b, and enjoys the following property.

Theorem 10. *The equation to determine b is of order N in b and hence it can be rearranged to have a polynomial equation of order N to determine b.*

A consequence of Theorem 10 is that one can look for symbolic solution up to $N = 4$ (matching 7 moments). Fitting more than 7 moments requires the use of numerical techniques.

The proofs of Theorem 9 and 10 are in Appendix A. We demonstrate the procedure for obtaining n_i, $2 \leq i \leq 2N - 3$, and a and b given n'_i, $2 \leq i \leq 2N - 1$, for $N = 3$ in Section 4.

3.3 Feasibility

The parameters of the APH$(N - 1)$-EXP structure, a and b, are positive real numbers with the following bounds, $0 < a < \infty$ and $0 < b = ap \leq a$ (since $0 < p \leq 1$). $p = 0$ is not feasible because it results in an exponential distribution, and $a = 0$ is not feasible because it implies $m_1 = 0$ or $\lambda = 0$.

Theorem 11. *A set of normalized moments $\{n'_2, \ldots, n'_{2N-1}\}$ is feasible with an APH(N) distribution if and only if there exists a solution of the reduction step, n_i, $2 \leq i \leq 2N-3$, and a and b such that*

- *a and b are real numbers such that $0 < b \ (= ap) \leq a$,*

- *the normalized moments $\{n_2, \ldots, n_{2N-3}\}$ are feasible with an APH(N − 1) distribution.*

Proof. The theorem is a consequence of Cumani's seminal result (Theorem 2). Since any APH(N) can be represented in canonical form, there is an APH(N) with the given moments if and only if there is an APH($N-1$)-EXP structure with the given constraints. □

The first condition of the theorem can be checked easily. The second condition has to be checked by repeated applications of the reduction step. In each step the number of moments is decreased by two. The whole procedure is described in the next section.

Theorem 11 has an important consequence. It allows us to check if a given set of moments is inside the moments bounds of the APH(N) class or not. Indeed, the procedure described hereinafter directly constructs an APH(N) distribution, if possible. To the best of our knowledge it is the first method to check this property for more than three moments.

Due to the simple stochastic interpretation of the APH distribution, it is easy to check if the result of the procedure is a proper APH distribution. This property is valuable with respect to the matrix exponential case where the matrix representation, calculated from the set of moments [10], does not indicate the feasibility of the solution.

3.4   Iterative approach of moments matching

Based on the reduction to one phase less presented in Section 3.2, the following iterative algorithm generates an APH(N) distributions given the first $2N-1$ moments.

1. We calculate the normalized moments, n_i, $2 \leq i \leq 2N-1$, based on the first $2N-1$ ordinary moments, m_i, $1 \leq i \leq 2N-1$, according to (1).

2. The reduction step can be applied according to the tree scheme depicted in Figure 3. At level zero we have n_i, $2 \leq i \leq 2N-1$, i.e., the normalized moments. At the first level the reduction step gives at most N different solutions (Theorems 9 and 10). The reduction step can be applied again to any of these solutions leading to at most $N(N-1)$ nodes at the second level. In order to avoid

10

cumbersome notation, indices of the different solutions for b, a and n_i are indicated on the top of the nodes. At the last level there are at most $N!$ nodes.

3. Taking any path from the root to a leaf gives a sequence of values for b and a. This sequence can be used to construct an APH(N) distribution with normalized moments n_i, $2 \leq i \leq 2N - 1$. We start with an APH(1) distribution with rate 1 and in each step the distribution is extended by one phase. The values for a and b are taken from bottom to top and are applied according to (10). Note that, as it will be illustrated later in this section, not all the paths lead to a proper APH(N) distribution.

4. Finally, we set the first ordinary moment (the mean) by properly scaling the intensities of the result.
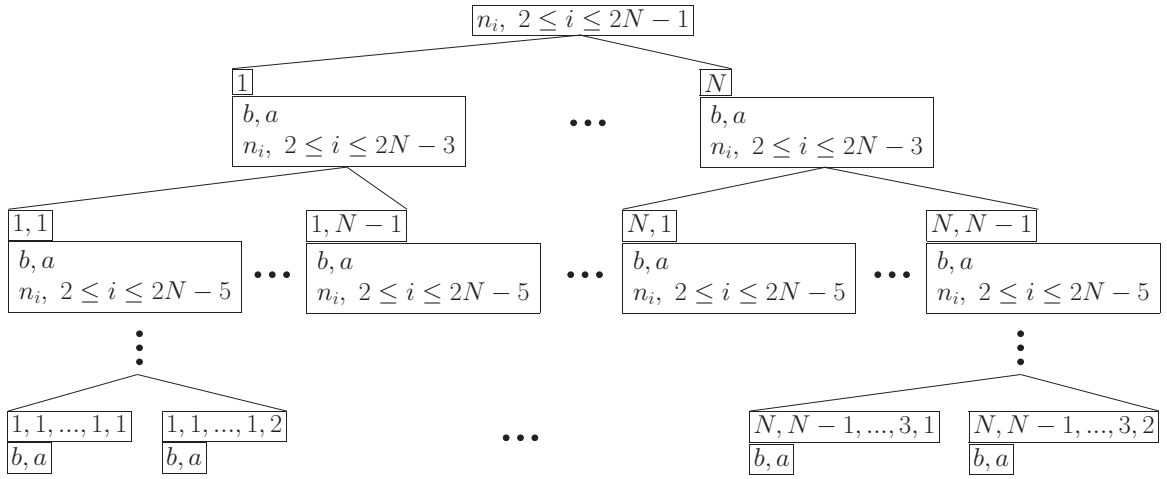

Figure 3: Tree scheme of the application of the reduction step

All the routes provide a series form (with coefficients which are not necessarily feasible). Due to fact that they solve equations (15) and (3), they all provide the given $2N - 1$ ordinary moments.

Corollary 12. *The $\underline{\lambda}$ vectors of the $N!$ solutions contain the same intensity values with all possible permutations.*

Proof. The eigenvalues of the PH generator is determined by the first $2N - 1$ moments of the distribution [10]. In case of APH distributions the eigenvalues are real and they determine the departure rate of the phases. $N!$ solutions are obtained by the permutations of these eigenvalues. □

Let $N = 3$, $m_1 = 1$, $m_2 = 42/25 \cong 1.68$, $m_3 = 61614/15625 \cong 3.94$, $m_4 = 4654584/390625 \cong 11.92$, $m_5 = 17249112/390625 \cong 44.16$. The normalized moments are $n_2 = 42/25 \cong 1.68$, $n_3 = 1467/625 \cong 2.35$,

11

a, b, n_2, n_3	a, b	$\{p_1,p_2,p_3\}, \{\lambda_1,\lambda_2,\lambda_3\}$
$\boxed{1}$ $\dfrac{7}{12}, \dfrac{7}{18}, \dfrac{100}{49}, \dfrac{78}{25}$	$\boxed{1,1}$ $\dfrac{3}{4}, -\dfrac{1}{8}$	$\dfrac{-1}{9}, \dfrac{7}{9}, \dfrac{1}{3}$, $\dfrac{25}{9}, \dfrac{25}{12}, \dfrac{25}{18}$
	$\boxed{1,2}$ $\dfrac{4}{3}, \dfrac{1}{6}$	$\dfrac{1}{12}, \dfrac{7}{12}, \dfrac{1}{3}$, $\dfrac{25}{12}, \dfrac{25}{9}, \dfrac{25}{18}$
$\boxed{2}$ $\dfrac{39}{28}, \dfrac{13}{12}, \dfrac{350}{169}, \dfrac{1029}{325}$	$\boxed{2,1}$ $\dfrac{1}{2}, -\dfrac{1}{14}$	$\dfrac{-1}{9}, \dfrac{8}{9}, \dfrac{2}{9}$, $\dfrac{25}{9}, \dfrac{25}{18}, \dfrac{25}{12}$
	$\boxed{2,2}$ $2, \dfrac{6}{7}$	$\dfrac{1}{3}, \dfrac{4}{9}, \dfrac{2}{9}$, $\dfrac{25}{18}, \dfrac{25}{9}, \dfrac{25}{12}$
$\boxed{3}$ $\dfrac{32}{15}, \dfrac{16}{9}, \dfrac{125}{64}, \dfrac{231}{80}$	$\boxed{3,1}$ $\dfrac{2}{3}, \dfrac{1}{15}$	$\dfrac{1}{12}, \dfrac{3}{4}, \dfrac{1}{6}$, $\dfrac{25}{12}, \dfrac{25}{18}, \dfrac{25}{9}$
	$\boxed{3,2}$ $\dfrac{3}{2}, \dfrac{3}{5}$	$\dfrac{1}{3}, \dfrac{1}{2}, \dfrac{1}{6}$, $\dfrac{25}{18}, \dfrac{25}{12}, \dfrac{25}{9}$

Table 1: The $N!$ different solutions of the iterative procedure for $N = 3$

$n_4 = 86196/28525 \cong 3.02$, $n_5 = 26619/7183 \cong 3.71$. Application of the iterative procedure is summarized in Table 1. The first column gives the three different solutions for a, b, n_2 and n_3 at the first application of the reduction step (first level of the tree in Figure 3). Then we apply the reduction step to all of these three solutions. At this level the procedure results in two solutions for a and b. These solutions, which are the leaves of the tree depicted in Figure 3, are listed in the second column. In the first two columns the indices of the solutions are indicated as in Figure 3. The corresponding APH(3) distributions are given in the third column by the initial probability vector and the vector of intensities.

As indicated by Corollary 12 the application of the procedure provides all permutations of the same set of intensities. In this particular case 4 permutations of the $3! = 6$ possible permutations result in a proper APH(3) distribution. These 4 different series representations correspond to the same APH(3) distribution, whose (unique) canonical representation is in the last row of the table.

The following corollary, whose proof is in Appendix B, indicates how to choose among the different solutions.

Corollary 13. *If the given moment set is feasible, selecting the largest solution for b in each step of the*

iterative procedure results in an APH distribution in CF, i.e., with non-decreasing intensities. If the given moment set is not feasible, the solutions provided by the procedure are improper for an APH distribution.

Corollary 13 allows us to reduce the $N!$ possibilities to one. This clearly simplifies the procedure. Furthermore, this solution is in CF which is a minimal unique representation of the APH distributions.

4 Constructing an APH(3) with given second to fifth normalized moments

We apply the iterative procedure for $N = 3$ with normalized moments n'_2, n'_3, n'_4 and n'_5. In order to construct an equation for a, as first step, we take (11) for $i = 4$:

$$n'_4 = \frac{4\left(1 + b\left(1 + \frac{1}{2}an_2 + \frac{1}{6}a^2 n_2 n_3 + \frac{1}{24}a^3 n_2 n_3 n_4\right)\right)}{(1+b)\left(1 + b\left(1 + \frac{1}{2}an_2 + \frac{1}{6}a^2 n_2 n_3\right)\right)}.$$

Then n_4 is substituted by $\mathcal{F}_4(n_2, n_3)$ (given in (4)) which results in

$$n'_4 = \frac{4\left(1 + b\left(1 + \frac{1}{2}an_2 + \frac{1}{6}a^2 n_2 n_3 + \frac{a^3 n_2^2\left(9n_2 + 2\left(n_3 - 6\right)n_3\right)}{36\left(n_2 - 2\right)}\right)\right)}{(1+b)\left(1 + b\left(1 + \frac{1}{2}an_2 + \frac{1}{6}a^2 n_2 n_3\right)\right)}.$$

By applying (15) to n_2 and n_3, and simplifying the right hand side we obtain

$$n'_4 = \left(2(9a(-2 + n'_2 + bn'_2)(4 + n'_2(-4 + (1+b)^2 n'_2)) - 12abn'_2(-2 - b + (1+b)^2 n'_2)n'_3 +\right.$$

$$\left. 2b(1+b)n'_2(-6n'_3 + n'_2(9 + (1+b)n'_3(-3 + n'_3 + bn'_3)))))\right) \Big/ \tag{17}$$

$$\left((3b(1+b)n'_2(-2(1+b+ab) + (1+b)^2 n'_2)n'_3)\right)$$

from which we have an expression for a as a function of n'_2, n'_3, n'_4 and b.

In order to construct the third order equation for b we take (11) for $i = 5$:

$$n'_5 = \frac{5\left(1 + b\left(1 + \frac{1}{2}an_2 + \frac{1}{6}a^2n_2n_3 + \frac{1}{24}a^3n_2n_3n_4 + \frac{1}{120}a^4n_2n_3n_4n_5\right)\right)}{(1+b)\left(1 + b\left(1 + \frac{1}{2}an_2 + \frac{1}{6}a^2n_2n_3 + \frac{1}{24}a^3n_2n_3n_4\right)\right)}.$$

Then n_5 is substituted by $\mathcal{F}_5(n_2, n_3, n_4)$ (given in (5)) which results in

$$n'_5 = \frac{5\left(1 + b\left(1 + \frac{1}{2}an_2 + \frac{1}{6}a^2n_2n_3 + \frac{1}{24}a^3n_2n_3n_4 + \frac{a^4n_2n_3^2\left(4n_2\left(2n_3 - 3n_4\right) + 3n_4^2\right)}{144\left(2n_3 - 3n_2\right)}\right)\right)}{(1+b)\left(1 + b\left(1 + \frac{1}{2}an_2 + \frac{1}{6}a^2n_2n_3 + \frac{1}{24}a^3n_2n_3n_4\right)\right)}.$$

By applying (15) to n_2, n_3 and n_4, and simplifying the right hand side we have

$$n'_5 = \Bigg(\big(5(-72n'_3n'_4 + 2n_2'^2(4(-3 + n'_3 + bn'_3)(9 + (1+b)^2n_3'^2) - 3(1+b)^2n'_3(-3 + 2(1+b)n'_3)n'_4) +$$

$$3n'_2n'_3(96 - 12bn'_4 + n'_3(-16(2+b) + 4(1+b)(2+b)n'_4 + b(1+b)^2n_4'^2)))\big)\Bigg) \Bigg/$$

$$\Bigg((6(1+b)n'_3(-12 - 3(1+b)^2n_2'^2 + 2n_2'^2(6 + b(3 + n'_3 + bn'_3)))n'_4\Bigg)$$

(18)

which provides a third order polynomial equation to determine b.

Solving (18) and substituting the solutions into (17) and (15) results in three solutions for n_2, n_3, a and b. According to Corollary 13 we can take that with the largest b, check the feasibility of a and b and look for an APH(2) with normalized moments n_2 and n_3 by applying the reduction step for $N = 2$.

5 Implementation and numerical accuracy

The presented method is implemented in Mathematica which allows us to check its properties by applying either symbolic calculations (with exact rational representation of the involved quantities) or numerical calculations with different precision. The experiments were carried out on a computer with 1.5GHz processor and 524MB RAM.

The exact rational arithmetic is applicable only to moment sets that correspond to solution vectors composed of rational numbers. We generated such moment sets by starting from an APH distribution with rational initial probabilities and intensities. Based on the obtained rational moments the procedure

	n_2	n_3	n_4	n_5	$\underline{p}$	$\underline{\lambda}$
D1	2.5	10	34.16	52.23	$2.412\ 10^{-3}, 9.975\ 10^{-1}$	$9.352\ 10^{-2}, 1.026$
D2	2.5	10	100	200	$3.349\ 10^{-5}, 1.205\ 10^{-1}, 8.794\ 10^{-1}$	$2.436\ 10^{-2}, 5.380\ 10^{-1}, 1.291$
D3	2.5	10	50	200	$4.383\ 10^{-7}, 2.884\ 10^{-3}, 9.971\ 10^{-1}$	$1.228\ 10^{-2}, 1.024\ 10^{-1}, 1.029$
D4	2.5	10	100	400	$9.187\ 10^{-7}, 4.958\ 10^{-3}, 9.950\ 10^{-1}$	$1.041\ 10^{-2}, 1.335\ 10^{-1}, 1.038$
D5	2.5	10	50	400	$9.500\ 10^{-9}, 2.575\ 10^{-3}, 9.974\ 10^{-1}$	$4.828\ 10^{-3}, 9.654\ 10^{-2}, 1.027$

Table 2: Normalized moments and descriptors of APH distributions for Example 1

calculated exactly the starting APH distribution for 8 phases (matching 15 moments) in 165 seconds.

We found that the standard floating point precision of Mathematica (16 digits) is usually sufficient up to $N = 5$, but the round-off error is already visible in the 8th digit of the normalized moments. Fitting higher order APH distributions requires higher numerical precision. For $N = 5$, using 32 digit accuracy floating point arithmetic, the computation time is approximately the same and the round-off error disappears. The computational complexity of matching higher order APH distributions increases exponentially with the order. The calculations for 5,6,7 and 8 phases takes about 2.5, 5.5, 32 and 165 seconds, respectively. For $N > 8$ we cannot perform the procedure because with the computer we use we are not able to provide the functions $\mathcal{F}_{2N}(n_2, ..., n_{2N-1})$ and $\mathcal{F}_{2N+1}(n_2, ..., n_{2N})$.

6   Examples

We demonstrate the properties of moments based APH distribution matching through some simple numerical examples.

Example 1: We construct APH distributions with $m_1 = 1$ and the sets of normalized moments listed in Table 2. Distribution D1 is the result of matching normalized moments $n_2 = 2.5$ and $n_3 = 10$ with an APH(2). All higher order moments of this distribution are determined by n_2 and n_3. All other cases are obtained by APH(3) distributions. Figure 4 presents the body and the tail of the pdf of the distributions. The queue length distribution of the M/APH/1 queue with $\rho = 0.8$ utilization is provided in Figure 5. Even if the second and third normalized moments are identical, the distributions and the associated

queue length distributions can differ significantly. In particular, we can recognize the difference between the role of the odd and the even moments. Comparing, for example, D2 and D4 ($\{n_2, n_3\} = \{100, 200\}$ and $\{n_2, n_3\} = \{100, 400\}$) indicates that a higher fifth moment results in a longer tail for the pdf and also for the queue length distribution. Instead, comparing D2 and D3 ($\{n_2, n_3\} = \{100, 200\}$ and $\{n_2, n_3\} = \{50, 200\}$) indicates that a higher fourth moment results in a shorter tail.
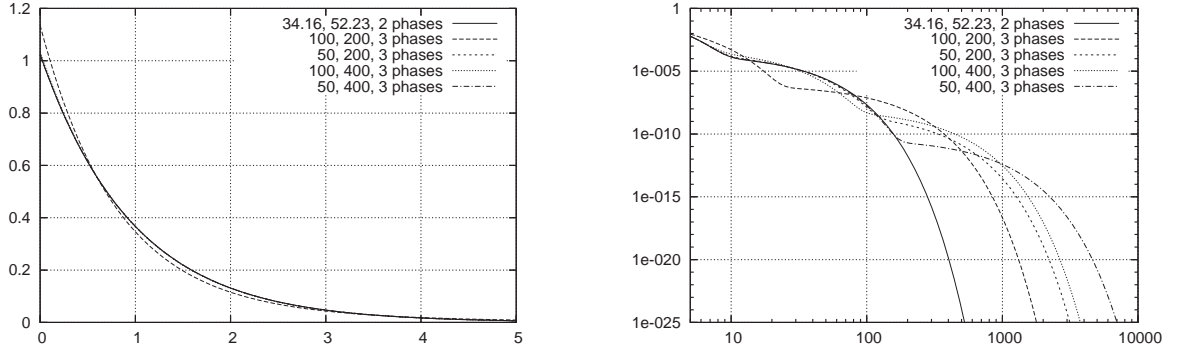

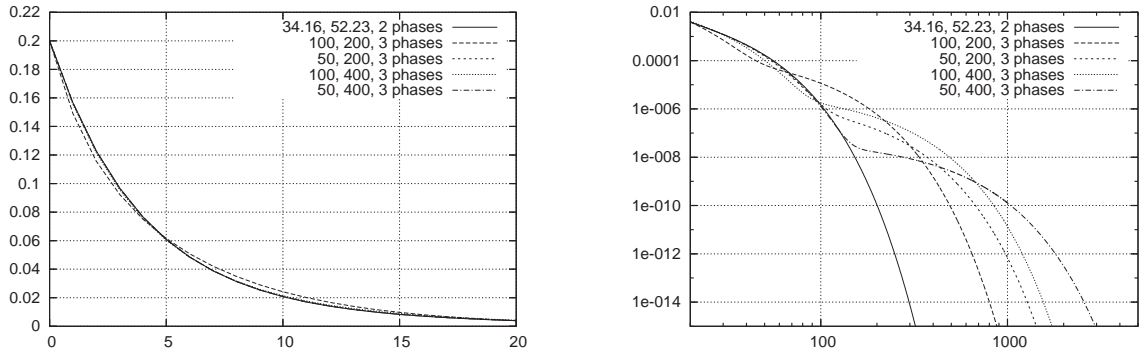Figure 4: The body and the tail of the pdf of the distributions D1-D5

Figure 5: The body and the tail of the queue length distribution of the M/APH/1 queue for distributions D1-D5

Example 2: We match an APH distribution with $m_1 = 1$ to sets of normalized moments listed in Table 3. Similar to Example 1, the first set of moments of this example (D6) can be realized with an APH(3) distribution. The other cases require 4 phases. The pdf and the queue length distribution of the M/APH/1 queue with $\rho = 0.8$ utilization are provided in Figures 6 and 7, respectively. The different behaviour of the odd and the even moments appears also in this case.

Example 3: We match APH distributions to a distribution with Pareto tail whose pdf is given by
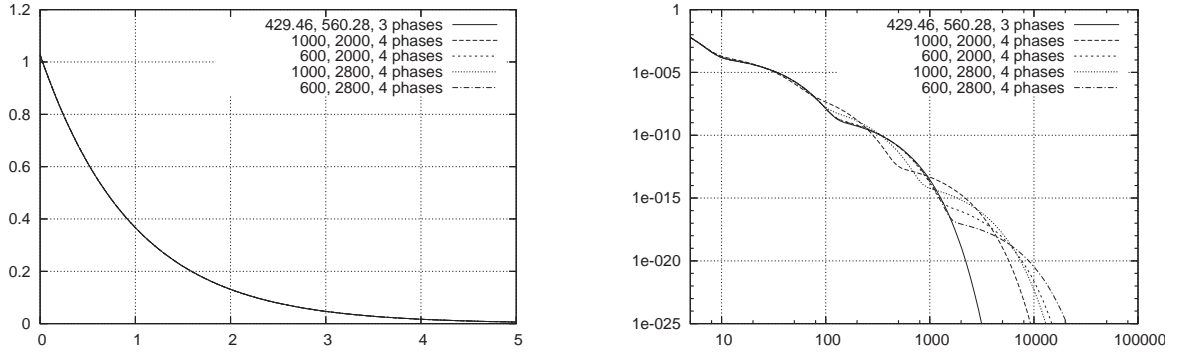
16

Figure 6: The body and the tail of the pdf of the distributions D6-D10

$$
f(t) = \begin{cases} \alpha B^{-1} e^{-\frac{\alpha}{B} t} & \text{for } t \leq B, \\ \alpha B^{\alpha} e^{-\alpha} t^{-(\alpha+1)} & \text{for } t > B, \end{cases}
$$

where $\alpha = 3.5$, $B = 4$ and the distribution is truncated at 10^4 to ensure finite higher moments. The matching is performed based on the distributions first 3, 5 or 7 moments.

The pdf and the resulting queue length distributions are depicted in Figures 8 and 9, respectively. The figures show that the approximate APH(3) and APH(4) distributions (matching 5 and 7 moments) captures the slowly decaying tail behaviour of the Pareto-like distribution.

Example 4: The normalized moments of the APH(5) distribution with

$$
\underline{p} = \left\{ \frac{7}{50}, \frac{2}{25}, \frac{11}{50}, \frac{9}{25}, \frac{1}{5} \right\} \quad \text{and} \quad \underline{\lambda} = \left\{ \frac{6}{5}, \frac{3}{2}, \frac{21}{5}, \frac{11}{2}, \frac{33}{5} \right\}
$$

	n_2	n_3	n_4	n_5	n_6	n_7
D6	2.5	10	50	200	429.46	560.28
D7	2.5	10	50	200	1000	2000
D8	2.5	10	50	200	600	2000
D9	2.5	10	50	200	1000	2800
D10	2.5	10	50	200	600	2800

Table 3: Normalized moments and descriptors of APH distributions of Example 2
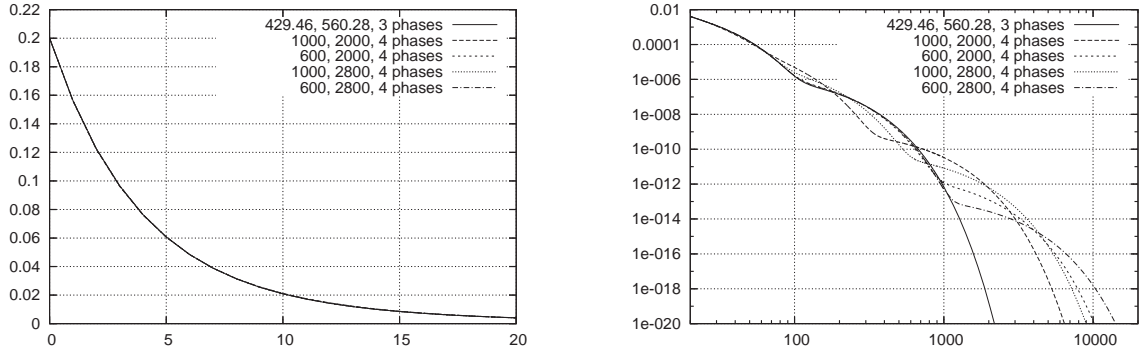
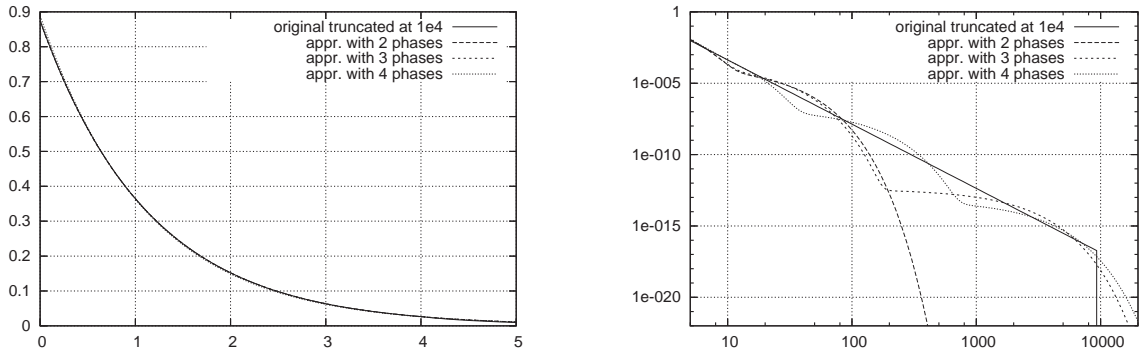Figure 7: The body and the tail of the queue length distribution of the M/APH/1 queue for distributions D6-D10

Figure 8: The pdf of the Pareto-like and the matching APH distributions

are

$$n_2 = \frac{591837100}{236021769} \ , \quad n_3 = \frac{127015173550}{30307977891} \ , \quad n_4 = \frac{73234144433000}{13008894074991} \ , \quad n_5 = \frac{23269903941935150}{3375288482772537} \ ,$$

$$n_6 = \frac{19293234891794731100}{2383303561732998063} \ , \quad n_7 = \frac{9169582135866851564950}{988006558808808179631} \ ,$$

$$n_8 = \frac{12639577877112590464966000}{1207476774457049490791373} \ , \quad n_9 = \frac{12579207440182804230357743 0}{1078787971811559596184848 1}$$

and its first moment is 5121/7700. If performed symbolically, i.e. the involved quantities are calculated exactly by rational fractions, the procedure returns an APH(5) distribution with exact rational fractions in 2.5 seconds. This distribution is the canonical representation of the original one and naturally has exactly the same moments. When performed numerically, with accuracy set to 16 effective digits (it is the normal machine precision), the procedure returns an APH(5) distribution whose normalized moments differ from the original normalized moments at most by 10^{-7}. Using 32 effective digits the error disappears.
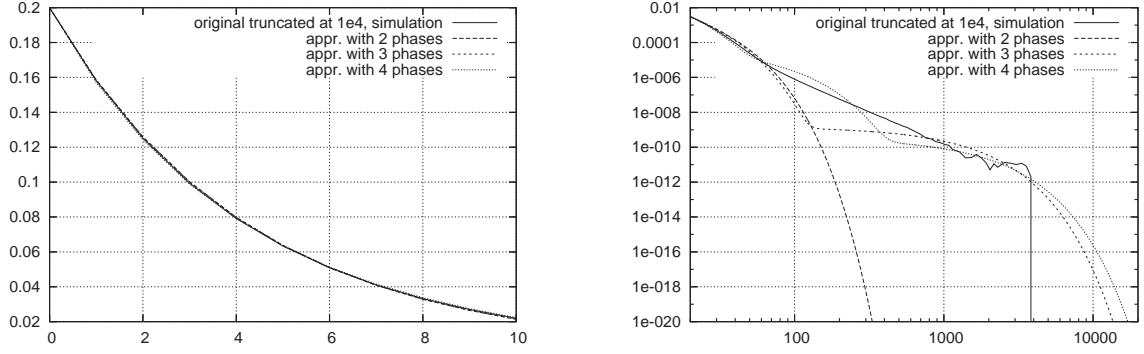
Figure 9: The queue length distribution of the M/•/1 queue with the Pareto-like and the matching APH distributions

Example 5: The normalized moments of the APH(6) distribution with

$$\underline{p} = \left\{ \frac{7}{50}, \frac{2}{25}, \frac{11}{50}, \frac{9}{25}, \frac{3}{25}, \frac{2}{25} \right\} \quad \text{and} \quad \underline{\lambda} = \left\{ \frac{6}{5}, \frac{3}{2}, \frac{16}{5}, \frac{21}{5}, \frac{11}{2}, \frac{33}{5} \right\}$$

are

$$n_2 = \frac{13910465950}{7071632649} \;,\quad n_3 = \frac{24046145958125}{7798485420889} \;,\quad n_4 = \frac{4466727822468564}{1078460027763523} \;,$$
$$n_5 = \frac{1435435363741658131625}{281715407081136714339} \;,\quad n_6 = \frac{192703588148997146178 9950}{32189350944833935270 0643} \;,$$
$$n_7 = \frac{7403583591151809030412389625}{10803348558809078009086 81769} \;,\quad n_8 = \frac{164660995794600341649657822009 20}{21345927597624996918667511339 49} \;,$$
$$n_9 = \frac{9897195470192916564642669334916 4975}{1153903093279610544195389602184 9713} \;,\quad n_{10} = \frac{3493903448416776027788733827992 88523250}{3699043816333035256313315521693 9245523} \;,$$
$$n_{11} = \frac{36383452218979323212746997371127 66471711375}{35257538722525433580580559255688 6837427947}$$

and its first moment is 28031/30800. With symbolic computations the procedure returns an APH(6) distribution with exact rational fractions in 5.5 seconds. When performed numerically with low accuracy (16 or 32 effective digits) the procedure fails. With 48 effective digits the algorithm returns an APH(6) whose normalized moments are equal to the original normalized moments (with respect to the first 48 digits) and whose parameters differ from the parameters of the original APH(6) at most by 10^{-15}.

Example 6: The normalized moments of the APH(7) distribution with

$$\underline{p} = \left\{ \frac{7}{50}, \frac{2}{25}, \frac{3}{100}, \frac{9}{25}, \frac{3}{25}, \frac{2}{25}, \frac{19}{100} \right\} \quad \text{and} \quad \underline{\lambda} = \left\{ \frac{6}{5}, \frac{3}{2}, \frac{16}{5}, \frac{21}{5}, \frac{11}{2}, \frac{33}{5}, 7 \right\}$$

19

are

$$n_2 = \frac{231370003400}{106993755801} \ , \quad n_3 = \frac{22323601720100}{6638675152819} \ , \quad n_4 = \frac{6126228628514858000}{1387385281818168081} \ ,$$

$$n_5 = \frac{5357085754374288325900}{1001941629079290768471} \ , \quad n_6 = \frac{363710029696853956408674 00}{5840991310566917790378547}$$

$$n_7 = \frac{14068555822667241811857073 5100}{1982819783396853871455347 6121} \ , \quad n_8 = \frac{15712181740896712706010238 73684000}{19722045175879851983214158 0916321} \ ,$$

$$n_9 = \frac{37886181842998848420401812 1271493140}{42828657793879781911860359 004846893} \ ,$$

$$n_{10} = \frac{20103343383358824457998073 326051014657000}{20654220324438467199108353 90829735576681} \ ,$$

$$n_{11} = \frac{69887740898933010255011787 1660953710324129 00}{65757835173532881213867117 86877960843290043} \ ,$$

$$n_{12} = \frac{79867798947162511649089160 31469341748175370 6000}{69273364122121480973951819 90982615081615523387} \ ,$$

$$n_{13} = \frac{18049786717034286902032065 42385535428146436 8290060}{14513709537676616887725230 6943199456471467 5304383}$$

and its first moment is 109033/123200. As in the previous example, with symbolic computations the procedure returns an APH(7) distribution with exact rational fractions in 32 seconds. With low accuracy the procedure returns incorrect values, while with higher accuracy the memory requirement of the computation is over the available amount of memory (524MB RAM).

Example 7: If the input of the procedure is the set of normalized moments of the APH(8) distribution with

$$\underline{p} = \left\{ \frac{7}{50}, \frac{2}{25}, \frac{3}{100}, \frac{9}{25}, \frac{3}{25}, \frac{2}{25}, \frac{11}{100}, \frac{2}{25} \right\} \quad \text{and} \quad \underline{\lambda} = \left\{ \frac{6}{5}, \frac{3}{2}, \frac{16}{5}, \frac{21}{5}, \frac{11}{2}, \frac{33}{5}, 7, \frac{36}{5} \right\}$$

then symbolic computations returns the original APH(8) in 165 seconds. As in the case of the APH(7) distribution, numerical computations fail either for low accuracy of for too high memory requirement.

7 Conclusion

The paper presents a moments matching method for the APH(N) class. According to the authors present knowledge, this is the first procedure that provides a "Markovian" distribution based on more than 3 moments, which has practical importance in applied PH fitting.

The procedure can provide symbolic result up to order 4 and numeric results for higher orders. The

practical applicability of the numerical procedure is limited by floating point errors and the complexity of the involved expressions. According to our experiences the procedure becomes instable around order $5 - 6$ with standard floating point arithmetic and requires the use of higher numerical precision.

The procedure always provides solution when the given $2N-1$ moments are the moments of an APH(N) distribution (apart of numerical errors in case of numerical computation). This way the procedure allows to investigate the symbolic/numeric moment bounds of the APH(N) class.

Acknowledgment

A   Proofs of Theorem 9 and 10

We start with a brief extract of the results presented in [10] which we use to determine the $2N$th and the $2N + 1$th normalized moments of an APH(N) distribution in terms of the previous ones.

For a given number of phases, N, the following two $N + 1 \times N + 1$ matrices of factorial moments, $r_i = m_i/i!$, are constructed

$$
M_{2N} = \begin{vmatrix} r_0 & r_1 & \cdots & r_N \\ r_1 & r_2 & \cdots & r_{N+1} \\ \vdots & & & \vdots \\ r_N & r_{N+1} & \cdots & r_{2N} \end{vmatrix}, \quad M_{2N+1} = \begin{vmatrix} r_1 & r_2 & \cdots & r_{N+1} \\ r_2 & r_3 & \cdots & r_{N+2} \\ \vdots & & & \vdots \\ r_{N+1} & r_{N+2} & \cdots & r_{2N+1} \end{vmatrix}. \tag{19}
$$

Then r_{2N} can be obtained as a function of $r_i, 0 \le i \le 2N - 1$ by solving the equation $\det(M_{2N}) = 0$. It follows from the definition of the determinant that

$$
\det(M_{2N}) = \sum_{i=1}^{N+1} r_{N+i-1}\det_{i,N+1}(M_{2N}) = r_{2N}\ \det(M_{2N-2}) + \sum_{i=1}^{N} r_{N+i-1}\ \det_{i,N+1}(M_{2N}),
$$

where $\det_{i,j}(M_{2N})$ is the (signed) subdeterminant of the element i, j. From which we have

$$
r_{2N} = -\frac{\sum_{i=1}^{N} r_{N+i-1}\ \det_{i,N+1}(M_{2N})}{\det(M_{2N-2})}\ . \tag{20}
$$

and, similarly, one can obtain

$$r_{2N+1} = -\frac{\sum_{i=1}^{N} r_{N+i} \det_{i,N+1}(M_{2N+1})}{\det(M_{2N-1})} \ . \tag{21}$$

Corollary 14. *The determinant of the $n \times n$ matrix,*

$$\begin{vmatrix} r_{a_1+1} & r_{a_2+1} & \cdots & r_{a_n+1} \\ r_{a_1+2} & r_{a_2+2} & \cdots & r_{a_n+2} \\ \vdots & & & \vdots \\ r_{a_1+n} & r_{a_2+n} & \cdots & r_{a_n+n} \end{vmatrix}, \tag{22}$$

is a sum of products $r_{i_1}^{p_1} r_{i_2}^{p_2} \cdots r_{i_m}^{p_m}$, where each products have the following property

$$\sum_{j=1}^{m} p_j = n, \quad and \quad \sum_{j=1}^{m} i_j p_j = \sum_{j=1}^{n} a_j + \frac{n(n+1)}{2} \tag{23}$$

Proof. The corollary holds for $n = 1$. Assuming that the corollary holds for $k > 1$ we evaluate the determinant of the matrix of (22) for $n = k+1$ using the subdeterminants of the last row and we get (23) for $k + 1$. □

Corollary 14 has the following consequences. The terms of $\det(M_{2N})$, $r_{i_1}^{p_1} r_{i_2}^{p_2} \cdots r_{i_m}^{p_m}$, as well as the terms of the numerator of (20) satisfy $\sum_{j=1}^{m} p_j = N+1$, and $\sum_{j=1}^{m} i_j p_j = N(N+1)$. The terms of $\det(M_{2N-2})$, which give the the denominator of (20), satisfy $\sum_{j=1}^{m} p_j = N$, and $\sum_{j=1}^{m} i_j p_j = N(N-1)$. The terms of $\det(M_{2N+1})$ and the terms of the numerator of (21) satisfy $\sum_{j=1}^{m} p_j = N+1$, and $\sum_{j=1}^{m} i_j p_j = (N+1)^2$. The terms of $\det(M_{2N-1})$, which give the denominator of (21) satisfy $\sum_{j=1}^{m} p_j = N$, and $\sum_{j=1}^{m} i_j p_j = N^2$.

As an example, for $N = 2$, (21) results in $r_5 = \dfrac{-r_3^3 + 2r_2 r_3 r_4 - r_1 r_4^2}{r_2^2 - r_1 r_3}$, where for any product in the numerator we have $\sum_{j=1}^{m} p_j = 3$, and $\sum_{j=1}^{m} i_j p_j = 9$, while for the denominator we have $\sum_{j=1}^{m} p_j = 2$, and $\sum_{j=1}^{m} i_j p_j = 4$.

From (20) and (21) one can obtain $\mathcal{F}_{2N}(n_2, ..., n_{2N-1})$ and $\mathcal{F}_{2N+1}(n_2, ..., n_{2N})$ using $r_0 = 1$, $r_i = \frac{1}{i!} \prod_{j=1}^{i} n_j$, for $1 \leq i \leq 2N - 1$. Based on the above discussed properties of r_{2N} and r_{2N+1}, we have the following properties of $\mathcal{F}_{2N}(n_2, ..., n_{2N-1})$ and $\mathcal{F}_{2N+1}(n_2, ..., n_{2N})$.

Property 1. *$\mathcal{F}_{2N}(n_2, ..., n_{2N-1})$ is a fraction of sums of products of normalized moments. For any product of the numerator, $n_{i_1}^{p_1} n_{i_2}^{p_2} \cdots n_{i_n}^{p_n}$, there exists a constant K such that $K \leq \sum_{j=1}^{n} p_j \leq K + 1$. For*

any product of the denominator, $n_{i_1}^{p_1} n_{i_2}^{p_2} \cdots n_{i_n}^{p_n}$, we have $K - 1 \leq \sum_{j=1}^{n} p_j \leq K$.

As an example consider (4) in which the some of the exponents of any product of the numerator (denominator) is either two or three (one or two).

Property 2. *$\mathcal{F}_{2N+1}(n_2, ..., n_{2N})$ is a fraction of sums of products of normalized moments. For any product of the numerator $n_{i_1}^{p_1} n_{i_2}^{p_2} \cdots n_{i_n}^{p_n}$, there exists a constant K such that $\sum_{j=1}^{n} p_j = K$. For any product of the denominator, $n_{i_1}^{p_1} n_{i_2}^{p_2} \cdots n_{i_n}^{p_n}$, we have $\sum_{j=1}^{n} p_j = K - 1$.*

As an example consider (5) in which the some of the exponents of any product of the numerator (denominator) is three (two).

The above two properties allows us to determine the structure of the two equations that are used to determine a and b.

Proof of Theorem 9. By Step 1 of the procedure to construct the equation for a we have

$$n'_{2N-2} = \frac{(2N-2)\left(1 + b \sum_{j=1}^{2N-2} \frac{a^{j-1}}{j!} \prod_{k=1}^{j} n_k\right)}{(1+b)\left(1 + b \sum_{j=1}^{2N-3} \frac{a^{j-1}}{j!} \prod_{k=1}^{j} n_k\right)} . \tag{24}$$

Then by Step 2 we obtain

$$n'_{2N-2} = \frac{(2N-2)\left(1 + b \sum_{j=1}^{2N-3} \frac{a^{j-1}}{j!} \prod_{k=1}^{j} n_k + \frac{ba^{2N-3}}{(2N-2)!} \mathcal{F}_{2N-2}(n_2, ..., n_{2N-3}) \prod_{k=1}^{2N-3} n_k\right)}{(1+b)\left(1 + b \sum_{j=1}^{2N-3} \frac{a^{j-1}}{j!} \prod_{k=1}^{j} n_k\right)} .$$

According to Property 1, $\mathcal{F}_{2N-2}(n_2, ..., n_{2N-3})$ is of the form

$$\frac{\sum_i \prod_{j=2}^{2N-3} n_j^{e_{i,j}}}{\sum_k \prod_{l=2}^{2N-3} n_l^{f_{k,l}}}$$

and there exists a constant, K such that

$$K \le \sum_{j=2}^{2N-2} e_{i,j} \le K+1, \ \forall i, \text{ and } K-1 \le \sum_{l=2}^{2N-2} f_{k,l} \le K, \ \forall k.$$

In Step 3, n_i is substituted according to (15). We obtain

$$n'_{2N-2} = (2N-2) \left(1 + b + b \sum_{j=2}^{2N-3} \frac{a^{j-1}}{j!} \prod_{k=2}^{j} \frac{n'_{k-1}(1+b)\left(n'_k(1+b) - k\right)}{a\left(n'_{k-1}(1+b) - (k-1)\right)} + \right. \tag{25}$$

$$\frac{ba^{2N-3}}{(2N-2)!} \frac{\sum_i \prod_{j=2}^{2N-3} \left(\frac{n'_{j-1}(1+b)\left(n'_j(1+b) - j\right)}{a\left(n'_{j-1}(1+b) - (j-1)\right)} \right)^{e_{i,j}}}{\sum_k \prod_{l=2}^{2N-3} \left(\frac{n'_{l-1}(1+b)\left(n'_l(1+b) - l\right)}{a\left(n'_{l-1}(1+b) - (l-1)\right)} \right)^{f_{k,l}}} \left. \prod_{k=2}^{2N-3} \frac{n'_{k-1}(1+b)\left(n'_k(1+b) - k\right)}{a\left(n'_{k-1}(1+b) - (k-1)\right)} \right) \Bigg/ \tag{26}$$

$$\left((1+b)\left(1 + b + b \sum_{j=2}^{2N-3} \frac{a^{j-1}}{j!} \prod_{k=2}^{j} \frac{n'_{k-1}(1+b)\left(n'_k(1+b) - k\right)}{a\left(n'_{k-1}(1+b) - (k-1)\right)} \right) \right) . \tag{27}$$

In the above equation, parts (25) and (27) do not depend on a. Part (26) instead depends on a in such a way that the whole equation (25-27) can be rearranged to the form

$$n'_{2N-2} = \frac{C_1 + a \dfrac{\dfrac{C_2}{a^K} + \dfrac{C_3}{a^{K+1}}}{\dfrac{C_4}{a^{K-1}} + \dfrac{C_5}{a^K}}}{C_6} \tag{28}$$

where the constants, $C_i, 1 \le i \le 6$ are independent of a. It is easy to see that (28) can be rearranged to an equation which is linear in a. $\qquad\square$

Proof of Theorem 10. Step 2 of the procedure to construct the equation for b leads to

$$n'_{2N-1} = \frac{(2N-1)\left(1 + b \sum_{j=1}^{2N-2} \frac{a^{j-1}}{j!} \prod_{k=1}^{j} n_k + \frac{ba^{2N-2}}{(2N-1)!} \mathcal{F}_{2N-1}(n_2, ..., n_{2N-2}) \prod_{k=1}^{2N-2} n_k \right)}{(1+b)\left(1 + b \sum_{j=1}^{2N-2} \frac{a^{j-1}}{j!} \prod_{k=1}^{j} n_k \right)} .$$

After application of step 3, by similar steps as in the proof of Theorem 9 and considering Property 2, we

obtain an equation of the form

$$n'_{2N-1} = \frac{D_1(1+b)^{2N-2} + (1+b)^{2N-2} \dfrac{\sum_i \dfrac{\prod_{j=2}^{2N-2}(n'_j(1+b)-j)^{g_{i,j}}}{D_{2,i}\, b^{\sum_{j=2}^{2N-2} g_{i,j}}}}{\sum_k \dfrac{\prod_{l=2}^{2N-3}(n'_l(1+b)-l)^{h_{k,l}}}{D_{3,k}\, b^{\sum_{l=2}^{2N-3} h_{k,l}}}}}{D_4(1+b)^{2N-1}} \qquad (29)$$

where the constants, $D_1, D_{2,i}, D_{3,k}, D_4$ are independent of b and a. Moreover

$$\sum_{j=2}^{2N-2} g_{i,j} \le N, \ \forall i, \ \text{and} \ \sum_{l=2}^{2N-2} h_{k,l} \le N-1, \ \forall k.$$

It is easy to verify that rearranging (29) we can obtain a polynomial equation for b which is of order N in b. $\qquad\square$

B Selecting the solution of interest

To be unique the canonical representation of an acyclic phase type distributions requires that $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_N$. The following theorem allows us to select the solution of b which results in an APH(N) distribution in CF, i.e., with non-decreasing intensities.

Theorem 15. *Among the APH(N − 1)-EXP representations of an APH(N) distribution the largest b value is associated with the one where the intensity of the additional phase is the largest intensity.*

Proof. Starting from an APH$^*(N-1)$-EXP representation of an APH(N) distribution where the intensity of the additional phase (λ^*) is not the largest intensity we apply an equivalent transformation into an APH$\hat{}(N-1)$-EXP structure such that the intensity of the additional phase becomes the largest intensity ($\hat{\lambda}$). The steps of the transformation are depicted in Figure 10. In the first step we replace the APH$^*(N-1)$ distribution with its CF (Fig. 10b). Consequently $\tilde{\lambda}_{N-1} = \hat{\lambda}$ is the largest intensity of the APH(N) distribution. The obtained structure is decomposed into an APH($N-2$) part and two additional exponential phases (Fig. 10c). After that the last two phases are interchanged such that the APH(N) distribution remains the same (Fig. 10d). This transformation is due to the fact that

$$\hat{p}\frac{\hat{\lambda}}{s+\hat{\lambda}}\frac{\lambda^*}{s+\lambda^*} + (1-p^*)\frac{\lambda^*}{s+\lambda^*} = \left(\hat{p} + (1-p^*)\frac{\hat{\lambda}-\lambda^*}{\hat{\lambda}}\right)\frac{\lambda^*}{s+\lambda^*}\frac{\hat{\lambda}}{s+\hat{\lambda}} + (1-p^*)\frac{\lambda^*}{\hat{\lambda}}\frac{\hat{\lambda}}{s+\hat{\lambda}} \ .$$
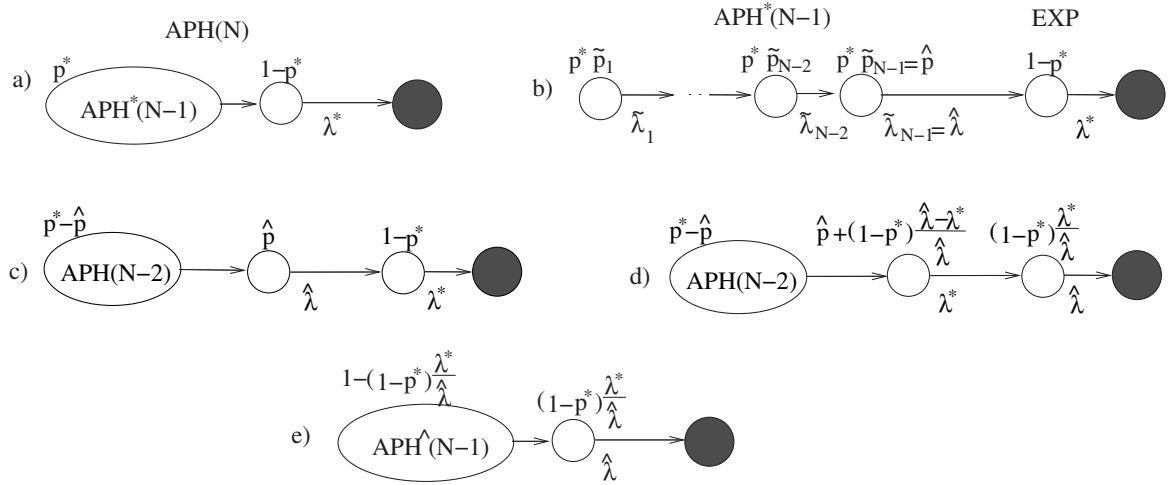
Figure 10: Equivalent transformation of APH$(N-1)$-EXP forms

The b parameter of the original APH$^*(N-1)$-EXP representation, b^* (Fig. 10a), and the final APHˆ$(N-1)$-EXP representation, $\hat{b}$ (Fig. 10e), are

$$b^* = p^* m_1^* \lambda^* \quad \text{and} \quad \hat{b} = \left(1 - (1-p^*)\frac{\lambda^*}{\hat{\lambda}}\right) \hat{m}_1 \hat{\lambda},$$

where m_1^* is the mean of APH$^*(N-1)$ and $\hat{m}_1$ is the mean of APHˆ$(N-1)$.

$$m_1^* = \sum_{i=1}^{N-1} \tilde{p}_i \sum_{j=i}^{N-1} \frac{1}{\tilde{\lambda}_j} = (1 - \tilde{p}_{N-1})m_1' + \frac{1}{\tilde{\lambda}_{N-1}} = \left(1 - \frac{\hat{p}}{p^*}\right) m_1' + \frac{1}{\hat{\lambda}},$$

$$\hat{m}_1 = \frac{p^* - \hat{p}}{1 - (1-p^*)\frac{\lambda^*}{\hat{\lambda}}} m_1' + \frac{1}{\hat{\lambda}},$$

where m_1' is the mean of the APH$(N-2)$ and $\hat{p} = p^* \tilde{p}_{N-1}$. Finally substituting m_1^* and $\hat{m}_1$ results:

$$
\begin{aligned}
\hat{b} - b^* &= \left(\hat{\lambda}(p^* - \hat{p})m_1' + 1 - (1-p^*)\frac{\lambda^*}{\hat{\lambda}}\right) - \left(\lambda^*(p^* - \hat{p})m_1' + p^*\frac{\lambda^*}{\hat{\lambda}}\right) \\
&= \underbrace{(\hat{\lambda} - \lambda^*)}_{>0} \underbrace{(p^* - \hat{p})}_{p^*(1-\tilde{p}_{N-1})>0} m_1' + \underbrace{1 - \frac{\lambda^*}{\hat{\lambda}}}_{>0} + \underbrace{p^*\frac{\lambda^*}{\hat{\lambda}} - p^*\frac{\lambda^*}{\hat{\lambda}}}_{0} > 0 \; .
\end{aligned}
$$

$\square$

Corollary 13 is a consequence of Theorems 11 and 15, and Corollary 12.

References

[1] S. Asmussen and O. Nerman. Fitting phase-type distributions via the EM algorithm. In *Proceedings: "Symposium i Advent Statistik"*, pages 335–346, Copenhagen, 1991.

[2] A. Bobbio, A. Horváth, and M. Telek. Matching three moments with minimal acyclic phase type distributions. *Stochastic Models*, 21:303–326, 2005.

[3] A. Cumani. On the canonical representation of homogeneous Markov processes modelling failure-time distributions. *Microelectronics and Reliability*, 22:583–602, 1982.

[4] Mark Fackrell. *Characterization of. Matrix-exponential distributions.* PhD thesis, The University of Adelaide, School of Applied Mathematics, 2003.

[5] A. Horváth and M. Telek. PhFit: A general purpose phase type fitting tool. In *Tools 2002*, pages 82–91, London, England, April 2002. Springer, LNCS 2324.

[6] A. Lang and J. L. Arthur. Parameter approximation for phase-type distributions. In S. R. Chakravarty and A. S. Alfa, editors, *Matrix-Analytic Methods in Stochastic Models*, Lecture Notes in Pure and Applied Mathematics, pages 151–206. Marcel Dekker, Inc., 1996.

[7] T. Osogami and M. Harchol-Balter. A closed form solution for mapping general distributions to minimal PH distributions. In *International Conference on Performance Tools – TOOLS 2003*, pages 200–217, Urbana, IL, USA, Sept 2003. Springer, LNCS 2794.

[8] T. Osogami and M. Harchol-Balter. Necessary and sufficient conditions for representing general distributions by Coxians. In *International Conference on Performance Tools – TOOLS 2003*, pages 182–199, Urbana, IL, USA, Sept 2003. Springer, LNCS 2794.

[9] M. Telek and A. Heindl. Matching moments for acyclic discrete and continuous phase-type distributions of second order. *International Journal of Simulation Systems, Science & Technology*, 3(3-4):47–57, dec. 2002. Special Issue on: Analytical & Stochastic Modelling Techniques.

[10] A. van de Liefvoort. The moment problem for continuous distributions. Technical report, University of Missouri, WP-CM-1990-02, Kansas City, 1990.