technische universität
dortmund

department of
computer science

TECHNICAL REPORTS IN COMPUTER SCIENCE

Technische Universität Dortmund

Equivalence Transformations for Acyclic Phase Type Distributions

Jan Kriege and Peter Buchholz

Informatik IV

Number: 827

October 2009

ABSTRACT

This short note introduces different equivalence transformations for acyclic phase type distributions. The goal of the first two transformations is to generate a representation of a phase type distribution which is more amenable for subsequent analysis steps or for an expansion of the distribution into a stochastic process. The third transformation is used to reduce the number of states of an acyclic phase type distribution.

1 Introduction

Phase type (PH) distributions are a powerful class of distributions which describe a random variable by means the absorption time of an absorbing Markov chain [5]. It is known [7] that every distribution with a continuous and strictly positive density on $(0, \infty)$ can be represented by a PH distribution. However, it is also known that the PH representation for a random variable is non unique [6]. In fact, as shown in [10], a PH distribution of order n has $n^2 - 1$ free parameters but $2n$ are sufficient to characterize the distribution. Canonical representations are only known for PH distributions of order 2 or 3 [3], for larger dimensions one probably has to work with different representations. For this reason, often acyclic or triangular phase type (APH) distributions [8] are used. Although this class is more restrictive than general PH distributions since it can only be used to model distribution where the poles of the Laplace transforms are real, it has some nice properties. In particular, a canonical representation with $2n - 1$ parameters exists [2]. Since a APH distribution of order n has $(n^2 + 3n - 2)/2$ parameters, the representation is redundant but [2] provides an algorithm to transform every APH in its canonical form.

In some situations the canonical representation of an APH distribution is not the best choice. This is for example the case when the APH distribution has to be expanded into a Markovian arrival process (MAP) to capture autocorrelations [1, 4]. In this case, the representation of the APH distribution determines the joint moments and lag k autocorrelation which can be reached by the MAP. Thus, methods to perform equivalence transformations for APH representations are important. In this note we present three different transformations. The first one has been proposed in [1] without giving proofs which are presented in this note. The second one is from [4] and the third is used in [9] to find APH representations with less states.

This report is structured as follows: Sec. 2 introduces (A)PH distributions, MAPs and some properties that will be used later. In Sec. 3 we give an overview of different equivalence transformations for APH distributions. The transformations in Sec. 3.1 and 3.2 aim at increasing the number of exit states of an APH and can be applied to prepare an APH distribution for a subsequent expansion into a MAP. The transformation in Sec. 3.3 is used for reducing the size of an APH representation. The paper ends with conclusions.

2 Phase Type distributions

A Phase type (PH) distribution of order n can be described by a CTMC with n transient states $(s_1, ..., s_n)$ and 1 absorbing state (s_{n+1}). The vector $\pi = (\pi(1), \pi(1), ..., \pi(n))$ describes the initial probabilities of the transient states. The initial probability of the absorbing state $\pi(n+1) = 1 - \pi\mathbf{e}^T$, where $\mathbf{e}$ is the unit row vector and $\mathbf{e}^T$ is its transposed, is usually assumed to be 0. Matrix $\mathbf{D}_0$ describes the transition rates between the transient states in the off-diagonal elements and has the negative sum of the transition rate of the ith state in the diagonal, i.e. $\mathbf{D}_0$ has the following properties:

- $\mathbf{D}_0(i,j) \geq 0$ for $i \neq j$
- $\mathbf{D}_0(i,i) \leq -\sum_{j=1, i\neq j}^{n} \mathbf{D}_0(i,j)$

The transition rates to the exit state are given by $-\mathbf{D}_0\mathbf{e}^T$. If $\mathbf{D}_0(i,i) < -\sum_{j=1, i\neq j}^{n} \mathbf{D}_0(i,j)$ state i is an exit state. If $\pi(i) > 0$ state i is an entry state. The distribution of the time to absorption in the CTMC defines the PH distribution and we call $(\pi, \mathbf{D}_0)$ a representation of the PH distribution. The cumulative distribution function of a PH distribution is given by

$$F(x) = 1 - \pi exp(\mathbf{D}_0 x)\mathbf{e}^T \text{ for } x \geq 0 \tag{1}$$

We will use the following notations for PH distributions: $\lambda_i = -\mathbf{D}_0(i,i)$ is the transition rate of the ith phase and $q(i,j)$ is the transition rate from phase i to j, i.e. $q(i,j) = \mathbf{D}_0(i,j), i \neq j$ and $q(i,i) = 0$. For transitions to the absorbing state we have $q(i, n+1) = \lambda_i - \sum_{j=1}^{n} q(i,j)$.

A subclass of PH distributions are acyclic Phase type (APH) distributions that have an upper triangular matrix $\mathbf{D}_0$, which implies that $q(i,j) = 0$ for $i > j$. APH distributions are interesting for fitting approaches because canonical forms exist for them [2]. Every APH distribution can be transformed into bidiagonal form with only $2n-1$ parameters (assuming that $\pi(n+1) = 0$), eliminating the redundancy in matrix $\mathbf{D}_0$. We will denote a bidiagonal matrix $\mathbf{D}_0$ as $Bi(\lambda_1, \lambda_2, \cdots, \lambda_n)$, i.e.

$$Bi(\lambda_1, \lambda_2, \cdots, \lambda_n) = \begin{bmatrix} -\lambda_1 & \lambda_1 & 0 & \cdots & 0 \\ 0 & -\lambda_2 & \lambda_2 & \cdots & 0 \\ 0 & 0 & -\lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -\lambda_n \end{bmatrix} \tag{2}$$

The canonical representation will be summarized in more detail in Sec. 2.1.

Markovian Arrival Processes (MAPs) are a generalization of PH distributions that introduce autocorrelations. Usually they are defined by the two $n \times n$ matrices $\mathbf{D}_0$ and $\mathbf{D}_1$ with the following properties:

- $\mathbf{D}_0(i,j) \geq 0$ for $i \neq j$
- $\mathbf{D}_0(i,i) \leq -\sum_{j=1,i\neq j}^{n} \mathbf{D}_0(i,j)$
- $\mathbf{D}_1(i,j) \geq 0$ and $\mathbf{D}_0\mathbf{e}^T = -\mathbf{D}_1\mathbf{e}^T$
- $\mathbf{D}_0$ is nonsingular and $\mathbf{D}_0 + \mathbf{D}_1$ is an irreducible generator matrix

Transitions in $\mathbf{D}_0$ are silent, while transitions in $\mathbf{D}_1$ generate an event. Every MAP describes an embedded PH distribution characterized by the pair $(\pi, \mathbf{D}_0)$ where the stationary distribution at arrival instants π is given by $\pi\mathbf{P} = \pi$, $\mathbf{P} = -\mathbf{D}_0^{-1}\mathbf{D}_1$ and $\pi\mathbf{e}^T = 1.0$. Moreover every PH distribution can be expanded into an equivalent MAP by defining $\mathbf{D}_1 = (-\mathbf{D}_0\mathbf{e}^T)\pi$. The resulting MAP permits no autocorrelation since the initial probability distribution after an arrival is independent of the state from where the arrival occurred. However, by modifying matrix $\mathbf{P}$, it is possible to keep the distribution and introduce autocorrelation. We come back to this point in subsection 2.2.

2.1 Canonical Representation of APH Distributions

In [2] an algorithm is presented that transforms any APH representation into canonical form. The structure of the canonical representation of APH distributions with n phases is shown in Fig. 1. The canonical representation

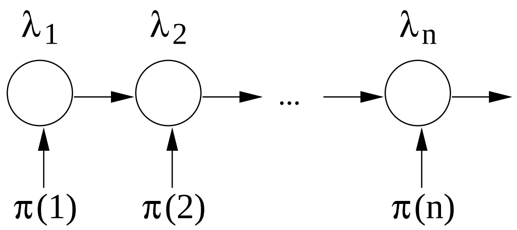

Figure 1: Canonical representation of APH distributions.

has a bidiagonal matrix $\mathbf{D}_0$ (cf. Eq. 2). It has 1 exit and up to n entry states. Furthermore, for the transition rates $\lambda_n \geq \lambda_{n-1} \geq \ldots \geq \lambda_1$ holds. The transformation of an APH representation into the canonical form is based on equivalent representations of the exponential distribution as shown in Fig. 2 and the fact that an APH can be represented by a set of elementary series as shown in Fig. 3. Each elementary series has a probability proportional to the product of the transition rates along the corresponding path and to the initial probability of the first state of the path [2]. Using the relation from Eq. 2 an elementary series containing a state with rate λ can be substituted by a mixture of two series, one containing a state with rate λ and one containing states with the rates λ and $\mu \geq \lambda$. Repeated application of this substitution

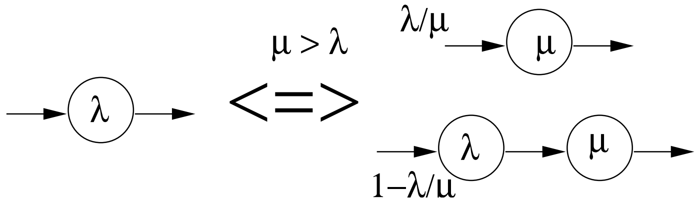

Figure 2: Equivalent representations of an exponential distribution.

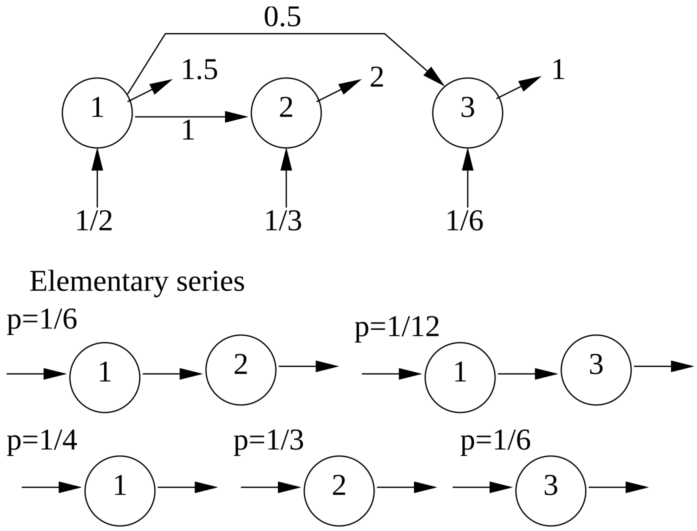

Figure 3: Example for an APH distribution an its elementary series.

results in a mixture of basic series, where a basic series is defined as $BS_i = (\lambda_n, \lambda_{n-1}, \cdots, \lambda_i)$. Together with appropriate initial probabilities this finally yields the canonical representation as shown in Fig. 1.

2.2 Expansion of APH Distributions into MAPs

Assume that an APH distribution $(\pi, \mathbf{D}_0)$ should be expanded into a MAP. This means that we have to find a matrix $\mathbf{D}_1$ such that $\pi \mathbf{D}_0^{-1} \mathbf{D}_1 = \pi$ and $-\mathbf{D}_0 \mathbf{e}^T = \mathbf{D}_1 \mathbf{e}^T$. One then may choose a matrix $\mathbf{D}_1$ that observes the preconditions and approximate some quantities that capture the autocorrelation structure of the process. E.g., the joint moments are given by

$$E(X^i, X^j) = i!j! \mathbf{D}_0^{-i} \left(\mathbf{D}_0^{-1} \mathbf{D}_1 \right) \mathbf{D}_0^{-j} \mathbf{e}^T \ . \tag{3}$$

The fitting approach proposed in [1] then tries to find a matrix $\mathbf{D}_1$ that observes the preconditions and approximates some joint moments as good as possible which results in a non negative lest squares problem.

However, the solution space depends on the initial representation $(\pi, \mathbf{D}_0)$ and since this representation is non-unique, different representation may result in different MAPs. If state i is not an exit state, then the corresponding row in $\mathbf{D}_1$ is zero and, similarly, if state i is not an entry state, then the corresponding column of $\mathbf{D}_1$ is zero. Since the row sums of $\mathbf{D}_1$ are determined by the row sums of $\mathbf{D}_0$, there are at most number of entry states minus one degrees of freedom to select the entries in one row of $\mathbf{D}_1$. However, this implies also that representations with one exit or one entry state allow no flexibility such that the resulting MAP is always identical to the APH distribution. Thus, equivalence transformations are necessary to increase the number of entry and exit states.

3 Equivalence Transformations

In the following we present three equivalence transformations for APH distributions. All three transformations use an APH in canonical form as input. The first and second transformations modify the APH representation such that it has more exit states and can be used as an intermediate step when expanding an APH into a MAP. The third transformation reduces the size of the APH representation by removing unnecessary states.

3.1 Increasing the Number of Exit States (1)

The transformation presented in this section has been proposed in [1] to transform the canonical form to a general APH representation with more than one exit state. As already mentioned, this is important when expanding an APH into a MAP, since an increase in the number of exit states allows for a higher flexibility for Matrix $\mathbf{D}_1$, i.e. the possible range for the joint moments is increased. The key idea of this transformation is to invert the steps described in [2] that lead to the canonical form. However, the transformation steps can be applied to APHs that are not in canonical form as well. The inversion is not unique and unfortunately, it is not clear yet how to find the most flexible APH representation for a distribution. In the following we will summarize the transformation steps and present proofs that have been omitted in [1] to show that the distribution is not altered by the approach.

The transformation consists of a sequence of steps and in each step two states i and j, $i < j$, that are connected by a transition, i.e. $q(i, j) > 0$, are chosen. For a transformation step only the incoming and outgoing transitions of i, the incoming transitions of j, and the probabilities $\pi(i)$ and $\pi(j)$ are considered. Hence, the PH distribution remains acyclic and the transition rates $\lambda_1, \ldots, \lambda_n$ are not changed by the transformation, i.e. $\lambda_i \leq \lambda_j$ holds for $i < j$.

Let i and j, $i < j$, be two states of the APH with transition rate $q(i,j) > 0$ and initial probabilities $\pi(i)$ and $\pi(j)$ and define $\delta \leq \delta^*$ where

$$\delta^* = \quad \min\left(\pi(j), \tfrac{\pi(i)q(i,j)}{\lambda_j - \lambda_i}, \min_{k<i, q(k,i)>0}\left(\pi(i)\tfrac{q(k,j)}{q(k,i)}\right)\right) \quad . \tag{4}$$

for $\lambda_j > \lambda_i$. If $\lambda_i = \lambda_j$, then the second term in the minimum of Eq. (4) becomes ∞ and does not count such that the minimum is computed according to the remaining two conditions.

If $\delta^* > 0$, we can choose $\delta > 0$ and compute new transition rates $q'(.,.)$ and initial probabilities $\pi'(.)$ for a different representation of the same distribution according to

$$\pi'(k) = \begin{cases} \pi(i) + \delta & \text{for } k = i \\ \pi(j) - \delta & \text{for } k = j \\ \pi(k) & \text{otherwise} \end{cases} \tag{5}$$

$$q'(k,l) = \begin{cases} q(i,j)\frac{\pi(i)}{\pi(i)+\delta} - \frac{(\lambda_j-\lambda_i)\delta}{\pi(i)+\delta} & \text{for } k = i \text{ and } l = j \\ q(i,l)\frac{\pi(i)}{\pi(i)+\delta} + q(j,l)\frac{\delta}{\pi(i)+\delta} & \text{for } k = i \text{ and } l \neq j \\ q(k,i)\frac{\pi(i)+\delta}{\pi(i)} & \text{for } k < i \text{ and } l = i \\ q(k,j) - q(k,i)\frac{\delta}{\pi(i)} & \text{for } k < j \text{ and } l = j \\ q(k,l) & \text{otherwise} \end{cases} \tag{6}$$

For the transformation the following properties hold:

Theorem 1. *If Eq. (5) and Eq. (6) are applied to an APH representation, then the resulting representation is still an APH representation and describes the same distribution.*

Proof. We have to show that the transformations made in Eqs. (5) and (6) do not alter the distribution. This will be done by showing that Laplace transform remains the same.
According to [2] (cf. Sec. 2.1) an APH distribution can be represented by a set of elementary series. An elementary series contains a subset of the phases of the APH distribution. Let $E = (\lambda_{i_1}, \ldots, \lambda_{i_m})$ be one series and denote by $p(E)$ the probability of E. For notational convenience we define $i_{m+1} = n+1$. The Laplace transform of E is given by

$$f_E^*(s) = \prod_{k=1}^{m} \frac{\lambda_{i_k}}{\lambda_{i_k} + s}$$

and

$$p(E) = \pi(i_1) \prod_{k=1}^{m} \frac{q(i_k, i_{k+1})}{\lambda_{i_k}} \; .$$

We denote by $\mathcal{E}$ the set of elementary series that are defined by the APH distribution. The Laplace transform of the APH distribution is given by

$$f^*_{APH}(s) = \prod_{E \in \mathcal{E}} p(E) f^*_E(s) \ .$$

Fig. 3 shows an example APH distribution and its elementary series. Outgoing arcs without a destination describe exit rates and incoming arrows describe the initial probabilities.

The equivalence shown in Fig. 2 can be easily seen in the Laplace domain since

$$\frac{\lambda}{\lambda + s} = \frac{\lambda}{\mu + s} + \frac{\lambda(\mu - \lambda)}{(\lambda + s)(\mu + s)} \tag{7}$$

for $\mu > \lambda$. It follows easily that an elementary series $E = (\lambda_1, \ldots, \lambda_r, \lambda, \lambda_{r+1}, \ldots, \lambda_{r+s})$ with probability p can be substituted by two series $E_1 = (\lambda_1, \ldots, \lambda_r, \mu, \lambda_{r+1}, \ldots, \lambda_{r+s})$ with probability $p\lambda/\mu$ and $E_2 = (\lambda_1, \ldots, \lambda_r, \lambda, \mu, \lambda_{r+1}, \ldots, \lambda_{r+s})$ with probability $p(1 - \lambda/\mu)$. The equivalence follows by comparison of the Laplace transforms.

We now show that the transformations used in theorem 1 are all based on the above equivalence. Only elementary series that contain λ_i or start with λ_j are affected by the transformations. The remaining series remain unmodified and need not be considered here. We have to distinguish the following cases.

1. Series that do not contain i and do not start in j.
2. Series that start in i or j.
3. Series that start in $k < i$ and contain i or j.

The first series are not affected by the transformation and need not be considered.

For the second set of series we consider E_s as some elementary series starting in j. Then there exists some series E_r which starts in i and results from E_s by adding λ_i at the beginning. Furthermore, there exists some series E_t which is identical to E_s when the first state j is substituted by i. Let k be the second state of E_s.

Let p_s be the probability of series E_s, then

$$p_r = p_s \frac{\pi(i)q(i,j)}{\pi(j)\lambda_i} \text{ and } p_t = p_s \frac{\pi(i)q(i,k)\lambda_j}{\pi(j)\lambda_i q(j,k)}$$

Observe that p_t is zero if $q(i,k) = 0$, i.e. E_t is not available before the transformation. Furthermore, $p_r = 0$ if $q(i,j) = 0$, i.e., a transition from i to j does not exist. Now consider the difference between the probabilities before and after the transformation step.

$$\begin{aligned}
p_s - p'_s &= p_s \left(1 - \frac{\pi(j)-\delta}{\pi(j)}\right) = p_s \frac{\delta}{\pi(j)} \\
p_r - p'_r &= p_s \left(\frac{\pi(i)q(i,j)}{\pi(j)\lambda_i} - \frac{(\pi(j)-\delta)(q(i,j)\pi(i)-\delta(\lambda_j-\lambda_i))}{\pi(j)(\pi(j)-\delta)\lambda_i}\right) = p_s \frac{(\lambda_j-\lambda_i)\delta}{\lambda_i \pi(j)} \\
p'_t - p_t &= p_s \left(\frac{\lambda_j(q(i,k)\pi(i)+\delta q(j,k))}{\pi(j)q(j,k)\lambda_i} - \frac{\pi(i)q(i,k)\lambda_j}{\pi(j)q(j,k)\lambda_i}\right) = p_s \left(\frac{\lambda_j \delta}{\lambda_i \pi(j)}\right)
\end{aligned}$$

Observe that $(\lambda_i/\lambda_j)^{-1}(p_s - p'_s) = (1-\lambda_i/\lambda_j)^{-1}(p_r - p'_r)$ and $(p_s - p'_s) + (p_r - p'_r) = (p'_t - p_t)$. Since E_r, E_s and E_t differ only at the beginning and $\lambda_j > \lambda_i$, we can substitute $(p_s - p'_s)$ of E_s and $(p_r - p'_r)$ of E_r by $(p'_t - p_t)$ of E_t (cf. Eq. (7)) which is done by the transformation. The transformation implies $\delta \leq \pi(j)$ and $\pi(i)q(i,j)/(\lambda_j - \lambda_i) > \delta$ otherwise negative values would result from the transformation.

For the third set of series the proof is similar to the previous proof. Let E_s be some series that starts in some state $m < i$ and contains the sequence $k \to j \to l$. Let p_s be the probability of E_s. Let E_r be the sequence that results from E_s by adding state i between k and j such that it contains the sequence $k \to i \to j \to l$. Finally, E_t is the series that results from E_s by substituting j by i, i.e. the series contains the sequence $k \to i \to l$. We have

$$p_r = p_s \frac{q(k,i)q(i,j)}{q(k,j)\lambda_i} \text{ and } p_t = p_s \frac{q(k,i)q(i,l)\lambda_j}{q(k,j)\lambda_i q(j,l)} \ .$$

Again consider the probabilities for the different series before and after the transformation.

$$\begin{aligned} p_s - p'_s &= p_s \tfrac{q(k,i)\delta}{q(k,j)\pi(i)} \\ p_r - p'_r &= p_s \tfrac{q(k,i)\delta(\lambda_j - \lambda_i)}{\pi(i)\lambda_i q(k,j)} \\ p'_t - p_t &= p_s \tfrac{q(k,i)\delta\lambda_j}{\pi(i)q(k,j)\lambda_i} \end{aligned}$$

Now we have $(\lambda_i/\lambda_j)^{-1}(p_s - p'_s) = -(1-\lambda_i/\lambda_j)^{-1}(p_r - p'_r)$and $(p'_s - p_s) - (p'_r - p_r) = (p'_t - p_t)$ such that E_r is substituted by E_s and E_t in the right proportion. □

The following theorem implies that if j is an exit state, then also i becomes an exit state:

Theorem 2. *If Eq. (5) and (6) are applied to an APH representation, then* $\sum_{l=k+1}^{n} q'(k,l) = \sum_{l=k+1}^{n} q(k,l)$ *for* $k \neq i$ *and*

$$\sum_{l=i+1}^{n} q'(i,l) = \sum_{l=i+1}^{n} q(i,l) + \frac{\delta}{\pi(i)+\delta}\left(q(i,n+1) - q(j,n+1)\right)$$

where $q(k,n+1) = \lambda_k - \sum_{l=k+1}^{n} q(k,l)$ $(k = i,j)$.

Proof. First consider some state $k \neq i$. If $k > i$, then no transition rates are changed such that the sum of outgoing transition rates remains the same. For $k < i$ transition rates $q'(k,l) = q(k,l)$ for $l \neq i,j$ such that we have to show $q'(k,i) + q'(k,j) = q(k,i) + q(k,j)$. We have

$$q'(k,i) + q'(k,j) = q(k,i)\frac{\pi(i)+\delta}{\pi(i)} + q(k,j) - q(k,i)\frac{\delta}{\pi(i)} = q(k,i) + q(k,j) \ .$$

For state i we obtain

$$\begin{aligned}
\sum_{k=i+1}^{n} q'(i,k) &= \tfrac{\pi(i)}{\pi(i)+\delta} \sum_{k=i+1}^{n} q(i,k) + \tfrac{\delta}{\pi(i)+\delta} \left(\sum_{k=i+1, k\neq j}^{n} q(j,k) - \lambda_j + \lambda_i \right) \\
&= \sum_{k=i+1}^{n} q(i,k) + \tfrac{\delta}{\pi(i)+\delta} \left(q(i,n+1) - q(j,n+1) \right)
\end{aligned}$$

where $q(k,n+1) = \lambda_k - \sum_{l=k+1}^{n} q(k,l)$. □

Under certain conditions repeated application of the transformation results in an APH representation with n exit states:

Theorem 3. *If we apply the transformation rules from Eqs. (5) and (6) with $\delta^* > \delta > 0$ consecutively to states $i = 1, 2, \ldots, n$ and $j = i+1, \ldots, n$ of an APH representation in canonical form with $\pi(i) > 0$ for all $i = 1, \ldots, n$, we obtain an APH representation of the same distribution where all states are exit states.*

Proof. Transitions in the matrix are transformed in the order $(1,2), \ldots, (1,n)$, $(2,3), \ldots, (n-1,n)$. We assume that we choose in each step δ such that $\delta^* > \delta > 0$. This implies that $\delta^* > 0$ which will be first assumed and proved later. Since the APH is in canonical form we have initially $q(i,i+1) > 0$ for $(i = 1, \ldots, n-1)$ and by assumption $\pi(i) > 0$ holds for $i = 1, \ldots, n$.

Now assume that transition (i,j) with $q(i,j)$ is handled. The transformations according to Eq. (6) result in non zero transition rates $q'(i,k)$ for all k with $q(j,k) > 0$. Since we choose $\delta < \delta^*$ all non zero transition rates before the transformation remain non zero after the transformation.

Thus, starting with $(1,2)$ the transformation generates a non-zero transition rate $(1,3)$ since $q(2,3) > 0$, $\pi(1) > 0$ and $\pi(3) > 0$. With similar arguments transition rates $q'(1,4), \ldots, q'(1,n) > 0$ are generated. The same argument can then be repeated for the rows 2 through $n-1$.

It remains to show that $\delta^* > 0$ holds in every step. By choosing $\delta < \delta^*$ $\pi(k) > 0 \Rightarrow \pi'(k) > 0$ and $q(k,l) > 0 \Rightarrow q'(k,l) > 0$. This implies $\delta^* > 0$ according to Eq. (4). □

Theorem 3 implies that in the canonical representation all states are entry states. For representations where this is not the case, the transformation may generate a representation with less exit states. An example for this is the Erlang n distribution, where no transformation is possible. In [1] this problem is solved by a slight modification of the APH fitting approach, that generated the canonical representation.

As one can see from the proofs the choice of δ in every step is important, since it has impact on the resulting APH representation and consequently, on the range of the joint moments that can be reached when expanding the

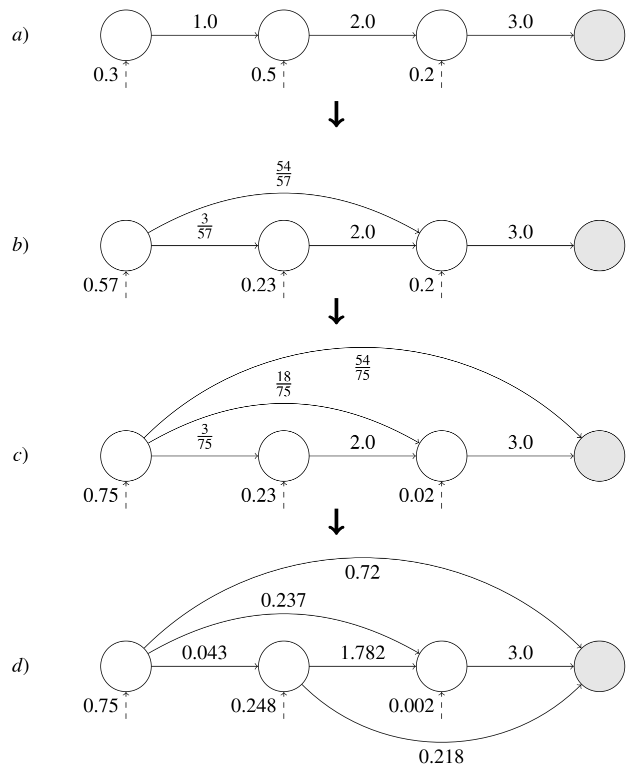

Figure 4: Steps of the APH transformation from the canonical form into a new representation with increased number of exit states

APH distribution into a MAP. [1] reports that $\delta = 0.9\delta^*$ is a good choice.

In the following we will give an example for the transformation: Consider the upper APH in canonical form of Fig. 4 with $\pi = (0.3, 0.5, 0.2)$ and $\mathbf{D}_0 = Bi(1.0, 2.0, 3.0)$. In the first transformation step states $i = 1$ and $j = 2$ with $q(1, 2) = 1.0$ are selected. From Eq. (4) we compute $\delta^* = 0.3$ and set $\delta = 0.9\delta^* = 0.27$. Application of Eq. (5) yields the new initial probability vector $\pi' = (0.57, 0.23, 0.2)$. Using Eq. (6) we get a new transition rate $q'(1, 2)$ and generate a new non-zero transition rate $q'(1, 3)$. The transition rate from state 2 to 3 is unmodified in this step. The resulting APH representation is shown in Fig. 4 b). We set $\pi = \pi'$ and $q(., .) = q'(., .)$ and continue with the second transformation step, that treats states $i = 1$ and $j = 3$ since we generated a transition rate $q(1, 3) > 0$ in the previous step. The step modifies the initial probabilities of states 1 and 3. Additionally, since state 3 is an exit state, state 1 becomes an exit state as well. The resulting representation is shown in Fig. 4 c). In the last step states 2 and 3 are handled and state 2 is transformed into an exit state. The final APH representation is shown in

Fig. 4 d).

3.2 Increasing the Number of Exit States (2)

The transformation presented in the following has been proposed in [4] to increase the number of exit states of an APH in canonical form for a later expansion into a MAP.
Let $(\pi, \mathbf{D}_0)$ be an APH representation defined by

$$\pi = \begin{bmatrix} p_n & p_{n-1} & \cdots & p_1 \end{bmatrix}, \qquad \mathbf{D}_0 = \begin{bmatrix} -\lambda_n & \lambda_n & & & \\ & -\lambda_{n-1} & \lambda_{n-1} & & \\ & & \ddots & \ddots & \\ & & & -\lambda_2 & \lambda_2 \\ & & & & -\lambda_1 \end{bmatrix}$$

Then $(\pi, \mathbf{D}_0)$ can be transformed into an equivalent representation $(\pi', \mathbf{D}_0')$ with

$$\pi' = \begin{bmatrix} x_n & x_{n-1} & \cdots & x_1 \end{bmatrix},$$
$$\mathbf{D}_0' = \begin{bmatrix} -\lambda_n & (1-a_n)\lambda_n & & & \\ & -\lambda_{n-1} & (1-a_{n-1})\lambda_{n-1} & & \\ & & \ddots & \ddots & \\ & & & -\lambda_2 & (1-a_2)\lambda_2 \\ & & & & -\lambda_1 \end{bmatrix}$$

that describes the same distribution.
The new initial probabilities $\pi' = \begin{bmatrix} x_n & x_{n-1} & \cdots & x_1 \end{bmatrix}$ can be computed by solving the following system of equations:

$$p_k = \sum_{n=k}^{m} x_n \sum_{i=0}^{k-1} \prod_{j=0}^{i-1} (1 - a_{n-j}) a_{n-i} d_n(i, 0, k)$$

where $d_n(i, l, k)$ is defined in a recursive manner by

$$\begin{aligned} d_n(i,l,k) &= \frac{\lambda_{n-i}}{\lambda_{l+1}} d_n(i-1, l+1, k) + \left(1 - \frac{\lambda_{n-i}}{\lambda_{l+1}}\right) d_n(i, l+1, k) \\ d_n(i, n-i, k) &= \begin{cases} 1 & k = n, \\ 0 & k \neq n \end{cases} \\ d_n(0,l,k) &= \begin{cases} \frac{\lambda_n}{\lambda_{l+1}} & k = l+1, \\ \left(1 - \frac{\lambda_n}{\lambda_{l+1}}\right) d_n(0, l+1, k) & k \neq l+1 \end{cases} \end{aligned}$$

Fig. 5 visualizes the transformation. The upper Markov chain shows the original APH $(\pi, \mathbf{D}_0)$ in canonical form and the lower Markov chain shows the result of the transformation with representation $(\pi', \mathbf{D}_0')$.

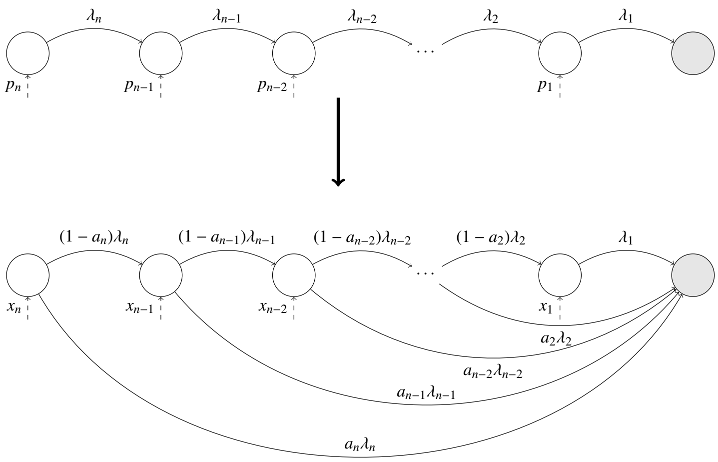

Figure 5: APH transformation from the canonical form into a new representation with increased number of exit states

3.3 Reduction of Acyclic Phase Type Representations

In [9] another equivalence transformation for APHs is presented. In contrast to the transformations summarized in Sec. 3.1 and 3.2 the approach from [9] aims at reducing the size of the matrix representation of an APH by removing unnecessary states.

For the transformation it is assumed that the APH is given in the bidiagonal canonical form as described in Sec. 2.1, i.e. it has representation $(\pi, Bi(\lambda_1, \lambda_2, \cdots, \lambda_n))$. Then the Laplace-Stieltjes transform (LST) of the APH can be written as

$$\begin{aligned} \overline{f}(s) &= \frac{\pi(1)}{L(\lambda_1)\cdots L(\lambda_n)} + \frac{\pi(2)}{L(\lambda_2)\cdots L(\lambda_n)} + \frac{\pi(n)}{L(\lambda_n)} \\ &= \frac{\pi(1) + \pi(2)L(\lambda_1) + \cdots + \pi(n)L(\lambda_1)L(\lambda_2)\cdots L(\lambda_{n-1})}{L(\lambda_1)L(\lambda_2)\cdots L(\lambda_n)} \end{aligned} \tag{8}$$

where $L(\lambda) = \frac{s+\lambda}{\lambda}$ is the reciprocal of the LST of an exponential distribution with rate $-\lambda$ called *L-term* in [9]. If both, numerator and denominator of Eq. 8, have a common L-term $L(\lambda_i)$, this term can be removed from the equation and the corresponding state can be removed from representation. If furthermore a new initial probability distribution π' can be found, i.e. π' is a sub-stochastic vector, then the two representations describe the same distribution:

$$PH(\pi, Bi(\lambda_1, \lambda_2, \cdots, \lambda_n)) = PH(\pi', Bi(\lambda_1, \lambda_2, \cdots, \lambda_{i-1}, \lambda_{i+1}, \cdots, \lambda_n)) \tag{9}$$

Thus, for reducing the number of states a L-term that divides numerator and denominator of Eq. 8 an a new valid initial probability have to be determined. An applicable L-term can be found by checking if

$$R(s) = \pi(1) + \pi(2)L(\lambda_1) + \cdots + \pi(i)L(\lambda_1)L(\lambda_2)\cdots L(\lambda_{i-1})$$

is divisible by $L(\lambda_i)$ which holds if $R(-\lambda_i) = 0$.
The new initial probability vector π' can be determined using the cdf of the APH. Since the two representations describe the same distribution, their cdf as given in Eq. 1 must be the same. To avoid computations of the matrix exponential in the expression of the cdf the i-th derivative of the cdf is evaluated for $x = 0$ for the determination of vector π' resulting in the following system of equations:

$$\pi Bi(\lambda_1, \lambda_2, \cdots, \lambda_n)^i \mathbf{e}^T = \pi' Bi(\lambda_1, \lambda_2, \cdots, \lambda_{i-1}, \lambda_{i+1}, \cdots, \lambda_n)^i \mathbf{e}^T$$

for $i = 0, \cdots, n-2$. The $n-1$ equations can be used to determine the $n-1$ components of vector π'. Additionally one has to verify that the resulting vector is indeed a sub-stochastic vector, i.e. $0 \leq \pi'(i) \leq 1$ for $i = 1, \cdots, n-1$ and $\pi'\mathbf{e}^T \leq 1$.
[9] presents an algorithm for the reduction of APHs that checks for all states of an APH if the corresponding L-term divides numerator and denominator of Eq. 8 and if a new valid initial probability distribution can be found and deletes the state if both conditions hold. However, [9] also reports about cases where a reduction with the described approach is not possible, although a smaller representation exists, because the algorithms ignores the interplay of total outgoing rates and the initial probability distribution.

4 Conclusions

In this report we summarized different approaches to perform equivalence transformations of acyclic phase type distributions. These transformations allow us to search for a representation which is more amenable for subsequent processing steps like the expansion of the distribution into a stochastic process that captures the autocorrelation. Since the transformations usually depend on free parameters, they provide some flexibility but it is usually not clear yet how to set the parameters to come to the best representation for subsequent processing.

References

[1] P. Buchholz and J. Kriege. A heuristic approach for fitting MAPs to moments and joint moments. In *Proc. of 6th International Conference on Quantitative Evaluation of SysTems (QEST09)*. IEEE, 2009.

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

[3] G. Horváth and M. Telek. On the canonical representation of phase type distributions. *Perform. Eval.*, 66(8):396–409, 2009.

[4] G. Horvath, M. Telek, and P. Buchholz. A MAP fitting approach with independent approximation of the inter-arrival time distribution and the lag-correlation. In *Proc. of 2nd Int. Conf. on the Quantitative Analysis of Systems*. IEEE CS Press, 2005.

[5] M. F. Neuts. A versatile Markovian point process. *Journal of Applied Probability*, 16:764–779, 1979.

[6] C. A. O'Cinneide. On non-uniqueness of represnetations of phase-type distributions. *Stochastic Models*, 5:247–259, 1989.

[7] C. A. O'Cinneide. Characterization of phase-type distributions. *Stochastic Models*, 6:1–57, 1990.

[8] C. A. O'Cinneide. Triangular order of triangular phase-type distributions. *Stochastic Models*, 9:507–529, 1993.

[9] R. Pulungan and H. Hermanns. Effective minimization of acyclic phase-type representations. In *ASMTA '08: Proceedings of the 15th international conference on Analytical and Stochastic Modeling Techniques and Applications*, pages 128–143, Berlin, Heidelberg, 2008. Springer-Verlag.

[10] M. Telek and G. Horváth. A minimal representation of Markov arrival processes and a moments matching method. *Perform. Eval.*, 64(9-12):1153–1168, 2007.