On Minimal Representations of Rational Arrival Processes

Peter Buchholz,
Informatik IV, TU Dortmund, D-44221 Dortmund, Germany
Miklós Telek
Department of Telecommunications, Technical University of Budapest, H-1521 Budapest, Hungary

September 15, 2011

Abstract

Rational Arrival Processes (RAPs) form a general class of stochastic processes which include Markovian Arrival Processes (MAPs) as a subclass. In this paper we study RAPs and their representations of different sizes. We show some transformation methods between different representations and present conditions to evaluate the size of the minimal representation. By using some analogous results from linear systems theory, a minimization approach is defined which allows one to transform a RAP (from a redundant high dimension) into one of its minimal representations. An algorithm for computing a minimal representation is also given. Furthermore, we extend the approach to RAPs with batch arrivals (BRAPs) and to RAPs with arrivals of different customer types (MRAPs).

1 Introduction

Markovian Arrival Processes (MAPs) [18, 20] are widely used in stochastic modeling. The advantage of MAPs is their nice stochastic interpretation as a continuous time Markov chain (CTMC) with some marked transitions and the possibility to use them as arrival or service processes in models that are analytically tractable. A superclass of MAPs is the class of rational arrival processes (RAPs) [3, 17]. RAPs, in contrast to MAPs, lack an intuitive stochastic description but the class is strictly larger than the class of MAPs and they may as well be used in analytically tractable models like quasi-birth-death processes (QBDs) as shown recently [4]. Nevertheless, in contrast to MAPs, RAPs have only rarely been considered in the literature on stochastic modeling.

In this paper, we present an approach to determine the equivalence of different representations of RAPs. In particular, we focus on the case when the size of the representations are different. This viewpoint allows us to apply methods from linear systems theory [9] to RAPs which to the best of our knowledge has been done in the literature only for phase type distributions [19, 7]. The contribution of the paper is summarized as follows. We present necessary and sufficient conditions for the size of minimal representations for RAPs. We introduce a numerically stable algorithm for computing a minimal representation and apply the approach to the computation of the output process of a BMAP/MAP/1 queue. Additionally, we study the class of RAPs generating arrivals of different type. They are commonly referred to as Marked RAPs, MRAPs [4]. The class of RAPs with batch arrivals, BRAPs, is isomorphic to the class of MRAPs. For these extended model classes the minimization approach is also adopted.

The paper is structured as follows. In the next section we introduce RAPs and some basic properties of these processes. Afterwards, in Sections 3 and 4, the equivalence of RAPs and minimal equivalent representations are defined. Then, in Section 5, an algorithm for minimizing the size of a RAP representation is introduced. Afterwards, we consider the approximation of the output process of a BMAP/MAP/1 queue. Finally, we introduce the extensions of the approach to BRAPs and MRAPs. The paper ends with the conclusion.

2 Rational Arrival Processes

A matrix exponential (ME) distribution is a distribution whose density function is a matrix exponential function of the variable and consequently its Laplace transform is a rational function of the transform variable [2]. Let ν be a row vector of size m and $\mathbf{D}_0$ be a $m \times m$ matrix. Then $(\nu, \mathbf{D}_0)$ defines a matrix exponential (ME) distribution, if and only if

$$F_{(\nu,\mathbf{D}_0)}(t) = 1 - \nu\, e^{\mathbf{D}_0\; t}\;\, \mathbb{1} \tag{1}$$

is a valid distribution function where $\mathbb{1}$ is the unit column vector of appropriate size. The row vector before and the column vector after the matrix exponential term are referred to as initial and closing vectors, respectively. In the literature (e.g., [2]) one can also find representations where the closing vector differs from the unit vector and is explicitly part of the representation. However, it is always to possible to transform the representation in an equivalent representation with a unit closing vector [15, 19] which is used here. In this paper we further assume that $F(0) = 0$, that is $\nu\;\, \mathbb{1} = 1$. If $F_{(\nu,\mathbf{D}_0)}(t)$ is a valid distribution function, then the real part of the eigenvalues of matrix $\mathbf{D}_0$ are negative, consequently the matrix is non-singular, and there is a real eigenvalue with maximal real part [15].

ME distributions have been used in stochastic modeling for some time and recently it has been shown that they may be used instead of PH distributions in stochastic models such as queuing networks or stochastic Petri nets, and numerical algorithms for computing the stationary or transient distributions may still be used to compute performance measures [1, 4, 6].

Different representations may exist that describe the same distribution. We denote two representations $(\nu, \mathbf{D}_0)$ and $(\phi, \mathbf{C}_0)$ as equivalent if and only if $F_{(\nu,\mathbf{D}_0)}(t) \equiv F_{(\phi,\mathbf{C}_0)}(t)$. This is equivalent to the condition that $(\nu, \mathbf{D}_0)$ of size m and $(\phi, \mathbf{C}_0)$ of size n are equivalent if and only if

$$\nu\,(\mathbf{D}_0)^k\;\, \mathbb{1}_m = \phi\,(\mathbf{C}_0)^k\;\, \mathbb{1}_n \tag{2}$$

for all $k \leq 0$ or all $k \geq 0$ [5, 13], where we intentionally indicate the size of the unit column vectors.

A rational arrival process (RAP) is a point process whose joint density of the consecutive interarrival times is a matrix exponential function of the variables and the double transform of the number of arrivals is a rational function of the transform variables. In [17] series of correlated matrix exponentially distributed arrivals are define, whereas [2] defines the class of RAPs. We use a slightly different definition of a RAP than [2]. Our definition is related to the commonly used definition of MAPs in contemporary literature [16, 20]. A pair of matrices, $(\mathbf{D}_0, \mathbf{D}_1)$, defines a RAP with representation $(\mathbf{D}_0, \mathbf{D}_1)$, denoted as RAP$(\mathbf{D}_0, \mathbf{D}_1)$, if and only if

$$f_{(\mathbf{D}_0,\mathbf{D}_1)}(t_1, \ldots, t_k) = \nu e^{\mathbf{D}_0 t_1} \mathbf{D}_1 e^{\mathbf{D}_0 t_2} \mathbf{D}_1 \ldots e^{\mathbf{D}_0 t_k} \mathbf{D}_1\, \mathbb{1} \tag{3}$$

is non-negative for all $k \geq 1$ and $t_1, t_2, \ldots, t_k \in \mathbb{R}^+$. In this case $f_{(\mathbf{D}_0,\mathbf{D}_1)}(t_1, \ldots, t_k)$ is the joint density function of the interarrival times and the reduced joint moments of the interarrival times are

$$\frac{E(T_1^{i_1} T_2^{i_2} \ldots T_k^{i_k})}{i_1! i_2! \ldots i_k!} = \nu(-\mathbf{D}_0)^{-i_1} \mathbf{P} (-\mathbf{D}_0)^{-i_2} \mathbf{P} \ldots (-\mathbf{D}_0)^{-i_k}\, \mathbb{1}\;, \tag{4}$$

with $i_1, \ldots, i_k \in (\mathbb{N}^+ \cup 0)$, $\mathbf{P} = (-\mathbf{D}_0)^{-1} \mathbf{D}_1$, $\nu \mathbf{P} = \nu$, $\nu\;\, \mathbb{1} = 1$, and the double transform of the number of arrivals in $(0, t)$, $N(t)$, is a rational function of the transform variables

$$\int_0^\infty e^{-st} E(z^{N(t)}) dt = \nu(s\mathbf{I} - \mathbf{D}_0 - z\mathbf{D}_1)^{-1}\, \mathbb{1}\;. \tag{5}$$

Similar to ME distributions there are other definitions of RAPs with different closing vectors, which are equivalent with the one above [15]. RAPs inherit several properties from ME distributions. The real parts of the eigenvalues of matrix $\mathbf{D}_0$ are negative, consequently the matrix is non-singular. There is a real eigenvalue with maximal real part. Additionally, $(\mathbf{D}_0 + \mathbf{D}_1)\;\, \mathbb{1} = \mathbf{0}$ to ensure that (3) is a density function.

Throughout this paper we consider only stationary RAPs and assume that the pair of matrices, $(\mathbf{D}_0, \mathbf{D}_1)$ uniquely defines the process. This requires that the solution of $\nu \mathbf{P} = \nu$, $\nu\;\, \mathbb{1} = 1$, is unique, that is, one is a unique eigenvalue of $\mathbf{P}$ (and, equivalently, zero is a unique eigenvalue of $\mathbf{D}_0 + \mathbf{D}_1$). If the transient analysis of RAPs is considered this restriction can be relaxed, but in that case the process should be defined by the governing matrices and an initial vector. The subsequent analysis identically applies to the transient case, only the stationary initial vector has to be replaced by the transient one.

It should also be noted that a pair of $(\mathbf{D}_0, \mathbf{D}_1)$ matrices does not necessarily represent a valid process, only if the function defined in (3) is non-negative for each valid set of parameters. Unfortunately, this condition is very hard to check.

The $(\mathbf{D}_0, \mathbf{D}_1)$ representation of a given RAP is non-unique. Several pairs of matrices can describe the same process [20]. This naturally raises the question for a minimal representation which will be answered in the following two sections.

3 Equivalent Representations of RAPs

We are interested in different representations of the same RAP and in particular in a representation with minimal size. This is formally defined below.

Definition 1 *Two RAPs, RAP*$(\mathbf{D}_0, \mathbf{D}_1)$ *and RAP*$(\mathbf{C}_0, \mathbf{C}_1)$, *are* equivalent *if and only if all joint density functions of the finite-dimensional distributions are identical (cf., (3)), which is equivalent to the condition that all reduced joint moments are identical (cf., (4)).*

Definition 2 *The* size *of the representation* $(\mathbf{D}_0, \mathbf{D}_1)$ *is the size of the square matrix* $\mathbf{D}_0$.

Definition 3 *A representation* $(\mathbf{D}_0, \mathbf{D}_1)$ *of size* m *is* minimal, *if no other equivalent representation* $(\mathbf{C}_0, \mathbf{C}_1)$ *of size* $n < m$ *exists.*

Let $(\mathbf{D}_0, \mathbf{D}_1)$ be a RAP of size m and $\mathbf{B}$ be a non-singular $m \times m$ matrix such that $\mathbf{B}\,\mathbb{1} = \mathbb{1}$. Then $(\mathbf{C}_0, \mathbf{C}_1)$ of size m with $\mathbf{C}_0 = \mathbf{B}^{-1}\mathbf{D}_0\mathbf{B}$ and $\mathbf{C}_1 = \mathbf{B}^{-1}\mathbf{D}_1\mathbf{B}$ is an equivalent representation resulting from a similarity transformation [20]. By means of similarity transformations infinitely many representations can be computed for a RAP. [20] focuses mainly on the equivalence of RAPs of minimal size. We now consider the existence of equivalent representations of different sizes.

Let $(\mathbf{D}_0, \mathbf{D}_1)$ of size m and $(\mathbf{C}_0, \mathbf{C}_1)$ of size n be two representations, where $m > n$. The following two theorems present sufficient conditions that both representations are equivalent.

Theorem 1 *If there is a matrix* $\mathbf{V} \in \mathbb{R}^{m,n}$ *such that* $\mathbb{1}_m = \mathbf{V}\,\mathbb{1}_n$, $\mathbf{D}_0\mathbf{V} = \mathbf{V}\mathbf{C}_0$, $\mathbf{D}_1\mathbf{V} = \mathbf{V}\mathbf{C}_1$ *and* $\nu\mathbf{V} = \phi$, *where* $\nu\left(-\mathbf{D}_0\right)^{-1}\mathbf{D}_1 = \nu$ *and* $\phi\left(-\mathbf{C}_0\right)^{-1}\mathbf{C}_1 = \phi$, *then* $(\mathbf{D}_0, \mathbf{D}_1)$ *and* $(\mathbf{C}_0, \mathbf{C}_1)$ *are equivalent.*

Proof. We start from the joint density function of representation $(\mathbf{D}_0, \mathbf{D}_1)$ and transform it by first using $\mathbb{1}_m = \mathbf{V}\,\mathbb{1}_n$, then $(\mathbf{D}_0)^i\mathbf{D}_1\mathbf{V} = \mathbf{V}(\mathbf{C}_0)^i\mathbf{C}_1$, and finally $\nu\mathbf{V} = \phi$. Therefore,

$$\begin{aligned} f_{(\mathbf{D}_0,\mathbf{D}_1)}(t_1,\ldots,t_k) &= \nu\left(\prod_{j=1}^{k}\left(\sum_{i_j=0}^{\infty}\frac{(t_j\mathbf{D}_0)^{i_j}}{i_j!}\mathbf{D}_1\right)\mathbb{1}_m\right) = \nu\left(\prod_{j=1}^{k}\left(\sum_{i_j=0}^{\infty}\frac{(t_j\mathbf{D}_0)^{i_j}}{i_j!}\mathbf{D}_1\right)\mathbf{V}\ \mathbb{1}_n\right) \\ &= \nu\mathbf{V}\left(\prod_{j=1}^{k}\left(\sum_{i_j=0}^{\infty}\frac{(t_j\mathbf{C}_0)^{i_j}}{i_j!}\mathbf{C}_1\right)\mathbb{1}_n\right) = \phi\left(\prod_{j=1}^{k}\left(\sum_{i_j=0}^{\infty}\frac{(t_j\mathbf{C}_0)^{i_j}}{i_j!}\mathbf{C}_1\right)\mathbb{1}_n\right) \\ &= f_{(\mathbf{C}_0,\mathbf{C}_1)}(t_1,\ldots,t_k)\ . \end{aligned}$$

This transformation results in the joint density function of representation $(\mathbf{C}_0, \mathbf{C}_1)$. ∎

Theorem 2 *If there is a matrix* $\mathbf{W} \in \mathbb{R}^{n,m}$ *such that* $\mathbb{1}_n = \mathbf{W}\,\mathbb{1}_m$, $\mathbf{W}\mathbf{D}_0 = \mathbf{C}_0\mathbf{W}$, $\mathbf{W}\mathbf{D}_1 = \mathbf{C}_1\mathbf{W}$ *and* $\nu = \phi\mathbf{W}$ *where* $\nu\left(-\mathbf{D}_0\right)^{-1}\mathbf{D}_1 = \nu$ *and* $\phi\left(-\mathbf{C}_0\right)^{-1}\mathbf{C}_1 = \phi$, *then* $(\mathbf{D}_0, \mathbf{D}_1)$ *and* $(\mathbf{C}_0, \mathbf{C}_1)$ *are equivalent.*

Proof. Using $\nu = \phi\mathbf{W}$, $\mathbf{W}(\mathbf{D}_0)^i\mathbf{D}_1 = (\mathbf{C}_0)^i\mathbf{C}_1\mathbf{W}$, and $\mathbb{1}_n = \mathbf{W}\,\mathbb{1}_m$, respectively, the following identity of the joint densities holds:

$$
\begin{aligned}
f_{(\mathbf{D}_0,\mathbf{D}_1)}(t_1,\ldots,t_k) &= \nu\left(\prod_{j=1}^{k}\left(\sum_{i_j=0}^{\infty}\frac{(t_j\mathbf{D}_0)^{i_j}}{i_j!}\mathbf{D}_1\right)\mathbb{1}_m\right) = \phi\mathbf{W}\left(\prod_{j=1}^{k}\left(\sum_{i_j=0}^{\infty}\frac{(t_j\mathbf{D}_0)^{i_j}}{i_j!}\mathbf{D}_1\right)\mathbb{1}_m\right)\\
&= \phi\left(\prod_{j=1}^{k}\left(\sum_{i_j=0}^{\infty}\frac{(t_j\mathbf{C}_0)^{i_j}}{i_j!}\mathbf{C}_1\right)\mathbf{W}\,\mathbb{1}_m\right) = \phi\left(\prod_{j=1}^{k}\left(\sum_{i_j=0}^{\infty}\frac{(t_j\mathbf{C}_0)^{i_j}}{i_j!}\mathbf{C}_1\right)\mathbb{1}_n\right)\\
&= f_{(\mathbf{C}_0,\mathbf{C}_1)}(t_1,\ldots,t_k)\ .
\end{aligned}
$$

■

The previous two theorems give sufficient conditions for reducing the size of the representation of a RAP. Unfortunately, these conditions are hard to check directly. The next sections provide explicit conditions for deciding the minimality of a representation.

4 Minimal Representation of RAPs

For notational convenience we often use the matrices $\mathbf{M} = -\mathbf{D}_0^{-1}$ and $\mathbf{P} = \mathbf{M}\mathbf{D}_1$, where $(\mathbf{M},\mathbf{P})$ and $(\mathbf{D}_0,\mathbf{D}_1)$ mutually define each other since $\mathbf{D}_0 = -\mathbf{M}^{-1}$ and $\mathbf{D}_1 = -\mathbf{D}_0\mathbf{P}$. For a RAP $(\mathbf{D}_0,\mathbf{D}_1)$ of size m we define the following matrices

$$
\begin{aligned}
\mathcal{C}_{\mathbf{D}_0} &= \mathcal{C}_{\mathbf{D}_0}(0) = \left(\mathbb{1}, \mathbf{M}\,\mathbb{1}, \ldots, \mathbf{M}^{m-1}\,\mathbb{1}\right),\\
\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1} &= \mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(0) = \left(\mathcal{C}_{\mathbf{D}_0}, \mathbf{P}\mathcal{C}_{\mathbf{D}_0}, \mathbf{P}^2\mathcal{C}_{\mathbf{D}_0}, \ldots, \mathbf{P}^{m-1}\mathcal{C}_{\mathbf{D}_0}\right),\\
\mathcal{C}_{\mathbf{D}_0}(n+1) &= \left(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n), \mathbf{M}\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n), \ldots, \mathbf{M}^{m-1}\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)\right),\\
\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n+1) &= \left(\mathcal{C}_{\mathbf{D}_0}(n+1), \mathbf{P}\mathcal{C}_{\mathbf{D}_0}(n+1), \mathbf{P}^2\mathcal{C}_{\mathbf{D}_0}(n+1), \ldots, \mathbf{P}^{m-1}\mathcal{C}_{\mathbf{D}_0}(n+1)\right).
\end{aligned}
$$

$\mathcal{C}_{\mathbf{D}_0}(n)$ is an $m \times m^{2n+1}$ matrix and $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)$ is an $m \times m^{2n+2}$ matrix. These matrices are extensions and generalizations of the controllability matrix known for linear systems [8, 10], and PH and ME distributions [7, 20]. Indeed $\mathcal{C}_{\mathbf{D}_0}$ is the controllability matrix used in linear systems theory, where, similar to the case of PH and ME distributions, a single matrix (and some additional vectors) characterizes the system behaviour. The introduced recursive definition of matrices is not present in the literature we know.

We define the rank of the generalized controllability matrix as $r_C = rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(m))$. This is a simple, but very redundant definition of r_C, which is relevant for the definition of a representation of minimal size as shown in the following theorems, but can be computed in a much more efficient way (see the RC_RANK$(\mathbf{D}_0,\mathbf{D}_1)$ procedure below). A cheap computation of r_C is possible based on the following corollaries.

Corollary 1 *Let* $\mathbf{H}$ *and* $\mathbf{G}$ *be matrices of size* $m \times m$ *and* $m \times r$*,* $R_i = rank(\mathbf{G}, \mathbf{H}\mathbf{G}, \ldots, \mathbf{H}^{i-1}\mathbf{G})$ *and* $\mathcal{S}_i = span(\mathbf{G}, \mathbf{H}\mathbf{G}, \ldots, \mathbf{H}^{i-1}\mathbf{G})$.

If $R_0 < R_1 < \ldots < R_i = R_{i+1}$ *then* $R_k = R_i$ *for* $k \geq i$.

Proof. From $R_i = R_{i+1}$ we have $\mathcal{S}_i = \mathcal{S}_{i+1}$ and consequently $\mathcal{S}_i = \mathcal{S}_{i+1} = \mathcal{S}_{i+2} = \ldots$, since

$$
\mathbf{H}^{i+2}\mathbf{G} = \underbrace{\mathbf{H}\underbrace{\mathbf{H}^{i+1}\mathbf{G}}_{\in\mathcal{S}_i}}_{\in\mathcal{S}_{i+1}=\mathcal{S}_i}.
$$

■

Assuming $\mathbf{G} = \mathcal{C}_{\mathbf{D}_0}(n)$ and $\mathbf{H} = \mathbf{P}$ or $\mathbf{G} = \mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)$ and $\mathbf{H} = \mathbf{M}$ Corollary 1 suggests a stepwise generation of $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)$ and $\mathcal{C}_{\mathbf{D}_0}(n+1)$, respectively, which can be terminated at the first time when the rank does not increase.

Corollary 2 *If* $rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)) = rank(\mathcal{C}_{\mathbf{D}_0}(n)) = r$ *then* $rank(\mathcal{C}_{\mathbf{D}_0}(j)) = rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(j)) = r$ *for* $j \geq n$.

Proof. $rank(\mathcal{C}_{\mathbf{D}_0}(n)) = r$ implies two consequences. The first one is that each column vector of $\mathcal{C}_{\mathbf{D}_0}(n)$ is a linear combination of r linearly independent vectors, say $\mathbf{z}_0, \ldots, \mathbf{z}_{r-1}$, and the second one is that a multiplication of any of these vectors with $\mathbf{M}^i$ is also a linear combination of $\mathbf{z}_0, \ldots, \mathbf{z}_{r-1}$. Additionally, $rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)) = rank(\mathcal{C}_{\mathbf{D}_0}(n)) = r$ implies that a multiplication of any of these vectors with $\mathbf{P}^i$ is also a linear combination of $\mathbf{z}_0, \ldots, \mathbf{z}_{r-1}$. Consequently, further multiplications with $\mathbf{M}$ or $\mathbf{P}$ remain a linear combination of $\mathbf{z}_0, \ldots, \mathbf{z}_{r-1}$, which preserves the rank of the subsequent matrices. ■

Corollary 3 *If* $rank(\mathcal{C}_{\mathbf{D}_0}(n)) = rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n-1)) = r$ *then* $rank(\mathcal{C}_{\mathbf{D}_0}(j)) = rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(j)) = r$ *for* $j \geq n$.

Proof. Similar to Corollary 2 $rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n-1)) = r$ implies two consequences. The first one is that each column vector of $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n-1)$ is a linear combination of r linearly independent vectors, say $\mathbf{z}_0, \ldots, \mathbf{z}_{r-1}$, and the second one is that a multiplication of any of these vectors with $\mathbf{P}^i$ is also a linear combination of $\mathbf{z}_0, \ldots, \mathbf{z}_{r-1}$. Additionally, $rank(\mathcal{C}_{\mathbf{D}_0}(n)) = rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n-1)) = r$ implies that a multiplication of any of these vectors with $\mathbf{M}^i$ is also a linear combination of $\mathbf{z}_0, \ldots, \mathbf{z}_{r-1}$. ■

Let $\mathcal{C}_{\mathbf{D}_0}(n) = \widetilde{\mathbf{U}}\widetilde{\mathbf{S}}\widetilde{\mathbf{T}}^*$ be the singular value decomposition (SVD) [11] of $\mathcal{C}_{\mathbf{D}_0}(n)$, where $\widetilde{\mathbf{U}}$ ($\widetilde{\mathbf{T}}$) is a unitary matrix of size $m \times m$ ($m^{2n+1} \times m^{2n+1}$), $\widetilde{\mathbf{U}}^*$ is the conjugate transpose of matrix $\widetilde{\mathbf{U}}$ (in our case $\widetilde{\mathbf{U}}^*$ is the transpose of the matrix $\widetilde{\mathbf{U}}$ since $\mathcal{C}_{\mathbf{D}_0}(n)$ is real), $\widetilde{\mathbf{S}}$ of size $m \times m^{2n+1}$ contains the singular values in decreasing order on the diagonal and we neglect the dependence on n to simplify the notation. Let $\mathbf{I}_{i,j}$ be the matrix of size $i \times j$ whose k, ℓ element equal to one if $k = \ell$ and zero otherwise.

Corollary 4 *If* $rank(\mathcal{C}_{\mathbf{D}_0}(n)) = r$ *then*

$$rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)) = rank\left(\widetilde{\mathbf{Z}}, \mathbf{P}\widetilde{\mathbf{Z}}, \mathbf{P}^2\widetilde{\mathbf{Z}}, \ldots, \mathbf{P}^{m-1}\widetilde{\mathbf{Z}}\right),$$

where $\widetilde{\mathbf{Z}} = \widetilde{\mathbf{U}}\widetilde{\mathbf{S}}\mathbf{I}_{m^{2n+1},r}$.

Proof. $rank(\mathcal{C}_{\mathbf{D}_0}(n)) = r$ implies that each column vector of $\mathcal{C}_{\mathbf{D}_0}(n)$ is a linear combination of r linearly independent vectors. Matrix $\widetilde{\mathbf{Z}}$ of size $m \times r$ contains such a set of vectors because according to the structure of the matrix $\widetilde{\mathbf{S}}$, the ith column vector of $\mathcal{C}_{\mathbf{D}_0}(n)$, $\mathcal{C}_{\mathbf{D}_0}(n)_i$, can be expressed as $\mathcal{C}_{\mathbf{D}_0}(n)_i = \sum_{j=1}^{r} \widetilde{\mathbf{T}}^*_{ij}\widetilde{\mathbf{Z}}_j$, where $\widetilde{\mathbf{Z}}_j$ is the jth column vector of $\widetilde{\mathbf{Z}}$ and $\widetilde{\mathbf{T}}^*_{ij}$ is the i,jth element of $\widetilde{\mathbf{T}}^*$.
A column vector of $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)$ has the form $\mathbf{P}^k\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)_i$ and it can be expressed as $\mathbf{P}^k\mathcal{C}_{\mathbf{D}_0}(n)_i = \mathbf{P}^k \sum_{j=1}^{r} \widetilde{\mathbf{T}}^*_{ij}\widetilde{\mathbf{Z}}_j$ from which the corollary follows. ■

Similarly, let $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n) = \widehat{\mathbf{U}}\widehat{\mathbf{S}}\widehat{\mathbf{T}}^*$ be the singular value decomposition of $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)$.

Corollary 5 *If* $rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)) = r$ *then*

$$rank(\mathcal{C}_{\mathbf{D}_0}(n+1)) = rank\left(\widehat{\mathbf{Z}}, \mathbf{M}\widehat{\mathbf{Z}}, \mathbf{M}^2\widehat{\mathbf{Z}}, \ldots, \mathbf{M}^{m-1}\widehat{\mathbf{Z}}\right),$$

where $\widehat{\mathbf{Z}} = \widehat{\mathbf{U}}\widehat{\mathbf{S}}\mathbf{I}_{m^{2n+2},r}$

Proof. The proof follows the same pattern as the one for Corollary 4. ■

Based on these results r_C can be computed with the following efficient procedure

1. RC_RANK($\mathbf{D}_0, \mathbf{D}_1$)
2. $r_C^o = 0$, $\mathbf{M} = -\mathbf{D}_0^{-1}$, $\mathbf{P} = \mathbf{M}\mathbf{D}_1$,
3. $r_C = 1$, $\mathbf{Z} = \mathbb{1}$, $\mathbf{\Theta} = \mathbf{Z}$, $\mathbf{H} = \mathbf{M}$,

FOR $n = 0$ TO m DO
  WHILE $r_c < rank(\mathbf{\Theta}, \mathbf{HZ})$ DO $\mathbf{\Theta} = (\mathbf{\Theta}, \mathbf{HZ})$, $\mathbf{H} = \mathbf{HM}$, $r_c = rank(\mathbf{\Theta})$;
  $/*\ r_C =$ rank of $\mathcal{C}_{\mathbf{D}_0}(n)$ $*/$
  IF $r_C == r_C^o$ OR $r_C == m$ THEN RETURN(r_C);
  $r_C^o = r_C$; $\mathbf{Z}$ =SPANNING($\mathbf{\Theta}$); $\mathbf{H} = \mathbf{P}$;
  WHILE $r_c < rank(\mathbf{\Theta}, \mathbf{HZ})$ DO $\mathbf{\Theta} = (\mathbf{\Theta}, \mathbf{HZ})$, $\mathbf{H} = \mathbf{HP}$, $r_c = rank(\mathbf{\Theta})$;
  $/*\ r_C =$ rank of $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)$ $*/$
  IF $r_C == r_C^o$ OR $r_C == m$ THEN RETURN(r_C);
  $r_C^o = r_C$; $\mathbf{Z}$ =SPANNING($\mathbf{\Theta}$); $\mathbf{H} = \mathbf{M}$;
ENDFOR

where

SPANNING($\mathbf{\Theta}$)
  $\{i, j\}$ =SIZE($\mathbf{\Theta}$);
  $\{\mathbf{U}, \mathbf{S}, \mathbf{T}\}$ =SVD($\mathbf{\Theta}$);
  r =RANK($\mathbf{\Theta}$);
  RETURN($\mathbf{U} * \mathbf{S} * \mathbf{I}_{i,r}$);

Note that, in this procedure the size of the matrix $\mathbf{\Theta}$ is never greater than $m \times m^2/4$. This computation of r_C helps us to show the following property.

Corollary 6

$$rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(\lceil m/2 \rceil)) = rank(\mathcal{C}_{\mathbf{D}_0}(\lceil m/2 \rceil)) \ . \tag{6}$$

Proof. The upper limit of the iterations, m, is also redundant, since on the one hand the rank is bounded above by m (the size of the representation) and, on the other hand, the rank is nondecreasing in each of the steps ($rank(\mathcal{C}_{\mathbf{D}_0}(n)) \Rightarrow rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n))$ and $rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)) \Rightarrow rank(\mathcal{C}_{\mathbf{D}_0}(n+1))$, since all matrices contain the previous ones) and this way r_C^o increases by at least 2 in each completed iteration. Consequently, the rank can not increase after $\lceil m/2 \rceil$ iterations. ■

Example 1 *Consider the MAP with representation*

$$\mathbf{D}_0 = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & -2 & 1 & 0 & 0 & 0 \\ 1 & 0 & -3 & 1 & 0 & 0 \\ 1 & 0 & 1 & -4 & 1 & 0 \\ 4 & 0 & 0 & 0 & -5 & 0 \\ 5 & 0 & 0 & 0 & 0 & -6 \end{pmatrix}, \quad \mathbf{D}_1 = \begin{pmatrix} 0.25 & 0.5 & 0 & 0.25 & 0 & 0 \\ 0 & 0 & 0.25 & 0 & 0.25 & 0 \\ 0.75 & 0 & 0 & 0.25 & 0 & 0 \\ 0 & 0.5 & 0 & 0.25 & 0.25 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0.25 & 0 & 0.25 \end{pmatrix}.$$

For this MAP

$$rank(\mathcal{C}_{\mathbf{D}_0}(0)) = 2,\ rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(0)) = 3,\ rank(\mathcal{C}_{\mathbf{D}_0}(1)) = 4,$$
$$rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(1)) = 5,\ rank(\mathcal{C}_{\mathbf{D}_0}(2)) = 5,\ rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(2)) = 5,$$

which means that for this MAP $r_C = 5$, but we need to perform 3 iterations in the r_C computing procedure to recognize this order.

Similarly, let

$$\mathcal{O}_{\mathbf{D}_0} = \mathcal{O}_{\mathbf{D}_0}(0) = \begin{pmatrix} \nu \\ \nu\mathbf{M} \\ \vdots \\ \nu\mathbf{M}^{m-1} \end{pmatrix}, \quad \mathcal{O}_{\mathbf{D}_0,\mathbf{D}_1} = \mathcal{O}_{\mathbf{D}_0,\mathbf{D}_1}(0) = \begin{pmatrix} \mathcal{O}_{\mathbf{D}_0} \\ \mathcal{O}_{\mathbf{D}_0}\mathbf{P} \\ \vdots \\ \mathcal{O}_{\mathbf{D}_0}\mathbf{P}^{m-1} \end{pmatrix},$$

$$\mathcal{O}_{\mathbf{D}_0}(n+1) = \begin{pmatrix} \mathcal{O}_{\mathbf{D}_0,\mathbf{D}_1}(n) \\ \mathcal{O}_{\mathbf{D}_0,\mathbf{D}_1}(n)\mathbf{M} \\ \vdots \\ \mathcal{O}_{\mathbf{D}_0,\mathbf{D}_1}(n)\mathbf{M}^{m-1} \end{pmatrix}, \quad \mathcal{O}_{\mathbf{D}_0,\mathbf{D}_1}(n+1) = \begin{pmatrix} \mathcal{O}_{\mathbf{D}_0}(n+1) \\ \mathcal{O}_{\mathbf{D}_0}(n+1)\mathbf{P} \\ \vdots \\ \mathcal{O}_{\mathbf{D}_0}(n+1)\mathbf{P}^{m-1} \end{pmatrix},$$

where $\nu\mathbf{P} = \nu$, $\nu\,\mathbb{1} = 1$. $\mathcal{O}_{\mathbf{D}_0}$ is the observability matrix used in linear systems theory [8, 10]. The other matrices are extensions and generalizations of the observability matrix. We define the rank of the generalized observability matrix as $r_O = rank(\mathcal{O}_{\mathbf{D}_0,\mathbf{D}_1}(m))$.

Theorem 3 *If for a RAP$(\mathbf{D}_0,\mathbf{D}_1)$ of size m $rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(m)) = r_C = n < m$, then there exists a non singular $m \times m$ transformation matrix $\mathbf{B}$ such that*

$$\mathbf{D}_0' = \mathbf{B}^{-1}\mathbf{D}_0\mathbf{B} = \begin{pmatrix} \mathbf{C}_0 & \star \\ \mathbf{0} & \star \end{pmatrix}, \quad \mathbf{D}_1' = \mathbf{B}^{-1}\mathbf{D}_1\mathbf{B} = \begin{pmatrix} \mathbf{C}_1 & \star \\ \mathbf{0} & \star \end{pmatrix} \text{ and } \mathbf{B}^{-1}\,\mathbb{1}_m = \begin{pmatrix} \mathbb{1}_n \\ \mathbf{0} \end{pmatrix},$$

where the considered vector and matrix blocks are of size $n, m-n$ in each dimension and $\star$ indicates irrelevant matrix block. $(\mathbf{C}_0,\mathbf{C}_1)$ is an equivalent representation of the same RAP of size n.

Proof. Let $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(m) = \mathbf{UST}^*$ be the singular value decomposition of $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(m)$. Since $rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(m)) = n$ the first n singular values are non-zero and the last $m-n$ are zero, that is, the last $m-n$ rows of $\mathbf{S}$ and $\mathbf{ST}^*$ are zero. From $\mathbf{U}^*\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(m) = \mathbf{ST}^*$ we have that $\mathbf{U}^*\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(m)$ is composed by m^{2m+2} column vectors of size m, whose last $m-n$ elements are zero. The very first column vector of $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(m)$, numbered as the zeroth column vector, is $\mathbb{1}$ by definition. From this we have

$$[\mathbf{U}^*\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(m)]_0 = \mathbf{U}^*\,\mathbb{1} = \begin{pmatrix} \mathbf{z}_0 \\ \mathbf{0} \end{pmatrix},$$

where $\mathbf{z}_0$ is a column vector of size n and $\mathbf{0}$ is a column vector of zeros of size $m-n$. For column vector number $\ell = \sum_{j=0}^{2m+1} m^j i_j$ (with $0 \le i_j \le m-1$) we have

$$[\mathbf{U}^*\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(m)]_\ell = \mathbf{U}^*\mathbf{P}^{i_{2m+1}}\mathbf{M}^{i_{2m}} \ldots \mathbf{P}^{i_1}\mathbf{M}^{i_0}\,\mathbb{1} = \tag{7}$$

$$\mathbf{U}^*\mathbf{P}^{i_{2m+1}}\mathbf{U}\ \ \mathbf{U}^*\mathbf{M}^{i_{2m}}\mathbf{U}\ \ \ldots \mathbf{P}^{i_1}\mathbf{U}\ \ \mathbf{U}^*\mathbf{M}^{i_0}\mathbf{U}\ \ \mathbf{U}^*\,\mathbb{1} = \begin{pmatrix} \mathbf{z}_\ell \\ \mathbf{0} \end{pmatrix}.$$

According to (6) $rank(\mathcal{C}_{\mathbf{D}_0}(m)) = n$. It means that the vectors $\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_{m^{2m+1}-1}$ of size n contain n linearly independent ones. Let $\mathbf{z}_0', \mathbf{z}_1', \ldots, \mathbf{z}_{n-1}'$ (column vectors of size n) be n linearly independent vectors from $\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_{m^{2m+1}-1}$ and compose the matrix $\begin{pmatrix} \mathbf{Z}' \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{z}_0' \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{z}_1' \\ \mathbf{0} \end{pmatrix}, \ldots, \begin{pmatrix} \mathbf{z}_{n-1}' \\ \mathbf{0} \end{pmatrix}$. From (7) we have

$$\mathbf{U}^*\mathbf{MU} \begin{pmatrix} \mathbf{Z}' \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \star \\ \mathbf{0} \end{pmatrix} \quad \text{and} \quad \mathbf{U}^*\mathbf{PU} \begin{pmatrix} \mathbf{Z}' \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \star \\ \mathbf{0} \end{pmatrix}.$$

Due to the fact that $\mathbf{Z}'$ is an $n \times n$ matrix of rank n, the rank of the lower left block of $\mathbf{U}^*\mathbf{MU}$ and $\mathbf{U}^*\mathbf{PU}$ are zero, i.e., $\mathbf{U}^*\mathbf{MU}$ and $\mathbf{U}^*\mathbf{PU}$ have the structure $\begin{pmatrix} \star & \star \\ \mathbf{0} & \star \end{pmatrix}$ with the zero block of size $m-n \times n$.

If $\mathbf{z}_0$ contains zero elements, then we introduce matrix $\mathbf{R} = \begin{pmatrix} \widehat{\mathbf{R}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$ such that $\widehat{\mathbf{R}}$, and consequently $\mathbf{R}$, are non-singular and $\hat{\mathbf{z}}_0 = \widehat{\mathbf{R}}\mathbf{z}_0$ is composed of non-zero elements. If $\mathbf{z}_0$ does not contain zero elements, then $\mathbf{R} = \mathbf{I}$. Since $\mathbf{R}$ and $\mathbf{R}^{-1}$ are block diagonal matrices a multiplication of $\mathbf{U}^*\mathbf{MU}$ ($\mathbf{U}^*\mathbf{PU}$) with $\mathbf{R}$ or $\mathbf{R}^{-1}$ preserves the zero block structure of $\mathbf{U}^*\mathbf{MU}$ ($\mathbf{U}^*\mathbf{PU}$).
Let $\mathbf{\Gamma}$ be the diagonal matrix composed of the n (non-zero) elements of the vector $\hat{\mathbf{z}}_0$ and $m-n$ ones, that is, $\mathbf{\Gamma} = \text{Diag}\langle \hat{\mathbf{z}}_0^T, \mathbb{1}_{m-n}^T \rangle$. Finally $\mathbf{B} = \mathbf{UR}^{-1}\mathbf{\Gamma}$ results in the required transformation matrix. Since $\mathbf{\Gamma}$ is a diagonal matrix $\mathbf{M}' = \mathbf{B}^{-1}\mathbf{MB} = \mathbf{\Gamma}^{-1}\mathbf{RU}^*\mathbf{MUR}^{-1}\mathbf{\Gamma}$ and also $\mathbf{P}' = \mathbf{B}^{-1}\mathbf{PB} = \mathbf{\Gamma}^{-1}\mathbf{RU}^*\mathbf{PUR}^{-1}\mathbf{\Gamma}$ have the same structure as $\mathbf{RU}^*\mathbf{MUR}^{-1}$ and $\mathbf{RU}^*\mathbf{PUR}^{-1}$, respectively, and

$$\mathbf{B}^{-1}\,\mathbb{1}_m = \mathbf{\Gamma}^{-1}\mathbf{RU}^*\,\mathbb{1}_m = \mathbf{\Gamma}^{-1}\mathbf{R}\begin{pmatrix} \mathbf{z}_0 \\ \mathbf{0} \end{pmatrix} = \mathbf{\Gamma}^{-1}\begin{pmatrix} \hat{\mathbf{z}}_0 \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbb{1}_n \\ \mathbf{0} \end{pmatrix}.$$

$\mathbf{D}'_0 = -\mathbf{M}'^{-1}$ and $\mathbf{D}'_1 = -\mathbf{D}'_0\mathbf{P}'$ has the required matrix structure, because the matrix inversion and the matrix multiplication maintains the zero block at the lower left corner. ■

Corollary 7 *If for a RAP$(\mathbf{D}_0, \mathbf{D}_1)$ of size m*

$$-\nu\mathbf{D}_0^{-1}\mathbf{D}_1 = \nu, \quad \nu\,\mathbb{1}_m = 1,$$

and

$$\mathbf{B}^{-1}\mathbf{D}_0\mathbf{B} = \begin{pmatrix} \mathbf{C}_0 & \star \\ \mathbf{0} & \star \end{pmatrix}, \ \mathbf{B}^{-1}\mathbf{D}_1\mathbf{B} = \begin{pmatrix} \mathbf{C}_1 & \star \\ \mathbf{0} & \star \end{pmatrix} \text{ and } \mathbf{B}^{-1}\,\mathbb{1}_m = \begin{pmatrix} \mathbb{1}_n \\ \mathbf{0} \end{pmatrix}$$

then $\nu\mathbf{B} = \begin{pmatrix} \phi & \star \end{pmatrix}$, *where*

$$-\phi\ \mathbf{C}_0^{-1}\mathbf{C}_1 = \phi, \phi\,\mathbb{1}_n = 1$$

Proof. From $\nu = -\nu\mathbf{D}_0^{-1}\mathbf{D}_1$ we have

$$\nu\mathbf{B} = -\nu\mathbf{B}\begin{pmatrix} \mathbf{C}_0^{-1} & \star \\ \mathbf{0} & \star \end{pmatrix}\begin{pmatrix} \mathbf{C}_1 & \star \\ \mathbf{0} & \star \end{pmatrix} = -\nu\mathbf{B}\begin{pmatrix} \mathbf{C}_0^{-1}\mathbf{C}_1 & \star \\ \mathbf{0} & \star \end{pmatrix},$$

and from $1 = \nu\,\mathbb{1}_m$ we have

$$1 = \nu\mathbf{B}\mathbf{B}^{-1}\,\mathbb{1}_m = \nu\mathbf{B}\begin{pmatrix} \mathbb{1}_n \\ \mathbf{0} \end{pmatrix}.$$

Since ϕ is the first n elements of $\nu\mathbf{B}$, we have $-\phi\ \mathbf{C}_0^{-1}\mathbf{C}_1 = \phi$ and $\phi\,\mathbb{1}_n = 1$. ■

Theorem 4 *If for a RAP $(\mathbf{D}_0, \mathbf{D}_1)$ of size m rank$(\mathcal{O}^m_{\mathbf{D}_0,\mathbf{D}_1}) = r_O = n < m$, then there exists a non singular $m \times m$ transformation matrix $\mathbf{B}$ such that*

$$\mathbf{B}^{-1}\mathbf{D}_0\mathbf{B} = \begin{pmatrix} \mathbf{C}_0 & \mathbf{0} \\ \star & \star \end{pmatrix}, \ \mathbf{B}^{-1}\mathbf{D}_1\mathbf{B} = \begin{pmatrix} \mathbf{C}_1 & \mathbf{0} \\ \star & \star \end{pmatrix} \text{ and } \nu\mathbf{B} = \begin{pmatrix} \phi & \mathbf{0} \end{pmatrix}$$

and $(\mathbf{C}_0, \mathbf{C}_1)$ is an equivalent representation of size n.

Proof. The proof follows the same pattern as the proof of Theorem 3. ■

Corollary 8 *If for a RAP$(\mathbf{D}_0, \mathbf{D}_1)$ of size m*

$$-\nu\mathbf{D}_0^{-1}\mathbf{D}_1 = \nu, \nu\,\mathbb{1}_m = 1,$$

and

$$\mathbf{B}^{-1}\mathbf{D}_0\mathbf{B} = \begin{pmatrix} \mathbf{C}_0 & \mathbf{0} \\ \star & \star \end{pmatrix}, \ \mathbf{B}^{-1}\mathbf{D}_1\mathbf{B} = \begin{pmatrix} \mathbf{C}_1 & \mathbf{0} \\ \star & \star \end{pmatrix} \text{ and } \nu\mathbf{B} = \begin{pmatrix} \phi & \mathbf{0} \end{pmatrix}$$

then $\mathbf{B}^{-1}\,\mathbb{1}_m = \begin{pmatrix} \mathbb{1}_n \\ \star \end{pmatrix}$ *and* $-\phi\ \mathbf{C}_0^{-1}\mathbf{C}_1 = \phi, \phi\,\mathbb{1}_n = 1$.

Proof. The proof follows the same pattern as the proof of Corollary 7. ■

An alternative way to obtain the relation of Theorem 3 and 4 is to recognize that $\mathcal{O}_{\mathbf{D}_0,\mathbf{D}_1}(m)$ of a RAP with representation $(\mathbf{D}_0, \mathbf{D}_1)$, initial vector ν, and closing vector $\mathbb{1}$ is the same as $\mathcal{C}_{\mathbf{D}_0^T,\mathbf{D}_1^T}(m)$ with representation $(\mathbf{D}_0^T, \mathbf{D}_1^T)$, initial vector $\mathbb{1}^T$, and closing vector ν^T, and vice versa, $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(m)$ of a RAP with representation $(\mathbf{D}_0, \mathbf{D}_1)$, initial vector ν, and closing vector $\mathbb{1}$ is the same as $\mathcal{O}_{\mathbf{D}_0^T,\mathbf{D}_1^T}(m)$ with representation $(\mathbf{D}_0^T, \mathbf{D}_1^T)$, initial vector $\mathbb{1}^T$, and closing vector ν^T.

Theorems 1 and 3 and similarly Theorems 2 and 4 are closely related. Matrix $\mathbf{V}$ of Theorem 1 results from the first n columns of matrix $\mathbf{B}$ of Theorem 3 and matrix $\mathbf{W}$ of Theorem 2 from the first n rows of matrix $\mathbf{B}$ of Theorem 4. Having all of these ingredients we can present the main theorem about the minimal representation of a RAP.

Theorem 5 $(\mathbf{D}_0, \mathbf{D}_1)$ *of size m is a minimal representation of a RAP if and only if $r_O = r_C = m$.*

Proof. First we prove that if $r_O = r_C = m$ then $(\mathbf{D}_0, \mathbf{D}_1)$ is minimal using contradiction.
Let us assume that $r_O = r_C = m$ and there exists an equivalent $(\mathbf{C}_0, \mathbf{C}_1)$ representation of the same RAP of size $n < m$, then $\mathcal{C}_{\mathbf{C}_0,\mathbf{C}_1}(m)$ is of size $n \times n^{2m+2}$ and consequently $rank(\mathcal{C}_{\mathbf{C}_0,\mathbf{C}_1}(m)) \leq n$, which is in contrast with $r_C = m$.
The other statement of the theorem says that if $(\mathbf{D}_0, \mathbf{D}_1)$ is minimal then $r_O = r_C = m$. This statement is a result of Theorem 3 which provides a procedure to represent the same RAP with size n if $r_C = n < m$, and Theorem 4 which provides a procedure to represent the same RAP with size n if $r_O = n < m$. ■

4.1 Computation of a minimal representation

The procedure of Theorem 3 can also be interpreted as a procedure to eliminate the redundancy of the representation caused by the closing vector, and the one of Theorem 4 as a procedure to eliminate the redundancy of the representation caused by the initial vector.

Due to the fact that a minimal representation does not allow reduction neither due to the initial nor due to the closing vector a general RAP with representation $(\mathbf{D}_0, \mathbf{D}_1)$ has to be minimized in two steps (in arbitrary order). In one step the procedure of Theorem 3 eliminates the redundancy of the representation caused by the closing vector and in the next step the procedure of Theorem 4 eliminates the redundancy of the representation caused by the initial vector. Example 3 below demonstrates a case when both steps are necessary for finding a minimal representation. More precisely the following procedure can be used to minimize the size of the representation of a RAP.

1. MINIMIZE$(\mathbf{D}_0, \mathbf{D}_1)$
2. $\{i, j\}$ =SIZE$(\mathbf{D}_0)$;
3. IF RC_RANK$(\mathbf{D}_0, \mathbf{D}_1) < i$ THEN $(\mathbf{D}_0, \mathbf{D}_1)$ =MINIMIZE_BY_THEOREM3$(\mathbf{D}_0, \mathbf{D}_1)$;
4. $\{i, j\}$ =SIZE$(\mathbf{D}_0)$;
5. IF RO_RANK$(\mathbf{D}_0, \mathbf{D}_1) < i$ THEN $(\mathbf{D}_0, \mathbf{D}_1)$ =MINIMIZE_BY_THEOREM4$(\mathbf{D}_0, \mathbf{D}_1)$;
6. RETURN$(\mathbf{D}_0, \mathbf{D}_1)$;

The proof of Theorem 3 and 4 suggests procedures to obtain a smaller representation of a RAP, but due to the redundancy of the definition of r_C and r_O this procedure might be inefficient. The following section proposes a computationally efficient procedure.

5 The Staircase Algorithm

In the previous section conditions for the minimality and the steps to compute a minimal representation have been introduced. Here we present an algorithm that computes a non-singular matrix $\mathbf{B}$ such that

$$\mathbf{B}^{-1}\,\mathbb{1}_m = \begin{pmatrix} \mathbb{1}_n \\ \mathbf{0} \end{pmatrix}, \; \mathbf{B}^{-1}\mathbf{D}_0\mathbf{B} = \begin{pmatrix} \mathbf{C}_0 & \star \\ \mathbf{0} & \star \end{pmatrix} \text{ and } \mathbf{B}^{-1}\mathbf{D}_1\mathbf{B} = \begin{pmatrix} \mathbf{C}_1 & \star \\ \mathbf{0} & \star \end{pmatrix}, \tag{8}$$

without computing the large matrices $\mathcal{C}_{\mathbf{D}_0}(m)$ and $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(m)$. By defining $\mathbf{B}^{-1} = \begin{pmatrix} \mathbf{W} \\ \mathbf{F} \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} \mathbf{V} & \mathbf{G} \end{pmatrix}$ where $\mathbf{V}$ is an $m \times n$ matrix and $\mathbf{W}$ is an $n \times m$ matrix we obtain the matrices to transform the representation $(\mathbf{D}_0, \mathbf{D}_1)$ of size m into the representation $(\mathbf{C}_0, \mathbf{C}_1)$ of size n (cf. Theorems 1 and 2). For the computation of $\mathbf{B}$ the staircase algorithm from linear system theory [10, 21] can be extended. The algorithm computes a transformation matrix $\mathbf{B}$ stepwise such that finally (8) holds. The algorithm is based on a SVD which has already been used in the proof of Theorem 3. For some matrix $\mathbf{A}$ of size $r \times c$ the SVD is $\mathbf{A} = \mathbf{UST}^*$. The singular values are in decreasing order and the first $rank(\mathbf{A})$ singular values are positive and the remaining are zero. In particular,

$$\mathbf{U}^*\mathbf{A} = \begin{pmatrix} \mathbf{F} \\ \mathbf{0} \end{pmatrix}$$

holds where $\mathbf{F}$ is a $rank(\mathbf{A}) \times c$ matrix. If all non-zero singular values are distinct, then matrix $\mathbf{U}^*$ is uniquely defined up to complex signs. If identical singular values exist, then $\mathbf{U}^*$ is unique up to complex signs and the ordering of rows belonging to identical singular values.

With these ingredients we define an algorithm that works on three matrices, two $m \times m$ matrices $\mathbf{X}$, $\mathbf{Y}$, and an $m \times c$ matrix $\mathbf{Z}$. The algorithm computes a unitary matrix $\mathbf{U}$ such that

$$\mathbf{U}^*\mathbf{X}\mathbf{U} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{X}_2 \\ \mathbf{0} & \mathbf{X}_4 \end{pmatrix}, \ \mathbf{U}^*\mathbf{Y}\mathbf{U} = \begin{pmatrix} \mathbf{Y}_1 & \mathbf{Y}_2 \\ \mathbf{0} & \mathbf{Y}_4 \end{pmatrix} \text{ and } \mathbf{U}^*\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{0} \end{pmatrix} \tag{9}$$

where $\mathbf{X}_1$ and $\mathbf{Y}_1$ are of size $n \times n$ and $\mathbf{Z}_1$ is of size $n \times c$.

The computation of a smaller representation of a RAP with representation $(\mathbf{D}_0, \mathbf{D}_1)$ according to Theorem 3 is done with the following algorithm which is called with square matrices $\mathbf{X} = \mathbf{D}_0$, $\mathbf{Y} = \mathbf{D}_1$ of size m and an $m \times c$ $(c = 1)$ matrix $\mathbf{Z} = \mathbb{1}_m$, that is, STAIRCASE$(\mathbf{D}_0, \mathbf{D}_1, \mathbb{1}_m)$.

STAIRCASE($\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$)

  $i = 0$; $\{m, j\}$ =SIZE($\mathbf{X}$); $\mathbf{U}^* = \mathbf{I}$;

  REPEAT

    $i = i + 1$; $r_i = rank(\mathbf{Z})$; $\{\mathbf{U}_i, \mathbf{S}_i, \mathbf{T}_i\}$ =SVD($\mathbf{Z}$) ;

    $\begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{0} \end{pmatrix} = \mathbf{U}_i^*\mathbf{Z}$; $\begin{pmatrix} \mathbf{X}_1 & \mathbf{X}_2 \\ \mathbf{X}_3 & \mathbf{X}_4 \end{pmatrix} = \mathbf{U}_i^*\mathbf{X}\mathbf{U}_i$; $\begin{pmatrix} \mathbf{Y}_1 & \mathbf{Y}_2 \\ \mathbf{Y}_3 & \mathbf{Y}_4 \end{pmatrix} = \mathbf{U}_i^*\mathbf{Y}\mathbf{U}_i$;
    /* $\mathbf{X}_1$ and $\mathbf{Y}_1$ are of size $r_i \times r_i$ */

    $\mathbf{U}^* = \begin{pmatrix} \mathbf{I}_{\sum_{j=1}^{i-1} r_j} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_i^* \end{pmatrix} \mathbf{U}^*$;

    $\mathbf{X} = \mathbf{X}_4$; $\mathbf{Y} = \mathbf{Y}_4$; $\mathbf{Z} = (\mathbf{X}_3 \ \mathbf{Y}_3)$;

  UNTIL $rank(\mathbf{Z}) = m - \sum_{j=1}^{i} r_j$ or $\mathbf{Z} = \mathbf{0}$;

  IF ($\mathbf{Z} = \mathbf{0}$) THEN

    $n = \sum_{j=1}^{i} r_j$;

    $\begin{pmatrix} \mathbf{x} \\ \mathbf{0}_{m-n} \end{pmatrix} = \mathbf{U}^* \mathbb{1}_m$;

    IF ($\mathbf{x} \neq \mathbf{0}$) THEN $\mathbf{R} = \mathbf{I}$ ELSE $\mathbf{R}$ = non-singular matrix such that $\mathbf{R}\mathbf{x} \neq \mathbf{0}$
    /* element-wise */

    $\mathbf{y} = \mathbf{R}\mathbf{x}$; $\mathbf{\Gamma} = diag(\mathbf{y})$; $\mathbf{B} = \left[\begin{pmatrix} \mathbf{\Gamma}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m-n} \end{pmatrix} \begin{pmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m-n} \end{pmatrix} \mathbf{U}^* \right]^{-1}$;

  ELSE $n = m$; $\mathbf{B} = \mathbf{I}$; /* no reduction is possible */

  RETURN($n, \mathbf{B}$) ;

The following corollary is needed for the proof of the subsequent theorem which shows the termination and correctness of the above algorithm.

Corollary 9 *Let $\mathbf{X}$ and $\mathbf{Y}$ be two $m \times m$ matrices and $\mathbf{Z}$ be a $m \times c$ matrix. If a non-singular matrix $\mathbf{B}$ exists such that*

$$\mathbf{B}^{-1}\mathbf{X}\mathbf{B} = \begin{pmatrix} \tilde{\mathbf{X}}_1 & \tilde{\mathbf{X}}_2 \\ \mathbf{0} & \tilde{\mathbf{X}}_4 \end{pmatrix}, \ \mathbf{B}^{-1}\mathbf{Y}\mathbf{B} = \begin{pmatrix} \tilde{\mathbf{Y}}_1 & \tilde{\mathbf{Y}}_2 \\ \mathbf{0} & \tilde{\mathbf{Y}}_4 \end{pmatrix} \text{ and } \mathbf{B}^{-1}\mathbf{Z} = \begin{pmatrix} \tilde{\mathbf{Z}}_1 \\ \mathbf{0} \end{pmatrix}, \tag{10}$$

then a unitary matrix $\mathbf{U}$ exists such that

$$\mathbf{U}^*\mathbf{X}\mathbf{U} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{X}_2 \\ \mathbf{0} & \mathbf{X}_4 \end{pmatrix}, \ \mathbf{U}^*\mathbf{Y}\mathbf{U} = \begin{pmatrix} \mathbf{Y}_1 & \mathbf{Y}_2 \\ \mathbf{0} & \mathbf{Y}_4 \end{pmatrix} \text{ and } \mathbf{U}^*\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{0} \end{pmatrix}, \tag{11}$$

where $\tilde{\mathbf{X}}_1, \mathbf{X}_1, \tilde{\mathbf{Y}}_1, \mathbf{Y}_1$ are $n \times n$ matrices and $\tilde{\mathbf{Z}}_1, \mathbf{Z}_1$ are $n \times c$ matrices.

Proof. The rank of $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(m)$ is n in this case and the result follows from the proof of Theorem 3 where the unitary matrix is constructed. ■

Theorem 6 *The algorithm computes in at most $m-1$ iterations a matrix $\mathbf{B}$ that observes (8), if such a matrix exists.*

Proof.
We first show that the number of iterations is at most $m-1$:
In the first iteration matrix $\mathbf{Z}$ has m rows, in the subsequent iterations the number of rows of $\mathbf{Z}$ equals the number of rows of $\mathbf{Z}$ from the previous iteration minus the rank of $\mathbf{Z}$ from the previous iteration. If $\mathbf{Z}$ becomes $\mathbf{0}$ or has full rank, the iteration stops. This implies that in each iteration $\mathbf{Z}$ has at least one row less such that after $m-1$ iterations a 1×1 matrix $\mathbf{Z}$ has been computed and this matrix is either $\mathbf{0}$ or has full rank.
After i iterations, the algorithm has computed the following matrices (see also [10])

$$\mathbf{H}_i^*\mathbf{Z} = \left(\begin{array}{c} \mathbf{Z}_1^{(1)} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \hline \mathbf{0} \\ \mathbf{0} \end{array}\right), \quad \mathbf{H}_i^*\mathbf{X}\mathbf{H}_i = \left(\begin{array}{ccccc|cc} \mathbf{X}_1^{(1)} & \star & \star & \cdots & & \cdots & \star \\ \mathbf{Z}_X^{(2)} & \mathbf{X}_1^{(2)} & \star & \star & & & \vdots \\ \mathbf{0} & \mathbf{Z}_X^{(3)} & \mathbf{X}_1^{(3)} & \star & & \star & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \star \\ \hline \vdots & & \mathbf{0} & \mathbf{Z}_X^{(i)} & & \mathbf{X}_1^{(i)} & \mathbf{X}_2^{(i)} \\ \mathbf{0} & \cdots & \cdots & \mathbf{0} & & \mathbf{X}_3^{(i)} & \mathbf{X}_4^{(i)} \end{array}\right),$$

$$\mathbf{H}_i^*\mathbf{Y}\mathbf{H}_i = \left(\begin{array}{cccc|cc} \mathbf{Y}_1^{(1)} & \star & \star & \cdots & \cdots & \star \\ \mathbf{Z}_Y^{(2)} & \mathbf{Y}_1^{(2)} & \star & \star & & \vdots \\ \mathbf{0} & \mathbf{Z}_Y^{(3)} & \mathbf{Y}_1^{(3)} & \star & \star & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \star \\ \hline \vdots & & \mathbf{0} & \mathbf{Z}_Y^{(i)} & \mathbf{Y}_1^{(i)} & \mathbf{Y}_2^{(i)} \\ \mathbf{0} & \cdots & \cdots & \mathbf{0} & \mathbf{Y}_3^{(i)} & \mathbf{Y}_4^{(i)} \end{array}\right), \tag{12}$$

where the size of the blocks are $r_1, r_2, \ldots, r_i, m-\rho_i$, $\rho_i = \sum_{j=1}^{i} r_i$, $\mathbf{H}_i$ is the unitary transformation matrix computed with the algorithm, $\mathbf{X}^{(i)}, \mathbf{Y}^{(i)}, \mathbf{Z}^{(i)}$ are the matrices $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ of the ith iterations, $\mathbf{X}_j^{(i)}, \mathbf{Y}_j^{(i)}$ ($j = 1, 2, 3, 4$) are the corresponding submatrices of $\mathbf{X}^{(i)}$ and $\mathbf{Y}^{(i)}$, and $\mathbf{Z}^{(i)} = \left(\mathbf{Z}_X^{(i)}, \mathbf{Z}_Y^{(i)}\right)$ for $i > 1$ where $\mathbf{Z}_X^{(i)}$ contains the reduced rows belonging to $\mathbf{X}_3^{(i-1)}$ and $\mathbf{Z}_Y^{(i)}$ contains the reduced rows belonging to $\mathbf{Y}_3^{(i-1)}$. By reduced row we mean that performing the unitary transformation of $\mathbf{X}_3^{(i-1)}$ ($\mathbf{Y}_3^{(i-1)}$) and removing the zero rows gives $\mathbf{Z}_X^{(i)}$ ($\mathbf{Z}_Y^{(i)}$).

The procedure terminates as $\mathbf{Z}^{(i+1)} = \mathbf{0}$:
It means that $\mathbf{X}_3^{(i)}$ and $\mathbf{Y}_3^{(i)}$ are zero. In this case the matrix has the required form and the relation with (11) is $n = \rho_i$, $\mathbf{U} = \mathbf{H}_i$, $\mathbf{X}_4 = \mathbf{X}_4^{(i)}$, $\mathbf{Y}_4 = \mathbf{Y}_4^{(i)}$.

The procedure terminates as $\mathbf{Z}^{(i+1)}$ has full rank:
We are going to show that if $\mathbf{Z}^{(i+1)}$ has full rank, then $m = n$. We use the result of Corollary 9 which shows that if a transformation of the required form is possible by some non-singular matrix $\mathbf{B}$, then a unitary matrix $\mathbf{U}$ exists which transforms the matrices $\mathbf{X}$ and $\mathbf{Y}$ into the same structure. In this proof we focus on the unitary matrix $\mathbf{U}$, as it is the primary result of SVD, but according to Corollary 9 the findings are valid for non-singular matrix $\mathbf{B}$ as well. Step 12 and 13 of STAIRCASE($\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$) generates $\mathbf{B}$ from $\mathbf{U}$.

We show the statement by contradiction. We assume that $\mathbf{Z}^{(i+1)}$ has full rank and a unitary matrix $\mathbf{U}$ exists which performs the transformation of (11). Since

$$\mathbf{U}^*\mathbf{X}\mathbf{U} = \underbrace{\mathbf{U}^*\mathbf{H}_i}_{\mathbf{G}^*}\mathbf{H}_i^*\mathbf{X}\mathbf{H}_i\underbrace{\mathbf{H}_i^*\mathbf{U}}_{\mathbf{G}} = \mathbf{G}^*\mathbf{H}_i^*\mathbf{X}\mathbf{H}_i\mathbf{G} \ ,$$

a unitary matrix $\mathbf{G}$ has to exist such that $\mathbf{U} = \mathbf{H}_i\mathbf{G}$. In the following we show in i steps that based on our assumptions $\mathbf{G}$ has the form

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_1^{(1)} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_1^{(i)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G}_4^{(i)} \end{pmatrix} , \tag{13}$$

where $\mathbf{G}_1$ is a $\rho_i \times \rho_i$ unitary matrix and $\mathbf{G}_4$ is a $(m-\rho_i) \times (m-\rho_i)$ unitary matrix. From

$$\mathbf{U}^*\mathbf{Z} = {\mathbf{G}^{(1)}}^*\mathbf{H}_i^*\mathbf{Z} = {\mathbf{G}^{(1)}}^* \begin{pmatrix} \mathbf{Z}_1^{(1)} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \star \\ \mathbf{0} \end{pmatrix},$$

and the fact that $\mathbf{Z}_1^{(1)}$ has full row rank, we have that the lower left $m - r_1 \times r_1$ size block of ${\mathbf{G}^{(1)}}^*$ is zero. Consequently, ${\mathbf{G}^{(1)}}^*$ and $\mathbf{G}^{(1)}$ have the form

$${\mathbf{G}^{(1)}}^* = \begin{pmatrix} {\mathbf{G}_1^{(1)}}^* & {\mathbf{G}_3^{(1)}}^* \\ \mathbf{0} & {\mathbf{G}_4^{(1)}}^* \end{pmatrix} \text{ and } \mathbf{G}^{(1)} = \begin{pmatrix} \mathbf{G}_1^{(1)} & \mathbf{0} \\ \mathbf{G}_3^{(1)} & \mathbf{G}_4^{(1)} \end{pmatrix} ,$$

where $\mathbf{G}_1^{(1)}$ is a $r_1 \times r_1$ matrix. Now from

$$\mathbf{I} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} = \mathbf{G}^{(1)}{\mathbf{G}^{(1)}}^* = \begin{pmatrix} \mathbf{I} & \mathbf{G}_1^{(1)}{\mathbf{G}_3^{(1)}}^* \\ \mathbf{G}_3^{(1)}{\mathbf{G}_1^{(1)}}^* & \star \end{pmatrix}$$

we have $\mathbf{G}_3^{(1)}{\mathbf{G}_1^{(1)}}^* = \mathbf{0} \Rightarrow \mathbf{G}_3^{(1)} = \mathbf{0} \Rightarrow {\mathbf{G}_3^{(1)}}^* = \mathbf{0}$, since $\mathbf{G}_1^{(1)}$ has full rank. As a result of the first step we have

$${\mathbf{G}^{(1)}}^* = \begin{pmatrix} {\mathbf{G}_1^{(1)}}^* & \mathbf{0} \\ \mathbf{0} & {\mathbf{G}_4^{(1)}}^* \end{pmatrix}$$

and we start the second step with

$${\mathbf{G}^{(2)}}^* = \begin{pmatrix} {\mathbf{G}_1^{(1)}}^* & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & {\mathbf{G}_1^{(2)}}^* & {\mathbf{G}_2^{(2)}}^* \\ \mathbf{0} & {\mathbf{G}_3^{(2)}}^* & {\mathbf{G}_4^{(2)}}^* \end{pmatrix} = \begin{pmatrix} \star & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \star & {\mathbf{G}_2^{(2)}}^* \\ \mathbf{0} & {\mathbf{G}_3^{(2)}}^* & \star \end{pmatrix} ,$$

where the size of the blocks are $r_1, r_2, m - \rho_2$, respectively.

From (11) we have $\mathbf{U}^*\mathbf{X}\mathbf{U} = {\mathbf{G}^{(2)}}^*\mathbf{H}_i^*\mathbf{X}\mathbf{H}_i\mathbf{G}^{(2)} = \begin{pmatrix} \star & \star & \star \\ \star & \star & \star \\ \mathbf{0} & \mathbf{0} & \star \end{pmatrix}$ and multiplying with ${\mathbf{G}^{(2)}}^*$ from the right it is ${\mathbf{G}^{(2)}}^*\mathbf{H}_i^*\mathbf{X}\mathbf{H}_i = \begin{pmatrix} \star & \star & \star \\ \star & \star & \star \\ \mathbf{0} & \mathbf{0} & \star \end{pmatrix}{\mathbf{G}^{(2)}}^*$. Substituting ${\mathbf{G}^{(2)}}^*$ and (12) into this equation we have

$$\begin{pmatrix} \star & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \star & {\mathbf{G}_2^{(2)}}^* \\ \mathbf{0} & {\mathbf{G}_3^{(2)}}^* & \star \end{pmatrix} \begin{pmatrix} \mathbf{X}_1^{(1)} & \star & \star \\ \mathbf{Z}_X^{(1)} & \star & \star \\ \mathbf{0} & \star & \star \end{pmatrix} = \begin{pmatrix} \star & \star & \star \\ \star & \star & \star \\ \mathbf{0} & \mathbf{0} & \star \end{pmatrix} \begin{pmatrix} \star & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \star & {\mathbf{G}_2^{(2)}}^* \\ \mathbf{0} & {\mathbf{G}_3^{(2)}}^* & \star \end{pmatrix} ,$$

where the relevant blocks of the right side are

$$\begin{pmatrix} \star & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \star & \mathbf{G}_2^{(2)^*} \\ \mathbf{0} & \mathbf{G}_3^{(2)^*} & \star \end{pmatrix} \begin{pmatrix} \mathbf{X}_1^{(1)} & \star & \star \\ \mathbf{Z}_X^{(1)} & \star & \star \\ \mathbf{0} & \star & \star \end{pmatrix} = \begin{pmatrix} \star & \star & \star \\ \star & \star & \star \\ \mathbf{0} & \star & \star \end{pmatrix} .$$

A completely analogous representation with matrix $\mathbf{Y}$ exists. To obtain the zero elements in the lower left corner of the matrix on the right side, we need $\mathbf{G}_3^{(2)^*}\mathbf{Z}_{\mathbf{X}}^{(2)} = \mathbf{G}_3^{(2)^*}\mathbf{Z}_{\mathbf{Y}}^{(2)} = \mathbf{0}$ which implies $\mathbf{G}_3^{(2)^*} = \mathbf{0}$ since $(\mathbf{Z}_{\mathbf{X}}^{(2)}, \mathbf{Z}_{\mathbf{Y}}^{(2)})$ has full row rank. Using the same arguments as above, from $\mathbf{I} = \mathbf{G}^{(2)}\mathbf{G}^{(2)^*}$ also $\mathbf{G}_2^{(2)^*}$ becomes $\mathbf{0}$ in $\mathbf{G}^{(2)^*}$. Using exactly the same line of argument we can continue with $\mathbf{Z}^{(3)}, \mathbf{Z}^{(4)}, \ldots$ until we finish with a structure of $\mathbf{G}$ as in (13). Since $\mathbf{G}$ is a unitary matrix its diagonal blocks have full rank. Using again the structure of $\mathbf{U}^*\mathbf{X}\mathbf{U}$ and $\mathbf{U}^*\mathbf{Y}\mathbf{U}$ from (11) for the block of size $(m - \rho_i) \times (r_i + r_i)$ we have

$$\mathbf{G}_4^{(i)^*} \left(\mathbf{X}_3^{(i)}, \mathbf{Y}_3^{(i)}\right) \mathbf{G}_1^{(i)} = \mathbf{0},$$

which cannot hold since $(\mathbf{X}_3^{(i)}\mathbf{Y}_3^{(i)})$ has full rank. Consequently, our assumptions are contradicted, and there is no matrix $\mathbf{U}$ that exists which transforms the $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$ matrices in the required from and therefore $n = m$ holds. ■

The computation of the matrix $\mathbf{B}$ from the matrices $(\mathbf{D}_0, \mathbf{D}_1)$ rather than from the matrices $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)$ and $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)$ is recommended since it is more efficient and numerically stable (see also [10, p. 116]). The algorithm can also be applied to compute a non-singular matrix $\mathbf{B}$ that observes the following relation.

$$\nu\mathbf{B} = (\phi, \mathbf{0}) \ , \ \mathbf{B}^{-1}\mathbf{D}_0\mathbf{B} = \begin{pmatrix} \mathbf{C}_0 & \mathbf{0} \\ \star & \star \end{pmatrix} \text{ and } \mathbf{B}^{-1}\mathbf{D}_1\mathbf{B} = \begin{pmatrix} \mathbf{C}_1 & \mathbf{0} \\ \star & \star \end{pmatrix} . \tag{14}$$

Instead of applying the algorithm to $(\mathbf{D}_0, \mathbf{D}_1, \mathbb{1}_m)$, it is called with $((\mathbf{D}_0)^T, (\mathbf{D}_1)^T, \nu^T)$ resulting in a matrix $\mathbf{B}$ that observes (14), if such a matrix exists.

Example 2 *We consider a MAP of size* 3 *that can be represented by an equivalent RAP of size* 2.

$$\mathbf{D}_0 = \begin{pmatrix} -5 & 1 & 0 \\ 3 & -3 & 0 \\ 1 & 1 & -5 \end{pmatrix} , \quad \mathbf{D}_1 = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} .$$

For this representation $r_C = 3$, *whereas* $r_O = 2$ *such that Theorem 4 applies. The embedded stationary initial vector equals* $\nu = (1/7, 1/7, 5/7)$. *With the staircase algorithm we obtain the following unitary matrix*

$$\mathbf{U} = \begin{pmatrix} -0.68041 & -0.19245 & -0.70711 \\ 0.68041 & -0.19245 & 0.70711 \\ -0.27217 & -0.96225 & 0 \end{pmatrix} ,$$

which is presented here with 5 *significant digits of accuracy. Since* $\mathbf{U}^* \, \mathbb{1}_3 = (1.08866, -1.34715, 0)^T$, *matrix* $\mathbf{U}^*$ *has to be normalized resulting in the transformation matrix*

$$\mathbf{B}^{-1} = \begin{pmatrix} 0.62500 & 0.62500 & -0.25001 \\ 0.14286 & 0.14286 & 0.71429 \\ -0.70711 & 0.70711 & 0 \end{pmatrix} \Leftrightarrow \mathbf{B} = \begin{pmatrix} 0.74074 & 0.25926 & -0.70710 \\ 0.74074 & 0.25926 & 0.70710 \\ -0.29629 & 1.2963 & 0 \end{pmatrix} ,$$

such that

$$\begin{aligned} \mathbf{B}^{-1}\mathbf{D}_0\mathbf{B} &= \begin{pmatrix} -2.5926 & 0.84261 & 0 \\ 1.6931 & -4.4074 & 0 \\ 2.0951 & 7.3331 & -6 \end{pmatrix} , & \mathbf{B}^{-1}\mathbf{D}_1\mathbf{B} &= \begin{pmatrix} -1.0370 & 2.7870 & 0 \\ 0.67724 & 2.0370 & 0 \\ 0.83805 & -3.6665 & 0 \end{pmatrix} \\ \nu\mathbf{B} &= (0, 1, 0) \ , & \mathbf{B}^{-1} \, \mathbb{1}_3 &= (1, 1, 0)^T \end{aligned}$$

and

$$\mathbf{C}_0 = \begin{pmatrix} -2.5926 & 0.84261 \\ 1.6931 & -4.4074 \end{pmatrix} , \quad \mathbf{C}_1 = \begin{pmatrix} -1.0370 & 2.7870 \\ 0.67725 & 2.0370 \end{pmatrix}$$

is the resulting equivalent RAP of minimal order. This example indicates that the STAIRCASE method is not meant for finding a MAP representation. For that purpose the procedure of [20] can be applied. In case of $(\mathbf{C}_0, \mathbf{C}_1)$ *this procedure finds the following equivalent MAP representation*

$$\mathbf{C}'_0 = \begin{pmatrix} -2.1713 & 0.421305 \\ 1.15009 & -4.82871 \end{pmatrix}, \quad \mathbf{C}'_1 = \begin{pmatrix} 0.3565 & 1.3935 \\ 3.035 & 0.6435 \end{pmatrix}.$$

Example 3 *We now modify the MAP slightly by modifying one element in* $\mathbf{D}_0$ *and* $\mathbf{D}_1$*.*

$$\mathbf{D}_0 = \begin{pmatrix} -5 & 1 & 0 \\ 3 & -3 & 0 \\ 1 & 1 & -4 \end{pmatrix}, \quad \mathbf{D}_1 = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

In this case $r_C = 2$ *and by the staircase algorithm we obtain*

$$\mathbf{C}_0 = \begin{pmatrix} -6 & 1.17157 \\ 0 & -2 \end{pmatrix} \text{ and } \mathbf{C}_1 = \begin{pmatrix} 0 & 4.82843 \\ 0 & 2 \end{pmatrix}.$$

Now for representation $(\mathbf{C}_0, \mathbf{C}_1)$ *we have* $r_O = 1$ *which means that the MAP is equivalent to a Poisson process with rate* 2*. Consequently,* $(\mathbf{D}_0, \mathbf{D}_1)$ *is a representation which contains redundancy according to both, the initial and the closing vector. The first step, STAIRCASE*$(\mathbf{D}_0, \mathbf{D}_1, \mathbb{1})$*, eliminates the redundancy due to the closing vector, while the second step STAIRCASE*$(\mathbf{C}_0^T, \mathbf{C}_1^T, \phi^T)$ *eliminates the redundancy due to the initial vector.*

6 Approximation of BMAP/MAP/1 Output Processes

In [14, 22] Heindl and his coworkers present an approach to approximate the output process of a MAP/MAP/1 and a BMAP/MAP/1 queue by a finite MAP. For these examples it has been found that the resulting MAPs or RAPs are often redundant and can be reduced with the algorithm proposed here. We consider in our example only the MAP/MAP/1 case but the approach can be easily extended to arrivals from a BMAP. Let the service process be MAP$(\mathbf{S}_0, \mathbf{S}_1)$ and the arrival process MAP$(\mathbf{A}_0, \mathbf{A}_1)$. The underlying Markov process has a quasi birth and death structure with generator matrix

$$\mathbf{Q} = \begin{pmatrix} \tilde{\mathbf{L}} & \mathbf{F} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{B} & \mathbf{L} & \mathbf{F} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{B} & \mathbf{L} & \mathbf{F} & \mathbf{0} & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix},$$

where $\tilde{\mathbf{L}} = \mathbf{A}_0 \otimes \mathbf{I}$, $\mathbf{F} = \mathbf{A}_1 \otimes \mathbf{I}$, $\mathbf{L} = \mathbf{A}_0 \oplus \mathbf{S}_0$ and $\mathbf{B} = \mathbf{I} \otimes \mathbf{S}_1$. The exact departure process can be described as MAP of infinite size with the matrices

$$\mathbf{D}_o = \begin{pmatrix} \tilde{\mathbf{L}} & \mathbf{F} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{L} & \mathbf{F} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{L} & \mathbf{F} & \mathbf{0} & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \text{ and } \mathbf{D}_1 = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{B} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{B} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}.$$

The following finite truncation up to population n has been defined in [22, 14] for the process.

$$\mathbf{D}_o^n = \begin{pmatrix} \tilde{\mathbf{L}} & \mathbf{F} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{L} & \mathbf{F} & \mathbf{0} & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{L} & \mathbf{F} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \vdots & \cdots & \mathbf{0} & \mathbf{L} & \mathbf{F} \\ \mathbf{0} & \cdots & \cdots & \cdots & \mathbf{0} & \mathbf{L}+\mathbf{F} \end{pmatrix} \text{ and } \mathbf{D}_1^n = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{B} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ddots & \vdots \\ \mathbf{0} & \mathbf{B} & \mathbf{0} & \mathbf{0} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \mathbf{0} & \mathbf{B} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \cdots & \cdots & \mathbf{0} & \mathbf{B}-\mathbf{FG} & \mathbf{FG} \end{pmatrix},$$

where $\mathbf{G}$ is the solution of the matrix quadratic equation $\mathbf{B} + \mathbf{L}\mathbf{G} + \mathbf{F}\mathbf{G}^2 = \mathbf{0}$. Observe that $(\mathbf{D}_0^n, \mathbf{D}_1^n)$ is not necessarily a MAP since the resulting matrices may contain negative elements outside the diagonal of $\mathbf{D}_0^n$. We now show an example where the resulting representation is non-minimal.

Example 4 *Consider a queue with arrival MAP*

$$\mathbf{A}_0 = \begin{pmatrix} -6.9375 & 0.9375 \\ 0.0625 & -0.1958 \end{pmatrix}, \quad \mathbf{A}_1 = \begin{pmatrix} 6 & 0 \\ 0 & 0.1333 \end{pmatrix}$$

and service MAP

$$\mathbf{S}_0 = \begin{pmatrix} -16 & 3 \\ 0 & -2 \end{pmatrix}, \quad \mathbf{S}_1 = \begin{pmatrix} 6.5 & 6.5 \\ 1 & 1 \end{pmatrix}.$$

For truncation level $n = 2$, the following representation for the output process is computed.

$$\mathbf{D}_0 = \begin{pmatrix} -6.9375 & 0 & 0.9375 & 0 & 6 & 0 & 0 & 0 \\ 0 & -6.9375 & 0 & 0.9375 & 0 & 6 & 0 & 0 \\ 0.0625 & 0 & -0.1958 & 0 & 0 & 0 & 0.1333 & 0 \\ 0 & 0.0625 & 0 & -0.1958 & 0 & 0 & 0 & 0.1333 \\ 0 & 0 & 0 & 0 & -16.938 & 3 & 0.9375 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2.9375 & 0 & 0.9375 \\ 0 & 0 & 0 & 0 & 0.0625 & 0 & -16.063 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0.0625 & 0 & -2.0625 \end{pmatrix},$$

$$\mathbf{D}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4.3570 & 4.3575 & -0.8575 & -0.8575 & 2.1425 & 2.1425 & 0.8575 & 0.8575 \\ -6.94e-2 & -6.94e-2 & -1.9306 & -1.9306 & 1.0694 & 1.0694 & 1.9306 & 1.9306 \\ -3.25e-4 & -3.25e-4 & 6.4337 & 6.4337 & 3.25e-4 & 3.25e-4 & 6.63e-2 & 6.63e-2 \\ -7.32e-4 & -7.32e-4 & 0.9341 & 0.9341 & 7.32e-4 & 7.32e-4 & 6.59e-2 & 6.59e-2 \end{pmatrix}.$$

The rank of this representation is 6 such it can be equivalently represented by the matrices.

$$\mathbf{C}_0 = \begin{pmatrix} -17.437 & -1.3856 & 1.2600 & 1.5045 & -2.2367 & 9.4778 \\ -0.7536 & -1.9889 & -1.6629 & -1.5543 & -5.1252 & 4.6997 \\ -2.7774 & 1.5484 & -6.5427 & -1.8290 & -5.8846 & 7.0416 \\ 0.00167 & 0.5878 & -3.6457 & -2.5127 & -5.7433 & 1.0085 \\ -0.1702 & -0.9751 & -1.3438 & 0.9071 & -14.863 & -1.2237 \\ 0.0000 & 0.0000 & 0.0000 & 0.1200 & -2.9253 & -1.7892 \end{pmatrix},$$

$$\mathbf{C}_1 = \begin{pmatrix} 1.0029 & -5.2946 & -5.1786 & 0.4482 & 2.1492 & 15.690 \\ -0.3269 & 1.46335 & 1.3392 & -4.4856 & -2.3315 & 10.726 \\ 0.1271 & -0.5211 & -0.4571 & -0.4191 & 0.01373 & 9.7002 \\ -0.1453 & 0.7544 & 0.7335 & -2.0075 & -1.0883 & 12.0418 \\ -0.1203 & 0.5500 & 0.5081 & -5.8666 & -2.5615 & 25.159 \\ 0.0000 & 0.0000 & 0.0000 & -0.9334 & -0.3760 & 5.9039 \end{pmatrix}.$$

7 Extension to batch and marked RAPs

MAPs and RAPs can be used to model processes with a single type of event, which result in single arrivals or services. It is natural to extend MAPs to generate batches of arrivals or multiple event types. This resulted in the definition of batch MAPs (BMAPs) [16, 18] and marked MAPs (MMAPs) [12].

Similar to the extension from MAPs to BMAPs, the class of *batch rational arrival processes* (BRAPs) has been introduced in [4]. We consider BRAPs with at most K arrivals at a time that are defined by a set of $K + 1$ matrices $(\mathbf{D}_0, \dots, \mathbf{D}_K)$, and also similar to the extension from MAPs to MMAPs, we define a *marked rational arrival process* (MRAP) with K event types by a set of $K + 1$ matrices $(\mathbf{D}_0, \dots, \mathbf{D}_K)$ such that

1. $(\mathbf{D}_0 + \sum_{k=1}^{K} \mathbf{D}_k) \mathbb{1} = \mathbf{0}$,
2. all eigenvalues of $\mathbf{D}_0$ have a negative real part which implies that the matrix is non-singular [15],
3. $\mathbf{P} = -\mathbf{D}_0^{-1} \left(\sum_{k=1}^{K} \mathbf{D}_k \right)$ has a unique eigenvalue 1 such that the solution $\nu \mathbf{P} = \nu$, $\nu \mathbb{1} = 1$ is unique, and
4. the function
$$f_{(\mathbf{D}_0,\ldots,\mathbf{D}_K)}(t_1, k_1, \ldots, t_j, k_j) = \nu e^{\mathbf{D}_0 t_1} \mathbf{D}_{k_1} e^{\mathbf{D}_0 t_2} \mathbf{D}_{k_2} \ldots e^{\mathbf{D}_0 t_j} \mathbf{D}_{k_j} \mathbb{1} \tag{15}$$
is a valid joint density for all $t_i \geq 0$ and $k_i \in \{1, \ldots, K\}$ $(i = 1, \ldots, j)$. That is $f_{(\mathbf{D}_0,\ldots,\mathbf{D}_K)}(t_1, k_1, \ldots, t_j, k_j) \geq 0$ and $\sum_{k_1} \cdots \sum_{k_j} \int_{t_1} \cdots \int_{t_j} f_{(\mathbf{D}_0,\ldots,\mathbf{D}_K)}(t_1, k_1, \ldots, t_j, k_j)\, dt_j \ldots dt_1 = 1$.

We assume a finite K in this paper, because even though our results are valid for $K = \infty$, the resulting procedure (for example, checking the rank of an infinite matrix) is impractical. Observe that the stationary vector of an MRAP, ν, and any matrix $\mathbf{D}_k$, $k = 0, \ldots, K$ may contain negative elements and the diagonal of $\mathbf{D}_0$ may contain positive elements. The class of BRAPs (MRAPs) contains BMAPs (MMAPs). The conditions of the Markovian behavior are $\nu \geq \mathbf{0}$, $\mathbf{D}_k \geq 0$ for $k = 1, \ldots, K$ and all non-diagonal elements of $\mathbf{D}_0$ are non-negative. To avoid duplicate references from now on we refer only to MRAPs, noting that the extensions presented apply to BRAPs in the same way. The extensions of the required definition and basic theorems to MRAPs is straightforward.

Definition 4 *Two MRAPs, MRAP$(\mathbf{D}_0, \ldots, \mathbf{D}_K)$ and MRAP$(\mathbf{C}_0, \ldots, \mathbf{C}_K)$, are* equivalent *if and only if all joint density functions are identical (cf., (15)).*

Definition 5 *The* size *of the representation $(\mathbf{D}_0, \ldots, \mathbf{D}_K)$ is the size of the square matrix $\mathbf{D}_0$.*

Definition 6 *A representation $(\mathbf{D}_0, \ldots, \mathbf{D}_K)$ of size m is* minimal*, if no other equivalent representation $(\mathbf{C}_0, \ldots, \mathbf{C}_K)$ of size $n < m$ exists.*

Theorem 7 *If there is a matrix $\mathbf{V} \in \mathbb{R}^{m,n}$ such that $\mathbb{1}_m = \mathbf{V} \mathbb{1}_n$, $\mathbf{D}_k \mathbf{V} = \mathbf{V} \mathbf{C}_k$, $k = 0, \ldots, K$ and $\nu \mathbf{V} = \phi$, where $\nu (-\mathbf{D}_0)^{-1} \left(\sum_{k=1}^{K} \mathbf{D}_k \right) = \nu$ and $\phi (-\mathbf{C}_0)^{-1} \left(\sum_{k=1}^{K} \mathbf{C}_k \right) = \phi$, then $(\mathbf{D}_0, \ldots, \mathbf{D}_K)$ and $(\mathbf{C}_0, \ldots, \mathbf{C}_K)$ are equivalent.*

Proof. The proof follows the same pattern as the one in Theorem 1. ■

Theorem 8 *If there is a matrix $\mathbf{W} \in \mathbb{R}^{n,m}$ such that $\mathbb{1}_n = \mathbf{W} \mathbb{1}_m$, $\mathbf{W} \mathbf{D}_k = \mathbf{C}_k \mathbf{W}$, $k = 0, \ldots, K$ and $\nu = \phi \mathbf{W}$ where $\nu (-\mathbf{D}_0)^{-1} \left(\sum_{k=1}^{K} \mathbf{D}_k \right) = \nu$ and $\phi (-\mathbf{C}_0)^{-1} \left(\sum_{k=1}^{K} \mathbf{C}_k \right) = \phi$, then $(\mathbf{D}_0, \ldots, \mathbf{D}_K)$ and $(\mathbf{C}_0, \ldots, \mathbf{C}_K)$ are equivalent.*

Proof. The proof follows the same pattern as the one in Theorem 2. ■

7.1 Minimal representation of MRAPs

Let $\mathbf{M} = -\mathbf{D}_0^{-1}$ and $\mathbf{P}_k = \mathbf{M} \mathbf{D}_k$, $k = 1, \ldots, K$, where $(\mathbf{M}, \mathbf{P}_1, \ldots, \mathbf{P}_K)$ and $(\mathbf{D}_0, \ldots, \mathbf{D}_K)$ mutually define each other since $\mathbf{D}_0 = -\mathbf{M}^{-1}$ and $\mathbf{D}_k = -\mathbf{D}_0 \mathbf{P}_k$. For an MRAP $(\mathbf{D}_0, \ldots \mathbf{D}_K)$ of size m define the following matrices which are natural extensions of the corresponding matrices defined in Section 4 for RAPs.

$$\mathcal{C}_{\mathbf{D}_0} = \mathcal{C}_{\mathbf{D}_0}(0) = \left(\mathbb{1}, \mathbf{M} \mathbb{1}, \ldots, \mathbf{M}^{m-1} \mathbb{1} \right),$$
$$\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_K} = \mathcal{C}_{\mathbf{D}_0,\mathbf{D}_K}(0) = \left(\mathcal{C}_{\mathbf{D}_0}, \mathbf{P}_1 \mathcal{C}_{\mathbf{D}_0}, \mathbf{P}_1^2 \mathcal{C}_{\mathbf{D}_0}, \ldots, \mathbf{P}_1^{m-1} \mathcal{C}_{\mathbf{D}_0}, \mathbf{P}_2 \mathcal{C}_{\mathbf{D}_0}, \ldots, \mathbf{P}_{K-1}^{m-1} \mathcal{C}_{\mathbf{D}_0}, \mathbf{P}_K \mathcal{C}_{\mathbf{D}_0}, \ldots, \mathbf{P}_K^{m-1} \mathcal{C}_{\mathbf{D}_0} \right),$$
$$\mathcal{C}_{\mathbf{D}_0}(n{+}1) = \left(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_K}(n), \mathbf{M} \mathcal{C}_{\mathbf{D}_0,\mathbf{D}_K}(n), \ldots, \mathbf{M}^{m-1} \mathcal{C}_{\mathbf{D}_0,\mathbf{D}_K}(n) \right),$$
$$\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_K}(n{+}1) = \left(\mathcal{C}_{\mathbf{D}_0}(n{+}1), \mathbf{P}_1 \mathcal{C}_{\mathbf{D}_0}(n{+}1), \ldots, \mathbf{P}_1^{m-1} \mathcal{C}_{\mathbf{D}_0}(n{+}1), \ldots, \mathbf{P}_K \mathcal{C}_{\mathbf{D}_0}(n{+}1), \ldots, \mathbf{P}_K^{m-1} \mathcal{C}_{\mathbf{D}_0}(n{+}1) \right),$$

$\mathcal{C}_{\mathbf{D}_0}(0)$ is of size $m \times m$, $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_K}(0)$ is of size $m \times m^2K$, $\mathcal{C}_{\mathbf{D}_0}(1)$ is of size $m \times m^3K$ and $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_K}(1)$ is of size $m \times m^4K^2$. In general, $\mathcal{C}_{\mathbf{D}_0}(n)$ is an $m \times \left(m^2K\right)^n m$ matrix and $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)$ is an $m \times \left(m^2K\right)^{n+1}$ matrix. The rank of the generalized controllability matrix is $r_C = rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_K}(m))$. Similar to the RAP case this is a simple, but very redundant definition of r_C, but it can be computed in a much more efficient way with the following algorithm.

RC_RANK_K($\mathbf{D}_0, \ldots, \mathbf{D}_K$)
  $r_C^o = 0$, $\mathbf{M} = -\mathbf{D}_0^{-1}$, FOR $k = 1$ TO K DO $\mathbf{P}_k = \mathbf{M}\mathbf{D}_k$,
  $r_C = 1$, $\mathbf{Z} = \mathbb{1}$, $\mathbf{\Theta} = \mathbf{Z}$, $\mathbf{H} = \mathbf{M}$,
  FOR $n = 0$ TO m DO
    WHILE $r_c < rank(\mathbf{\Theta}, \mathbf{HZ})$ DO $\mathbf{\Theta} = (\mathbf{\Theta}, \mathbf{HZ})$, $\mathbf{H} = \mathbf{HM}$, $r_c = rank(\mathbf{\Theta})$;
    /* r_C = rank of $\mathcal{C}_{\mathbf{D}_0}(n)$ */
    IF $r_C == r_C^o$ OR $r_C == m$ THEN RETURN(r_C);
    $r_C^o = r_C$;
    FOR $k = 1$ TO K DO
      $\mathbf{Z}$ =SPANNING($\mathbf{\Theta}$); $\mathbf{H} = \mathbf{P}_k$;
      WHILE $r_c < rank(\mathbf{\Theta}, \mathbf{HZ})$ DO $\mathbf{\Theta} = (\mathbf{\Theta}, \mathbf{HZ})$, $\mathbf{H} = \mathbf{HP}_k$, $r_c = rank(\mathbf{\Theta})$;
    /* r_C = rank of $\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1}(n)$ */
    IF $r_C == r_C^o$ OR $r_C == m$ THEN RETURN(r_C);
    $r_C^o = r_C$; $\mathbf{Z}$ =SPANNING($\mathbf{\Theta}$); $\mathbf{H} = \mathbf{M}$;
  ENDFOR

Similarly, let

$$\mathcal{O}_{\mathbf{D}_0} = \mathcal{O}_{\mathbf{D}_0}(0) = \begin{pmatrix} \nu \\ \nu\mathbf{M} \\ \vdots \\ \nu\mathbf{M}^{m-1} \end{pmatrix}, \quad \mathcal{O}_{\mathbf{D}_0,\mathbf{D}_K} = \mathcal{O}_{\mathbf{D}_0,\mathbf{D}_K}(0) = \begin{pmatrix} \mathcal{O}_{\mathbf{D}_0} \\ \mathcal{O}_{\mathbf{D}_0}\mathbf{P}_1 \\ \vdots \\ \mathcal{O}_{\mathbf{D}_0}\mathbf{P}_1^{m-1} \\ \vdots \\ \mathcal{O}_{\mathbf{D}_0}\mathbf{P}_K^{m-1} \end{pmatrix},$$

$$\mathcal{O}_{\mathbf{D}_0}(n+1) = \begin{pmatrix} \mathcal{O}_{\mathbf{D}_0,\mathbf{D}_K}(n) \\ \mathcal{O}_{\mathbf{D}_0,\mathbf{D}_K}(n)\mathbf{M} \\ \vdots \\ \mathcal{O}_{\mathbf{D}_0,\mathbf{D}_K}(n)\mathbf{M}^{m-1} \end{pmatrix}, \quad \mathcal{O}_{\mathbf{D}_0,\mathbf{D}_K}(n+1) = \begin{pmatrix} \mathcal{O}_{\mathbf{D}_0}(n+1) \\ \mathcal{O}_{\mathbf{D}_0}(n+1)\mathbf{P}_1 \\ \vdots \\ \mathcal{O}_{\mathbf{D}_0}(n+1)\mathbf{P}_1^{m-1} \\ \vdots \\ \mathcal{O}_{\mathbf{D}_0}(n+1)\mathbf{P}_K^{m-1} \end{pmatrix},$$

where $\nu\mathbf{P} = \nu$, $\nu\,\mathbb{1} = 1$. The rank of the generalized observability matrix is $r_O = rank(\mathcal{O}_{\mathbf{D}_0,\mathbf{D}_K}(m))$. With these definition of r_C and r_O Theorem 5 remains valid for MRAPs.

7.2 The staircase algorithm for MRAPs

The computation of a smaller representation of an MRAP with representation $(\mathbf{D}_0, \ldots, \mathbf{D}_K)$ according to Theorem 3 can be done with the following version of the staircase algorithm. It should be called as STAIRCASE_K($\mathbf{D}_0, \mathbf{D}_1, \ldots, \mathbf{D}_K, \mathbb{1}_m$) to remove redundancy according to the closing vector or as STAIRCASE_K($(\mathbf{D}_0)^T, (\mathbf{D}_1)^T, \ldots, (\mathbf{D}_K)^T, \nu^T$) to remove redundancy according to the initial vector.

STAIRCASE_K($\mathbf{X}_0$, $\mathbf{X}_1$, …, $\mathbf{X}_K$, $\mathbf{Z}$)

  $i = 0$; $\{m, j\}$ =SIZE($\mathbf{X}_0$); $\mathbf{U}^* = \mathbf{I}$;

  REPEAT

    $i = i + 1$; $r_i = rank(\mathbf{Z})$; $\{\mathbf{U}_i, \mathbf{S}_i, \mathbf{T}_i\}$ =SVD($\mathbf{Z}$) ;

    $\begin{pmatrix} \hat{\mathbf{Z}} \\ \mathbf{0} \end{pmatrix} = \mathbf{U}_i^* \mathbf{Z}$; FOR $k = 0$ TO K DO $\begin{pmatrix} \hat{\mathbf{X}}_k & \check{\mathbf{X}}_k \\ \tilde{\mathbf{X}}_k & \bar{\mathbf{X}}_k \end{pmatrix} = \mathbf{U}_i^* \mathbf{X}_k \mathbf{U}_i$;

    /* $\hat{\mathbf{X}}_k$ are of size $r_i \times r_i$ */

    $\mathbf{U}^* = \begin{pmatrix} \mathbf{I}_{\sum_{j=1}^{i} r_j} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_i^* \end{pmatrix} \mathbf{U}^*$;

    $\mathbf{Z} = \begin{pmatrix} \tilde{\mathbf{X}}_0 \; \tilde{\mathbf{X}}_1 \; \dots \; \tilde{\mathbf{X}}_K \end{pmatrix}$; FOR $k = 0$ TO K DO $\mathbf{X}_k = \bar{\mathbf{X}}_k$;

  UNTIL $rank(\mathbf{Z}) = m - \sum_{j=1}^{i} r_j$ or $\mathbf{Z} = \mathbf{0}$;

  IF ($\mathbf{Z} = \mathbf{0}$) THEN

    $n = \sum_{j=1}^{i-1} r_j$;

    $\begin{pmatrix} \mathbf{x} \\ \mathbf{0}_{m-n} \end{pmatrix} = \mathbf{U}^* \mathbb{1}_m$;

    IF ($\mathbf{x} \neq \mathbf{0}$) THEN $\mathbf{R} = \mathbf{I}$ ELSE $\mathbf{R}$ = non-singular matrix such that $\mathbf{Rx} \neq \mathbf{0}$

    /* element-wise */

    $\mathbf{y} = \mathbf{Rx}$; $\mathbf{\Gamma} = diag(\mathbf{y})$; $\mathbf{B} = \left[\begin{pmatrix} \mathbf{\Gamma}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m-n} \end{pmatrix} \begin{pmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m-n} \end{pmatrix} \mathbf{U}^* \right]^{-1}$;

  ELSE $n = m$; $\mathbf{B} = \mathbf{I}$; /* no reduction is possible */

  RETURN($n, \mathbf{B}$) ;

Example 5 *We consider the following MMAP with 3 states and the matrices*

$$\mathbf{D}_0 = \begin{pmatrix} -6 & 0 & 0 \\ 0 & -5 & 2 \\ 0 & 0 & -3 \end{pmatrix}, \; \mathbf{D}_1 = \begin{pmatrix} 0 & 1 & 5 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \; \mathbf{D}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

After removing the redundancy due to the closing vector we obtain the following MRAP of size 2

$$\mathbf{D}_0 = \begin{pmatrix} -5.8165 & -0.63299 \\ -0.8165 & -3.18350 \end{pmatrix}, \; \mathbf{D}_1 = \begin{pmatrix} -2.8777 & 9.47698 \\ -0.5443 & 3.87766 \end{pmatrix}, \; \mathbf{D}_2 = \begin{pmatrix} 0.061168 & -0.2110 \\ -0.27217 & 0.9388 \end{pmatrix}.$$

8 Conclusions

This paper presents explicit conditions to decide if a RAP representation is minimal and a computational method for creating a minimal representation if it is not the case.

It turns out that, in contrast with the preliminary expectations based on [20], the rank of a RAP cannot be decided based on $rank(\mathcal{C}_{\mathbf{D}_0,\mathbf{D}_1})$ and $rank(\mathcal{O}_{\mathbf{D}_0,\mathbf{D}_1})$ alone, both matrices are necessary to compute a minimal representation. The minimization of RAPs is important because then they can be used in a more efficient system analysis. As shown by the example of finding the minimal representation of RAPs is underlined by an example in which the output process of an MAP/MAP/1 queue, it is sometimes possible to find a reduced equivalent representation which can then be used in further analysis steps for example as input process of another queue in tandem system. In a final step we extended the class of RAPs to RAPs with multiple entity types denoted as MRAPs or BRAPs. It is shown that the conditions for minimality representations and also the minimization algorithms can be naturally extended to this more general case.

The paper presents one step in the more general context of using RAPs rather than only Markovian processes in performance and dependability analysis. This extension is attractive since RAPs are more general in particular for a fixed size of the state space. In this paper we showed how to reduce a MAP to a RAP such that the state space size is minimal, other aspects are exact and approximate analysis techniques for the resulting processes as they are available for Markov processes but still require research work in the case of RAPs.

References

[1] N. Akar. Solving the ME/ME/1 queue with state-space methods and the matrix sign function. *Performance Evaluation*, 63:131–145, 2006.

[2] S. Asmussen and M. Bladt. Renewal theory and queueing algorithms for matrix-exponential distributions. In *Matrix-analytic methods in stochastic models*, volume 183 of *Lecture Notes in Pure and Appl. Math.*, pages 313–341, 1997.

[3] S. Asmussen and M. Bladt. Point processes with finite-dimensional conditional probabilities. *Stochastic Processes and their Application*, 82:127–142, 1999.

[4] N. G. Bean and B. F. Nielsen. Quasi-birth-and-death processes with rational arrival process components. *Stochastic Models*, 26(3):309–334, 2010.

[5] L. Bodrog, A. Horvath, and M. Telek. Moment characterization of matrix exponential and Markovian arrival processes. *Annals of Operations Research*, 160:51–58, 2008.

[6] P. Buchholz and M. Telek. Stochastic Petri nets with matrix exponentially distributed firing times. *Performance Evaluation*, 67(12):1373–1385, 2010.

[7] C. Commault and S. Mocanu. Phase-type distributions and representations: some open problems for system theory. *Int. J. Control*, 76(6):566–580, 2003.

[8] J. M. Davis, I. A. Gravagne, B. J. Jackson, and R. J. Marks II. Controllability, observability, realizability and stability of dynamic linear systems. *Electronic Journal of Differential Equations*, (37):1–32, 2009.

[9] B. de Schutter. Minimal state-space realization in linear system theory: An overview. *Journal of Computational and Applied Mathematics*, 121(1-2):331–354, 2000.

[10] P. M. Van Dooren. The generalized eigenstructure on linear system theory. *IEEE Trans. on Automatic Control*, 26(1):111–129, 1981.

[11] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, 3 edition, 1996.

[12] Q. M. He and M. Neuts. Markov arrival processes with marked transitions. *Stochastic Processes and their Applications*, 74:37–52, 1998.

[13] Q. M. He and H. Zhang. On matrix exponential distributions. *Advances in Applied Probability*, 39:271–292, 2007.

[14] A. Heindl, Q. Zhang, and E. Smirni. ETAQA truncation models for the MAP/MAP/1 departure process. In *QEST*, pages 100–109, 2004.

[15] L. Lipsky. *Queueing Theory: A Linear Algebraic Approach*. Springer, 2008.

[16] D. M. Lucantoni. New results on the single server queue with a batch Markovian arrival process. *Stochastic Models*, 7:1–46, 1991.

[17] K. Mitchell. Constructing a correlated sequence of matrix exponentials with invariant first order properties. *Operations Research Letters*, 28:27–34, 2001.

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

[19] C. Art O'Cinneide. Phase-type distributions: Open problems and a few properties. *Stochastic Models*, 15(4):731–757, 1999.

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

[21] A. Varga. Minimal realization procedures based on balancing and related techniques. In F. Pichler and R. Moreno-Díaz, editors, *EUROCAST*, volume 585 of *Lecture Notes in Computer Science*, pages 733–761. Springer, 1991.

[22] Q. Zhang, A. Heindl, and E. Smirni. Characterizing the BMAP/MAP/1 departure process via the ETAQA truncation. *Stochastic Models*, 21(2-3):821–846, 2005.