J. KSIAM Vol.20, No.3, 225–242, 2016 http://dx.doi.org/10.12941/jksiam.2016.20.225

AN OPTIMIZATION APPROACH FOR COMPUTING A SPARSE MONO-CYCLIC POSITIVE REPRESENTATION

KYUNGSUP KIM

DEPARTMENT OF COMPUTER ENGINEERING, CHUNGNAM NATIONAL UNIVERSITY, DAEJEON, 305-764, KOREA

E-mail address: kskim@cse.cnu.ac.kr

ABSTRACT. The phase-type representation is strongly connected with the positive realization in positive system. We attempt to transform phase-type representation into sparse mono-cyclic positive representation with as low order as possible. Because equivalent positive representations of a given phase-type distribution are non-unique, it is important to find a simple sparse positive representation with lower order that leads to more effective use in applications. A Hypo-Feedback-Coxian Block (HFCB) representation is a good candidate for a simple sparse representation. Our objective is to find an HFCB representation with possibly lower order, including all the eigenvalues of the original generator. We introduce an efficient nonlinear optimization method for computing an HFCB representation from a given phase-type representation. We discuss numerical problems encountered when finding efficiently a stable solution of the nonlinear constrained optimization problem. Numerical simulations are performed to show the effectiveness of the proposed algorithm.

1. INTRODUCTION

The positive realization problem has been independently studied in the system theory and probability theory communities [1, 2, 3, 4]. The phase-type representation (so called, positive representation) is strongly connected with the positive realization in system theory [5]. The relation between the Laplace-Stieltjes transform (LST) of a probability distribution and a corresponding phase-type representation is similar to that between a transfer function and a corresponding state space realization.

Because equivalent positive representations of a given phase-type(PH) distribution are non-unique [6], finding smaller and simpler matrix representations for PH-distributions has become an important theoretical and practical issue. In modeling and simulation applications, the importance of simple sparse representations comes from the fact that the computational complexity of PH-distributed random-variate generation depends on the representation [7, 8]. Finding

Received by the editors August 3 2016; Revised September 9 2016; Accepted in revised form September 9 2016; Published online September 13 2016.

2000 *Mathematics Subject Classification.* 60J27, 93B15.

Key words and phrases. Phase-type representation, Hypo-feedback-Coxian block, Constrained nonlinear optimization, Positive realization, Positive system.

smaller, simpler and structured representations with lower order can lead to more effective use of PH-distribution random number generators in stochastic modeling and simulation.

There are some important approaches for minimal and specially structured positive realization such as the triangular, bi-diagonal, mono-cyclic, and uni-cyclic representations [5, 9]. Every phase-type distribution has a mono-cyclic representation with larger order [9]. The mono-cyclic representation is a natural extension of a generator with a bi-diagonal matrix. The mono-cyclic representations are referred to as feed-back Erlang representations (or feed-back Coxian representations) due to their structures [9]. We introduce a Hypo-Feedback-Coxian Block (HFCB) representation covering Erlang, Coxian, Feedback Erlang (FE) and Feedback Coxian (FC) representations. These approach can be applied to the sparse positive realization of positive system since it has the same structure [3, 5]. Similar constructive methods of simple compact positive realizations for general transfer functions have been extensively discussed in the positive system [2, 4, 10, 11]. The main objective of this paper is to compute an HFCB positive representation with a lower order using nonlinear optimization method.

Finding the roots of polynomials is an important task for finding a proper HFCB representation. A coefficient relation function between the roots and coefficients of a polynomial can be used to find all of the roots of a given polynomial simultaneously and in parallel, instead of using a deflation method finding the roots one by one such as Laguerre's and Newton-Raphson's [12]. A constrained nonlinear optimization problem computing a sparse mono-cyclic representation is formulated by using coefficient relation function. We discuss the numerical method to find a optimal solution of the constrained nonlinear optimization problem depending on how to choose an initial value.

The rest of this paper is organized as follows. Section 2 provides some relevant background material, including definitions and preliminary results. In Section 3, various structured phase-type representations classified into acyclic and cyclic forms and their properties are considered. Numerical optimization methods for finding a sparse mono-cyclic representation with possibly small order are discussed in Section 4. The numerical experiments are presented in Section 5. Finally, Section 6 concludes the paper.

2. PHASE-TYPE DISTRIBUTIONS

Before proceeding, we introduce some basic notations. An $n \times n$ non-negative matrix A is denoted by $A \geq 0$ if its entries are non-negative and at least one entry is positive. A strict positive matrix $A = [a_{ij}]$ is denoted $A > 0$ if all entries $a_{ij} > 0$. We discuss the phase-type distribution for a non-negative random variable X in terms of a continuous-time Markov process. A continuous-time Markov process is defined on an $n + 1$ finite state space. The row vector α gives the initial probability vector. A phase-type (PH) distribution is defined as the distribution of the time needed for absorption in a Continuous-Time Markov Chain (CTMC) with one absorbing state. If the $n+1$ state is an absorbing state and all other states are transient, the infinitesimal generator matrix of the Markov chain in the form of the augmented matrix

tuple $(\bar{T}, \bar{b}, \bar{\alpha})$, called the augmented phase-type representation, can be defined by

$$\begin{aligned} \bar{T} &= \begin{bmatrix} T & -T\mathbb{1} \\ \mathbf{0} & 0 \end{bmatrix}, \quad \bar{b} = \mathbf{e}_{n+1}, \\ \bar{\alpha} &= \begin{bmatrix} \alpha & \alpha_{n+1} \end{bmatrix}, \end{aligned} \tag{2.1}$$

where $\mathbf{0}$ refers to the column vector, row vector or matrix with all entries equal to zero in the case without ambiguity, $\mathbf{e}_k$ is a k-th standard basis vector with zeros, except the k-th one, which is 1, and $\mathbb{1}$ is the column vector with all entries being one. We can see that $\alpha_{n+1} = 0$ if $\alpha\mathbb{1} = 1$, and $\alpha_{n+1} = 1 - \alpha\mathbb{1}$ otherwise. PH distributions are commonly represented by a vector-matrix tuple (α, T) that describes the transient part of the CTMC. The vector-matrix tuple (α, T) denotes a phase-type representation (or a phase-type generator) of a phase-type distribution if α and T have following properties: $T_{ii} < 0$, $T_{ij} \geq 0$ for $i \neq j$, $\alpha\mathbb{1} \leq 1$, $\alpha \geq 0$, and $T\mathbb{1} \leq 0$. Equivalently, a phase-type representation is regarded as the positive realization (representation) of positive system [5].

The probability density function (PDF), cumulative distribution function (CDF), and Laplace-Stieltjes Transform (LST) of the PDF, respectively, are defined by

$$\begin{aligned} f(x) &= \alpha \exp(Tx)(-T\mathbb{1}) \\ F(x) &= 1 - \alpha \exp(Tx)\mathbb{1} \\ F^*(s) &= E(\exp(sX)) = \alpha(sI - T)^{-1}(-T\mathbb{1}) \end{aligned}$$

where $E(X)$ means an expectation of a random variable X.

The particular structured representation problem has been studied as an important work in the theory of phase-type representations [5]. Depending on their structures, we can typically divide a phase-type representation into two classes: an acyclic subclass and a cyclic subclass. In the concept of graph theory, the acyclic class consists of a vector-matrix pair (α, T), which the generator T represents an acyclic transition graph and α is a positive vector.

The concepts of PH-simplicity and PH-majorization are useful tools in the study of PH-distributions [6, 13]. For a given T, $PH(T)$ denotes the set of all distributions with a phase-type representation (α, T). For two generators T and S, S is said to PH-majorize T if $PH(T) \subset PH(S)$. T is called by PH-simple if for any different α and β, (α, T) and (β, T) represent different distributions. S PH-majorizes T if and only if there exists a non-negative matrix P such that

$$TP = PS, \quad P\mathbb{1} = \mathbb{1}. \tag{2.2}$$

This means that (α, T) and $(\alpha P, S)$ represent the same phase-type distribution. We note that a non-negative condition of P is not necessary for the existence of a new phase-type representation $(\alpha P, S)$. The non-negative condition of P is replaced by $\alpha P \geq 0$.

A multiset is defined by a collection of elements in which certain elements may occur more than once. Let $|A|$ be the cardinality of A. A multiset $A \uplus B$ denotes an additive union (counting multiplicities) in multiset such that $|A \uplus B| = |A| + |B|$.

3. Specially structured positive representations

3.1. **Coxian representation.** Any PH-distribution with a triangular representation (α, T) has a bi-diagonal representation $(\beta, S(\vec{\lambda}))$ with the same order. Every phase-type representation whose LST has only real poles has an ordered Coxian representation with larger order than the original [13]. By using an optimization method, several algorithms have been introduced for finding a Coxian generator, $S(\lambda)$ PH-majorizing a PH-generator [14].

We consider the computation problem for PH-majorizing $S(\lambda)$ for a PH-generator (α, T) such that for a given vector $\vec{\lambda} = (\lambda_1, \cdots, \lambda_m)$, $TP = PS(\vec{\lambda})$ and a bi-diagonal matrix $S(\vec{\lambda})$ is defined by

$$S(\vec{\lambda}) = \begin{bmatrix} -\lambda_1 & \lambda_1 & \cdots & 0 & 0 \\ 0 & -\lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & -\lambda_{m-1} & \lambda_{m-1} \\ 0 & 0 & \cdots & 0 & -\lambda_m \end{bmatrix}.$$

A spectral polynomial algorithm for computing a canonical bi-diagonal representation of a state space representation has been proposed [14]. We assume that all eigenvalues of T are nonzero and real-valued. The spectral polynomial algorithm is given as follows:

(1) Find the descending ordered multiset $\{-\lambda_1, -\lambda_2, \cdots, -\lambda_m\}$ including all the eigenvalues (counting multiplicities) of T.
(2) Let $p_m = -T\mathbb{1}/\lambda_m$. For $2 \le k \le m$, recursively, set $k = k - 1$ and compute p_k such that
$$p_k = (\lambda_{k+1} I + T) p_{k+1} / \lambda_k.$$
(3) Construct the bi-diagonal matrix $S(\vec{\lambda})$ for $\vec{\lambda} = (\lambda_1, \cdots, \lambda_m)$ and $n \times m$ matrix $P = \begin{bmatrix} p_1 & p_2 & \cdots & p_m \end{bmatrix}$.
(4) Check that $TP - PS(\vec{\lambda}) = \mathbf{0}$ and $\beta = \alpha P \ge 0$.

PH-generators with only real eigenvalues can be transformed into Coxian generators by the spectral algorithm [14]. By using the spectral polynomial algorithm, a Coxian representation is induced based on the next theorem.

Theorem 3.1. *We assume that a vector-matrix pair (α, T) is a PH-generator such that all the eigenvalues of T are real and $\alpha \ge 0$. Then, a Coxian representation $(\beta, S(\vec{\lambda}))$ is an equivalent bi-diagonal representation of a vector-matrix pair (α, T) such that $TP = PS(\vec{\lambda})$, $\mathbb{1} = P\mathbb{1}$ and $\beta = \alpha P$.*

3.2. **Monocyclic positive representation.** We attempt to find sparse representations representing phase-type distributions with complex and real poles. The presence of complex poles in LST implies the presence of backward transitions in the associated Markov chain. Several candidates for canonical sparse positive realization, such as feedback Erlang representation [13], mono-cyclic Coxian representation [9] and unicyclic representation [15, 16], have been introduced.

An $n \times n$ matrix A is said to be reducible if there exists a permutation matrix P such that $P^tAP = \begin{bmatrix} B & C \\ \mathbf{0} & D \end{bmatrix}$ where B and D are square matrices and P^t is the transpose of P. Otherwise, A is said to be irreducible . Define E_{ij} be the $n \times n$ matrix that has only a one in the position (i, j) and zeros otherwise. If A is irreducible and for any $a_{ij} \neq 0$, $A - a_{ij}E_{ij}$ is reducible, then A is said to be nearly reducible. Suppose A is an $n \times n$ nearly reducible matrix [17]. Then, there exist permutation matrices P and Q and an integer $s > 1$ such that

$$PAQ = \begin{bmatrix} A_1 & B_1 & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_2 & \cdots & \mathbf{0} & \mathbf{0} \\ & & \ddots & & \\ \mathbf{0} & \mathbf{0} & \cdots & A_{s-1} & B_{s-1} \\ B_s & \mathbf{0} & \cdots & \mathbf{0} & A_s \end{bmatrix} \tag{3.1}$$

where each B_i has exactly one entry equal to one and each A_i is nearly reducible. When all B_i's are 1×1, this is a simplified form for nearly reducible matrices. We note that this form is a good candidate as a simple sparse representation of a phase-type distribution with complex poles.

The simplest cyclic representations are obtained by adding a feedback on an Erlang representation, which is called a Feedback Erlang (FE) representation or mono-cyclic representation [16, 9]. A new type of representation can be obtained by adding a feedback to the Coxian representation. It is specially denoted as a Feedback Coxian (FC) representation if the transition matrix of a vector matrix pair can be obtained by adding feedback to the Coxian representation. Thus, for a given vector $\vec{\lambda} = (\lambda_1, \cdots, \lambda_m)$ and $0 \leq z < 1$, an FC representation $(\alpha, M(\vec{\lambda}, z))$ is defined by

$$M(\vec{\lambda}, z) = \begin{bmatrix} -\lambda_1 & \lambda_1 & \cdots & 0 & 0 \\ 0 & -\lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\lambda_{m-1} & \lambda_{m-1} \\ z\lambda_m & 0 & \cdots & 0 & -\lambda_m \end{bmatrix} \tag{3.2}$$

We can see that the matrix $M(\vec{\lambda}, z)$ is a simple nearly reducible matrix.

Lemma 3.2. *Let P be a permutation matrix and $\vec{\lambda}' = \vec{\lambda}P$. Then an FC representation $(\alpha, M(\vec{\lambda}', z))$ has the same eigenvalue multiset.*

Proof. The characteristic polynomial $p(s) = \det(sI - M(\vec{\lambda}, z))$ is given by $p(s) = (s + \lambda_1)(s + \lambda_2) \cdots (s + \lambda_m) - \theta$ where $\theta = z\lambda_1\lambda_2 \cdots \lambda_m$. Since the characteristic polynomial is independent of the order of the entries of $\vec{\lambda}$, the FC representation $(\alpha, M(\vec{\lambda}', z))$ has the same eigenvalue multiset. □

We can resort $\vec{\lambda}$ in ascending order representing the given characteristic polynomial. FC representations have more free parameters than FE representations. We define a Hypo-Feedback-Coxian Block (HFCB) representation as a generalized form, covering Erlang, Coxian, FE and FC block representation.

Definition 3.1. *A Hypo-Feedback-Coxian Block (HFCB) representation is defined as a transient generator having the following structure*

$$\mathbf{M} = \begin{bmatrix} M_1 & M_1^* & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & M_2 & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & M_{J-1} & M_{J-1}^* \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & M_J \end{bmatrix} \tag{3.3}$$

where M_i is a feedback Coxian $M(\vec{\lambda}, z)$ or Coxian $S(\lambda)$ for each i and M_j^* is defined by $M_j^* = (-M_j)\mathbb{1}\mathbf{e}_1$, where $\mathbf{e}_1$ is the first standard basis vector.

In the above definition, if cyclic M_i is an FE, it is denoted by a Hypo-Feedback-Erlang Block (HFEB) as a special case of HFCB [8]. It has been proven that every phase-type representation can be transformed into a mixture of HFEB representations [5, 9].

Algorithm 1 A cyclic spectral polynomial algorithm

Input: (α, T)
Output: P and $\beta = \alpha P$

(1) Choose an appropriate HFCB generator $\mathbf{M}$ such that the eigenvalue multiset of $\mathbf{M}$ includes all eigenvalues of T counting multiplicities.
(2) Give an initial column vector $p_m = -T\mathbb{1}$.
(3) ζ_k is defined by the number of nonzero elements of the k-th column of $\mathbf{M}$ and $c_{j,k}$ is the (j,k)-th element of $\mathbf{M}$. For a given p_k and $1 < k \leq m$, recursively, set $k = k - 1$ and compute p_k as follows
   (a) if $\zeta_k = 2$,
   $$p_k = (\lambda_{k+1} I + T) p_{k+1} / \lambda_k \tag{3.4}$$
   (b) if $\zeta_k = 3$ and $c_{k+n',k} \neq 0$,
   $$p_k = ((\lambda_{k+1} I + T) p_{k+1} - c_{k+n',k+1} p_{k+n'}) / c_{k,k+1} \tag{3.5}$$
(4) Construct $P = \begin{bmatrix} p_1 & p_2 & \cdots & p_m \end{bmatrix}$
(5) Finally, check that $TP = P\mathbf{M}$ and $\alpha P \geq 0$, and return $\beta = \alpha P$ and P.

We propose a cyclic spectral polynomial algorithm to compute a transform matrix P generating $\mathbf{M}$ from a representation (α, T) in Algorithm 1. Finally, check two facts: $TP = P\mathbf{M}$ and $\alpha P \geq 0$. In the next section, we will discuss how to find an appropriate HFCB generator $\mathbf{M}$ with lower order in the first step in Algorithm 1.

Theorem 3.3. *We assume that a vector-matrix pair (α, T) is a PH-generator such that all the eigenvalues of T are real and complex numbers and $\alpha \geq 0$. Then the matrix P satisfies that $TP = P\mathbf{M}$ and the eigenvalue multiset (counting multiplicities) of T is included all eigenvalues of the HFCB matrix $\mathbf{M}$ in Eq. (3.3) .*

Proof. First, we assume that an eigenvalue multiset (counting multiplicities) of T is included in the multiset of eigenvalues of the HFCB matrix $\mathbf{M}$. However, it is postponed how to find $\mathbf{M}$ satisfying the assumption to the next section. A matrix P is computed by Algorithm 1. We will show that we can obtain P such that $TP = P\mathbf{M}$. We have $TP = \begin{bmatrix} Tp_1 & Tp_2 & \cdots & Tp_{m-1} & Tp_m \end{bmatrix}$ and set $p_m = -T\mathbb{1}$. The last column of $\mathbf{M}$ has only two nonzero values. From Algorithm 1, we obtain $Tp_m = \lambda_{m-1}p_{m-1} - \lambda_m p_m$. For $1 < k < m$, since k-th column of $\mathbf{M}$ has two or three nonzero elements, we have two cases $Tp_{k+1} = c_{k,k+1}p_k - \lambda_{k+1}p_{k+1} + c_{k+n',k+1}p_{k+n'}$ or $Tp_{k+1} = \lambda_k p_k - \lambda_{k+1}p_{k+1}$ in each step. For the first column of $\mathbf{M}$, we have two cases $Tp_1 = -\lambda_1 p_1 + c_{1+n',1}p_{1+n'}$ or $Tp_1 = -\lambda_1 p_1$. Therefore, we can see that $TP = P\mathbf{M}$. □

4. Optimization method for finding a sparse HFCB representation

We consider the properties of FE blocks and FC blocks. For a given FE block $M(\vec{\lambda}, z)$ with $\lambda_i = \lambda > 0$ and a constant λ, the eigenvalue multiset of $M(\vec{\lambda}, z)$ can be computed explicitly in a closed form [9].

Lemma 4.1. [9] *Assume that we have a real root $\mu_0 = -\lambda + \lambda z^{1/m}$ and a complex root $\mu_1 = a + ib$. Then each n-th root μ_k of its characteristic polynomial equation of an FE representation $M(\vec{\lambda}, z)$ is simply obtained by*

$$\mu_k = -\lambda(1 - z^{\frac{1}{m}} \cos\frac{2k\pi}{m}) + iz^{\frac{1}{m}}\lambda \sin\frac{2k\pi}{m} \tag{4.1}$$

for $0 \leq k < m$, where $\vec{\lambda} = (\lambda, \cdots, \lambda)$, and the parameters of the representation can uniquely be determined by

$$\lambda = \frac{1}{2}\left(2a - b\tan\frac{\pi}{m} + b\cot\frac{\pi}{m}\right) \tag{4.2}$$

$$z = \frac{b(\tan\frac{\pi}{m} + \cot\frac{\pi}{m})}{2\lambda} \tag{4.3}$$

for a given μ_1 and an integer $m > 2$.

For a complex root $\mu_1 = a + ib$, we can obtain an FE block $M(\vec{\lambda}, z)$ such that $\vec{\lambda} = (\lambda, \cdots, \lambda)$ and z are defined by Eqs (4.2) and (4.3). We consider a lower bound problem of the order of a phase-type generator T including a given complex eigenvalue. It is related to the continuous version of famous Kolmogorov's remark: "for a fixed integer m, what is the multiset of all complex numbers that are eigenvalues of an order-n stochastic matrix?" [13].

Lemma 4.2. [13] *Let (α, T) be a generator of phase-type distribution with a maximal real eigenvalue μ_0 and with any complex eigenvalue $\mu_1 = -a + bi$ $(a > 0)$. Then the order m of T satisfies*

$$\frac{|b|}{a - |\mu_0|} \leq \cot \frac{\pi}{m} \tag{4.4}$$

It has been shown that a wedge as in (4.4) contains all eigenvalues of T using an invariant polytope argument [13]. By using the result of Lemma 4.2, we can determine the lower bound of the order of an FC block including μ_1 because an FC block is regarded as a special case of a phase-type representation. Since the eigenvalue multiset of FE block computed by using Lemma 4.1 covers at most two or three eigenvalues of the original matrix T, the total order of $\mathbf{M}$ gets larger. On other hand, because FC blocks have more free parameters, we can find an FC block including more common eigenvalues of T than that of FE block. So we can reduce the total order of $\mathbf{M}$.

The relationship of the roots of a polynomial equation and its coefficients has been studied in the complex analysis field [18]. It is closely related to a typical theorem of [18], which shows a connection between the roots of the polynomial and the generator $(\vec{\lambda}, z)$ of an FC matrix $M(\vec{\lambda}, z)$.

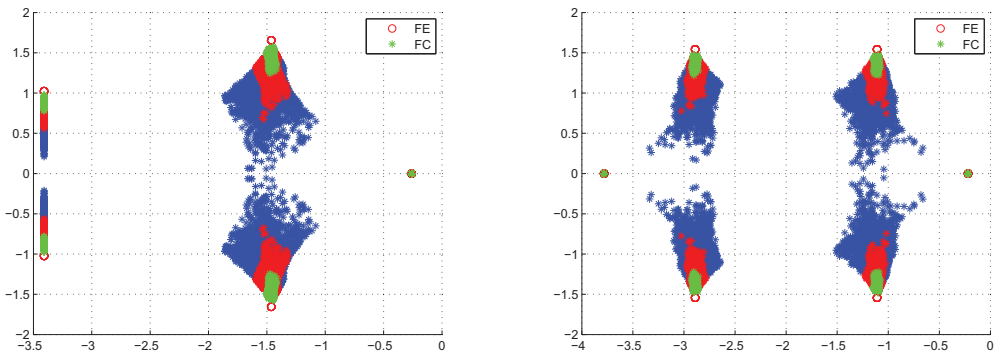
Lemma 4.3. *Assume that a diagonal vector $\vec{\lambda} = (\lambda_1, \lambda_2, \cdots, \lambda_m)$ of the mono-cyclic matrix $M(\vec{\lambda}, z)$ in (3.2) lies in a circular region whose center is $\bar{\lambda}$ and a given radius $r > 0$.*

1. *Then, all of the roots $\{\mu_1, \cdots, \mu_m\}$ of the polynomial equation $p(s) = (s + \lambda_1)(s + \lambda_2) \cdots (s + \lambda_m) - \theta = 0$ lie on or within one of the n circles, which have the common radius r and whose centers, $\bar{\mu}_k$'s, are given by $\bar{\mu}_k = -\bar{\lambda} + \exp(\frac{2ik\pi}{m})|\theta|^{1/m}$, where $\theta = z\lambda_1\lambda_2 \cdots \lambda_m$.*
2. *If these m circles are mutually external, each circle contains precisely one of the roots $\{\mu_1, \cdots, \mu_m\}$ of the polynomial equation $p(s)$.*
3. *If $\bar{\lambda} = E(\lambda_k)$, then we have $|\theta|^{1/m} \leq |\bar{\lambda}| z^{\frac{1}{n}}$, and equality holds if and only if $\lambda_k = \bar{\lambda}$ for all k.*

Proof. Proofs for the first and the second can be shown by using the results of [18]. The arithmetic mean is larger than or equal to the geometric mean and the equality only holds if $\lambda_k = \bar{\lambda}$ for all k. The third is trivial. □

Lemma 4.3 shows the connection between $\{\lambda_k\}$ in $M(\vec{\lambda}, z)$ and the root multiset $\{\mu_k\}$ of $p(s)$. With the help of a numerical procedure, we can depict the feasible regions of the eigenvalues of $M(\vec{\lambda}, z)$ for a given r.

Example 4.1. *By using a random number generator, for given $0 \leq z < 1$ and $r > 0$, we generate a random number vector $\vec{\lambda}$ such that $\bar{\lambda} - r \leq \lambda_k \leq \bar{\lambda} + r$ and $\bar{\lambda} = E(\lambda_k)$. We compute all the roots $\{\mu_k\}$ of the characteristic polynomial $p(s)$ depending on a random number vector $\vec{\lambda}$ with respect to $z = 0.5$ and $\bar{\lambda} = 2$. Figure 1 shows a relationship between the eigenvalues $\{\mu_k\}$ of an FC block $M(\vec{\lambda}, z)$ and a random number vector $\vec{\lambda}$ for $n = 5$ or $n = 6$. In Fig. 1,*



(a) $n = 5$ (b) $n = 6$

FIGURE 1. The location of the eigenvalues $\{\mu_k\}$'s of FE, and FC for orders $n = 5$ and $n = 6$

the green dots, the red dots and the blue dots denote the eigenvalues points corresponding to $0.8 \leq r \leq 1$, $1.12 \leq r \leq 1.4$, *and* $1.44 \leq r \leq 1.8$, *respectively. The diamond points present the eigenvalues of an FE block computed by Eq. (4.1).*

From Example 4.1, we can predict the moving pattern of the eigenvalues corresponding to the change of $(\vec{\lambda}, z)$. We can verify that there will be just one root in each circle with a given radius r, remarked in Lemma 4.3. From the simulation, we can see that the eigenvalue points $\{\mu_k\}$ are distributed to the inside direction from the vertices of the polygons.

An eigenvalue multiset $\mathcal{A} = \{\nu_1, \cdots, \nu_n\}$ (counting multiplicities) is a multiset of eigenvalues of T. An eigenvalue multiset $\mathcal{B} = \{\mu_1, \cdots, \mu_m\}$ (counting multiplicities) is the multiset of eigenvalues of the HFCB matrix $\mathbf{M}$. Let $\vec{\nu} = (\nu_1, \cdots, \nu_n)$ be an eigenvalue vector such that ν_i for each i is an eigenvalue of T. The multiset $\mathcal{B}$ must include the eigenvalue multiset $\mathcal{A}$. An eigenvalue multiset $\mathcal{B}^{(i)} = \{\mu_1^{(i)}, \cdots, \mu_{n_i}^{(i)}\}$ (counting multiplicities) denotes the eigenvalue multiset of an FC block $M(\vec{\lambda}^{(i)}, z^{(i)})$ for each $1 \leq i \leq r$. A vector pair $(\vec{\lambda}^{(i)}, z^{(i)})$ denotes a generator matrix of $M(\vec{\lambda}^{(i)}, z^{(i)})$ for each $1 \leq i \leq r$. Let a generator $\Upsilon = \{(\vec{\lambda}^{(1)}, z^{(1)}), \cdots, (\vec{\lambda}^{(r)}, z^{(r)})\}$ generate the HFCB matrix $\mathbf{M}$. The problem can be reformulated as follows.

Problem 4.1. *Find a generator* $\Upsilon = \{(\vec{\lambda}^{(1)}, z^{(1)}), \cdots, (\vec{\lambda}^{(r)}, z^{(r)})\}$ *to minimize the order of the HFCB matrix* $\mathbf{M}$ *such that*

$$\mathcal{A} \subset \mathcal{B} = \biguplus_{i=1}^{r} \mathcal{B}^{(i)} \tag{4.5}$$

where $\mathcal{B}^{(i)}$ *is the eigenvalue multiset of* $M(\vec{\lambda}^{(i)}, z^{(i)})$.

We note that the problem is closely related to the simultaneous root finding problem of a polynomial [12]. Finding the roots of polynomials is an important task for various areas of signal processing, such as filter and wavelet design, spectral estimation, phase unwrapping, and communication. The constrained relation between the roots and coefficients of a polynomial can be used to find all of the roots of a given polynomial simultaneously, instead of using sequential methods, such as Laguerre's and Newton-Raphson's [12]. Since $p_i(s)$ the characteristic is polynomial of each FC block matrix $M(\vec{\lambda}^{(i)}, z^{(i)})$, the characteristic function $p(s)$ of the HFCB matrix $\mathbf{M}$ is $p(s) = p_1(s)p_2(s)\cdots p_r(s)$.

Lemma 4.4. *For any generator $(\vec{\lambda}^{(i)}, z^{(i)}) \in \Upsilon$, assume that $\mathcal{B}^{(i)}$ denotes the eigenvalue multiset (counting multiplicities) of the sub-block $M(\vec{\lambda}^{(i)}, z^{(i)})$ of the HFCB matrix $\mathbf{M}$. Then it satisfies the coefficient relation equations such that each ϕ_k is defined by*

$$\begin{aligned}
\phi_1 &\triangleq \sum_{k=1}^{m_i} \lambda_k^{(i)} - \sum_{k=1}^{m_i} (-\mu_k^{(i)}) = 0 \\
\phi_2 &\triangleq \sum_{k<j}^{m_i} \lambda_k^{(i)} \lambda_j^{(i)} - \sum_{i<j}^{m_i} (-1)^2 \mu_k^{(i)} \mu_j^{(i)} = 0 \\
&\vdots \\
\phi_{m_i} &\triangleq (1 - z^{(i)}) \lambda_1^{(i)} \cdots \lambda_{m_i}^{(i)} - (-1)^{m_i} \mu_1^{(i)} \cdots \mu_{m_i}^{(i)} = 0
\end{aligned} \tag{4.6}$$

for each $\mu_k^{(i)} \in \mathcal{B}^{(i)}$.

Proof. We will show that for given generator Υ and $\mathcal{B}^{(i)}$, the coefficient relation equations between roots and coefficients of a polynomial as (4.6) can be defined. The characteristic polynomial $p_i(s) = \det(sI - M(\vec{\lambda}^{(i)}, z^{(i)})))$ for each i is given by

$$p_i(s) = (s + \lambda_1^{(i)})(s + \lambda_2^{(i)}) \cdots (s + \lambda_{m_i}^{(i)}) - \theta \tag{4.7}$$

where $\theta = z^{(i)} \lambda_1^{(i)} \lambda_2^{(i)} \cdots \lambda_{m_i}^{(i)}$ and m_i is an order of $M(\vec{\lambda}^{(i)}, z^{(i)})$. The eigenvalue multiset of $M(\vec{\lambda}^{(i)}, z^{(i)})$ is equal to the root multiset of the polynomial $p_i(s)$. The polynomial $p_i(s)$ can be rewritten in the form

$$p_i(s) = \prod_{k=1}^{m_i} (s - \mu_k^{(i)}). \tag{4.8}$$

All coefficient relation functions ϕ_i's can be defined by comparing with coefficients of (4.7) and (4.8) with respect to $(\vec{\lambda}^{(i)}, z^{(i)}) \triangleq (\lambda_1^{(i)}, \cdots, \lambda_{m_i}^{(i)}, z^{(i)})$ and $\vec{\mu}^{(i)} \triangleq (\mu_1^{(i)}, \mu_2^{(i)} \cdots, \mu_{m_i}^{(i)})$, respectively. Therefore we can obtain the relation equations, ϕ_k's, between roots and coefficients of a polynomial, comparing the coefficients induced from the eigenvalue multiset $\mathcal{B}^{(i)}$ with the coefficients of the characteristic polynomial $p_i(s)$ of each FC block matrix $M(\vec{\lambda}^{(i)}, z^{(i)})$. □

$\mathcal{B}^{(i)}$ is partitioned into a fixed multiset $\mathcal{F}^{(i)}$ and a flexible multiset $\mathcal{U}^{(i)} = \mathcal{B}^{(i)} - \mathcal{F}^{(i)}$. Let us set $\vec{\mathbf{s}} = (\vec{\lambda}^{(i)}, z^{(i)})$. For each $\mathcal{U}^{(i)}$, we define a vector $\vec{\mathbf{u}}$ by $\vec{\mathbf{u}} = (u_1, \cdots, u_d)$ for $u_k \in \mathcal{U}^{(i)}$

and $d = |\mathcal{U}^{(i)}|$ where $|A|$ is defined by the cardinality of A. A free parameter vector $\vec{\mathbf{x}}$ is defined by $\vec{\mathbf{x}} = (Re(\vec{\mathbf{u}}), Im(\vec{\mathbf{u}}))$. We define a fixed vector $\vec{\mathbf{x}}_c = (Re(\vec{\mathbf{v}}), Im(\vec{\mathbf{v}}))$, where $\vec{\mathbf{v}} = (v_1, \cdots, v_{d'})$ for each $v_k \in \mathcal{F}^{(i)}$.

For all integers with $1 \leq k \leq m_i$, the coefficient relation functions ϕ_k can be defined in the form, $\phi_k(\vec{\mathbf{x}}, \vec{\mathbf{x}}_c, \vec{\mathbf{s}})$. Set $\vec{\phi} = (\phi_1, \cdots, \phi_{m_i})$ as in (4.6). The coefficient relation vector function $\vec{\phi} : R^{\tilde{n}} \to R^{m_i}$ represents smooth nonlinear functions defined by Eq. (4.6), where $\tilde{n}$ is the dimension of the vector $(\vec{\mathbf{x}}, \vec{\mathbf{x}}_c, \vec{\mathbf{s}})$. We formulate a constrained nonlinear optimization problem to compute an FC block $M(\vec{\lambda}^{(i)}, z^{(i)})$ whose eigenvalue multiset includes $\mathcal{F}^{(i)}$.

Problem 4.2. *For each multiset $\mathcal{F}^{(i)}$, find a solution $(\vec{\mathbf{x}}, \vec{\mathbf{s}})$ of a nonlinear objective function $f(\vec{\mathbf{x}}, \vec{\mathbf{s}})$ subject to equation or inequality constraints such that*

$$\begin{aligned} \underset{\vec{\mathbf{x}}, \vec{\mathbf{s}}}{\textit{min}} \quad & f(\vec{\mathbf{x}}, \vec{\mathbf{s}}) \\ \textit{s.t.} \quad & b_l \leq \vec{\mathbf{x}} \leq b_u \\ & 0 < \vec{\mathbf{s}} \leq c_u \\ & \vec{\phi}(\vec{\mathbf{x}}, \vec{\mathbf{x}}_c, \vec{\mathbf{s}}) = 0. \end{aligned} \tag{4.9}$$

for a fixed parameter $\vec{\mathbf{x}}_c$ where b_l is a lower bound vector, b_u is an upper bound vector, and the objective function $f(\vec{\mathbf{x}}, \vec{\mathbf{s}})$ in (4.9) is given by

$$f(\vec{\mathbf{x}}, \vec{\mathbf{s}}) = -\sum_{k=1}^{m_i} \ln \lambda_k. \tag{4.10}$$

Now we consider what reasonable constraints are given in the constrained optimization problem in (4.9). Because $M(\vec{\lambda}^{(i)}, z^{(i)})$ and T are phase-type, all the real values of eigenvalues are negative, i.e., $Re(\mu_k^{(i)}) < 0$ for any $\mu_k^{(i)} \in \mathcal{B}^{(i)}$. We can set the upper bounds zero $b_u = 0$ for the real numbers, because $Re(\mu_k^{(i)}) < 0$. The other bounds, b_l and b_u, are given in a proper bound from initial values for the algorithm to be convergence. We can see that the entries $\vec{\lambda}^{(i)}$ are positive and $0 \leq z^{(i)} < 1$. The upper bound c_u for $\vec{\lambda}^{(i)}$ are given in an appropriate bound from initial values.

Remark 4.5.
- *We can assume that first and second derivatives of the objective functions and constraints are available because $\vec{\phi}$ consists of polynomials.*
- *We note that roots of the polynomial in (4.8) depend continuously on its coefficients. That is well-known as the use of Rouche's Theorem* [19]*. A small change in the coefficients of $p(s)$ causes only a small change in the roots of the polynomial.*
- *Using the result of Lemma 4.4, Problem 4.1 can be reformulated in the form of a constrained nonlinear optimization problem for the coefficient relation functions with respect to $(\vec{\mathbf{x}}, \vec{\mathbf{s}})$ in Problem 4.2. By solving a constrained nonlinear optimization in Problem 4.2, we find a generator $(\vec{\lambda}^{(i)}, z^{(i)})$ and $\mathcal{U}^{(i)}$ for a fixed $\mathcal{F}^{(i)}$.*

Our objective is to find a HFCB representation with possible lower order, whose eigenvalue multiset includes all eigenvalues of T. We will discuss numerical problems for finding a HFCB

representation with lower order by using a nonlinear optimization method. Our strategy is to obtain the optimal solutions by extending some partially known solutions recursively as follows:

(1) Initially, compute an FE block $M(\vec{\lambda}^{(i)}, z^{(i)})$ such that $\mathcal{F}^{(i)} = \mathcal{A} \bigcap \mathcal{B}^{(i)}$ is nonempty, using Lemma 4.1.
(2) We choose a target multiset $\mathcal{T}$ such that $\mathcal{A} \supset \mathcal{T} \supset \mathcal{F}^{(i)}$. Set $\mathcal{F}_k = \mathcal{F}^{(i)}$
(3) We choose a new multiset $\mathcal{F}_k (\subset \mathcal{T})$ corresponding to $\vec{\mathbf{x}}_c$ after $\vec{\mathbf{x}}_c$ moves a little and solve Problem 4.2 for a new multiset $\mathcal{F}_k$, recursively. If Problem 4.2 is solved finally, then we can obtain a generator $(\vec{\lambda}^{(i)}, z^{(i)})$ such that $\mathcal{F}^{(i)} = \mathcal{T} = \mathcal{A} \bigcap \mathcal{B}^{(i)}$.
(4) If there is more additional target multiset $\mathcal{T}$, then go to Step 2. Otherwise, go to the next step.
(5) Finally, we obtain a $M(\vec{\lambda}^{(i)}, z^{(i)})$ whose eigenvalue multiset includes $\mathcal{F}^{(i)} = \mathcal{A} \bigcap \mathcal{B}^{(i)}$

We need to discuss the efficient numerical algorithm at the third step in the above method. While a nonlinear optimization method is a powerful tool to find an optimal solution, there are several potential difficulties to solve some problems such as divergence and instability. Most of the numerical methods for solving the nonlinearly constrained optimization problems meet a danger converging to a local minimal point or diverging, depending on initial values.

We propose a **linear search algorithm** in Algorithm 2. We can generate a sequence of iterates that converges to an optimal solution provided that the initial value is sufficiently close to that optimal solution. A basic idea is to make shift by following a straight line to a final point from an initial point knowing a solution, and find a solution of the next stage using the previous value as an initial value.

We choose a target multiset $\mathcal{T} \subset \mathcal{A}$, to which the multiset $\mathcal{F}_k$ expands in each iteration. Initially, assume that an optimal solution $\vec{\mathbf{x}}_{c,0}$ of Problem 4.2 is already computed and can be used as an initial vector, because the μ_k's in (4.8) are explicitly computed by using Lemma 4.1. n_d denotes the number of segments. It is related with the length of change. If the procedure fails, then in order to avoid diverging we can try to execute the algorithm again after increasing n_d recursively until some threshold. For $0 \leq k \leq n_d - 1$, compute $\vec{\mathbf{x}}_{c,k+1}$ such that

$$\vec{\mathbf{x}}_{c,k+1} = \vec{\mathbf{x}}_{c,k} + \vec{\mathbf{d}} \tag{4.11}$$

where $\vec{\mathbf{d}} = \frac{\vec{\mathbf{x}}_c - \vec{\mathbf{x}}_{c,0}}{n_d}$. Let $\mathcal{F}_k$ correspond to $\vec{\mathbf{x}}_{c,k}$. We find $(\vec{\mathbf{x}}, \vec{\mathbf{s}})$ corresponding to $\mathcal{F}_k$ by solving Problem 4.2 recursively, until arriving at the final destination. If the multiset $\mathcal{F}_k$ arrives at the final destination multiset $\mathcal{T}$ successfully, then the adaptive process succeeds. We can obtain a vector $(\vec{\mathbf{x}}, \vec{\mathbf{s}})$ for a target multiset $\mathcal{T}$. Additionally, if the procedure fails, then, in order to avoid diverging, we can try to execute the algorithm again after increasing n_d recursively until a given threshold.

Finally, we implement a complete algorithm transforming from a given phase-type representation to an HFCB representation $(\beta, \mathbf{M})$, which has as minimal order as possible in Algorithm 3. First, we compute an eigenvalue multiset $\mathcal{A}$ of T. Recursively, we find an FC block $M(\vec{\lambda}^{(i)}, z^{(i)})$ including a chosen eigenvalue $\mu \in \mathcal{A}$. We typically have two cases: an FC block for a complex eigenvalue and a Coxian block for a real eigenvalue. For a complex eigenvalue

Algorithm 2 LineSearch($\mathcal{B}$, $\mathcal{F}$, $\mathcal{T}$, $\vec{\mathbf{s}}$)

Input: $\mathcal{B}$, $\mathcal{F}$, $\mathcal{T}$, $\vec{\mathbf{s}}$
Output: $\vec{\mathbf{s}}$, $Flag$, $\mathcal{B}$

Compute a vector $\vec{\mathbf{x}}$ with respect to $\mathcal{U} = \mathcal{B} - \mathcal{F}$
Compute an initial vector $\vec{\mathbf{x}}_{c,0}$ with respect to $\mathcal{F}$
Compute $\vec{\mathbf{x}}_c$ with respect to a target multiset $\mathcal{T}$
Set a segment number n_d
$\vec{\mathbf{d}} = \frac{\vec{\mathbf{x}}_c - \vec{\mathbf{x}}_{c,0}}{n_d}$
for $0 \leq k < n_d$ **do**
  Set $\vec{\mathbf{x}}_{c,k+1} = \vec{\mathbf{x}}_{c,k} + \vec{\mathbf{d}}$
  Compute a new solution $(\vec{\mathbf{x}}, \vec{\mathbf{s}})$ of Problem 4.2 for an old value $(\vec{\mathbf{x}}, \vec{\mathbf{s}})$ and $\vec{\mathbf{x}}_{c,k+1}$
  Compute $\mathcal{F}$ corresponding to $\vec{\mathbf{x}}_{c,k+1}$
  if No solution **then**
    return $Flag = False$
  else
    $Flag = True$
  end if
  Resort $\vec{\lambda}^{(i)}$ in ascending order
  update $\vec{\mathbf{s}}$ using $\vec{\lambda}^{(i)}$
end for
$\mathcal{B}$ is updated by using $(\vec{\mathbf{x}}, \vec{\mathbf{x}}_c)$
return $\vec{\mathbf{s}}$, $Flag$, $\mathcal{B}$

μ, an FC induced from μ possibly includes other elements of $\mathcal{A}$. Choose a subset collection $\mathfrak{P} = \{\mathcal{S}_1, \cdots, \mathcal{S}_N\}$ such that $\mathcal{S}_k \subset \Omega$, $\mathcal{S}_k \cap \mathcal{F} = \emptyset$, and $|\mathcal{S}_k| \leq |\mathcal{S}_j|$ for $k > j$. If $w \in \mathcal{S}_k$, then $\bar{w} \in \mathcal{S}_k$. Using $\mathfrak{P}$, we can choose a new target multiset $\mathcal{T}$ ($\supset \mathcal{F}$) as large as possible. Compute $(\vec{\lambda}^{(i)}, z^{(i)})$ corresponding to $\mathcal{T}$ by applying the linear search algorithm. We can reduce the total order of the HFCB matrix $\mathbf{M}$ depending on how to choose a target multiset $\mathcal{T}$. The process continues until all complex numbers from Ω are removed. Next, for a selected real eigenvalue $\mu \in \Omega$, we compute $\vec{\lambda}^{(i)} = \mu$, $m_i = 1$ and $z = 0$, which generate the Coxian block. We then resort a generator Υ in the ascending order of the minimal number of $\vec{\lambda}^{(i)}$. Finally, we compute an HFCB representation $\mathbf{M}$ and P using a generator Υ such that $TP = P\mathbf{M}$. We can obtain a transformation matrix P and $\mathbf{M}$ satisfying $TP = P\mathbf{M}$.

Finally, we note the initial vector β corresponding to the obtained mono-cyclic representation can be negative. Even though T has a Coxian generator $S(\vec{\lambda})$ as Theorem 3.1, the positivity of β is not guaranteed. The order of the resulting Coxian generator can be larger than that of the original one. The process for finding such P and S is called the Coxianization of PH-generator T. The Coxianization of a PH-generator with only real eigenvalues has been proven to be feasible; three numerical methods for Coxianization have been introduced and their performances were compared [14].

Algorithm 3 Compute an HFCB representation.

Input: (α, T)
Output: $(\beta, \mathbf{M})$

Compute the eigenvalue multiset $\mathcal{A}$ of T
Index $\mathcal{A}$ in descending order of their real values
$\Omega = \mathcal{A}$, $i = 0$
while Ω is nonempty **do**
  $i = i + 1$
  if Ω has any complex number **then**
    Select a complex number $\mu \in \Omega$ with a maximal real value
    Choose a proper number m_i satisfying the inequality in (4.4)
    Compute $(\vec{\lambda}^{(i)}, z^{(i)})$ and an eigenvalue multiset $\mathcal{B} = \{\mu_0, \cdots, \mu_{m_i-1}\}$ of $M(\vec{\lambda}^{(i)}, z^{(i)})$ by using Lemma 4.1.
    $\mathcal{F} = \{\mu, \bar{\mu}\}$
    Choose a subset collection $\mathfrak{P} = \{\mathcal{S}_1, \cdots, \mathcal{S}_N\}$ such that $\mathcal{S}_k \subset \Omega$, $\mathcal{S}_k \cap \mathcal{F} = \emptyset$, $|\mathcal{S}_k| \leq |\mathcal{S}_j|$ for $k > j$, and if $w \in \mathcal{S}_k$, then $\bar{w} \in \mathcal{S}_k$.
    for $t = 1 : N$ **do**
      $\mathcal{T} = \mathcal{F} \cup \mathcal{S}_t$
      [$\vec{\mathbf{s}}$, $Flag$, $\mathcal{B}$]= LineSearch($\mathcal{B}$, $\mathcal{F}$, $\mathcal{T}$, $\vec{\mathbf{s}}$)
      if $Flag == True$ **then**
        Compute $(\vec{\lambda}^{(i)}, z^{(i)})$ using $\vec{\mathbf{s}}$, and insert it into Υ.
        Break
      end if
    end for
  else
    Select a maximal real value $\mu \in \Omega$ with
    $\mathcal{B} = \{\mu\}$
    $\vec{\lambda}^{(i)} = \mu$, $m_i = 1$ and $z_i = 0$.
  end if
  $\Omega = \Omega - \mathcal{B} \cap \mathcal{A}$
end while
Resort a generator Υ in the ascending order of the minimal number of $\vec{\lambda}^{(i)}$
By using Algorithm 1, compute a HFCB representation $\mathbf{M}$ and P using a generator Υ such that $TP = \mathbf{M}P$
Compute $\beta = \alpha P$
return $(\beta, \mathbf{M})$

To make the initial vector non-negative, additional states have to be added to the distribution similarly. If β contains at least one negative element then a further transformation is required to obtain a PH representation. We can extend the PH generator by adding an Erlang tail as

follows. For sufficiently large λ, we define a vector $\vec{\lambda} = (\lambda, \cdots, \lambda)$ with repeat times N. Set a new generator $\mathbf{M}_1 = \begin{bmatrix} \mathbf{M} & -\mathbf{M}\mathbb{1} \\ \mathbf{0} & S(\vec{\lambda}) \end{bmatrix}$. We have $TP = P\mathbf{M}_1$, $\mathbb{1} = P\mathbb{1}$ and $\beta = \alpha P$. For sufficiently large N, the matrix αP is non-negative.

5. Numerical results and discuss

We have implemented the proposed optimization methods to compute an HFCB representation (β, M) by using Matlab. In order to solve constrained nonlinear problems, we use an interior point option in the function 'fmincon', consisting of sequential quadratic programming and trust region techniques [20]. Extensive numerical experiments have been carried out. The orders of the PH-generators are between 4 and 6 in our numerical simulation. This section presents typical numerical examples to illustrate the effects of the proposed method.

Example 5.1. *Consider a phase-type distribution with the PH-representation* $((1,0,0,0), T)$*, where*

$$T = \begin{bmatrix} -4 & 0 & 0 & 0 \\ 1 & -4 & 0 & 2.7 \\ 0 & 1.5 & -2 & 0 \\ 0.1 & 0 & 0.9 & -1 \end{bmatrix}$$

as introduced in [21]*. By using Algorithm 3, we can obtain a* 4×4 *transformation matrix* P *and* $\mathbf{M}$ *satisfying* $TP = P\mathbf{M}$*, where* $\mathbf{M}$ *is constructed in the form (3.3) by using* $\vec{\lambda}^{(1)} = \begin{bmatrix} 1.435 & 1.435 & 3.716 \end{bmatrix}$ *and* $z^{(1)} = 0.611$*,* $\vec{\lambda}^{(2)} = 4.412$*,* $z^{(2)} = 0$ *and*

$$P = \begin{bmatrix} 1.325 & 0.514 & 0.571 & 2.000 \\ 1.685 & 0.954 & 41.472 & 0.300 \\ -0.146 & 2.910 & 1.148 & 0.500 \\ 2.349 & 1.611 & 0.450 & 0 \end{bmatrix}.$$

Because the computed matrix P *is not non-negative, the other constraint as* $\alpha P \geq 0$ *should be needed instead of* $P \geq 0$*. The dimension is less than that of* [9]*. Therefore, we obtain an FC representation* $(\alpha P, M(\vec{\lambda}, z)$ *with the same order of 4, while Mocanu and Commault's method* [9] *can transform an HFEB representation with an order of 5.*

We note that the positivity of P is not necessary for the existence of $\beta = \alpha P > 0$ from Example 5.1. The PH-majorization of T is a too strong condition. Even for Coxianization algorithm, making P a positive matrix enforces the order to increase [14]. If β contains at least one negative element then a further transformation is required to obtain a monocyclic representation. We can resort the diagonal elements of the HFCB be in an ascending order. The ordering of blocks affects whether the vector β will be non-negative. We can extend the monocyclic generator by adding an Erlang tail for sufficient large λ similar to that of [8]. We give a counterexample condition for the nonexistence of $\beta = \alpha P > 0$. We present a counterexample condition, $\max(\mathbb{1}^t\mathbf{M}) > 0$ for $\beta > 0$ not to exist. In Algorithm 1, let us observe the recursion $p_k = ((\lambda_{k+1}I + T)p_{k+1} - c_{k+n',k+1}p_{k+n'})/c_{k,k+1}$ for $\max(\mathbb{1}^t\mathbf{M}) > 0$.

When p_{k+1} and $p_{k+n'}$ are positive, and $c_{k+n',k} > \lambda_k$ (or $\mathbb{1}^t\mathbf{M} > 0$), negative entries in p_{k-1} can appear. Therefore, we cannot obtain $\alpha P \geq 0$ for any $\alpha \geq 0$. In the next example, we can verify different properties of two cases, $\mathbb{1}^t\mathbf{M} \leq 0$ and $\mathbb{1}^t\mathbf{M} > 0$ by using numerical simulations.

Example 5.2. *Consider a phase-type distribution with the PH-representation (α, T) such that*

$$T = \begin{bmatrix} -6.0 & 0.6 & 0.6 & 0.6 & 3 \\ 2 & -4.5 & 0 & 0 & 0.1 \\ 0 & 2 & -3.5 & 0 & 0 \\ 0 & 0 & 2 & -3 & 0 \\ 0 & 0 & 0 & 2 & -2 \end{bmatrix}$$

While Mocanu and Commault's method [9] *can transform T into an HFEB representation with order of 8, we can obtain a 5×6 matrix P and an HFCB generator $\mathbf{M}$ with an order of 6 by using Algorithm 3. We can see that the matrix P is not non-negative. The proposed method can derive multiple solutions depending on an given initial value. For $\max(\mathbb{1}^t\mathbf{M}) \leq 0$, we could obtain $\alpha \geq 0$ such that $\alpha P \geq 0$ after attempting several times. For example, there are two typical cases after applying the proposed algorithm.*

(1) *The first case has $\vec{\lambda}^{(1)} = 1.120$, $z^{(1)} = 0$, $\vec{\lambda}^{(2)} = (2.673, 2.678, 3.116, 4.49201, 6.366)$ and $z^{(2)} = 0.0591$. For $\max(\mathbb{1}^t\mathbf{M}) \leq 0$, we obtained $\alpha \geq 0$ such that $\alpha P \geq 0$ after several trials.*

(2) *In the second case, we have $\vec{\lambda}^{(1)} = 1.12084$, $z^{(1)} = 0$, $\vec{\lambda}^{(2)} = (1.517, 3.660, 3.660, 3.660, 6.420)$ and $z^{(2)} = 0.096$. For $\max(\mathbb{1}^t\mathbf{M}) > 0$, we compute P for several times. Typically, we can obtain P such as*

$$P = \begin{bmatrix} -0.202 & 2.084 & 0.920 & 0.858 & 0.940 & 1.2 \\ -0.127 & 1.043 & 0.481 & 0.090 & 1.914 & 2.4 \\ -0.107 & 0.644 & 0.099 & 1.156 & 2.507 & 1.5 \\ -0.113 & 0.539 & 0.936 & 1.686 & 1.753 & 1.0 \\ -0.259 & 2.840 & 1.468 & 1.206 & 0.5463 & 0 \end{bmatrix}.$$

Because all entries of the first column are negative, we can see that $\min(\alpha P) < 0$ for any $\alpha \geq 0$.

In order to show that the proposed method has good performance, a large set of random PH generators samples were generated in the next example. The random matrices T are dense, but the canonical representations obtained by the proposed method are sparse.

Example 5.3. *A 31 samples of T with order 5 and 4 complex eigenvalues are generated randomly. Then for each sample, the proposed method and Mocanu and Commault's method* [9] *are applied to obtain $HFCB$ and $HFEB$, respectively. For each algorithm, we determine $\mathbf{M}$ and $\beta = \alpha P$ and check non-negativity of β. We compared the orders of the HFCB generators and HFEB generators. Let d_{FC} and d_{FE} be the order of $HFCB$ and $HFEB$, respectively.*

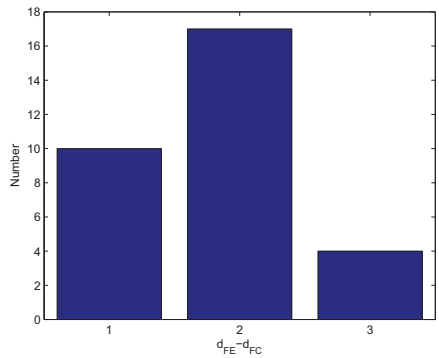


FIGURE 2. Histogram of the order d_{FE} minus the order d_{FC}

In Fig. 2, we plot $d_{FE} - d_{FC}$. We could see that the order of the solution obtained by using the proposed method is smaller than that of the solution obtained by using Mocanu and Commault's method [9].

We note that there are several similar approaches to find compact PH representations with a unicyclic structure [15, 21]. A unicyclic block has more flexibility than FE block or FC block, since the representation has more free parameter and is denser. Since a unicyclic block has too many free parameters ($2m - 1$) comparing to the number of equations (m), it is impossible to obtain a stable optimal solution using numerical method. So a modified unicyclic block was introduced [15]. A modified unicyclic block approach reduces the free parameter number to m. Thus, the values of m free parameters can be obtained by solving m equations. When the given eigenvalue set frequently has no feasible solution, it cannot be applied. On the other hand, for the case, the proposed method can obtain a feasible solution by solving Problem 4.2.

Furthermore, the importance of simple sparse representations comes from the fact that the computational complexity of PH-distributed random-variate generation depends on the representation in simulation applications [8]. The FC block approach can lead to more effective random number generation of PH-distributions in stochastic modeling and analysis because the number of its feedback loop can be reduced.

6. CONCLUSION

We tried to transform a phase-type representation into a sparse HFCB representation with an order as low as possible. We introduced a numerical method for computing an HFCB representation by solving a nonlinear optimization problem. To find a sparse HFCB representation with lower order, including the multiset of eigenvalues of the original generator, we developed a numerical algorithm solving a constrained nonlinear optimization program. We discussed some numerical issues in efficiently finding a stable solution of the nonlinear constrained optimization problem. The proposed method can be applied to the sparse positive realization of positive system.

ACKNOWLEDGEMENTS

The author thanks the editor and the referees for their careful reading of the manuscript. This work was supported by research fund of Chungnam National University.

REFERENCES

[1] C. Altafini. Minimal eventually positive realizations of externally positive systems. *Automatica*, **68** (2016), 140–147.
[2] L. Benvenuti, L. Farina, and B. D. O. Anderson. Filtering through combination of positive filters. *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, **46**(12) (1999), 1431–1440.
[3] L. Farina and S. Rinaldi. *Positive Linear Systems: Theory and Applications*. Wiley Interscience, 2000.
[4] B. Nagy, M. Matolcsi, and M. Szilvasi. Order Bound for the Realization of a Combination of Positive Filters. *IEEE Transactions on Automatic Control*, **52**(4) (2007), 724–729.
[5] C. Commault and S. Mocanu. Phase-type distributions and representations: Some results and open problems for system theory. *International Journal of Control*, **76**(6) (2003), 566–580.
[6] C. A. O'Cinneide. On non-uniqueness of representations of phase-type distributions. *Communications in Statistics. Stochastic Models*, **5**(2) (1989), 247–259.
[7] G. Horváth, P. Reinecke, M. Telek, and K. Wolter. Efficient Generation of PH-distributed Random. *Lecture Notes in Computer Science*, **7314** (2012), 271–285.
[8] G. Horváth, P. Reinecke, M. Telek, and K. Wolter. Heuristic representation optimization for efficient generation of PH-distributed random variates. *Annals of Operations Research*, (2014), 1–23.
[9] S. Mocanu and C. Commault. Sparse representations of phase-type distributions. *Communications in Statistics. Stochastic Models*, **15**(4) (1999), 759–778.
[10] K. Kim. A construction method for positive realizations with an order bound. *Systems & Control Letters*, **61**(7) (2012), 759–765.
[11] K. Kim. A constructive positive realization with sparse matrices for a continuous-time positive linear system. *Mathematical Problems in Engineering*, **2013** (2013), 1–9.
[12] D. S. Huang. A constructive approach for finding arbitrary roots of polynomials by neural networks. *IEEE Transactions on Neural Networks*, **15**(2) (2004), 477–491.
[13] C. A. O'Cinneide. Phase-Type Distributions and Invariant Polytopes. *Advances in Applied Probability*, **23**(3) (1991), 515–535.
[14] Q.-M. He, H. Zhang, and J. Xue. Algorithms for Coxianization of Phase-Type Generators. *INFORMS J. on Computing*, **23**(1) (2011), 153–164.
[15] I. Horváth and M. Telek. A heuristic procedure for compact Markov representation of PH distributions. In *ValueTools*, 2014.
[16] C. A. O'Cinneide. Phase-type distributions: open problems and a few properties. *Stochastic Models*, **15**(4) (1999), 731–757.
[17] D. J. Hartfiel. A simplified form for nearly reducible and nearly decomposable matrices, 1970.
[18] J. L. Walsh. On the location of the roots of certain types of polynomials. *Transactions of American Mathematic Society*, **24** (1922), 163–180.
[19] J. M. Ortega. *Numerical Analysis: A Second Course*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1990.
[20] R. a. Waltz, J. L. Morales, J. Nocedal, and D. Orban. An interior algorithm for nonlinear optimization that combines line search and trust region steps. *Mathematical Programming*, **107**(3) (2006), 391–408.
[21] Q.-M. He and H. Zhang. A Note on Unicyclic Representations of Phase Type Distributions. *Stochastic Models*, **21**(2-3) (2005), 465–483.