Low-rank solutions of Riccati equations for positive real balanced truncations of RLC networks

Yuichi Tanji

Faculty of Engineering and Design, Kagawa University
2217–20 Hayashi-cho, Takamatsu, Kagawa 761–0396, Japan.

tanji@eng.kagawa-u.ac.jp

Received September 15, 2017; Revised March 28, 2018; Published October 1, 2018

Abstract: This paper presents an efficient iterative method to solve standard and generalized algebraic Riccati equations for RLC networks. The proposed method generates a low-rank solution of the Riccati equation for a positive real balanced truncation. Linear passive RLC networks are index-1 or -2 systems; the proposed method generates a low-rank solution of the standard algebraic Riccati equation for an index-1 system and that of a generalized Riccati equation for an index-2 system. To generate accurate reduced-order models at low and high frequencies, the parameters of the iterative method for solving the Riccati equations are investigated. In general, the balanced truncation accuracy at low frequencies is not necessarily satisfactory, compared to that of (Krylov) projection-based model-order reduction. The accuracy of a reduced-order model at low frequencies can be improved by considering a constant shift parameter.

Key Words: Signal/Power Integrity, interconnects, Riccati equations, positive real balanced truncation

1. Introduction

Many researchers in the design automation community have studied model reduction over the past 30 years. In an integrated circuit system, there are numerous interconnects; hence, the assessment of the interconnect effects entails prohibitive computational costs. Therefore, a compact model of the interconnects is required. Model reduction algorithms offer efficient models, with orders that are considerably lower than those of the original system and their frequency-domain responses fit into the original, within a frequency range.

When the physical effects of the interconnects are adequately modeled by RC or RLC networks, the networks become passive [25]. Therefore, the passivity should be preserved for the stable execution of the post time-domain simulations. Krylov subspace methods [1–4], time-constant equilibration reduction [5], and positive real balanced truncation [16] have been proposed as passivity-preserving model reduction algorithms. The Krylov subspace methods and the time-constant equilibration reduction generate reduced-order models efficiently; although the positive real balanced truncation is
more accurate than these algorithms, its application is restricted to middle-class problems.

Positive real balanced truncation cannot be applied to a large-scale problem because of the difficulty in obtaining a solution for the algebraic Riccati equation,

$$A^T X + X A + X B B^T X + C^T C = 0.$$  \hspace{1cm} (1)

This is explained in detail later. Although the algebraic Riccati equation is a nonlinear matrix equation, it can be solved by eigenvalue decomposition; thus, large-scale problems cannot be easily solved. The Newton method was introduced [21] as a powerful algorithm for obtaining the algebraic Riccati equation solution.

At each Newton step, the algebraic Riccati equation is replaced by a Lyapunov equation. Then, an alternating direction implicit iteration is introduced [20]. The solution can be accurately approximated by a low-rank solution, expressed as the Cholesky factor [18, 19]; this method is called a Cholesky factorized alternating direction implicit iteration. As this iteration is faster than the alternating direction implicit iteration, by utilizing it, the algebraic Riccati equation solution is efficiently obtained by the Newton method. In addition, this concept is used for the solution of a generalized algebraic Riccati equation [11, 29],

$$A^T Y E + E^T Y A + E^T Y B B^T Y E + C^T C = 0.$$ \hspace{1cm} (2)

Linear passive RLC networks are index-1 or -2 systems. When the networks are index-1, the algebraic Riccati equations are solved; whereas, when they are index-2, the generalized algebraic Riccati equations are solved.

The alternating direction implicit iteration for a direct algebraic Riccati equation solution is called a quadratic alternating direction implicit iteration [6]. A Cholesky factorized version of the quadratic alternating direction implicit iteration is presented for accelerating the quadratic alternating direction implicit iteration. The real and complex conjugate pair shift parameters must be considered in order to enhance its convergence [18–20]. Multishift Cholesky factorized quadratic alternating direction implicit iteration [8] uses real and complex conjugate pair shift parameters; then, a Cholesky factor is obtained, in which the number of columns is larger than that of the rows. Although the solution obtained by multishift Cholesky factorized quadratic alternating direction implicit iteration is low-rank and a real matrix, it cannot be efficiently used for positive real balanced truncation.

In this paper, an improved quadratic alternating direction implicit iteration and a Cholesky factorized one for the solutions of algebraic Riccati equation and generalized one are presented. First, the algebraic Riccati equation and generalized one are obtained for index-1 and -2 systems, respectively, starting from the modified nodal analysis equations of RLC networks. Then, the positive real balanced truncation is presented for each case. In [11], a special projector [15] is used for obtaining the generalized algebraic Riccati equation, with sparse coefficient matrices. As the treatment of sparse matrices is not completely known, it is difficult to obtain a generalized algebraic Riccati equation with sparse matrices. In [14], a passivity test algorithm is presented, in which a descriptor system with sparse matrices, whose index is more than two, is treated. This method does not provide a generalized algebraic Riccati equation; hence, the proposed method is useful for obtaining the generalized algebraic Riccati equation, with sparse matrices. In [6], the well-posedness of the Cholesky factorized quadratic alternating direction implicit iteration is proved. However, it is incomplete and Cholesky factorized quadratic alternating direction implicit iteration cannot be always applied to an algebraic Riccati equation associated with a multiport network. Hence, the Cholesky factorized quadratic alternating direction implicit iteration is improved without contradicting the well-posedness of the algebraic Riccati equation and generalized one.

Recently, sophisticated algebraic Riccati equation solvers have been proposed [28, 30]. These methods present a better algebraic Riccati equation solution, efficiently. The main purpose of model order reduction for RLC networks is to establish a low-frequency accurate model; however, the balanced truncation is not satisfactory [16]. With numerical examples, we demonstrate that a low-frequency accurate model can be obtained by considering a small shift parameter. Moreover, the constant shift
2. Positive real balanced truncation

2.1 Realization of a descriptor system for RLC networks

Linear passive RLC networks are generally formulated using modified nodal analysis (MNA) equations. Assuming the impedance matrix, the dynamics of the system can be expressed by

\[ E_0 \frac{dx(t)}{dt} = A_0 x(t) + B_0 u(t), \quad z(t) = B_0^T x(t), \] (3)

where

\[ E_0 = \begin{bmatrix} A_G C A_T & 0 \\ 0 & -\mathcal{L} \end{bmatrix}, \quad A_0 = \begin{bmatrix} -A_G G A_T & -A_L \\ -A_L^T & 0 \end{bmatrix}, \quad B_0 = \begin{bmatrix} -A_L \\ 0 \end{bmatrix}. \] (4)

In (4), \( G \), \( L \), and \( C \) are the conductance, inductance, and capacitance matrices, respectively; \( A_G \), \( A_L \), and \( A_T \) are the incidence matrices on the conductors, capacitors, inductors, and independent current sources, respectively. \( x(t) \in R^m \) is a vector of the node voltages and inductor currents; \( u(t) \in R^m \) are the independent current sources.

In the control community, Eq. (3) is called a descriptor system. To apply positive real balanced truncation (PRBT), the descriptor system must be converted into a state equation. As the first step in obtaining the state equation, the Weierstrass canonical form is considered:

\[ E_0 = W \begin{bmatrix} I_q & 0 \\ 0 & N \end{bmatrix} T, \quad A_0 = W \begin{bmatrix} J & 0 \\ 0 & I_{n-q} \end{bmatrix} T, \] (5)

where \( I_k \) is a \( k \times k \) identity matrix; \( J \) is a Jordan form, whose eigen values correspond to the finite eigen values of the generalized eigen value problem, \((E,A)\), whereas \( N \) is a nilpotent, whose eigen values are zero. When \( N^\mu = 0 \), \( N^{\mu'} \neq 0 \), and \( \mu' = \mu + 1 \), \( \mu \) is called the index. For RLC networks, the index is at most two [16].

With the Weierstrass canonical form (5), relationships \( C_0 T^{-1} = [C_p \ C_\infty] \) and \( W^{-1}B = [B_p \ B_\infty]^T \) are defined. Then, the transfer function, \( Z(s) \), which is the impedance matrix, is obtained by

\[ Z(s) = B_0^T \left(sE_0 - A_0\right)^{-1} B_0 \]
\[ = \left[ \begin{array}{cc} C_p & C_\infty \end{array} \right] \left[ \begin{array}{cc} (sI_q - J)^{-1} & 0 \\ 0 & (sN - I_{n-q})^{-1} \end{array} \right] \left[ \begin{array}{c} B_p \\ B_\infty \end{array} \right] \]
\[ = C_p (sI_q - J)^{-1} B_p + M_0 + sM_1, \] (6)

where \( M_0 = -C_\infty B_\infty \) and \( M_1 = -C_\infty NB_\infty \) are positive semidefinite [16].

As the Weierstrass canonical form is computationally expensive [26], alternative methods are presented for index-1 and -2 systems, respectively, in this paper.

2.2 Index-1 system

The descriptor system (3) is expressed in a singular value decomposition (SVD) canonical form for converting MNA Eq. (3) into a state equation. Although SVD is necessary, it is computationally
expensive for a large system. Hence, instead of SVD, we use LDL factorization\(^1\), which is a Cholesky-like factorization for a semidefinite case: \(P^T \alpha P = LDL^T\) where \(\alpha\) is symmetric, \(P\) is a permutation, \(L\) is a lower triangular matrix, and \(D\) is a diagonal matrix with a rank less than \(n\). Applying LDL factorization to \(A_CCA_C\) and \(L\) in (4), matrix, \(E_0\), can be expressed as

\[
E_0 = V \begin{bmatrix} I_{r_1} & 0 & 0 \\ 0 & -I_{r_2} & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T = V \begin{bmatrix} I_r' & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T,
\]

where the ranks of \(A_CCA_C\) and \(L\) are assumed to be \(r_1\) and \(r_2\), respectively and \(r = r_1 + r_2\).

Then, Eq. (3) is converted into

\[
\begin{bmatrix} I_r' & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{d}{dt} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(t)
\]

\[
z = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix},
\]

which is SVD canonical form of (3), where \(V^T x(t) = [x_1(t)^T \ x_2(t)^T]^T\).

Assuming \(A_{22}\) to be nonsingular, from (8), we can obtain the state equation as

\[
\dot{x}_1(t) = \dot{A}_0 x_1(t) + \dot{B}_0 u(t)
\]

\[
z(t) = C_0 x_1(t) + \dot{D}_0 u(t)
\]

where

\[
\dot{A}_0 = I_r' (A_{11} - A_{12} A_{22}^{-1} A_{21}), \quad \dot{B}_0 = I_r' (B_1 - A_{12} A_{22}^{-1} B_2),
\]

\[
\dot{C}_0 = C_1 - C_2 A_{22}^{-1} A_{21}, \quad \dot{D}_0 = -C_2 A_{22}^{-1} B_2.
\]

The index of (8) is one, if and only if \(A_{22}\) is nonsingular [27] and the descriptor system is always converted into a state equation.

The algebraic Riccati equation (ARE) is obtained using Lur’e equation associated with (9) [25]. Then, the following two AREs are solved for PRBT\(^2\):

\[
A^T X + XA + XBB^T X + C^T C = 0
\]

\[
AY + Y A^T + Y C^T CY + B B^T = 0
\]

where \(DD^T = (\dot{D}_0 + \dot{D}_0^T)^{-1}\), \(B = \dot{B}_0 D\), \(C = D^T \dot{C}_0\), and \(A = \dot{A}_0 - BC\).

Consider the similarity transformations to be \(\dot{A}_0 = T^{-1} A_0 T\), \(\dot{B}_0 = T^{-1} B_0\), and \(\dot{C}_0 = CT_0\). The solutions of the transformed AREs are both diagonalized as

\[
T^T X T = \Sigma = T^{-1} Y T^{-T}
\]

The diagonal elements of \(\Sigma\) are called Hankel singular values. Small values exhibit weak effects on the input-output behavior [16]. Therefore, the components corresponding to small Hankel singular values can be removed; thereby, the reduced-order models are obtained. PRBT algorithm for an index-1 system is called **PRBT-1**.

**PRBT-1**

1. Solve (11) for \(X\).
2. Solve (12) for \(Y\).
3. Compute the Cholesky factors, \(X = L_x L_x^T\) and \(Y = L_y L_y^T\).

---

\(^1\)A MATLAB function `ldl()` is available.

\(^2\)When matrix, \((\dot{D}_0 + \dot{D}_0^T)^{-1}\), is singular, it is approximated as \(DD^T = (\dot{D}_0 + \dot{D}_0^T + \epsilon I_m)^{-1}\), where \(\epsilon\) is a small positive number.

482
4. Compute SVD as $U \Sigma V^T = L_T^T L_q$.
5. Compute the transform matrices as $T = L_q V \Sigma^{-1/2}$ and $T^{-1} = \Sigma^{-1/2} U^T L_T^T$.
6. Form the balanced realizations as $\tilde{A}_0 = T^{-1} A_0 T$, $\tilde{B}_0 = T^{-1} B_0$, and $\tilde{C}_0 = C_0 T$.
7. Partition $\tilde{A}_0$, $\tilde{B}_0$, and $\tilde{C}_0$ as
   \[
   \tilde{A}_0 = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix}, \quad \tilde{B}_0 = \begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{bmatrix}, \quad \tilde{C}_0 = \begin{bmatrix} \tilde{C}_1 & \tilde{C}_2 \end{bmatrix}.
   \]
   8. Truncate $\tilde{A}_0$, $\tilde{B}_0$, and $\tilde{C}_0$ to form the reduced realizations $\tilde{A}_{11}$, $\tilde{B}_1$, and $\tilde{C}_1$; the reduced-order transfer function is obtained by $\tilde{Z}(s) = \tilde{C}_1(sI - \tilde{A}_{11})^{-1} \tilde{B}_1 + \tilde{D}_0$.

It should be noted that the Cholesky factorization of step 3 is not required because the Cholesky factors are directly obtained by the ARE solver presented in Section 4.

### 2.3 Index-2 system

When matrix, $A_{22}$, in (8) is singular, the index becomes two for passive RLC networks. Therefore, all the terms of (6) must be calculated. Then, we introduce the right and left spectral projectors associated with the deflated invariant subspaces of the matrix pencil, $\lambda E_0 - A_0$. For obtaining the spectral projectors, LDL decomposition is applied to $A_{22}$ and the following relationship is obtained:

\[
A_{22} = Q \begin{bmatrix} S_r & 0 \\ 0 & 0 \end{bmatrix} Q^T
\]

where matrix, $S_r$, is nonsingular. In (8), expressing $Q^T x_2(t) = y(t) = [y_1(t)^T \ y_2(t)^T]^T$,

\[
I_r \frac{dx_1(t)}{dt} = A_{11} x_1(t) + A_{12} y_1(t) + A_{21} y_2(t) + B_1 u(t)
\]

0 = \begin{bmatrix} 0 \\ A_{21,1} \end{bmatrix} + \begin{bmatrix} S_r y_1(t) \\ 0 \end{bmatrix} + \begin{bmatrix} B_{21} \\ B_{22} \end{bmatrix} u(t) \quad (15)

\[
z(t) = C_1 x_1(t) + C_{21} y_1(t) + C_{22} y_2(t).
\]

Eliminating $y_1(t)$ from (15), the following relationship is obtained:

\[
\begin{bmatrix} I_r \\ 0
\end{bmatrix} \frac{d}{dt} \begin{bmatrix} x_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ y_2(t) \end{bmatrix} + \begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{bmatrix} u(t)
\]

\[
z(t) = \tilde{C}_1 x_1(t) + \tilde{C}_{22} y_2(t) + \tilde{D}_0 u(t),
\]

where $\tilde{A}_{11} = A_{11} - A_{12} S_r^{-1} A_{21}$, $\tilde{A}_{12} = A_{12}$, $\tilde{A}_{21} = A_{21}$, $\tilde{B}_1 = B_1 - A_{12} S_r^{-1} B_{21}$, $\tilde{B}_2 = B_{22}$, $\tilde{C}_1 = C_1 - C_{21} S_r^{-1} A_{21}$, $\tilde{C}_2 = C_{22}$, and $\tilde{D}_0 = -C_{21} S_r^{-1} B_{21}$. Equation (16) is called a Stokes-type index-2 system, whose left and right spectral projectors, $P_l$ and $P_r$, are explicitly expressed as

\[
P_l = \begin{bmatrix} \Pi_l & 0 \\ 0 & -\Pi_l \tilde{A}_{11} I_r \tilde{A}_{12} (\tilde{A}_{21} I_r \tilde{A}_{12})^{-1} \end{bmatrix}, \quad P_r = \begin{bmatrix} \Pi_r \\ (\tilde{A}_{21} I_r \tilde{A}_{12})^{-1} \tilde{A}_{21} I_r \tilde{A}_{11} \Pi_r \\ 0 \\ 0 \end{bmatrix} \quad (17)
\]

where $\Pi_l = I_r - \Pi_l (\tilde{A}_{21} I_r \tilde{A}_{12})^{-1} \tilde{A}_{21} I_r$ and $\Pi_r = I_r - \Pi_r (\tilde{A}_{21} I_r \tilde{A}_{12})^{-1} \tilde{A}_{21}$. Note that $\Pi_l$ is a projector on the kernel of $\tilde{A}_{21} I_r$ along the image of $\tilde{A}_{12}$ [10].

To apply PRBT, the descriptor system (16) is rewritten as

\[
\tilde{E}_0 \tilde{x}(t) = \tilde{A}_0 \tilde{x}(t) + \tilde{B}_0 u(t), \quad z(t) = \tilde{C}_0 \tilde{x}(t) + \tilde{D}_0 u(t).
\]

Using the right projector, we can express $M_0$ and $M_1$ of (6) as [14],

\[
M_0 = \tilde{C}_0 (I_n - P_r) \tilde{A}_0^{-1} \tilde{B}_0 + \tilde{D}_0, \quad M_1 = -\tilde{C}_0 \tilde{A}_0^{-1} \tilde{E}_0 (I_n - P_l) \tilde{A}_0^{-1} \tilde{B}_0.
\]
The following two generalized AREs (GAREs) are solved for PRBT [13]:

\[
\begin{align*}
A^T_1 XE + E^T XA_1 + E^T XBB^T XE + P^T r C^T CP_r &= 0, \\
A_2 Y E^T + E Y A^T_2 + E Y C^T CY E^T + P BB^T P^T_l &= 0
\end{align*}
\]

where \( DD^T = (M_0 + M_0^T)^{-1} \), \( A_1 = \bar{A}_0 - BCP_r \), \( A_2 = \bar{A}_0 - P_l BC \), \( C = D^T C_0 \), and \( B = \bar{B}_0 D \). Then, PRBT algorithm for an index-2 system [11–13] is called **PRBT-2**.

**PRBT-2**

1. Solve (20) for \( X \).
2. Solve (21) for \( Y \).
3. Compute the Cholesky factors, \( X = L_x L_x^T \) and \( Y = L_y L_y^T \).
4. Compute SVD as \( U \Sigma V^T = L_x^T E_0 L_y \).
5. Compute the transform matrices as \( T = L_y V \Sigma^{-1/2} \) and \( W = \Sigma^{-1/2} U^T L_x^T \).
6. Form the balanced realizations as \( \bar{E}_0 = \bar{W} \bar{E}_0 \bar{T} \), \( \bar{A}_0 = \bar{W} A_0 \bar{T} \), \( \bar{B}_0 = \bar{W} \bar{B}_0 \), and \( \bar{C}_0 = \bar{C}_0 \bar{T} \).
7. Partition \( \bar{E}_0 \), \( \bar{A}_0 \), \( \bar{B}_0 \), and \( \bar{C}_0 \) as

\[
\bar{E}_0 = \begin{bmatrix} \bar{E}_{11} & \bar{E}_{12} \\ \bar{E}_{21} & \bar{E}_{22} \end{bmatrix}, \quad \bar{A}_0 = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix}, \quad \bar{B}_0 = \begin{bmatrix} \bar{B}_1 \\ \bar{B}_2 \end{bmatrix}, \quad \bar{C}_0 = \begin{bmatrix} \bar{C}_1 \\ \bar{C}_2 \end{bmatrix}.
\]

8. Truncate \( \bar{E}_0 \), \( \bar{A}_0 \), \( \bar{B}_0 \), and \( \bar{C}_0 \) to form the reduced realizations \( \tilde{E}_{11} \), \( \hat{A}_{11} \), \( \tilde{B}_1 \), and \( \hat{C}_1 \); the reduced-order transfer function is obtained by \( \tilde{Z}(s) = \hat{C}_1 \left( s \tilde{E}_{11} - \hat{A}_{11} \right)^{-1} \tilde{B}_1 + M_0 + s M_1 \).

As in **PRBT-1**, the Cholesky factorization of step 3 is not required because it is obtained by the GARE solver. In [11–13], the projected discrete-time Lyapunov equations are solved for obtaining the improper components of the transfer matrix. However, the terms, \( M_0 \) and \( M_1 \), are directly calculated by (19); thus, the Lyapunov equations need not be solved.

**3. Review**

We review the quadratic alternating direction implicit iteration (QADI) [6], for solving the algebraic Riccati equation (ARE) and modify it slightly. QADI, which is an extension of the alternating direction implicit iteration (ADI), for solving the ARE (1), is expressed as

\[
\begin{align}
(A^T + X_{j-1} T BB^T + p_j I_n) X_{j-1/2} &= -C^T C - X_{j-1} (A - p_j I_n), \\
(A^T + X_{j-1} T BB^T + p_j I_n) X_j &= -C^T C - X_{j-1/2} (A - p_j I_n),
\end{align}
\]

where \( p_j \) is a shift parameter and \( X_{j-1/2} \) is an auxiliary matrix. \( X_j \) is the solution at the \( j \)-th iteration. Equations (22) and (23) are alternately solved, starting from \( X_0 = 0 \); then, the solution at finite steps gives a good approximate stable solution of ARE (1). Cholesky factorized QADI (CFQADI) was developed for accelerating QADI and the obtained Cholesky factor (CF) can be directly used for positive real balanced truncation (PRBT). Linear fractional transformation (LFT) is used for obtaining CFQADI and is defined by the transfer matrix from \( w \) to \( z \), as shown in Fig. 1; it is represented by

\[
F_l(P, \Delta) = P_{11} + P_{12} \Delta (I - P_{22} \Delta)^{-1} P_{21}.
\]

In [6], the respective solutions of (22) and (23) are \( X_{j-1/2} = F_l(P, X_{j-1}) \) and \( X_j = F_l(P^T, X_{j-1/2}) \). However, the latter cannot be obtained, when \( X_{j-1/2} \) is singular.

Alternatively, we use another relationship [7]. If

\[
(\alpha + \beta Q)(\gamma + \theta Q)^{-1} = F_l(M, Q),
\]

\[
(\gamma + Q \theta)^{-1}(\alpha + Q \beta) = F_l(N, Q),
\]

484
matrices $M$ and $N$ are expressed as

$$M = \begin{bmatrix} \alpha \gamma^{-1} & \beta - \alpha \gamma^{-1} \theta \\ \gamma^{-1} & -\gamma^{-1} \theta \end{bmatrix}, \quad N = \begin{bmatrix} \gamma^{-1} \alpha & \gamma^{-1} \\ \beta - \theta \gamma^{-1} \alpha & -\theta \gamma^{-1} \end{bmatrix}. \quad (27)$$

The solutions of (22) and (23) are $X_{j-1/2} = F_i(M, X_{j-1})$ and $X_j = F_i(N, X_{j-1/2})$. Then, we can obtain the nested relationship $X_j = F_i(N, F_i(M, X_{j-1}))$, which is called a Redheffer star product.

With $S_j = (A + p_j I_n)^{-1}$ and $T_j = A - p_j I_n$, the relationship, $S_j T_j = I_n - 2p_j S_j$, is obtained. Assume that $B^T S_j^T C^T$ is nonsingular. Then, the solution, $X_j$, can be represented in the following form:

$$X_j = M_{11} + M_{12} X_{j-1} (I_n - M_{22} X_{j-1})^{-1} M_{12}^T$$

where

$$M_{11} = -2p_j S_j^T C^T (I_n - B^T S_j^T C^T C S_j B)^{-1} C S_j,$$

$$M_{22} = -2p_j S_j B (I_n - B^T S_j^T C^T C S_j B)^{-1} B^T S_j^T,$$

$$M_{12} = I_n - 2p_j S_j^T + S_j^T C^T C M_{22}. $$

As the solution $X_{j-1} = Z_{j-1} Z_{j-1}^T$, it can also be expressed as

$$X_j = M_{11} + M_{12} Z_{j-1} (I_{(j-1)m} - Z_{j-1}^T M_{22} Z_{j-1})^{-1} Z_{j-1}^T M_{12}.$$ \quad (29)

Note that the inverse matrix lemma [7] is used for obtaining (29). The two Cholesky factorizations are $R_{1j}^T R_{1j} = (I_m - B^T S_j^T C^T C S_j B)^{-1}$ and $R_{2j}^T R_{2j} = (I_{(j-1)m} - Z_{j-1}^T M_{22} Z_{j-1})^{-1}$. Then, CF, $Z_j$, at the $j$-th step of CFQADI is obtained as

$$Z_j = [\sqrt{-2p_j S_j^T C^T} R_{1j}^T \{ Z_{j-1} - 2p_j S_j^T Z_{j-1} \\ -2p_j S_j^T C^T C S_j B R_{1j}^T R_{1j} B^T S_j^T Z_{j-1} \} R_{2j}^T]. \quad (30)$$

The orders of matrices $(I_m - B^T S_j^T C^T C S_j B)^{-1}$ and $(I_{(j-1)m} - Z_{j-1}^T M_{22} Z_{j-1})^{-1}$ are considerably smaller than that of $A$. Although CFQADI is faster than QADI, it is applicable only when the shift parameter, $p_j$, is real. This is because Cholesky factorization cannot be applied to $(I_m - B^T S_j^T C^T C S_j B)^{-1}$, with a complex $p_j$.

The optimum shift parameters for $L$ runs are obtained by solving the minimax problem [17] defined as

$$\min_{\{p_1, p_2, \ldots, p_L\}} \max_{\lambda_i \in \text{spec}(A)} \left| \prod_{j=1}^L \frac{p_j - \lambda_j}{p_j + \lambda_j} \right|. \quad (31)$$

where $\tilde{A} = A + B B^T X$. Although $\tilde{A}$ is self-referential to $X$, its spectrum can be obtained a priori [6]; $\text{spec}(\tilde{A}) = \text{spec}(H) \cap \mathcal{C}_-$, in which $\mathcal{C}_-$ is the open left half complex plane and

$$H = \begin{bmatrix} A & B B^T \\ -C^T C & -A^T \end{bmatrix}. \quad (32)$$
is the Hamiltonian matrix. Therefore, the shift parameters are obtained as real and complex conjugate pairs. In [8], multishift CFQADI is proposed. We can obtain a real and low-rank CF by multishift CFQADI; however, the number of columns are more than the rows. As CF cannot be efficiently used for PRBT, we improve it using multishift parameters, which are real and complex conjugate pairs.

4. Improved QADI and CFQADI

4.1 Algebraic riccati equation

The quadratic alternating direction implicit iteration (QADI) is improved by utilizing the complex conjugate of the shift parameter. Steps (22) and (23) are modified as

\[
(A^T + X_{j-1}^TB^T + p_jI) X_{j-1/2} = -C^T C - X_{j-1}^T (A - p_jI),
\]

\[
(A^T + X_{j-1/2}B^T + \bar{p}_jI) X_{j-1/2} = -C^T C - X_{j-1/2} (A - \bar{p}_jI),
\]

where \(\bar{p}_j\) is the complex conjugate of \(p_j\).

In the Cholesky factorized QADI (CFQADI), the term, \(B^T S_j^H C^T\), is assumed to be nonsingular. However, this does not always hold for a multiport network. The well-posedness of CFQADI is not guaranteed, because the inverse matrix lemma cannot be directly used. Hence, we express \(B^T S_j^H C^T = U_j D_j V_j^H\), using singular value decomposition (SVD); \(D_j \in R^{q \times q}\) is a nonsingular diagonal matrix and \(q\) is less than \(m\), indicating that zero singular values are eliminated in SVD. The inverse matrix lemma, associated with \(U_j D_j V_j^H\) rather than \(B^T S_j^H C^T\), is applied. Then, the solution, \(X_j\), can be obtained in the following form:

\[
X_j = M_{11} + M_{12} X_{j-1} (I_n - M_{22} X_{j-1})^{-1} M_{12}^T
\]

where

\[
M_{11} = -2 \text{Re} p_j S_j^H C^T \left\{ I_q + V_j \left( D_j^{-2} - I_q \right)^{-1} V_j^H \right\} C S_j,
\]

\[
M_{22} = -2 \text{Re} p_j S_j B \left\{ I_q + U_j \left( D_j^{-2} - I_q \right)^{-1} U_j^H \right\} B^T S_j^H,
\]

\[
M_{12} = I_n - 2 \text{Re} p_j S_j^H - 2 \text{Re} p_j S_j^H C^T V_j \left( I_q - D_j^2 \right)^{-1} D_j U_j^H B^T S_j^H.
\]

The two Cholesky factorizations are represented by \(R_{1,j}^H R_{1,j} = \left\{ I_q + V_j \left( D_j^{-2} - I_q \right)^{-1} V_j^H \right\}\) and \(R_{2,j}^H R_{2,j} = (I_{(j-1)m} - Z_{j-1}^H M_{22} Z_{j-1})^{-1}\). Unlike CFQADI in Section 3, the Cholesky factorizations can always be done, even if the complex conjugate shift parameters are considered. Then, the Cholesky factor (CF), \(Z_j\), at the \(j\)-th step of CFQADI is obtained as

\[
Z_j = \left[ \sqrt{-2 \text{Re} p_j S_j^H C^T R_{1,j}^H} \left\{ Z_{j-1} - 2 \text{Re} p_j S_j^H Z_{j-1} \right. \right.

\left. -2 \text{Re} p_j S_j^H C^T V_j \left( I_q - D_j^2 \right)^{-1} D_j U_j^H B^T S_j^H Z_{j-1} \left. \right\} \right] R_{2,j}^H
\]

When a complex shift parameter is used, \(Z_j\) becomes complex. However, a real CF is required for positive real balanced truncation (PRBT). The following lemma is useful for obtaining a real CF.

**Lemma 1** If real and complex conjugate pair shift parameters are used, \(X_{j+1}\) is symmetric and positive semidefinite.

**Proof:** See Appendix. \(\blacksquare\)

As \(X_{j+1} = Z_{j+1}^H Z_{j+1}\), we can obtain a real CF, \([\text{Re} Z_{j+1} \quad \text{Im} Z_{j+1}]\), using Lemma 1. This indicates that the obtained CF is not a full column rank. In PRBT, the components corresponding to the zero Hankel singular values are eliminated; hence, it is not necessary for CF to be a full column rank.
4.2 Generalized algebraic riccati equation

CFQADI is presented for the generalized algebraic Riccati equation (GARE) solution, as in the case of ARE. QADI for GARE solution is obtained by

\[
\begin{align*}
(A^T + E^TY_{j-1}BB^T + p_j E^T)Y_{j-1/2}E &= -C^TC - E^TY_{j-1}^T(A - p_j E), \quad (37) \\
(A^T + E^TY_{j-1/2}BB^T + \tilde{p}_j E^T)Y_j E &= -C^TC - E^TY_{j-1/2}(A - \tilde{p}_j E). \quad (38)
\end{align*}
\]

Expressing \( \tilde{S}_j = (A + p_j E)^{-1} \) and \( \tilde{T}_j = A - p_j E \), the relationship, \( \tilde{S}_j \tilde{T}_j = I - 2p_j \tilde{S}_j E \), is obtained. As in the case of ARE, expressing \( B^T \tilde{S}_j^H C^T = U_j D_j V_j^H \), the solutions, \( Y_j E \), can be obtained by

\[
Y_j E = \tilde{M}_{11} + \tilde{M}_{12}Y_{j-1}E \left( I_n - \tilde{M}_{22}Y_{j-1}E \right)^{-1} \tilde{M}_{22}^H, \quad (39)
\]

where

\[
\begin{align*}
\tilde{M}_{11} &= -2\text{Rep}_j \tilde{S}_j^H C^T \left\{ I_q + V_j \left( D_j^{-2} - I_q \right)^{-1} V_j^H \right\} C \tilde{S}_j E, \\
\tilde{M}_{12} &= -2\text{Rep}_j \tilde{S}_j B \left\{ I_q + U_j \left( D_j^{-2} - I_q \right)^{-1} U_j^H \right\} B^T \tilde{S}_j^H E^T, \\
\tilde{M}_{22} &= I_n - 2\text{Rep}_j \tilde{S}_j^H E^T - 2\text{Rep}_j \tilde{S}_j^H C^T V_j \left( I_q - D_j^{-2} \right)^{-1} D_j U_j^H B^T \tilde{S}_j^H E^T.
\end{align*}
\]

Then, CF, \( Z_j \), at the \( j \)-th step is obtained as

\[
Z_j = \left[ \sqrt{-2\text{Rep}_j \tilde{S}_j^H C^T \hat{R}_{1,j}^H} \{ Z_{j-1} - 2\text{Rep}_j \tilde{S}_j^H E^T Z_{j-1} \} + \right. \\
2\text{Rep}_j \tilde{S}_j^H C^T V_j \left( I_q - D_j^{-2} \right)^{-1} D_j U_j^H B^T \tilde{S}_j^H E^T Z_{j-1}\left. \right] \hat{R}_{2,j}^H, \quad (40)
\]

in which Cholesky factorizations, \( \hat{R}_{1,j}^H \hat{R}_{1,j} = \left\{ I_q + V_j \left( D_j^{-2} - I_q \right)^{-1} V_j^H \right\} \) and \( \hat{R}_{2,j}^H \hat{R}_{2,j} = (I_{(j-1)m} - Z_{j-1}^H E \tilde{M}_{22} Z_{j-1})^{-1} \), are used.

The shift parameters must be determined for executing CFQADI. Assuming that \( E \) is nonsingular, GARE (2) is represented as

\[
E^{-T} A^T Y + YAE^{-1} + YBB^T Y + E^{-T} C^T C E^{-1} = 0. \quad (41)
\]

Hence, the eigen value problem related to the Hamiltonian matrix is expressed as

\[
\lambda \begin{bmatrix} \alpha \\ \gamma \end{bmatrix} = \begin{bmatrix} A E^{-1} & BB^T \\ -E^{-T} C^T C E^{-1} & -E^{-T} A^T \end{bmatrix} \begin{bmatrix} \alpha \\ \gamma \end{bmatrix}, \quad (42)
\]

Expressing \( E^{-1} \alpha = \beta \), the eigen value problem can be represented as

\[
\lambda \begin{bmatrix} E & 0 \\ 0 & E^T \end{bmatrix} - \begin{bmatrix} A & BB^T \\ -C^T C & -A \end{bmatrix} \begin{bmatrix} \beta \\ \gamma \end{bmatrix} = 0. \quad (43)
\]

Therefore, with \( Q = \text{diag}(E, E^T) \), the Hamiltonian matrix for GARE becomes \( Q^{-1} H \), if \( Q \) is nonsingular. However, as matrix, \( E \), is singular, we cannot solve the problem as done for ARE. In the next section, the procedures for calculating the eigenvalues are presented.

5. Shift-parameter determination

The alternating direction implicit iteration (ADI) converges rapidly, when the eigenvalues of the coefficient matrix and the reciprocal of the eigenvalues of the inverse coefficient matrix extracted by the Krylov subspace method [22] are used as the shift parameters [18]. Hence, we use certain stable eigen values of the Hamiltonian matrix, \( H \), of (32) and the reciprocal eigen values of \( H^{-1} \) extracted by the Krylov subspace method for the Cholesky factorized quadratic alternating direction implicit
iteration (CFQADI) of the algebraic Riccati equation (ARE). The Krylov subspace method tends to produce certain large eigen values. The eigen value problem is expressed as $Hx = \lambda x$. We rewrite it as $H^{-1}x = (1/\lambda)x = tx$. Hence, if the Krylov subspace method is applied to $H^{-1}$, certain small eigen values of $H$ are obtained by $\lambda = 1/t$.

Applying the same procedure to the generalized ARE (GARE), the generalized eigen problem, $(Q, H)$, is represented as $\lambda Qx = Hx$. We rewrite it as $H^{-1}Qx = (1/\lambda)x = tx$. Hence, if the Krylov subspace method is applied to $H^{-1}Q$, certain small generalized eigenvalues, $\lambda = 1/t$, are obtained. As the eigen value, $t$, obtained by the Krylov subspace method is relatively large, the generalized eigen value, $\lambda$, is relatively small.

Large generalized eigenvalues cannot be obtained because $Q$ is singular. In [4], an expansion point other than the origin is introduced to overcome the singularity issue for suscetptance-based model reduction. We use an expansion point other than the origin, for obtaining certain large generalized eigenvalues of $(Q, H)$. The corresponding generalized eigen value problem is expressed as $Qx = (1/\lambda)Hx = tHx$. Assuming an expansion point, $s_0$, the problem is again rewritten as $(Q - s_0H)x = (t - s_0)Hx$. As the matrix pencil, $(Q, H)$, is regular, $Q - s_0H$ is nonsingular. Thus, the problem is described by $(Q - s_0H)^{-1}Hx = 1/(t - s_0)x = \xi x$. Using the Krylov subspace method, we obtain an eigen value, $\lambda = 1/(1/\xi + s_0)$, which is large. The Krylov subspace method determines certain eigen values around the expansion point, $s_0$. As $t$ should be small, $s_0$ is assumed to be a small negative value.

After obtaining small and large eigen values by applying the two Krylov subspace methods presented in this section, we can apply the selection procedure, [18] based on the minimax function (31). However, as this procedure is not always effective [23], we use a fixed number of small and large eigenvalues, in addition, without applying the selection procedure.

The above discussion is effective in decreasing the residual error of ARE or GARE. It is not necessarily sufficient for obtaining an accurate reduced-order model. The main purpose of model order reduction for RLC networks is to obtain a low-frequency accurate model. Most balanced truncation methods are more accurate than model order reduction by the Krylov subspace methods; however, the accuracy at low frequencies is worse than that of the Krylov subspace methods. As an alternative approach, model reduction of the reciprocal system is proposed; however, this method sacrifices the accuracy, at high frequencies [16]. Therefore, the key issue for balanced truncation for RLC networks is the accuracy with which the responses at high frequencies can be approximated, after ensuring the accuracy at low frequencies.

Through numerical examples, we demonstrate that low-frequency accurate models can be obtained by considering small exact eigen values as the shift parameters. However, exact eigen values cannot be used for actual problems because eigen value decomposition is very expensive for large-scale problems. Fortunately, a certain small value can be used as one of the shift parameters, instead of the exact small eigen value, to complement the accuracy at low frequencies. With numerical examples, we demonstrate that this technique is also effective in the Newton method, which uses ADI for the related Lyapunov-equation solutions.

6. Results
The proposed method was evaluated using circuits, whose fundamental structures are shown in Figs. 2(a) and 2(b); these circuits are index-1 and -2, respectively, when the impedance matrices are considered. Although their indices differ, their structures are nearly the same. This implies that RLC networks with index-2 are common. In the numerical examples, we increased the number of RLC $(1 \Omega, 1 [nH], 1 [nF])$ sections shown in Figs. 2(a) and 2(b).

6.1 Index-1 system
The (improved) Cholesky factorized quadratic alternating direction implicit iteration (CFQADI) presented in Section 3 was applied to algebraic Riccati equation (ARE) for an index-1 system in which one hundred RLC sections of Fig. 2(a) were considered; the dimensions of the matrices in (10) are $\hat{A}_0 \in \mathbb{R}^{200,200}$, $\hat{B}_0 \in \mathbb{R}^{200,2}$, and $\hat{C}_0 \in \mathbb{R}^{2,200}$. Figure 3 shows CFQADI convergence curves; ‘eig’
Fig. 2. Fundamental circuit structures with (a)index-1 and (b)index-2.

Fig. 3. (Improved) CFQADI convergence curves.

indicates that the 20 shift parameters were obtained by eigen value decomposition; for the ‘lrg’, corresponding to the original CFQADI, the largest eigen value was used as the shift parameter, ‘kry_lrg’ and ‘kry_sml’ indicate that the 20 eigen values were extracted by the Krylov subspace method for the Hamiltonian matrix, \( H \), of (32) and the inverse matrix, \( H^{-1} \), respectively. The ten shift parameters obtained from \( H \) and \( H^{-1} \), each, (20 in total) were used in ‘kry_mixed’; for the ‘minimax’, the 20 shift parameters were selected by Penzl’s procedure [18]. In Fig. 3, the residual error is evaluated by the Frobenius norm\(^3\) on the left-side of ARE (1). Note that the convergence is evaluated by the Frobenius norm on the left-side of ARE (1) or generalized ARE (GARE) (2), which is expressed as ‘error(ARE)’ or ‘error(GARE)’, throughout this paper.

ARE solution is the steady state solution of the differential equation,

\[
\frac{dX}{dt} = A^T X + XA + XB^T X + C^T C.
\]

QADI corresponds to the application of the backward Euler method to the differential equation. Therefore, a large shift parameter corresponds to a small time-step size of the Euler method. Using a small time-step size, we can reduce the residual error considerably in the initial steps of QADI; however, the small step-size becomes insufficient for reducing the error, as QADI iterations increase. Therefore, we need both small and large shift parameters; the ‘kry_mixed’ case with small and large shift parameters exhibits good performance. However, the selection procedure [18] is not effective, as shown in Fig. 3. For the ‘eig’, the 20 largest eigen values were considered and the convergence curve nearly overlapped with that of the ‘lrg’. This implies that eigen values with various magnitudes are necessary for the rapid reduction of the residual error.

To investigate the influence of the shift parameters on the frequency response of the reduced-order model obtained by positive real balanced truncation (PRBT), we considered the 20 largest and

\[^3\text{The Frobenius norm is defined by } \|A\|_F = \sum_{i=1}^{m} \sum_{j=1}^{m} |a_{ij}|^2, \text{ in which } a_{ij} \text{ is the } (i, j) \text{ element of the } m \times n \text{ matrix, } A.\]

489
smallest eigen values, each, obtained by eigen value decomposition. Figures 4(a) and 4(b) show the convergence curves of CFQADI and the relative errors of the frequency responses of the 15-th reduced-order model obtained by PRBT with CFQADI, respectively. Note that all the frequency responses are obtained using the 15-th reduced-order model, throughout this paper and the relative error is evaluated by \( \| \hat{Z}(j\omega) - Z(j\omega) \|_2 / \| Z(j\omega) \|_2 \), in which \( Z(j\omega) \) is the response of the network, which is not reduced. In Figs. 4(a) and 4(b), ‘eig_0’ and ‘eig_1’ indicate that 20 smallest and largest eigen values, each, are used, respectively. From Fig. 4(a), it can be observed that large shift parameters are effective in decreasing the residual error significantly; however, the accuracy at low frequencies is not improved, as shown in Fig. 4(b); large shift parameters mainly depend on the accuracy at high frequencies.

The relative errors of the frequency responses by CFQADI, with shift parameters obtained from the Krylov subspace method, are displayed in Fig. 5(a). The responses are not accurate at low frequencies, as depicted by ‘kry_lrg’ and ‘kry_sml’ in which the Krylov subspace method is applied to the Hamiltonian matrix, \( H \) and its inverse, \( H^{-1} \), respectively. We can improve the accuracy using a constant shift parameter, \(-10^{-3}\). With the 15 shift parameters obtained from the Krylov subspace method and the same five constant shift parameters, 30 CFQADI iterations were conducted. The responses of ‘kry_lrg_c’ and ‘kry_sml_c’ correspond to ‘kry_lrg’ and ‘kry_sml’, in the case, where a constant shift parameter was appended. Using the constant shift parameter, the accuracy at low frequencies is improved. CFQADI with a constant shift parameter only, provided the response, as denoted by ‘const’ in Fig. 5(a). It is highly accurate at low frequencies, improving the accuracies of ‘kry_lrg’ and ‘kry_sml’ at low frequencies. Figure 5(b) shows CFQADI convergence curves. From Figs. 4(a) and 4(b), if the curve attains a small error, the accuracy at high frequencies should be good. Although curves ‘kry_lrg_c’ and ‘kry_sml_c’ attain small errors, in Fig. 5(b), the corresponding responses are less accurate at high frequencies, compared to ‘kry_sml’. Moreover, the convergence curve of CFQADI with a constant parameter only, is nearly flat, as indicated by the ‘const’ of Fig. 5(b). This implies that an accurate response is not necessarily obtained, even if better CFQADI convergence is demonstrated.

Appending a constant shift parameter is also effective in the Newton method with Cholesky factorized ADI (CFADI) [23]. At each step of the Newton method, a Lyapunov equation is solved. Then, CFADI is applied to the solution of the Lyapunov equation. With five iterations of the Newton method and 30 CFADI iterations with 20 shift parameters, the frequency response is obtained by PRBT and the relative errors are shown in Fig. 5(c); the legends are the same as those in Fig. 5(a).
except that ‘QADI’ is the ‘kry_lrg_c’ of Fig. 5(a). The accuracy at low frequencies is improved by appending a constant shift parameter. It is proved in [28] that certain fast Riccati solvers, including CFQADI, are mathematically identical. Therefore, if the frequency response obtained from one of these methods is inaccurate at low frequencies, appending a constant shift parameter may improve its accuracy.

6.2 Index-2 system
As an index-2 system, the circuit with hundred RLC sections in Fig. 2(b) was analyzed. The two-port impedance matrix was reduced by PRBT-2, where $A_0 \in \mathbb{R}^{301,301}$, $B_0 \in \mathbb{R}^{301,2}$, and $C_0 \in \mathbb{R}^{2,301}$ in (3). The Krylov subspace method for $(Q - s_0 H)^{-1} H$ with $s_0 = -10^{-5}$ and $H^{-1}Q$ was applied, corresponding to ‘kry_lrg’ and ‘kry_sml’, respectively.

Figure 6(a) shows the relative errors of the frequency responses in which 20 eigen values were considered; the response labels are the same as those in the previous examples. From Fig. 4(b), the accuracy at low and high frequencies depends upon the small and large eigen values, respectively because the 20 smallest and largest eigen values, each, were used for CFQADI, corresponding to ‘eig_0’
Fig. 6. Relative errors of the reduced-order model frequency responses for an index-2 system in which (a) 20 shift parameters are used for CFQADI and (b) a constant shift parameter is appended, and (c) CFQADI convergence curves.

and ‘eig_1’, respectively. However, the shift parameters extracted by the Krylov subspace method are not sufficient to capture accurate responses at low frequencies, as depicted by ‘kry_lrg’ and ‘kry_sml’.

With the same five small shift parameters, $-10^{-3}$ and the 15 shift parameters obtained by the Krylov subspace method, 30 CFQADI iterations were conducted. Figure 6(b) shows the relative errors of the frequency responses, in this case. The accuracy at low frequencies is improved by appending a constant shift parameter. When only a constant shift parameter is used for CFQADI, a low-frequency accurate reduced-order model is obtained, as depicted by ‘const’. However, as the middle-to-high frequency accuracy is not improved, this model cannot be used, suggesting that the shift parameters extracted by the Krylov subspace method are necessary. Figure 6(c) shows CFQADI convergence curves. From the profile of the convergence curves, it cannot be determined whether an accurate reduced-order model is obtained.

For comparison, we applied the Newton method, in addition, for solving GARE with CFADI [29]. Figure 7(a) shows the relative errors of the frequency responses obtained by the Newton method. In this example, the relationship between ‘eig_0’ and ‘eig_1’ is opposite to that in Fig. 4(b). For the Newton method, a generalized projected Lyapunov equation is solved by ADI [11] in which matrix,
$A + EXBB^T$, of (2) is inverted to generate a standard Lyapunov equation. Therefore, the large shift parameter corresponds to the small shift parameter of the other examples.

The frequency responses of the Newton method with CFADI, using $-10^{-5}$ ('const-') and $-10^{+5}$ ('const+') as the shift parameters, were calculated for investigating the effects of the constant shift parameters. Figure 7(b) shows the relative errors of the frequency responses. A large shift parameter renders the low-frequency region accurate, whereas a small shift parameter renders the high-frequency accurate. Therefore, we should use a large shift parameter to compensate for the accuracy at low frequencies.

With the same five shift parameters, $-10^{+5}$ and the 15 eigen values obtained from the Krylov subspace method, CFADI was carried out. Figure 7(c) shows the relative errors of the frequency responses, in this case. The accuracy of the response, 'kry_sml', at low frequencies is improved, as depicted by 'kry_sml_c'. 'kry_lrg' is not improved at low frequencies; however, the accuracy at high frequencies is marginally improved, as depicted by 'kry_lrg_c'. Therefore, we should further examine the determination of the constant shift parameter.
Table I. CFQADI and RADI computational time for an index-1 system.

| size         | RADI [s] | QADI [s] |
|--------------|----------|----------|
| 3,001        | 0.11     | 0.51     |
| 300,001      | 12.25    | 48.41    |
| 30,000,001   | 1,289.85 | 5,399.45 |

Table II. CFQADI and RADI computational time for an index-2 system.

| size         | RADI [s] | QADI [s] |
|--------------|----------|----------|
| 3,001        | 0.16     | 0.73     |
| 300,001      | 18.39    | 76.73    |
| 30,000,001   | 3,005.42 | 19,702.38|

6.3 Computational time

Finally, we compare the computational time of CFQADI with RADI, in which RADI is a state- 
of-the-art ADI method for solving both AREs and GAREs [28]. For the comparison, 30 iterations 
were carried out and the 10 shift parameters extracted by the Krylov subspace method were used, 
respectively. The simulations were carried out on a computer equipped with a 3.7-GHz Intel Xeon 
E5-1620 on a 32-GB memory.

Tables I and II show the computational time for the index-1 and -2 systems, respectively, in which ‘size’ indicates the order of the solution vector of the descriptor system (3). RADI is maximally 4.6 times faster than CFQADI for an index-1 system and is maximally 6.6 times faster than for an index-2 one. As proved in [28], the RADI convergence curve associated with the residual norm of ARE was identical to that of CFQADI. Moreover, the RADI curve for GARE was also identical to that of CFQADI. Therefore, using RADI instead of the proposed CFQADI, we can obtain low frequency accurate reduced-order models efficiently for index-1 and -2 systems.

7. Conclusions

We have presented an efficient iterative method for solving standard and generalized algebraic Riccati equations for RLC networks. This method provides a low-rank Riccati equation solution for positive real balanced truncation; the low-rank solution of a standard algebraic Riccati equation for an index-1 system and that of a generalized equation for an index-2 system are obtained. For the reduced-order modeling of RLC networks, a low-frequency accurate model is required. A small shift parameter is used to obtain this model. Recently, fast algebraic Riccati-equation solvers have been proposed [28, 30]. These methods focus on the rapid convergence of the residual norm of the algebraic Riccati equation, which can be effective in rendering the frequency response accurate at high frequencies. However, rapid convergence is not necessarily sufficient for low-frequency accuracy. To compensate for the accuracy at low frequencies, the small shift parameter examined in this paper is effective in these methods.

Acknowledgments

This work was supported in part by the JSPS KAKENHI Grant Number: 16K00073. The author also thanks the anonymous reviewers for their useful suggestions.

Appendix

Proof of Lemma 1: The difference, $X - X_{j+1}$, is Hermitian and positive semidefinite, and the solution, $X$, is symmetric and positive semidefinite. Thus, we only need to prove that the difference is real. Write $\bar{S}_j = (\bar{A} + p_j I)^{-1}$ and $\bar{T}_j = \bar{A} - p_j I$, where $\bar{A} = \bar{A} + XBB^T$. Then, the solutions are

$$X_{j-1/2} = F_l \left( \begin{bmatrix} 0 & -\bar{T}_j^T \\ \bar{S}_j & \bar{S}_j BB^T \end{bmatrix}, X_{j-1} \right), \quad X_j = F_l \left( \begin{bmatrix} 0 & \bar{S}_j^T \\ -\bar{T}_j & BB^T \bar{S}_j^T \end{bmatrix}, X_{j-1/2} \right). \quad (A-1)$$

Taking the star product, we have
\[ X_j = F_l \left( \begin{bmatrix} 0 & -\tilde{S}_j^T \tilde{T}_j^T \\ -\tilde{S}_j \tilde{T}_j & 2 \text{Rep}_j \tilde{S}_j \tilde{B} \tilde{B}^T \tilde{S}_j^T \end{bmatrix}, X_{j-1} \right). \] (A-2)

Therefore, if the shift parameter, \( p_{j+1} = \bar{p}_j \), the solution, \( X_{j+1} \), is

\[ X_{j+1} = F_l \left( \begin{bmatrix} 0 & -\tilde{S}_j^T \tilde{T}_j^T \\ -\tilde{S}_j \tilde{T}_j & 2 \text{Rep}_j \tilde{S}_j \tilde{B} \tilde{B}^T \tilde{S}_j^T \end{bmatrix}, X_{j-1} \right) \]

\[ = F_l \left( \begin{bmatrix} 0 & -\tilde{S}_j^T \tilde{T}_j^T \\ -\tilde{S}_j \tilde{T}_j & 2 \text{Rep}_j \tilde{S}_j \tilde{B} \tilde{B}^T \tilde{S}_j^T + 2 \text{Rep}_j \tilde{S}_j \tilde{T}_j \tilde{S}_j^T \tilde{T}_j^T \end{bmatrix}, X_{j-1} \right). \] (A-3)

From the commutative properties of \( \tilde{S}_j \) and \( \tilde{T}_j \) [19], the (1,2) and (2,1) blocks of the inner matrix of (A-3) are real. The (2,2) block is represented as

\[ 2 \text{Rep}_j \left( \tilde{S}_j \tilde{B} \tilde{B}^T \tilde{S}_j^T + \tilde{S}_j \tilde{B} \tilde{B}^T \tilde{S}_j^T \right) - 4 (\text{Rep}_j)^2 \tilde{S}_j \tilde{S}_j \tilde{B} \tilde{B}^T \tilde{S}_j^T \tilde{S}_j^T A^T - 4 (\text{Rep}_j)^2 A \tilde{S}_j \tilde{S}_j \tilde{B} \tilde{B}^T \tilde{S}_j^T \tilde{S}_j^T. \]

We can see that the (2,2) part is also real.

References

[1] A. Odabasioglu, M. Celik, and L.T. Pileggi, “PRIMA: Passive reduced-order interconnect macromodeling algorithm,” IEEE Trans. Comput. -Aided Design Integr. Circuit Syst., vol. 17, no. 8, pp. 645–654, 1998.

[2] R.W. Freund, “Passive reduced-order models for interconnect simulation and their computation via Krylov-subspace algorithms,” Numerical Analysis Manuscript, Bell Laboratories, 98-3-06, 1998.

[3] K.J. Kerns and A.T. Yang, “Stable and efficient reduction of large multiport RC networks by pole analysis via congruence transformation,” IEEE Trans. Comput. -Aided Design Integr. Circuit Syst., vol. 16, no. 7, pp. 734–744, 1997.

[4] B.N. Sheehan, “ENOR: Model order reduction of RLC circuits using nodal equations for efficient factorization,” Proc. Des. Autom. Conf., pp. 17–21, 1999.

[5] B.N. Sheehan, “Realizable reduction of RC networks,” IEEE Trans. Comput. -Aided Design Integr. Circuit Syst., vol. 26, no. 8, pp. 1393–1407, 2007.

[6] N. Wong and V. Balakrishnan, “Fast positive-real balanced truncation via quadratic alternating direction implicit iteration,” IEEE Trans. Comput. -Aided Design Integr. Circuit Syst., vol. 26, no. 9, pp. 1725–1731, 2007.

[7] K. Zhou, J.C. Doyle, and K. Glover, Robust and Optimal Control, NJ: Prentice Hall, 1995.

[8] N. Wong and V. Balakrishnan, “Multi-shift quadratic alternating direction implicit iteration for high-speed positive-real balanced truncation,” Proc. IEEE Des. Autom. Conf., pp. 257–260, 2006.

[9] N. Wong, “Efficient positive-real balanced truncation of symmetric systems via cross-Riccati equations,” IEEE Trans. Comput. -Aided Design Integr. Circuit Syst., vol. 27, no. 3, pp. 470–480, 2008.

[10] T. Stykel, “Low-rank iterative methods for projected generalized Lyapunov equations,” Electronic Trans. Numerical Analysis, vol. 30, pp. 187–202, 2008.

[11] T. Reis and T. Stykel, “PABTEC: Passivity-preserving balanced truncation of electrical circuits,” IEEE Trans. Comput. -Aided Design Integr. Circuit Syst., vol. 29, no. 9, pp. 1354–1367, 2010.

[12] T. Stykel, “Gramian based model reduction for descriptor systems,” Mathematics of Control, Signals, and Systems, vol. 16, pp. 297–319, 2004.

[13] T. Reis and T. Stykel, “Positive real and bounded real balancing for model reduction of descriptor systems,” Int. J. Control, vol. 83, pp. 74–88, 2008.

[14] Z. Zhang and N. Wong, “An efficient projector-based passivity test for descriptor systems,” IEEE Trans. Comput. -Aided Design Integr. Circuit Syst., vol. 29, no. 8, pp. 1203–1214, 2010.
[15] R. Lamour, R. Marz, and C. Tischendorf, “Differential-Algebraic Equations: A Projector Based Analysis,” Springer, 2013.

[16] J.R. Phillips, L. Daniel, and L.M. Silveira, “Guaranteed passive balancing transformations for model order reduction,” *IEEE Trans. Comput. -Aided Design Integr. Circuit Syst.*, vol. 22, no. 8, pp. 1–15, 2003.

[17] E. Wachspress, The ADI Model Problem, Springer, 2013.

[18] T. Penzl, “A cyclic low-rank smith method for large sparse Lyapunov equations,” *SIAM J. Sci. Comput.*, vol. 21, no. 4, pp. 1401–1418, 2000.

[19] J.R. Li and J. White, “Low-rank solution of Lyapunov equations,” *SIAM Review*, vol. 46, no. 4, pp. 693–713, 2004.

[20] P. Benner, J.R. Li, and T. Penzl, “Numerical solution of large-scale Lyapunov equations, Riccati equations, and linear-quadratic optimal control problems,” *Numer. Linear Algebra Appl.*, vol. 15, pp. 755–777, 2008.

[21] D.L. Kleinman, “On an iterative technique for Riccati equation computations,” *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 114–115, 1968.

[22] G.H. Golub and C.F. Von Loan, “Matrix Computation 4th Ed., Baltimore: The Johns Hopkins University Press,” 2013.

[23] Y. Tanji, “Real Cholesky factor-ADI method for low-rank solution of projected generalized Lyapunov equations,” *IEICE Fundamentals*, vol. E99-A, pp. 702–709, 2016.

[24] D.G. Luenberger, “Dynamic equations in descriptor form,” *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 312–321, 1977.

[25] B.D.O. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis*, NY: Dover, 2006.

[26] M. Gerdin, “Computation of a canonical form for linear differential-algebraic equations,” *Technical Reports of Linköpings Universitet*, no. LiTh-ISY-R-2002, 2004.

[27] T. Katayama, *Senkei-System-no-Saiteki-Seigyo*, Kindai Kagaku Sya (in Japanese).

[28] P. Benner, Z. Bujanović, P. Kürschner, J. Saak, “RADI: A low-rank ADI-type algorithm for large scale algebraic Riccati equations,” *Numerische Mathematik*, DOI 10.1007/s00211-017-0907-5, 2017.

[29] P. Benner and T. Stykel, “Numerical solution of projected algebraic Riccati equations,” *SAIM J. Numer. Anal.*, vol. 52, no. 2, pp. 581–600, 2014.

[30] A. Massoudi, M. Opmeer, and T. Reis, “The ADI method for bounded real and positive real Lur’e equations,” *Numerische Mathematik*, vol. 135, no. 2, pp. 431–458, 2017.

[31] N. Wong, V. Balakrishnan, C.K. Koh, and T.S. Ng, “Two algorithms for fast and accurate passivity-preserving model order reduction,” *IEEE Trans. Comput. -Aided Design Integr. Circuit Syst.*, vol. 25, no. 10, pp. 2062–2075, 2006.