CONVERGENCE ACCELERATION OF SHIFTED LR TRANSFORMATIONS FOR TOTALLY NONNEGATIVE HESSENBERG MATRICES

Akiko Fukuda, Saitama, Yusaku Yamamoto, Tokyo,
Masashi Iwasaki, Kyoto, Emiko Ishiwata, Tokyo,
Yoshimasa Nakamura, Kyoto

Received December 30, 2019. Published online September 7, 2020.

Abstract. We design shifted LR transformations based on the integrable discrete hungry Toda equation to compute eigenvalues of totally nonnegative matrices of the banded Hessenberg form. The shifted LR transformation can be regarded as an extension of the extension employed in the well-known dqds algorithm for the symmetric tridiagonal eigenvalue problem. In this paper, we propose a new and effective shift strategy for the sequence of shifted LR transformations by considering the concept of the Newton shift. We show that the shifted LR transformations with the resulting shift strategy converge with order $2-\varepsilon$ for arbitrary $\varepsilon > 0$.

Keywords: LR transformation; totally nonnegative matrix; Newton shift; convergence rate

MSC 2020: 34B16, 34C25

1. Introduction

Rutishauser [13] presented the quotient-difference (qd) algorithm, which has a recursion formula incorporating the quotient and the difference, for computing eigenvalues of symmetric tridiagonal matrices. Fernando and Parlett [2] showed that the qd algorithm can be applied to compute singular values of bidiagonal matrices. The differential form of the qd (dqd) algorithm is a subtraction-free version, and the dqd with shift (dqds) algorithm was formulated by introducing a shift of origin to the

This research was partially supported by Grants-in-Aid for Scientific Research © Number 19K03624 from the Japan Society for the Promotion of Science.

DOI: 10.21136/AM.2020.0378-19

dqd algorithm to accelerate convergence [2], [10], [13]. The well-known linear algebra package LAPACK [8] adopts the dqds algorithm as a solver for singular values.

The original qd algorithm repeatedly employs the following recursion formula:

$$
\begin{cases}
q_k^{(n+1)} = q_k^{(n)} - e_{k-1}^{(n+1)} + e_k^{(n)}, \quad k = 1, 2, \ldots, m, \\
e_k^{(n+1)} = \dfrac{q_{k+1}^{(n)}}{q_k^{(n+1)}} e_k^{(n)}, \quad k = 1, 2, \ldots, m-1, \\
e_0^{(n)} \equiv 0, \quad e_m^{(n)} \equiv 0, \quad n = 0, 1, \ldots,
\end{cases}
\tag{1.1}
$$

where $q_k^{(n)}$ and $e_k^{(n)}$ are variables that define a symmetric tridiagonal matrix and the superscript n refers to the iteration number. The qd recursion formula (1.1) generates a similarity transformation, known as the LR transformation, from a symmetric tridiagonal matrix defined by $q_k^{(n)}$ and $e_k^{(n)}$ to one defined by $q_k^{(n+1)}$ and $e_k^{(n+1)}$. It corresponds to computing the LR decomposition of the tridiagonal matrix and multiplying the L and R factors in the reverse order. Note that the qd recursion formula for the LR transformation is simply the integrable discrete Toda equation, which is a representative discrete integrable system. In the case of the discrete Toda equation, the superscript n and subscript k are the discrete time and spatial variables, respectively.

One extension of the discrete Toda equation (1.1) is the discrete hungry Toda (dhToda) equation:

$$
\begin{cases}
Q_k^{(n+M)} = Q_k^{(n)} + E_k^{(n)} - E_{k-1}^{(n+1)}, \quad k = 1, 2, \ldots, m, \\
E_k^{(n+1)} = \dfrac{Q_{k+1}^{(n)}}{Q_k^{(n+M)}} E_k^{(n)}, \quad k = 1, 2, \ldots, m-1, \\
E_0^{(n)} \equiv 0, \quad E_m^{(n)} \equiv 0, \quad n = 0, 1, \ldots
\end{cases}
\tag{1.2}
$$

Here, M is a positive integer. The dhToda equation (1.2) was derived in the study of box and ball systems (BBS), see [15], and differs from the discrete Toda equation (1.1) in that it has an additional parameter M. The dhToda equation (1.2) with $M = 1$ coincides with the discrete Toda equation (1.1). In a previous work, we designed an algorithm for computing eigenvalues of totally nonnegative (TN) matrices of the banded Hessenberg form, where a TN matrix is a matrix with all minors nonnegative [3]. Since this algorithm is based on the dhToda equation (1.2), it is called the *dhToda algorithm*. The positive integer M corresponds to the bandwidth of the target TN matrix in the dhToda algorithm. Algorithms for computing eigenvalues of TN matrices have been also designed using the discrete hungry Lotka-Volterra system [3] and the discrete Bogoyavlensky lattice [14].

Symmetric positive-definite tridiagonal matrices belong to a class of TN matrices, and so the dhToda algorithm can be regarded as a generalization of the qd algorithm. As in the case of the qd algorithm, it is easy to derive the differential form of the dhToda algorithm [3]. However, the discussion in this paper makes no distinction between the original form and the differential form of the dhToda algorithm, because they are mathematically equivalent. Like the qd algorithm, the dhToda algorithm can be interpreted as a recursion formula for generating the LR transformation of the TN banded Hessenberg matrix. To accelerate convergence, we developed the shifted dhToda equation by introducing a shift of origin into this LR transformation [5]. Further, we showed that the shifted dhToda algorithm acts without breakdown if the shift at each step is chosen to be smaller than the minimum eigenvalue of the target TN matrix [5]. We proved that the shifted dhToda algorithm is numerically stable in floating point arithmetic [4]. However, no concrete shift strategy has been presented in the literature yet. The main purpose of this paper is to propose an effective shift strategy for the sequence of shifted LR transformations for the TN banded Hessenberg matrix, and then to analyze its advantages with respect to convergence rate.

The remainder of this paper is organized as follows. In Section 2, we briefly review the shifted LR transformation for TN matrices of the banded Hessenberg form, which is derived from the study of the dhToda equation (1.2). In Section 3, we propose a shift strategy based on the Newton shift. In Section 4, we clarify the properties of the minimum eigenvalue of the TN matrix, and in Section 5 we focus on the bottom-right entry of the TN matrix and its neighboring entries. In Section 6, we investigate the convergence rate of the sequence of shifted LR transformations under the proposed shift strategy. We also numerically verify the convergence acceleration through some examples. Finally, we provide concluding remarks in Section 7.

2. The shifted LR transformation based on the dhToda equation

This section briefly reviews our previous papers [3], [5] concerning the shifted LR transformation for a TN matrix based on the dhToda equation (1.2).

We begin by relating the dhToda equation (1.2) to the LR transformation for a TN matrix. We showed in [3] that the dhToda equation (1.2) has the matrix representation

$$L^{(n+M)}R^{(n+1)} = R^{(n)}L^{(n)}, \quad n = 0, 1, \ldots, \tag{2.1}$$

where $L^{(n)}$ and $R^{(n)}$ are the lower and upper bidiagonal matrices involving the dhToda variables $Q_k^{(n)}$ and $E_k^{(n)}$, which are defined as:

$$L^{(n)} := \begin{pmatrix} Q_1^{(n)} & & & \\ 1 & Q_2^{(n)} & & \\ & \ddots & \ddots & \\ & & 1 & Q_m^{(n)} \end{pmatrix}, \quad R^{(n)} := \begin{pmatrix} 1 & E_1^{(n)} & & \\ & 1 & \ddots & \\ & & \ddots & E_{m-1}^{(n)} \\ & & & 1 \end{pmatrix}.$$

Here, we introduce an m-by-m lower Hessenberg matrix $A^{(n)}$ as the product of $L^{(n)}$, $L^{(n+1)}$, …, $L^{(n+M-1)}$ and $R^{(n)}$, namely:

$$A^{(n)} := L^{(n)}L^{(n+1)} \dots L^{(n+M-1)}R^{(n)}. \tag{2.2}$$

If $Q_1^{(n)} > 0,\ Q_2^{(n)} > 0, \dots, Q_m^{(n)} > 0$ for $n = 0, 1, \dots, M-1$ and $E_1^{(0)} > 0$, $E_2^{(0)} > 0, \dots, E_{m-1}^{(0)} > 0$, then it holds that $Q_1^{(n)} > 0,\ Q_2^{(n)} > 0, \dots, Q_m^{(n)} > 0$ and $E_1^{(n)} > 0,\ E_2^{(n)} > 0, \dots, E_{m-1}^{(n)} > 0$ for any n. In other words, both $L^{(n)}$ and $R^{(n)}$ always have positive bidiagonal entries. Since both $L^{(n)}$ and $R^{(n)}$ are TN for any n, the lower Hessenberg matrix $A^{(n)}$ is also TN, see [11]. Regarding (2.1) as the LR transformation and using it repeatedly, we can rewrite $A^{(n+M)}$ as:

$$\begin{aligned} A^{(n+M)} &= L^{(n+M)}L^{(n+M+1)} \dots L^{(n+2M-1)}R^{(n+M)} \\ &= L^{(n+M)}L^{(n+M+1)} \dots R^{(n+M-1)}L^{(n+M-1)} \\ &\ \ \vdots \\ &= R^{(n)}L^{(n)}L^{(n+1)} \dots L^{(n+M-1)}. \end{aligned}$$

This implies that the dhToda equation (1.2) generates an LR transformation from $A^{(n)}$ to $A^{(n+M)}$, where

$$\begin{cases} A^{(n)} = (L^{(n)}L^{(n+1)} \dots L^{(n+M-1)})R^{(n)}, \\ A^{(n+M)} = R^{(n)}(L^{(n)}L^{(n+1)} \dots L^{(n+M-1)}). \end{cases} \tag{2.3}$$

Since it follows from (2.3) that $R^{(n)}A^{(n)}(R^{(n)})^{-1} = A^{(n+M)}$, we see that $A^{(n+M)}$ is similar to $A^{(n)}$.

Moreover, our previous paper [3] showed that the dhToda variables $Q_k^{(n)}$ and $E_k^{(n)}$ have the following asymptotic behaviors as $n \to \infty$:

$$\lim_{n\to\infty} \prod_{p=0}^{M-1} Q_k^{(n-p)} = c_k, \quad k = 1, 2, \dots, m, \tag{2.4}$$

$$\lim_{n\to\infty} E_k^{(n)} = 0, \quad k = 1, 2, \dots, m-1, \tag{2.5}$$

where $c_1, c_2, \ldots, c_m$ are positive constants such that $c_1 \geqslant c_2 \geqslant \ldots \geqslant c_m$. Considering (2.4) and (2.5) in the entries of $A^{(n)}$, we observe that $A^{(n)}$ converges to a lower triangular matrix with diagonal entries $c_1, c_2, \ldots, c_m$ as $n \to \infty$. Therefore, $c_1, c_2, \ldots, c_m$ coincide with the eigenvalues of $A^{(0)}$. Our previous paper [3] presented the dhToda algorithm for computing eigenvalues of the TN matrix $A^{(0)} = (L^{(0)}L^{(1)} \ldots L^{(M-1)})R^{(0)}$ based on the above properties of the dhToda equation (1.2).

To accelerate the convergence of the dhToda algorithm, we introduced (see [5]) a shift of origin into the LR transformation (2.3) as

$$\begin{cases} A^{(n)} - s^{(n)}I = L^{(n)}L^{(n+1)} \ldots L^{(n+M-1)}R^{(n)} - s^{(n)}I = \bar{L}^{(n)}R^{(n,0)}, \\ A^{(n+M)} = R^{(n,0)}\bar{L}^{(n)} + s^{(n)}I, \end{cases} \tag{2.6}$$

where $s^{(n)}$ is a shift of origin, $\bar{L}^{(n)}$ is a lower triangular matrix, and $R^{(n,0)}$ is an upper bidiagonal matrix whose diagonal entries are all 1. The shifted LR transformation (2.6) immediately leads to $A^{(n+M)} = R^{(n,0)}A^{(n)}(R^{(n,0)})^{-1}$, which implies that $A^{(n+M)}$ is similar to $A^{(n)}$. Our previous paper [5] proved that the shifted LR transformation (2.6) does not fail if $s^{(n)}$ is smaller than the minimum eigenvalue of $A^{(n)}$. This shift strategy simultaneously guarantees the TN property of $A^{(n+M)}$ if $A^{(n)}$ is TN.

In the non-shifted case, the LR transformation from $A^{(n)}$ to $A^{(n+M)}$ is performed as a sequence of LR transformations (2.1) involving only bidiagonal matrices. We showed that this structure is essential for computing small eigenvalues of $A^{(0)}$ with high relative accuracy [4]. We reformulate the shifted LR transformation (2.6) also as a sequence of bidiagonal LR and RR transformations [5]. Assume for the moment that $R^{(n,0)}$ has been computed in some way. Then compute the bidiagonal matrices $L^{(n+M)}$, $L^{(n+M+1)}$, …, $L^{(n+2M-1)}$, $R^{(n,1)}, R^{(n,2)}, \ldots, R^{(n,M)}$ and $R^{(n+M)}$ by

$$L^{(n+M+p)}R^{(n,p+1)} = R^{(n,p)}L^{(n+p)}, \quad p = 0, 1, \ldots, M-1, \tag{2.7}$$

$$R^{(n+M)}R^{(n,0)} = R^{(n,M)}R^{(n)}. \tag{2.8}$$

From (2.7) and (2.8), we derive

$$\begin{aligned} A^{(n+M)} &= R^{(n,0)}A^{(n)}(R^{(n,0)})^{-1} \\ &= R^{(n,0)}L^{(n)}L^{(n+1)} \ldots L^{(n+M-2)}L^{(n+M-1)}R^{(n)}(R^{(n,0)})^{-1} \\ &= L^{(n+M)}R^{(n,1)}L^{(n+1)} \ldots L^{(n+M-2)}L^{(n+M-1)}R^{(n)}(R^{(n,0)})^{-1} \\ &\ \ \vdots \\ &= L^{(n+M)}L^{(n+M+1)}L^{(n+M+2)} \ldots R^{(n,M-1)}L^{(n+M-1)}R^{(n)}(R^{(n,0)})^{-1} \\ &= L^{(n+M)}L^{(n+M+1)}L^{(n+M+2)} \ldots L^{(n+2M-1)}R^{(n,M)}R^{(n)}(R^{(n,0)})^{-1} \\ &= L^{(n+M)}L^{(n+M+1)}L^{(n+M+2)} \ldots L^{(n+2M-1)}R^{(n+M)}. \end{aligned}$$

This implies that the bidiagonal LR transformations (2.7) and the bidiagonal RR transformation (2.8) generate a similarity transformation from $A^{(n)}$ to $A^{(n+M)}$. Moreover, by letting $E_k^{(n,p)}$ denote the $(k, k+1)$ entry of $R^{(n,p)}$, we obtain the following recursion formula for giving the LR transformations (2.7) and the RR transformation (2.8), respectively:

$$
\text{(2.9)} \qquad \begin{cases} Q_k^{(n+M+p)} = Q_k^{(n+p)} + E_k^{(n,p)} - E_{k-1}^{(n,p+1)}, \\ \qquad\qquad k = 1, 2, \ldots, m, \quad p = 0, 1, \ldots, M-1, \\ E_k^{(n,p+1)} = \dfrac{Q_{k+1}^{(n+p)}}{Q_k^{(n+M+p)}} E_k^{(n,p)}, \\ \qquad\qquad k = 1, 2, \ldots, m-1 \quad p = 0, 1, \ldots, M-1, \end{cases}
$$

$$
\text{(2.10)} \qquad \begin{cases} E_k^{(n+M)} = E_k^{(n)} + E_k^{(n,M)} - E_k^{(n,0)}, \quad k = 1, 2, \ldots, m-1, \\ E_{k+1}^{(n,0)} = \dfrac{E_{k+1}^{(n)}}{E_k^{(n+M)}} E_k^{(n,M)}, \quad k = 1, 2, \ldots, m-2. \end{cases}
$$

A close examination of (2.9) and (2.10) reveals that we need not compute the entries $E_2^{(n,0)}, E_3^{(n,0)}, \ldots, E_{m-1}^{(n,0)}$ in $R^{(n,0)}$ to start the LR and RR transformations. In fact, if only $E_1^{(n,0)}$ in $R^{(n,0)}$ is given, then (2.9) and (2.10) allow us to compute all the entries of $R^{(n,0)}$, $L^{(n+M)}, \ldots, L^{(n+2M-1)}$ and $R^{(n+M)}$ in the following order:

$$
\begin{aligned}
&\{Q_1^{(n+M+p)}, E_1^{(n,p+1)}\}_{p=0,1,\ldots,M-1}, E_1^{(n+M)}, E_2^{(n,0)}, \\
&\{Q_2^{(n+M+p)}, E_2^{(n,p+1)}\}_{p=0,1,\ldots,M-1}, E_2^{(n+M)}, E_3^{(n,0)}, \\
&\quad \vdots \\
&\{Q_{m-1}^{(n+M+p)}, E_{m-1}^{(n,p+1)}\}_{p=0,1,\ldots,M-1}, E_{m-1}^{(n+M)}, E_m^{(n,0)}, \\
&\{Q_m^{(n+M+p)}\}_{p=0,1,\ldots,M-1}.
\end{aligned}
$$

The formula for $E_1^{(n,0)}$ can be obtained by observing the $(1,1)$ and $(1,2)$ entries of $L^{(n)}L^{(n+1)}\ldots L^{(n+M-1)}R^{(n)} - s^{(n)}I$ and $\bar{L}^{(n)}R^{(n,0)}$ in the first equality of (2.6):

$$
\text{(2.11)} \qquad E_1^{(n,0)} = \frac{Q_1^{(n)}Q_1^{(n+1)}\ldots Q_1^{(n+M-1)}}{Q_1^{(n)}Q_1^{(n+1)}\ldots Q_1^{(n+M-1)} - s^{(n)}} E_1^{(n)}.
$$

Therefore, the shifted LR transformation (2.6) is completed by employing (2.9), (2.10), and (2.11).

In the actual algorithm, we modify (2.9) and (2.10) to the differential form without subtraction by introducing auxiliary variables [4], as is done in the dqd algorithm [2]. The differential form is mathematically equivalent to the original (2.9) and (2.10) but has better stability properties. For details and error analysis of the shifted dhToda algorithm in the differential form, we refer to [4].

3. Newton shift strategy

The Newton shift is known to be an effective shift for the shifted QR and dqds algorithms [1], [12]. The Newton shift $s_{\mathrm{N}}^{(n)}$ in the shifted LR transformation (2.6) from $A^{(n)}$ to $A^{(n+M)}$ is defined by

$$s_{\mathrm{N}}^{(n)} = [\mathrm{tr}((A^{(n)})^{-1})]^{-1}, \tag{3.1}$$

where $\mathrm{tr}(\cdot)$ denotes the sum of all diagonal entries of a matrix. Since the TN matrix $A^{(n)}$ has distinct positive eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ satisfying $\lambda_1 > \lambda_2 > \ldots > \lambda_m > 0$, we easily derive

$$0 < s_{\mathrm{N}}^{(n)} = \Big(\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \ldots + \frac{1}{\lambda_m}\Big)^{-1} = \lambda_m\Big(1 + \frac{\lambda_m}{\lambda_1} + \frac{\lambda_m}{\lambda_2} + \ldots + \frac{\lambda_m}{\lambda_{m-1}}\Big)^{-1} < \lambda_m.$$

Thus, we expect that the Newton shift $s_{\mathrm{N}}^{(n)}$ is useful in the shifted LR transformation (2.6). However, we emphasize that $s_{\mathrm{N}}^{(n)}$ computed by (3.1) is a constant that does not depend on n, because the eigenvalues of $A^{(n)}$ are equal to those of $A^{(0)}$. This causes the convergence rate of the sequence $\{A^{(n+lM)}\}_{l=0,1,\ldots}$ to be at most linear.

In this section, we explain how to utilize the idea of the Newton shift more efficiently for the shifted LR transformation (2.6). Suppose $s^{(n)}$ has been computed in some way, and we need to determine the next shift $s^{(n+M)}$ for $A^{(n+M)}$. The key idea here is to apply the Newton shift to $\bar{A}^{(n)} := A^{(n)} - s^{(n)}I$ instead of $A^{(n+M)}$. Let $\bar{s}_{\mathrm{N}}^{(n)} := 1/\mathrm{tr}((\bar{A}^{(n)})^{-1})$ be the Newton shift for $\bar{A}^{(n)}$ and let $s^{(n+M)} := s^{(n)} + \bar{s}_{\mathrm{N}}^{(n)}$. Then, since $0 < \bar{s}_{\mathrm{N}}^{(n)} < \lambda_m - s^{(n)}$, it follows that $0 < s^{(n+M)} < \lambda_m$ and $s^{(n+M)}$ can be used as a valid shift. Moreover, it can easily be verified that $s + [\mathrm{tr}(A^{(n)} - sI)^{-1}]^{-1}$ is an increasing function of s when $0 \leqslant s < \lambda_m$. Hence, $s^{(n+M)}$ is a better shift than $s_{\mathrm{N}}^{(n+M)} = s_{\mathrm{N}}^{(n)}$. The difficulty with this approach is that the matrix $\bar{A}^{(n)}$ is not computed explicitly in the algorithm. To address this, we consider performing the following two steps of the shifted LR transformations using the same shift $s^{(n)}$,

$$\begin{cases} A^{(n)} - s^{(n)}I = \bar{L}^{(n)}R^{(n,0)}, \\ A^{(n+M)} - s^{(n)}I = R^{(n,0)}\bar{L}^{(n)}, \\ A^{(n+M)} - s^{(n)}I = \bar{L}^{(n+M)}R^{(n+M,0)}, \\ A^{(n+2M)} - s^{(n)}I = R^{(n+M,0)}\bar{L}^{(n+M)}. \end{cases} \tag{3.2}$$

Then we can compute $\bar{s}_{\mathrm{N}}^{(n)}$ efficiently from the quantities appearing in the twofold shifted LR transformations (3.2), as will be clarified in the following paragraphs. Thus, in the next two steps, we can employ the shift $s^{(n+2M)}$ as

$$s^{(n+2M)} = s^{(n)} + \bar{s}_{\mathrm{N}}^{(n)}. \tag{3.3}$$

To estimate the value of $\bar{s}_{\mathrm{N}}^{(n)}$, we hereinafter compute the Newton shift $\bar{s}_{\mathrm{N}}^{(n+M)} = 1/\mathrm{tr}((\bar{A}^{(n+M)})^{-1})$ instead of $\bar{s}_{\mathrm{N}}^{(n)}$, where $\bar{A}^{(n+M)} = A^{(n+M)} - s^{(n)}I$. Since eigenvalues of $\bar{A}^{(n+M)}$ are equal to those of $\bar{A}^{(n)}$, it is obvious that $\bar{s}_{\mathrm{N}}^{(n+M)}$ can be used instead of $\bar{s}_{\mathrm{N}}^{(n)}$. Since $\det(\bar{A}^{(n)}) = \det(\bar{A}^{(n+M)})$, we can express the (i,i) entry of $(\bar{A}^{(n+M)})^{-1}$, denoted by $((\bar{A}^{(n+M)})^{-1})_{i,i}$ as

$$((\bar{A}^{(n+M)})^{-1})_{i,i} = \frac{\mathrm{cof}(\bar{A}_{i,i}^{(n+M)})}{\det(\bar{A}^{(n)})}, \tag{3.4}$$

where $\mathrm{cof}(\bar{A}_{i,i}^{(n+M)})$ is the (i,i) *cofactor* of $\bar{A}^{(n+M)}$ and denotes the determinant of the submatrix obtained by deleting the ith row and column of $\bar{A}^{(n+M)}$. According to the first equation of (3.2), we can decompose $\bar{A}^{(n)} = A^{(n)} - s^{(n)}I$ using the lower triangular matrix $\bar{L}^{(n)}$ and the upper bidiagonal matrix $R^{(n,0)}$ as $\bar{A}^{(n)} = \bar{L}^{(n)}R^{(n,0)}$. Noting that all the diagonal entries of $R^{(n,0)}$ are 1, we see that the denominator $\det(\bar{A}^{(n)})$ is equal to the product of all diagonal entries of $\bar{L}^{(n)}$, namely

$$\det(\bar{A}^{(n)}) = (\bar{L}^{(n)})_{1,1}(\bar{L}^{(n)})_{2,2}\dots(\bar{L}^{(n)})_{m,m}. \tag{3.5}$$

We can also examine the numerator $\mathrm{cof}(\bar{A}_{i,i}^{(n+M)})$ in terms of the principal submatrices formed by gathering the $i_1, i_1+1, \dots, i_2$th rows and $j_1, j_1+1, \dots, j_2$th columns of $\bar{A}^{(n+M)}$, denoted $\bar{A}^{(n+M)}(i_1 : i_2; j_1 : j_2)$. Since $\bar{A}^{(n+M)}$ is a lower Hessenberg matrix, we easily derive

$$\mathrm{cof}(\bar{A}_{i,i}^{(n+M)}) = \det(\bar{A}^{(n+M)}(1 : i-1; 1 : i-1))\det(\bar{A}^{(n+M)}(i+1 : m; i+1 : m)). \tag{3.6}$$

The third equation of (3.2) immediately leads to $\bar{A}^{(n+M)}(1 : i-1; 1 : i-1) = \bar{L}^{(n+M)}(1 : i-1; 1 : i-1)R^{(n+M,0)}(1 : i-1; 1 : i-1)$. Noting that $\bar{L}^{(n+M)}(1 : i-1; 1 : i-1)$ is lower triangular and $\det(R^{(n+M,0)}(1 : i-1; 1 : i-1)) = 1$, we obtain

$$\det(\bar{A}^{(n+M)}(1 : i-1; 1 : i-1)) = (\bar{L}^{(n+M)})_{1,1}(\bar{L}^{(n+M)})_{2,2}\dots(\bar{L}^{(n+M)})_{i-1,i-1}. \tag{3.7}$$

Similarly, from the second equation of (3.2), we derive

$$\begin{aligned} &\det(\bar{A}^{(n+M)}(i+1 : m; i+1 : m)) \\ &\quad = \det(R^{(n,0)}(i+1 : m; i+1 : m))\det(\bar{L}^{(n)}(i+1 : m; i+1 : m)) \\ &\quad = (\bar{L}^{(n)})_{i+1,i+1}(\bar{L}^{(n)})_{i+2,i+2}\dots(\bar{L}^{(n)})_{m,m}. \end{aligned} \tag{3.8}$$

Therefore, it follows from (3.6), (3.7), and (3.8) that

$$\begin{aligned}
\text{(3.9)} \qquad \operatorname{cof}(\bar{A}^{(n+M)}_{i,i}) &= (\bar{L}^{(n+M)})_{1,1}(\bar{L}^{(n+M)})_{2,2}\dots(\bar{L}^{(n+M)})_{i-1,i-1} \\
&\quad \times (\bar{L}^{(n)})_{i+1,i+1}(\bar{L}^{(n)})_{i+2,i+2}\dots(\bar{L}^{(n)})_{m,m}.
\end{aligned}$$

Combining (3.5) and (3.9) with (3.4), we have

$$\begin{aligned}
\text{(3.10)} \qquad \bar{s}^{(n)}_{\mathrm{N}} &= \left(\sum_{i=1}^{m}((\bar{A}^{(n+M)})^{-1})_{i,i}\right)^{-1} \\
&= \left(\sum_{i=1}^{m}\frac{(\bar{L}^{(n+M)})_{1,1}(\bar{L}^{(n+M)})_{2,2}\dots(\bar{L}^{(n+M)})_{i-1,i-1}}{(\bar{L}^{(n)})_{1,1}(\bar{L}^{(n)})_{2,2}\dots(\bar{L}^{(n)})_{i,i}}\right)^{-1}.
\end{aligned}$$

To compute values of the diagonal entries of $\bar{L}^{(n)}$ and $\bar{L}^{(n+M)}$, we employ the first three equations of (3.2). Since the $(k,k+1)$ entries of $A^{(n)}-s^{(n)}I = L^{(n)}L^{(n+1)}\dots L^{(n+M-1)}R^{(n)}-s^{(n)}I$ and $\bar{L}^{(n)}R^{(n,0)}$ are $Q^{(n)}_kQ^{(n+1)}_k\dots Q^{(n+M-1)}_kE^{(n)}_k$ and $(\bar{L}^{(n)})_{k,k}E^{(n,0)}_k$, respectively, we derive the following from the first equation of (3.2)

$$\text{(3.11)} \qquad (\bar{L}^{(n)})_{k,k} = \frac{Q^{(n)}_kQ^{(n+1)}_k\dots Q^{(n+M-1)}_kE^{(n)}_k}{E^{(n,0)}_k}, \quad k=1,2,\dots,m-1.$$

On the other hand, the equality of the $(m-1,m)$ entries on both sides of the second equation of (3.2) leads to

$$\text{(3.12)} \qquad (\bar{L}^{(n)})_{m,m} = \frac{Q^{(n+M)}_{m-1}Q^{(n+M+1)}_{m-1}\dots Q^{(n+2M-1)}_{m-1}E^{(n+M)}_{m-1}}{E^{(n,0)}_{m-1}}.$$

Finally, we compute the diagonal entries of $\bar{L}^{(n+M)}$ from the third equation of (3.2)

(3.13)

$$(\bar{L}^{(n+M)})_{k,k} = \frac{Q^{(n+M)}_kQ^{(n+M+1)}_k\dots Q^{(n+2M-1)}_kE^{(n+M)}_k}{E^{(n+M,0)}_k}, \quad k=1,2,\dots,m-1.$$

At the beginning of the iteration, we determine the shift $s^{(0)} < \lambda_m$ in some way; for example, using the Newton shift (3.1). Then we perform two steps of shifted LR transformations using $s^{(0)}$ as in (3.2), and obtain $L^{(M)}, L^{(M+1)}, \dots, L^{(2M-1)}$, $R^{(n)}$, $R^{(M)}$, $R^{(0,0)}$, and $R^{(M,0)}$. By substituting the entries of these matrices into (3.11), (3.12), and (3.13), we can compute $(\bar{L}^{(0)})_{1,1}, (\bar{L}^{(0)})_{2,2}, \dots, (\bar{L}^{(0)})_{m,m}$ and $(\bar{L}^{(M)})_{1,1}, (\bar{L}^{(M)})_{2,2}, \dots, (\bar{L}^{(M)})_{m-1,m-1}$. Then, we can compute $\bar{s}^{(n)}_{\mathrm{N}}$ from (3.10) and the shift to be used in the next two steps by $s^{(2M)} = s^{(0)} + \bar{s}^{(n)}_{\mathrm{N}}$. This process is repeated and the shift is updated at each step of the twofold shifted LR transformations.

4. Minimum eigenvalue

To examine the convergence rate of the sequence of shifted LR transformations, we need to understand the behavior of the minimum eigenvalue λ_m of the TN matrix $A^{(n)}$. In this section, we prepare three lemmas for the value of λ_m.

We first clarify the relationship between the minimum eigenvalues of the TN matrix $A^{(n)}$ and its principal submatrix $A^{(n)}(1:m-1;1:m-1)$. Here, we express $A^{(n)}(1:m-1;1:m-1)$ using the leading principal submatrices of $L^{(n)}$, $L^{(n+1)},\dots,L^{(n+M-1)}$ and $R^{(n)}$ as

$$\begin{aligned}
(4.1)\quad & A^{(n)}(1:m-1;1:m-1)\\
&= \begin{pmatrix} I_{m-1} & \mathbf{0}_{m-1} \end{pmatrix} L^{(n)}L^{(n+1)}\dots L^{(n+M-1)}R^{(n)} \begin{pmatrix} I_{m-1} \\ \mathbf{0}_{m-1}^{\top} \end{pmatrix}\\
&= \begin{pmatrix} L^{(n)}(1:m-1;1:m-1) & \mathbf{0}_{m-1} \end{pmatrix} L^{(n+1)}L^{(n+2)}\dots L^{(n+M-1)}\\
&\quad\times \begin{pmatrix} R^{(n)}(1:m-1;1:m-1) \\ \mathbf{0}_{m-1}^{\top} \end{pmatrix}\\
&\quad\vdots\\
&= \begin{pmatrix} \prod_{p=0}^{M-1} L^{(n+p)}(1:m-1;1:m-1) & \mathbf{0}_{m-1} \end{pmatrix}\\
&\quad\times \begin{pmatrix} R^{(n)}(1:m-1;1:m-1) \\ \mathbf{0}_{m-1}^{\top} \end{pmatrix}\\
&= L^{(n)}(1:m-1;1:m-1)L^{(n+1)}(1:m-1;1:m-1)\\
&\quad\times\dots\times L^{(n+M-1)}(1:m-1;1:m-1)\times R^{(n)}(1:m-1;1:m-1),
\end{aligned}$$

where I_k and $\mathbf{0}_k$ denote the k-by-k identity matrix and the k-dimensional zero column vector, respectively. Since bidiagonal entries of $L^{(n)}(1:m-1;1:m-1)$ and $R^{(n)}(1:m-1;1:m-1)$ are positive, $A^{(n)}(1:m-1;1:m-1)$ is a nonsingular TN matrix and its eigenvalues are positive. The interlacing theorem [7], [9] immediately leads to the inequality $\lambda_m \leqslant \mu_{m-1}$, where μ_{m-1} denotes the minimum eigenvalue of $A^{(n)}(1:m-1;1:m-1)$. Noting the TN property of $A^{(n)}$ and $A^{(n)}(1:m-1;1:m-1)$, we derive the following lemma for a stricter inequality:

Lemma 4.1. *The minimum eigenvalues λ_m and μ_{m-1} satisfy:*

$$\lambda_m < \mu_{m-1}. \tag{4.2}$$

P r o o f. We define two matrices that are similar to $L^{(n)}$ and $R^{(n)}$, respectively, by $\tilde{L}^{(n)} := J_m L^{(n)} J_m$ and $\widetilde{R}^{(n)} := J_m R^{(n)} J_m$, where $J_m := \operatorname{diag}(1,-1,\dots,(-1)^{m-1})$.

Then $\tilde{L}^{(n)}$ and $\tilde{L}^{(n)}(1:m-1;1:m-1)$ have positive diagonal entries and negative lower subdiagonal entries, so we easily see that all lower triangular entries, including diagonals, of $(\tilde{L}^{(n)})^{-1}$ and $(\tilde{L}^{(n)}(1:m-1;1:m-1))^{-1}$ are positive. Similarly, all upper triangular entries, including diagonals, of $(\widetilde{R}^{(n)})^{-1}$ and $(\widetilde{R}^{(n)}(1:m-1;1:m-1))^{-1}$ are positive.

Next, we express $\tilde{L}^{(n)}$ and $\widetilde{R}^{(n)}$ as 2-by-2 block matrices:

$$\tilde{L}^{(n)} = \begin{pmatrix} \tilde{L}^{(n)}(1:m-1;1:m-1) & \mathbf{0}_{m-1} \\ -\boldsymbol{e}_{m-1}^{\top} & Q_m^{(n)} \end{pmatrix},$$

$$\widetilde{R}^{(n)} = \begin{pmatrix} \widetilde{R}^{(n)}(1:m-1;1:m-1) & -E_{m-1}^{(n)}\boldsymbol{e}_{m-1} \\ \mathbf{0}_{m-1}^{\top} & 1 \end{pmatrix},$$

where $\boldsymbol{e}_k$ denotes the k-dimensional unit column vector whose kth entry is 1. With the formula for the inverse of a block matrix, we obtain

$$(4.3)\quad (\tilde{L}^{(n)})^{-1} = \begin{pmatrix} (\tilde{L}^{(n)}(1:m-1;1:m-1))^{-1} & \mathbf{0}_{m-1} \\ (Q_m^{(n)})^{-1}\boldsymbol{e}_{m-1}^{\top}(\tilde{L}^{(n)}(1:m-1;1:m-1))^{-1} & (Q_m^{(n)})^{-1} \end{pmatrix},$$

$$(4.4)\quad (\widetilde{R}^{(n)})^{-1} = \begin{pmatrix} (\widetilde{R}^{(n)}(1:m-1;1:m-1))^{-1} & E_{m-1}^{(n)}(\widetilde{R}^{(n)}(1:m-1;1:m-1))^{-1}\boldsymbol{e}_{m-1} \\ \mathbf{0}_{m-1}^{\top} & 1 \end{pmatrix}.$$

Equations (4.3) and (4.4) suggest that

$$\begin{pmatrix} (\tilde{L}^{(n)}(1:m-1;1:m-1))^{-1} & \mathbf{0}_{m-1} \\ \mathbf{0}_{m-1}^{\top} & 0 \end{pmatrix} \preceq (\tilde{L}^{(n)})^{-1},$$

$$\begin{pmatrix} (\widetilde{R}^{(n)}(1:m-1;1:m-1))^{-1} & \mathbf{0}_{m-1} \\ \mathbf{0}_{m-1}^{\top} & 0 \end{pmatrix} \preceq (\tilde{R}^{(n)})^{-1},$$

where $X \preceq Y$ signifies that $Y - X$ is nonnegative. For nonnegative square matrices X_1, Y_1, X_2, and Y_2, if $X_1 \preceq Y_1$ and $X_2 \preceq Y_2$, then $X_1X_2 \preceq Y_1Y_2$ [7]. Therefore, we derive

$$(4.5)\quad \begin{aligned} &\begin{pmatrix} (\widetilde{R}^{(n)}(1:m-1;1:m-1))^{-1} & \mathbf{0}_{m-1} \\ \mathbf{0}_{m-1}^{\top} & 0 \end{pmatrix} \\ &\quad\times \begin{pmatrix} (\tilde{L}^{(n+M-1)}(1:m-1;1:m-1))^{-1} & \mathbf{0}_{m-1} \\ \mathbf{0}_{m-1}^{\top} & 0 \end{pmatrix} \\ &\quad\times \begin{pmatrix} (\tilde{L}^{(n+M-2)}(1:m-1;1:m-1))^{-1} & \mathbf{0}_{m-1} \\ \mathbf{0}_{m-1}^{\top} & 0 \end{pmatrix} \\ &\quad\times \dots \times \begin{pmatrix} (\tilde{L}^{(n)}(1:m-1;1:m-1))^{-1} & \mathbf{0}_{m-1} \\ \mathbf{0}_{m-1}^{\top} & 0 \end{pmatrix} \\ &\preceq (\widetilde{R}^{(n)})^{-1}(\tilde{L}^{(n+M-1)})^{-1}(\tilde{L}^{(n+M-2)})^{-1}\dots(\tilde{L}^{(n)})^{-1}. \end{aligned}$$

We introduce a matrix similar to $A^{(n)}$ as

$$\tilde{A}^{(n)} := J_m A^{(n)} J_m = \tilde{L}^{(n)} \tilde{L}^{(n+1)} \dots \tilde{L}^{(n+M-1)} \widetilde{R}^{(n)}.$$

Similarly to (4.1), it follows that $\tilde{A}^{(n)}(1:m-1;1:m-1) = \tilde{L}^{(n)}(1:m-1;1:m-1)\tilde{L}^{(n+1)}(1:m-1;1:m-1)\dots\tilde{L}^{(n+M-1)}(1:m-1;1:m-1)\widetilde{R}^{(n)}(1:m-1;1:m-1)$. Thus, we can rewrite (4.5) as

$$\begin{pmatrix} (\tilde{A}^{(n)}(1:m-1;1:m-1))^{-1} & \mathbf{0}_{m-1} \\ \mathbf{0}_{m-1}^{\top} & 0 \end{pmatrix} \preceq (\tilde{A}^{(n)})^{-1}.$$

From the positivity of $(\tilde{L}^{(n)})^{-1}$, $(\widetilde{R}^{(n)})^{-1}$ and their principal matrices, it is obvious that $(\tilde{A}^{(n)})^{-1}$ and $(\tilde{A}^{(n)}(1:m-1;1:m-1))^{-1}$ are positive matrices.

Using the Perron-Frobenius theorem [7], we see that the normalized eigenvector of $\tilde{A}^{(n)}(1:m-1;1:m-1)$ corresponding to the maximum eigenvalue μ_{m-1}^{-1} has only positive entries. For $\varepsilon > 0$, $\eta > 0$, and the normalized eigenvector $\boldsymbol{x}_{m-1}$, we prepare the positive matrix:

$$\begin{pmatrix} (\tilde{A}^{(n)}(1:m-1;1:m-1))^{-1} & \varepsilon \boldsymbol{x}_{m-1} \\ \mu_{m-1}^{-1}\eta \boldsymbol{x}_{m-1}^{\top} & \varepsilon\eta \end{pmatrix}. \tag{4.6}$$

Then, we derive

$$\begin{aligned} &\begin{pmatrix} (\tilde{A}^{(n)}(1:m-1;1:m-1))^{-1} & \varepsilon \boldsymbol{x}_{m-1} \\ \mu_{m-1}^{-1}\eta \boldsymbol{x}_{m-1}^{\top} & \varepsilon\eta \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_{m-1} \\ \eta \end{pmatrix} \\ &\qquad = \begin{pmatrix} \mu_{m-1}^{-1}\boldsymbol{x}_{m-1} + \varepsilon\eta \boldsymbol{x}_{m-1} \\ \mu_{m-1}^{-1}\eta \|\boldsymbol{x}_{m-1}\|^2 + \varepsilon\eta^2 \end{pmatrix} = (\mu_{m-1}^{-1} + \varepsilon\eta) \begin{pmatrix} \boldsymbol{x}_{m-1} \\ \eta \end{pmatrix}, \end{aligned}$$

which implies that the nonnegative matrix in (4.6) has an eigenvalue $\mu_{m-1}^{-1} + \varepsilon\eta$. Since $(\tilde{A}^{(n)})^{-1}$ is a positive matrix, there exist some ε and η satisfying

$$\begin{pmatrix} (\tilde{A}^{(n)}(1:m-1;1:m-1))^{-1} & \varepsilon \boldsymbol{x}_{m-1} \\ (\mu_{m-1})^{-1}\eta {\boldsymbol{x}_{m-1}}^{\top} & \varepsilon\eta \end{pmatrix} \preceq (\tilde{A}^{(n)})^{-1}. \tag{4.7}$$

According to Horn and Johnson [7], p. 491, for positive square matrices X and Y, the spectral radius of X is smaller than or equal to that of Y if $X \preceq Y$. Thus, we obtain

$$\mu_{m-1}^{-1} < \mu_{m-1}^{-1} + \varepsilon\eta \leqslant \lambda_m^{-1},$$

which immediately yields (4.2). $\square$

For convenience, we hereinafter use the abbreviations $\bar{A}_{\lambda_m}^{(n)} := A^{(n)} - \lambda_m I$ and $\bar{A}_{\lambda_m}^{(n)}(1:k;1:k) := A^{(n)}(1:k;1:k) - \lambda_m I_k$. We next present an implicit expression of the minimum eigenvalue λ_m using the entries of $A^{(n)}$.

Lemma 4.2. *The entries and minimum eigenvalue of $A^{(n)}$ satisfy*

$$\lambda_m = (A^{(n)})_{m,m} - (A^{(n)})_{m-1,m} A^{(n)}(m;1:m-1)(\bar{A}_{\lambda_m}^{(n)}(1:m-1;1:m-1))^{-1}\boldsymbol{e}_{m-1}. \tag{4.8}$$

P r o o f. It is obvious that $\bar{A}_{\lambda_m}^{(n)}$ has an eigenvalue of 0. Let $\boldsymbol{z}$ be the eigenvector corresponding to the 0 eigenvalue. Moreover, let $z(i)$ denote the ith entry of $\boldsymbol{z}$. Then, from $\bar{A}_{\lambda_m}^{(n)}\boldsymbol{z} = \boldsymbol{0}$, it follows that

$$\bar{A}_{\lambda_m}^{(n)}(1:m-1;1:m-1)\boldsymbol{z}(1:m-1) + (A^{(n)})_{m-1,m} z(m)\boldsymbol{e}_{m-1} = \boldsymbol{0}, \tag{4.9}$$

$$A^{(n)}(m;1:m-1)\boldsymbol{z}(1:m-1) + ((A^{(n)})_{m,m} - \lambda_m) z(m) = 0. \tag{4.10}$$

With the help of Lemma 4.1, we see that $\bar{A}_{\lambda_m}^{(n)}(1:m-1;1:m-1)$ is nonsingular. Therefore, (4.9) leads to

$$\boldsymbol{z}(1:m-1) = -(A^{(n)})_{m-1,m} z(m)(\bar{A}_{\lambda_m}^{(n)}(1:m-1;1:m-1))^{-1}\boldsymbol{e}_{m-1}. \tag{4.11}$$

Combining (4.10) with (4.11), we derive

$$\begin{aligned}[-(A^{(n)})_{m-1,m} A^{(n)}(m;1:m-1)(\bar{A}_{\lambda_m}^{(n)}(1:m-1;1:m-1))^{-1}\boldsymbol{e}_{m-1} \\ + (A^{(n)})_{m,m} - \lambda_m] z(m) = 0.\end{aligned} \tag{4.12}$$

Let us assume here that $z(m) = 0$. Then, we can simplify (4.9) as

$$\bar{A}_{\lambda_m}^{(n)}(1:m-1;1:m-1)\boldsymbol{z}(1:m-1) = \boldsymbol{0},$$

which implies that λ_m is an eigenvalue of $A^{(n)}(1:m-1;1:m-1)$. This contradicts Lemma 4.1. Consequently, noting that $z(m) \neq 0$ in (4.12), we have (4.8). $\square$

We also present a lemma for an implicit expression of the minimum eigenvalue λ_m by considering the LU decomposition of $\bar{A}_{\lambda_m}^{(n)}$.

Lemma 4.3. *There exists a lower triangular matrix $\breve{L}^{(n)}$ with $(\breve{L}^{(n)})_{m,m} = 0$ and a unit upper bidiagonal matrix $\breve{R}^{(n,0)}$ such that $\bar{A}_{\lambda_m}^{(n)} = \breve{L}^{(n)}\breve{R}^{(n,0)}$. Let $\breve{E}_k^{(n,0)}$ be the upper diagonal entries of $\breve{R}^{(n,0)}$. Moreover, let $\mathcal{L}^{(n)} := L^{(n)}L^{(n+1)}\dots L^{(n+M-1)}$. Then,*

$$\lambda_m = \left[\sum_{j=1}^{m}(\breve{E}_j^{(n,0)}\breve{E}_{j+1}^{(n,0)}\dots\breve{E}_{m-1}^{(n,0)})((\mathcal{L}^{(n)})^{-1})_{m,j}\right]^{-1}. \tag{4.13}$$

P r o o f. Using Lemma 4.1 recursively, we observe that all eigenvalues of principal submatrices $A^{(n)}(1:m-1;1:m-1)$, $A^{(n)}(1:m-2;1:m-2)$, …, $A^{(n)}(1;1)$ are larger than λ_m. In other words, $\det(\bar{A}^{(n)}_{\lambda_m}(1:k;1:k)) \neq 0$ for $k = 1,2,\dots,m-1$. Hence, $\bar{A}^{(n)}_{\lambda_m}(1:m-1;1:m-1)$ admits the LU decomposition $\bar{A}^{(n)}_{\lambda_m}(1:m-1;1:m-1) = \breve{L}^{(n)}(1:m-1;1:m-1)\breve{R}^{(n,0)}(1:m-1;1:m-1)$, where $\breve{L}^{(n)}(1:m-1;1:m-1)$ and $\breve{R}^{(n,0)}(1:m-1;1:m-1)$ are nonsingular. Using the block LU decomposition of a 2-by-2 block matrix [6], we obtain

$$\begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix} = \begin{pmatrix} L_{1,1} & \boldsymbol{O} \\ X_{2,1}R_{1,1}^{-1} & X_{2,2} - X_{2,1}X_{1,1}^{-1}X_{1,2} \end{pmatrix} \times \begin{pmatrix} R_{1,1} & (L_{1,1})^{-1}X_{1,2} \\ \boldsymbol{O} & I \end{pmatrix}, \tag{4.14}$$

which holds when the $(1,1)$ block $X_{1,1}$ admits the LU decomposition $X_{1,1} = L_{1,1}R_{1,1}$. We can obtain the LU decomposition of $\bar{A}^{(n)}_{\lambda_m}$ by

$$\bar{A}^{(n)}_{\lambda_m} = \begin{pmatrix} \breve{L}^{(n)}(1:m-1;1:m-1) & \mathbf{0} \\ A^{(n)}(m;1:m-1)(\breve{R}^{(n,0)}(1:m-1;1:m-1))^{-1} & (\breve{L}^{(n)})_{m,m} \end{pmatrix} \times \begin{pmatrix} \breve{R}^{(n,0)}(1:m-1;1:m-1) & (A^{(n)})_{m-1,m}(\breve{L}^{(n)}(1:m-1;1:m-1))^{-1}\boldsymbol{e}_{m-1} \\ \mathbf{0} & 1 \end{pmatrix}. \tag{4.15}$$

Since $\bar{A}^{(n)}_{\lambda_m}$ is singular and both $\breve{L}^{(n)}(1:m-1;1:m-1)$ and $\breve{R}^{(n,0)}(1:m-1;1:m-1)$ are nonsingular, $(\breve{L}^{(n)})_{m,m}$ must be 0. Therefore, we can adopt the first and second matrices on the right-hand side of (4.15) as $\breve{L}^{(n)}$ and $\breve{R}^{(n,0)}$, respectively.

Noting that $A^{(n)} = \mathcal{L}^{(n)}R^{(n)}$, we derive the following equality for $\breve{L}^{(n)}$ and $\breve{R}^{(n,0)}$:

$$\breve{R}^{(n,0)}\breve{L}^{(n)} + \lambda_m I = \breve{R}^{(n,0)}\mathcal{L}^{(n)}R^{(n)}(\breve{R}^{(n,0)})^{-1}. \tag{4.16}$$

Since the mth column of $\breve{L}^{(n)}$ is $\mathbf{0}$, the (m,m) entry of $\breve{A}^{(n+M)} := \breve{R}^{(n,0)}\breve{L}^{(n)} + \lambda_m I$ is λ_m. Equation (4.16) implies that the (m,m) entry of $\breve{R}^{(n,0)}\mathcal{L}^{(n)}R^{(n)}(\breve{R}^{(n,0)})^{-1}$ is equal to λ_m. Considering the $(m-1)$-by-$(m-1)$ principal submatrix in the matrix equality $\breve{A}^{(n+M)} = \breve{R}^{(n,0)}\mathcal{L}^{(n)}R^{(n)}(\breve{R}^{(n,0)})^{-1}$, we obtain

$$\begin{aligned} \breve{A}^{(n+M)}(1:m-1;1:m-1) &= \breve{R}^{(n,0)}(1:m-1;1:m)\mathcal{L}^{(n)}(1:m;1:m-1) \\ &\quad \times R^{(n)}(1:m-1;1:m-1)(\breve{R}^{(n,0)})^{-1}(1:m-1;1:m-1). \end{aligned} \tag{4.17}$$

Noting here that all the diagonal entries of $R^{(n)}$ and $(\breve{R}^{(n,0)})^{-1}$ are 1 in (4.17), we derive

$$\operatorname{cof}(\breve{A}^{(n+M)}_{m,m}) = \det(\breve{R}^{(n,0)}(1:m-1;1:m)\mathcal{L}^{(n)}(1:m;1:m-1)).$$

With the help of the Cauchy-Binet formula [7], p. 22, we derive

$$\det(\breve{R}^{(n,0)}(1:m-1;1:m)\mathcal{L}^{(n)}(1:m;1:m-1)) = \sum_{j=1}^{m} \mathrm{cof}(\breve{R}^{(n,0)}_{m,j})\mathrm{cof}(\mathcal{L}^{(n)}_{j,m}).$$

Recalling the matrix form of $R^{(n)}$ and $\breve{R}^{(n,0)}$, we see that

$$\mathrm{cof}(\breve{R}^{(n,0)}_{m,j}) = \breve{E}^{(n,0)}_{j}\breve{E}^{(n,0)}_{j+1}\dots\breve{E}^{(n,0)}_{m-1}$$

and

$$\begin{aligned}\mathrm{cof}(\mathcal{L}^{(n)}_{j,m}) &= ((\mathcal{L}^{(n)})^{-1})_{m,j}\det(\mathcal{L}^{(n)}) \\ &= ((\mathcal{L}^{(n)})^{-1})_{m,j}\det(A^{(n)}) = ((\mathcal{L}^{(n)})^{-1})_{m,j}\det(A^{(n+M)}).\end{aligned}$$

It follows that

$$\mathrm{cof}(\breve{A}^{(n+M)}_{m,m}) = \sum_{j=1}^{m}(\breve{E}^{(n,0)}_{j}\breve{E}^{(n,0)}_{j+1}\dots\breve{E}^{(n,0)}_{m-1})((\mathcal{L}^{(n)})^{-1})_{m,j}\det(A^{(n+M)}). \tag{4.18}$$

Since

$$\det(A^{(n+M)}) = \det(\breve{A}^{(n+M)}) \text{ and } \mathrm{cof}(\breve{A}^{(n+M)}_{m,m})/\det(\breve{A}^{(n+M)}) = ((\breve{A}^{(n+M)})^{-1})_{m,m},$$

we can rewrite (4.18) as

$$((\breve{A}^{(n+M)})^{-1})_{m,m} = \sum_{j=1}^{m}(\breve{E}^{(n,0)}_{j}\breve{E}^{(n,0)}_{j+1}\dots\breve{E}^{(n,0)}_{m-1})((\mathcal{L}^{(n)})^{-1})_{m,j}. \tag{4.19}$$

Noting that $\breve{A}^{(n+M)}$ is a 2×2 block lower triangular matrix with the $(2,2)$ block λ_m, we know that $((\breve{A}^{(n+M)})^{-1})_{m,m} = \lambda_m^{-1}$. Substituting this into the left-hand side of (4.19), we obtain (4.13). □

5. Bottom-right entry and its neighboring entries

We can observe the asymptotic convergence of the shifted LR transformation (2.6) with $s^{(n)} < \lambda_m$ from $A^{(n)}$ to $A^{(n+M)}$ as $n \to \infty$ by comparing the (m,m) and $(m-1,m)$ entries of $A^{(n+M)}$ with those of $A^{(n)}$. In this section, we present an expression of the (m,m) and $(m-1,m)$ entries of $A^{(n+M)}$ using the entries of $A^{(n)}$ and the shift $s^{(n)}$.

We first give a lemma for a relationship of the (m,m) entry of $A^{(n+M)}$ to that of $A^{(n)}$ involving the other entries of $A^{(n)}$, the minimum eigenvalue λ_m, and the shift $s^{(n)}$.

Lemma 5.1. *Under the shifted LR transformation* (2.6) *from* $A^{(n)}$ *to* $A^{(n+M)}$ *with* $s^{(n)} < \lambda_m$ *, it holds that*

$$(5.1)\qquad \begin{aligned}&(A^{(n+M)})_{m,m} - s^{(n)}\\ &\quad = (\lambda_m - s^{(n)})[1 + (A^{(n)}(m;1:m-1))(\bar{A}^{(n)}_{\lambda_m}(1:m-1;1:m-1))^{-1}\\ &\qquad \times (\bar{A}^{(n)}(1:m-1;1:m-1))^{-1}\boldsymbol{e}_{m-1}(A^{(n)})_{m-1,m}].\end{aligned}$$

P r o o f. Since $\bar{A}^{(n)}$ with $s^{(n)} < \lambda_m$ has the LU decomposition $\bar{A}^{(n)} = \bar{L}^{(n)}R^{(n,0)}$, we can also decompose $\bar{A}^{(n)}(1:m-1;1:m-1)$ as $\bar{A}^{(n)}(1:m-1;1:m-1) = \bar{L}^{(n)}(1:m-1;1:m-1)R^{(n,0)}(1:m-1;1:m-1)$. By letting $X_{1,1} = \bar{A}^{(n)}(1:m-1;1:m-1)$, $X_{1,2} = (A^{(n)})_{m-1,m}\boldsymbol{e}_{m-1}$, $X_{2,1} = A^{(n)}(m;1:m-1)$, and $X_{2,2} = (\bar{A}^{(n)})_{m,m} = (A^{(n)})_{m,m} - s^{(n)}$ in (4.14), we obtain

(5.2)

$$\bar{A}^{(n)} = \begin{pmatrix} \bar{L}^{(n)}(1:m-1;1:m-1) & \mathbf{0}\\ A^{(n)}(m;1:m-1)(R^{(n,0)}(1:m-1;1:m-1))^{-1} & (\bar{L}^{(n)})_{m,m}\end{pmatrix}$$
$$\times\begin{pmatrix} R^{(n,0)}(1:m-1;1:m-1) & (A^{(n)})_{m-1,m}(\bar{L}^{(n)}(1:m-1;1:m-1))^{-1}\boldsymbol{e}_{m-1}\\ \mathbf{0} & 1\end{pmatrix},$$

where

$$(5.3)\qquad \begin{aligned}(\bar{L}^{(n)})_{m,m} &= (A^{(n)})_{m,m} - s^{(n)}\\ &\quad - (A^{(n)})_{m-1,m}A^{(n)}(m;1:m-1)(\bar{A}^{(n)}(1:m-1;1:m-1))^{-1}\boldsymbol{e}_{m-1}.\end{aligned}$$

Equation (2.6) immediately leads to

$$(5.4)\qquad (A^{(n+M)})_{m,m} = (R^{(n,0)}\bar{L}^{(n)})_{m,m} + s^{(n)} = (\bar{L}^{(n)})_{m,m} + s^{(n)}.$$

Combining (5.4) with (5.3), we derive

$$(5.5)\qquad \begin{aligned}(A^{(n+M)})_{m,m} &= (A^{(n)})_{m,m}\\ &\quad - (A^{(n)})_{m-1,m}A^{(n)}(m;1:m-1)(\bar{A}^{(n)}(1:m-1;1:m-1))^{-1}\boldsymbol{e}_{m-1}.\end{aligned}$$

Using Lemma 4.2, we can rewrite (5.5) as

$$(5.6)\qquad \begin{aligned}(A^{(n+M)})_{m,m} &= \lambda_m + (A^{(n)})_{m-1,m}A^{(n)}(m;1:m-1)\\ &\quad \times [(\bar{A}^{(n)}_{\lambda_m}(1:m-1;1:m-1))^{-1} - (\bar{A}^{(n)}(1:m-1;1:m-1))^{-1}]\boldsymbol{e}_{m-1}.\end{aligned}$$

Since $Y^{-1} - X^{-1} = X^{-1}(X-Y)Y^{-1}$ for nonsingular matrices X and Y in (5.6), we have (5.1). $\square$

Similarly to the case of the (m,m) entry of $A^{(n+M)}$, we can derive a lemma for the case of the $(m-1,m)$ entry of $A^{(n+M)}$ with the help of the block LU decomposition of the principal submatrix.

Lemma 5.2. *Under the shifted LR transformation* (2.6) *with* $s^{(n)} < \lambda_m$ *from* $A^{(n)}$ *to* $A^{(n+M)}$*, it holds that*

$$\begin{aligned}(5.7)\quad (A^{(n+M)})_{m-1,m} &= (A^{(n)})_{m-1,m}(\lambda_m - s^{(n)})\\ &\times [(A^{(n)})_{m-1,m-1} - s^{(n)} - (A^{(n)})_{m-2,m-1}A^{(n)}(m-1;1:m-2)\\ &\times (\bar{A}^{(n)}(1:m-2;1:m-2))^{-1}\boldsymbol{e}_{m-2}]^{-1}\\ &\times [1 + (A^{(n)})_{m-1,m}A^{(n)}(m;1:m-1)(\bar{A}^{(n)}_{\lambda_m}(1:m-1;1:m-1))^{-1}\\ &\times (\bar{A}^{(n)}(1:m-1;1:m-1))^{-1}\boldsymbol{e}_{m-1}].\end{aligned}$$

Proof. From (2.6), we easily observe that

$$(5.8)\qquad (A^{(n+M)})_{m-1,m} = (R^{(n,0)})_{m-1,m}(\bar{L}^{(n)})_{m,m},$$

$$(5.9)\qquad (\bar{L}^{(n)})_{m,m} = (A^{(n+M)})_{m,m} - s^{(n)}.$$

The block LU decomposition (5.2) also leads to

$$(5.10)\qquad (R^{(n,0)})_{m-1,m} = (A^{(n)})_{m-1,m}((\bar{L}^{(n)})_{m-1,m-1})^{-1}.$$

Considering the resulting block LU decomposition:

$$\begin{aligned}\bar{A}^{(n)}(1:m-1;1:m-1) &= \bar{L}^{(n)}(1:m-1;1:m-1)R^{(n,0)}(1:m-1;1:m-1)\\ &= \begin{pmatrix}\bar{A}^{(n)}(1:m-2;1:m-2) & (A^{(n)})_{m-2,m-1}\boldsymbol{e}_{m-2}\\ A^{(n)}(m-1;1:m-2) & (A^{(n)})_{m-1,m-1} - s^{(n)}\end{pmatrix},\end{aligned}$$

we derive

$$\begin{aligned}(5.11)\quad (\bar{L}^{(n)})_{m-1,m-1} &= (A^{(n)})_{m-1,m-1} - s^{(n)}\\ &\quad - (A^{(n)})_{m-2,m-1}A^{(n)}(m-1;1:m-2)(\bar{A}^{(n)}(1:m-2;1:m-2))^{-1}\boldsymbol{e}_{m-2}.\end{aligned}$$

Substituting (5.9), (5.10), and (5.11) into (5.8) and using Lemma 5.1, we obtain (5.7). □

The following lemma also gives an expression of the product $Q_m^{(n+M)}Q_m^{(n+M+1)}\dots Q_m^{(n+2M-1)}$, which plays an important role in Section 6.

Lemma 5.3. *The product* $Q_m^{(n+M)}Q_m^{(n+M+1)}\dots Q_m^{(n+2M-1)}$ *satisfies*

$$(5.12)\qquad Q_m^{(n+M)}Q_m^{(n+M+1)}\dots Q_m^{(n+2M-1)} = \left[\sum_{j=1}^{m}(E_j^{(n,0)}E_{j+1}^{(n,0)}\dots E_{m-1}^{(n,0)})((\mathcal{L}^{(n)})^{-1})_{m,j}\right]^{-1}.$$

P r o o f. It is obvious that $Q_m^{(n+M)} Q_m^{(n+M+1)} \dots Q_m^{(n+2M-1)} = (\mathcal{L}^{(n+M)})_{m,m}$. According to the formula concerning the bottom-right entry of the lower triangular matrix in the LU decomposition, we can rewrite $(\mathcal{L}^{(n+M)})_{m,m}$ as $\det(A^{(n+M)})/\operatorname{cof}(A_{m,m}^{(n+M)})$. It follows that

$$Q_m^{(n+M)} Q_m^{(n+M+1)} \dots Q_m^{(n+2M-1)} = \frac{\det(A^{(n+M)})}{\operatorname{cof}(A_{m,m}^{(n+M)})}. \tag{5.13}$$

Similarly to the proof of Lemma 4.3, we derive

$$\operatorname{cof}(A_{m,m}^{(n+M)}) = \sum_{j=1}^{m} (E_j^{(n,0)} E_{j+1}^{(n,0)} \dots E_{m-1}^{(n,0)})((\mathcal{L}^{(n)})^{-1})_{m,j} \det(A^{(n+M)}). \tag{5.14}$$

Therefore, by combining (5.13) with (5.14), we have (5.12). $\square$

6. Convergence rate

In this section, we clarify the convergence rate of the sequences $\{Q_k^{(2lM)}\}_{l=0,1,\dots}$, $\{E_k^{(2lM)}\}_{l=0,1,\dots}$, and $\{s^{(2lM)}\}_{l=0,1,\dots}$ appearing in the shifted LR transformation (2.6) as $l \to \infty$ under the shift strategy (3.3). We also provide a numerical example to check the convergence acceleration.

We first investigate the convergence rate of $s^{(2lM)}$ as $l \to \infty$. Rewriting the shift strategy (3.3) in terms of the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_m$, we obtain

$$s^{(2(l+1)M)} = s^{(2lM)} + \frac{\lambda_m - s^{(2lM)}}{1 + (\lambda_m - s^{(2lM)}) \sum_{j=1}^{m-1} (\lambda_j - s^{(2lM)})^{-1}}. \tag{6.1}$$

From this expression, it is clear that if $s^{(2lM)} < \lambda_m$, the denominator in the second term on the right-hand side is larger than 1. Therefore, we have

$$s^{(2(l+1)M)} < s^{(2lM)} + \lambda_m - s^{(2lM)} = \lambda_m.$$

Since $s^{(0)}$ is chosen to be smaller than λ_m, we know by induction that $s^{(2lM)} < \lambda_m$ holds for all $l \geqslant 0$. Furthermore, we see that the sequence $\{s^{(2lM)}\}_{l=0,1,\dots}$ is monotonically increasing, since the second term on the right-hand side of (6.1) is positive. That is, it is a monotonically increasing sequence bounded above and it converges to some constant smaller than or equal to λ_m. Now, we rewrite (6.1) as

$$\lambda_m - s^{(2(l+1)M)} = (\lambda_m - s^{(2lM)}) \left\{ 1 - \left[1 + \sum_{j=1}^{m} \left(1 - \frac{\lambda_j - \lambda_m}{\lambda_j - s^{(2lM)}} \right) \right]^{-1} \right\}. \tag{6.2}$$

Since $\{s^{(2lM)}\}_{l=0,1,\dots}$ is monotonically increasing, there is a constant $0 < c < \lambda_m$ such that $s^{(2lM)} \geqslant c$ holds for any sufficiently large l. For a sufficiently large l, the following inequality holds:

$$0 < 1 - \left[1 + \sum_{j=1}^{m}\left(1 - \frac{\lambda_j - \lambda_m}{\lambda_j - s^{(2lM)}}\right)\right]^{-1} \leqslant 1 - \left[1 + \sum_{j=1}^{m}\left(1 - \frac{\lambda_j - \lambda_m}{\lambda_j - c}\right)\right]^{-1} < 1.$$

Therefore, the sequence $\{\lambda_m - s^{(2lM)}\}_{l=0,1,\dots}$ converges at least geometrically to 0 and $\lim\limits_{l\to\infty} s^{(2lM)} = \lambda_m$. We can further rewrite (6.2) as

$$(6.3)\quad \lambda_m - s^{(2(l+1)M)} = (\lambda_m - s^{(2lM)})^2 \frac{\sum_{j=1}^{m-1}(\lambda_j - s^{(2lM)})^{-1}}{1 + (\lambda_m - s^{(2lM)})\sum_{j=1}^{m-1}(\lambda_j - s^{(2lM)})^{-1}}.$$

Noting that the second factor on the right-hand side converges to a positive constant as $l \to \infty$, we can conclude that $\{s^{(2lM)}\}_{l=0,1,\dots}$ converges to λ_m quadratically as follows:

$$(6.4)\quad \begin{aligned}\lim_{l\to\infty} \frac{\lambda_m - s^{(2(l+1)M)}}{(\lambda_m - s^{(2lM)})^2} &= \lim_{l\to\infty} \frac{\sum_{j=1}^{m-1}(\lambda_j - s^{(2lM)})^{-1}}{1 + (\lambda_m - s^{(2lM)})\sum_{j=1}^{m-1}(\lambda_j - s^{(2lM)})^{-1}} \\ &= \sum_{j=1}^{m-1} \frac{1}{\lambda_j - \lambda_m}.\end{aligned}$$

With respect to the convergence rate of the sequence $E_k^{(2lM)}$ as $l \to \infty$, we present the following theorem.

Theorem 6.1. *Under the shifted LR transformation* (2.6) *with the shift strategy* (3.3), $E_{m-1}^{(2lM)}$ *converges to* 0 *with order* $2 - \varepsilon$ *for arbitrary* $\varepsilon > 0$, *that is,*

$$(6.5)\quad \lim_{l\to\infty} \frac{E_{m-1}^{(2(l+1)M)}}{(E_{m-1}^{(2lM)})^{2-\varepsilon}} = 0.$$

Proof. It is easy to check that

$$\begin{aligned}(A^{(2lM)})_{m-1,m} &= (Q_{m-1}^{(2lM)} Q_{m-1}^{(2lM+1)} \dots Q_{m-1}^{((2l+1)M-1)}) E_{m-1}^{(2lM)}, \\ (A^{((2l+1)M)})_{m-1,m} &= (Q_{m-1}^{((2l+1)M)} Q_{m-1}^{((2l+1)M+1)} \dots Q_{m-1}^{(2(l+1)M-1)}) E_{m-1}^{((2l+1)M)}.\end{aligned}$$

Combining these with Lemma 5.2, we derive

$$(6.6)\quad E_{m-1}^{((2l+1)M)} = \varrho^{(2lM)}(\lambda_m - s^{(2lM)}) E_{m-1}^{(2lM)},$$

where

(6.7)

$$\varrho^{(2lM)} := \frac{Q_{m-1}^{(2lM)} Q_{m-1}^{(2lM+1)} \cdots Q_{m-1}^{((2l+1)M-1)}}{Q_{m-1}^{((2l+1)M)} Q_{m-1}^{((2l+1)M+1)} \cdots Q_{m-1}^{(2(l+1)M-1)}}$$
$$\times \frac{1 + (A^{(2lM)})_{m-1,m} A^{(2lM)}(m; 1:m-1)(\bar{A}_{\lambda_m}^{(2lM)}(1:m-1; 1:m-1))^{-1} \bar{\boldsymbol{e}}_{m-1}}{(A^{(2lM)})_{m-1,m-1} - s^{(2lM)} - (A^{(2lM)})_{m-2,m-1} A^{(2lM)}(m-1; 1:m-2) \bar{\boldsymbol{e}}_{m-2}},$$
$$\bar{\boldsymbol{e}}_{m-1}^{(2lM)} := (\bar{A}^{(2lM)}(1:m-1; 1:m-1))^{-1} \boldsymbol{e}_{m-1},$$
$$\bar{\boldsymbol{e}}_{m-2}^{(2lM)} := (\bar{A}^{(2lM)}(1:m-2; 1:m-2))^{-1} \boldsymbol{e}_{m-2}.$$

Since we proved in [5] that $A^{(2lM)}$ converges to a lower triangular matrix with diagonals $\lambda_1, \lambda_2, \ldots, \lambda_m$ as $l \to \infty$, the term $(A^{(2lM)})_{m-1,m-1}$ in (6.7) converges to λ_{m-1} and both $(A^{(2lM)})_{m-1,m}$ and $(A^{(2lM)})_{m-2,m-1}$ converge to 0. The row vectors $A^{(2lM)}(m; 1:m-1)$ and $A^{(2lM)}(m-1; 1:m-2)$ converge to some constant vectors. Since $\bar{A}^{(2lM)} = A^{(2lM)} - s^{(2lM)} I$ is nonsingular, the inverse matrices $(\bar{A}^{(2lM)}(1:m-1; 1:m-1))^{-1}$ and $(\bar{A}^{(2lM)}(1:m-2; 1:m-2))^{-1}$ converge to some constant matrices. Moreover, we showed in [5] that $Q_{m-1}^{(2lM)} Q_{m-1}^{(2lM+1)} \cdots Q_{m-1}^{((2l+1)M-1)}$ converges to λ_{m-1} as $l \to \infty$, so the ratio $Q_{m-1}^{(2lM)} Q_{m-1}^{(2lM+1)} \cdots Q_{m-1}^{((2l+1)M-1)} / (Q_{m-1}^{((2l+1)M)} Q_{m-1}^{((2l+1)M+1)} \cdots Q_{m-1}^{((2(l+1)M-1)})$ converges to 1. Substituting all of these into (6.7), we obtain

$$\lim_{l \to \infty} \varrho^{(2lM)} = \lim_{l \to \infty} \frac{1}{\lambda_{m-1} - s^{(2lM)}} = \frac{1}{\lambda_{m-1} - \lambda_m}. \tag{6.8}$$

We now consider the ratio $E_{m-1}^{(2(l+1)M)} / (E_{m-1}^{(2lM)})^{2-\varepsilon}$ where ε is arbitrary positive. First, by using (6.6) twice and noting that $s^{((2l+1)M)} = s^{(2lM)}$ in (3.2), we have

$$E_{m-1}^{(2(l+1)M)} = \varrho^{(2lM)} \varrho^{((2l+1)M)} (\lambda_m - s^{(2lM)})^2 E_{m-1}^{(2lM)}. \tag{6.9}$$

From (6.9), we immediately derive

(6.10)

$$\frac{E_{m-1}^{(2(l+1)M)}}{(E_{m-1}^{(2lM)})^{2-\varepsilon}} = \frac{\varrho^{(2lM)} \varrho^{((2l+1)M)}}{(\varrho^{((2l-2)M)} \varrho^{((2l-1)M)})^{2-\varepsilon}} \left(\frac{\lambda_m - s^{(2lM)}}{(\lambda_m - s^{(2(l-1)M)})^2} \right)^2$$
$$\times (\lambda_m - s^{(2(l-1)M)})^{2\varepsilon} \frac{E_{m-1}^{(2lM)}}{(E_{m-1}^{(2(l-1)M)})^{2-\varepsilon}}$$
$$= \prod_{j=0}^{l-1} \left[\frac{\varrho^{((2j+2)M)} \varrho^{((2j+3)M)}}{(\varrho^{(2jM)} \varrho^{((2j+1)M)})^{2-\varepsilon}} \left(\frac{\lambda_m - s^{(2(j+1)M)}}{(\lambda_m - s^{(2jM)})^2} \right)^2 (\lambda_m - s^{(2jM)})^{2\varepsilon} \right] \frac{E_{m-1}^{(2M)}}{(E_{m-1}^{(0)})^{2-\varepsilon}}.$$

From (6.8) and (6.4), we see that the rates $\varrho^{((2j+2)M)} \varrho^{((2j+3)M)} / (\varrho^{(2jM)} \varrho^{((2j+1)M)})^{2-\varepsilon}$ and $(\lambda_m - s^{(2(j+1)M)}) / (\lambda_m - s^{(2jM)})^2$ converge to some constants as $j \to \infty$. Moreover, $(\lambda_m - s^{(2jM)})^{\varepsilon} \to 0$ as $j \to \infty$. Thus, the bracketed part on the right-hand side of (6.10) converges to 0 as $j \to \infty$. Therefore, their product also approaches 0 as $l \to \infty$. This immediately leads to (6.5). $\square$

Before analyzing the convergence of the sequence $Q_k^{(2lM)}$ as $l \to \infty$, we present the relationships of $E_k^{(2lM,0)}$ and $\breve{E}_k^{(2lM,0)}$ to $E_k^{(2lM)}$.

Lemma 6.1. *The variables* $E_k^{(2lM,0)}, \breve{E}_k^{(2lM,0)}$, *and* $E_k^{(2lM)}$ *satisfy*

$$E_k^{(2lM,0)} = \beta_k^{(2lM)} E_k^{(2lM)}, \quad k = 1, 2, \ldots, m-1, \tag{6.11}$$

$$\breve{E}_k^{(2lM,0)} = \breve{\beta}_k^{(2lM)} E_k^{(2lM)}, \quad k = 1, 2, \ldots, m-1, \tag{6.12}$$

where

$$\begin{aligned}\beta_k^{(2lM)} &:= [(A^{(2lM)})_{k,k} - s^{(2lM)} - (A^{(2lM)})_{k-1,k} A^{(2lM)}(k; 1:k-1) \\ &\quad \times (\bar{A}^{(2lM)}(1:k-1; 1:k-1))^{-1} \boldsymbol{e}_{k-1}]^{-1} (Q_k^{(2lM)} Q_k^{(2lM+1)} \ldots Q_k^{((2l+1)M-1)}) \\ \breve{\beta}_k^{(2lM)} &:= [(A^{(2lM)})_{k,k} - \lambda_m - (A^{(2lM)})_{k-1,k} A^{(2lM)}(k; 1:k-1) \\ &\quad \times (\bar{A}_{\lambda_m}^{(2lM)}(1:k-1; 1:k-1))^{-1} \boldsymbol{e}_{k-1}]^{-1} (Q_k^{(2lM)} Q_k^{(2lM+1)} \ldots Q_k^{((2l+1)M-1)}).\end{aligned}$$

Moreover, $\beta_k^{(2lM)} \to \lambda_k/(\lambda_k - \lambda_m)$ *and* $\breve{\beta}_k^{(2lM)} \to \lambda_k/(\lambda_k - \lambda_m)$ *as* $l \to \infty$.

P r o o f. Similarly to the derivation of (5.11), we derive for $k = 1, 2, \ldots, m-1$:

$$\begin{aligned}(\bar{L}^{(2lM)})_{k,k} &= (A^{(2lM)})_{k,k} - s^{(2lM)} \\ &\quad - (A^{(2lM)})_{k-1,k} A^{(2lM)}(k; 1:k-1)(\bar{A}^{(2lM)}(1:k-1; 1:k-1))^{-1} \boldsymbol{e}_{k-1}.\end{aligned} \tag{6.13}$$

Combining this with (3.11), we have (6.11). By replacing $s^{(2lM)}$ with λ_m and repeating the same argument, we obtain (6.12). The limits of $\beta_k^{(2lM)}$ and $\breve{\beta}_k^{(2lM)}$ as $l \to \infty$ are easily checked by using $(A^{(2lM)})_{k,k} \to \lambda_k$, $(A^{(2lM)})_{k-1,k} \to 0$, and $Q_k^{(2lM)} Q_k^{(2lM+1)} \ldots Q_k^{((2l+1)M-1)} \to \lambda_k$ as $l \to \infty$. $\square$

Combining Lemma 6.1 with the lemmas from the previous sections, we obtain the following theorem concerning the convergence rate of $Q_k^{(2lM)}$ as $l \to \infty$.

Theorem 6.2. *Under the shifted LR transformation* (2.6) *with the shift strategy* (3.3), *the product* $Q_m^{(2lM)} Q_m^{(2lM+1)} \ldots Q_m^{((2l+1)M-1)}$ *converges to the minimum eigenvalue* λ_m *with order* $2 - \varepsilon$ *for arbitrary* $\varepsilon > 0$, *that is,*

$$\lim_{l \to \infty} \frac{\lambda_m - Q_m^{(2(l+1)M)} Q_m^{(2(l+1)M+1)} \ldots Q_m^{((2l+3)M-1)}}{(\lambda_m - Q_m^{(2lM)} Q_m^{(2lM+1)} \ldots Q_m^{((2l+1)M-1)})^{2-\varepsilon}} = 0. \tag{6.14}$$

P r o o f. From Lemmas 4.3 and 5.3, we easily derive

$$
\begin{aligned}
(6.15)\qquad & (\lambda_m - Q_m^{((2l+1)M)} Q_m^{((2l+1)M+1)} \dots Q_m^{(2(l+1)M-1)}) \\
& \times \left[\sum_{j=1}^{m} (E_j^{(2lM,0)} E_{j+1}^{(2lM,0)} \dots E_{m-1}^{(2lM,0)})((\mathcal{L}^{(2lM)})^{-1})_{m,j}\right] \\
& \times \left[\sum_{j=1}^{m} (\breve{E}_j^{(2lM,0)} \breve{E}_{j+1}^{(2lM,0)} \dots \breve{E}_{m-1}^{(2lM,0)})((\mathcal{L}^{(2lM)})^{-1})_{m,j}\right] \\
= & \sum_{j=1}^{m-1} ((\mathcal{L}^{(2lM)})^{-1})_{m,j} (E_j^{(2lM,0)} E_{j+1}^{(2lM,0)} \dots E_{m-1}^{(2lM,0)} \\
& - \breve{E}_j^{(2lM,0)} \breve{E}_{j+1}^{(2lM,0)} \dots \breve{E}_{m-1}^{(2lM,0)}) \\
= & \sum_{j=1}^{m-1} [((\mathcal{L}^{(2lM)})^{-1})_{m,j} \sum_{k=j}^{m-1} (\breve{E}_j^{(2lM,0)} \breve{E}_{j+1}^{(2lM,0)} \dots \breve{E}_{k-1}^{(2lM,0)}) \\
& \times (E_{k+1}^{(2lM,0)} E_{k+2}^{(2lM,0)} \dots E_{m-1}^{(2lM,0)})(E_k^{(2lM,0)} - \breve{E}_k^{(2lM,0)})].
\end{aligned}
$$

Using Lemma 6.1, we can rewrite $\breve{E}_k^{(2lM,0)} - E_k^{(2lM,0)}$, appearing in (6.15), as

$$
(6.16)\qquad E_k^{(2lM,0)} - \breve{E}_k^{(2lM,0)} = (\lambda_m - s^{(2lM)})\gamma_k^{(2lM)} E_k^{(2lM)},
$$

where

$$
\begin{aligned}
(6.17)\quad \gamma_k^{(2lM)} := & -[1 + A^{(2lM)}(k; 1:k-1)(\bar{A}^{(2lM)}(1:k-1; 1:k-1))^{-1} \\
& \times (\bar{A}_{\lambda_m}^{(2lM)}(1:k-1; 1:k-1))^{-1} \boldsymbol{e}_{k-1} (A^{(2lM)})_{k-1,k}] \beta_k^{(2lM)} \breve{\beta}_k^{(2lM)} \\
& \times (Q_k^{(2lM)} Q_k^{(2lM+1)} \dots Q_k^{((2l+1)M-1)})^{-1}.
\end{aligned}
$$

Recalling that $A^{(2lM)}$ converges to a lower triangular matrix as $l \to \infty$, we see that the $(k-1)$-dimensional row vector $A^{(2lM)}(k; 1:k-1)$ converges to some constant vector as $l \to \infty$. Similarly, the inverse matrices $(\bar{A}^{(2lM)}(1:k-1; 1:k-1))^{-1}$ and $(\bar{A}_{\lambda_m}^{(2lM)}(1:k-1; 1:k-1))^{-1}$ converge to some constant matrices as $n \to \infty$. Considering this result and using $(A^{(2lM)})_{k-1,k} \to 0$, $\beta_k^{(2lM)} \to \lambda_k/(\lambda_k - \lambda_m)$, $\breve{\beta}_k^{(2lM)} \to \lambda_k/(\lambda_k - \lambda_m)$ and $Q_k^{(2lM)} Q_k^{(2lM+1)} \dots Q_k^{((2l+1)M-1)} \to \lambda_k$ as $l \to \infty$ in (6.17), we see that

$$
\lim_{l\to\infty} \gamma_k^{(2lM)} = -\frac{\lambda_k}{(\lambda_k - \lambda_m)^2}.
$$

Using Lemma 6.1 and (6.16), we can rewrite (6.15) as

$$
\begin{aligned}
(6.18)\qquad &(\lambda_m - Q_m^{((2l+1)M)} Q_m^{((2l+1)M+1)} \dots Q_m^{(2(l+1)M-1)}) \\
&\quad \times \left[\sum_{j=1}^{m} (E_j^{(2lM,0)} E_{j+1}^{(2lM,0)} \dots E_{m-1}^{(2lM,0)})((\mathcal{L}^{(2lM)})^{-1})_{m,j}\right] \\
&\quad \times \left[\sum_{j=1}^{m} (\breve{E}_j^{(2lM,0)} \breve{E}_{j+1}^{(2lM,0)} \dots \breve{E}_{m-1}^{(2lM,0)})((\mathcal{L}^{(2lM)})^{-1})_{m,j}\right] \\
&= (\lambda_m - s^{(2lM)}) \delta^{(2lM)} E_{m-1}^{(2lM)},
\end{aligned}
$$

where

$$
\begin{aligned}
(6.19)\qquad \delta^{(2lM)} := &\sum_{j=1}^{m-1} [(E_j^{(2lM)} E_{j+1}^{(2lM)} \dots E_{m-2}^{(2lM)})((\mathcal{L}^{(2lM)})^{-1})_{m,j}] \\
&\times \sum_{k=j}^{m-1} (\breve{\beta}_j^{(2lM)} \breve{\beta}_{j+1}^{(2lM)} \dots \breve{\beta}_{k-1}^{(2lM)})(\beta_{k+1}^{(2lM)} \beta_{k+2}^{(2lM)} \dots \beta_{m-1}^{(2lM)}) \gamma_k^{(2lM)}.
\end{aligned}
$$

Since $A^{(2lM)}$ converges to a lower triangular matrix with positive diagonals λ_1, $\lambda_2, \dots, \lambda_m$, from the continuity of the LU decomposition, its lower triangular factor $\mathcal{L}^{(2lM)}$ also converges to the same nonsingular matrix. Thus, the limit of $((\mathcal{L}^{(2lM)})^{-1})_{m,k}$ as $l \to \infty$ exists. Letting $\sigma_k := \lim_{l\to\infty} ((\mathcal{L}^{(2lM)})^{-1})_{m,k}$ and noting that $E_k^{(2lM)} \to 0$, $\beta_k^{(2lM)} \to \lambda_k/(\lambda_k - \lambda_m)$, $\hat{\beta}_k^{(2lM)} \to \lambda_k/(\lambda_k - \lambda_m)$, and $\gamma_k^{(2lM)} \to -\lambda_k/(\lambda_k - \lambda_m)^2$ as $l \to \infty$ in (6.19), we obtain

$$
(6.20)\qquad \lim_{l\to\infty} \delta^{(2lM)} = \lim_{l\to\infty} ((\mathcal{L}^{(2lM)})^{-1})_{m,m-1} \gamma_{m-1}^{(2lM)} = -\frac{\sigma_{m-1}\lambda_{m-1}}{(\lambda_{m-1} - \lambda_m)^2}.
$$

Considering (6.18) again, we derive

$$
(6.21)\quad \lambda_m - Q_m^{((2l+1)M)} Q_m^{((2l+1)M+1)} \dots Q_m^{(2(l+1)M-1)} = (\lambda_m - s^{(2lM)}) \tau^{(2lM)} E_{m-1}^{(2lM)},
$$

where

$$
\begin{aligned}
(6.22)\qquad \tau^{(2lM)} := \delta^{(2lM)} &\left[\sum_{j=1}^{m} (E_j^{(2lM,0)} E_{j+1}^{(2lM,0)} \dots E_{m-1}^{(2lM,0)})((\mathcal{L}^{(2lM)})^{-1})_{m,j}\right]^{-1} \\
&\times \left[\sum_{j=1}^{m} (\breve{E}_j^{(2lM,0)} \breve{E}_{j+1}^{(2lM,0)} \dots \breve{E}_{m-1}^{(2lM,0)})((\mathcal{L}^{(2lM)})^{-1})_{m,j}\right]^{-1}.
\end{aligned}
$$

Combining (6.22) with $E_k^{(2lM,0)} \to 0$, $\breve{E}_k^{(2lM,0)} \to 0$, and $((\mathcal{L}^{(2lM)})^{-1})_{m,m} = ((\mathcal{L}^{(2lM)})_{m,m})^{-1} \to ((A^{(2lM)})_{m,m})^{-1} \to \lambda_m^{-1}$ as $l \to \infty$ with (6.20) leads to

$$\lim_{l\to\infty} \tau^{(2lM)} = -\frac{\sigma_{m-1}\lambda_{m-1}\lambda_m^2}{(\lambda_{m-1}-\lambda_m)^2}. \tag{6.23}$$

From (6.21), it follows that, for arbitrary positive ε,

$$\begin{aligned} &\frac{\lambda_m - Q_m^{((2l+3)M)} Q_m^{((2l+3)M+1)} \ldots Q_m^{(2(l+2)M-1)}}{\lambda_m - (Q_m^{((2l+1)M)} Q_m^{((2l+1)M+1)} \ldots Q_m^{(2(l+1)M-1)})^{2-\varepsilon}} \\ &\quad = \frac{\tau^{(2(l+1)M)}}{(\tau^{(2lM)})^{2-\varepsilon}} \cdot \frac{\lambda_m - s^{(2(l+1)M)}}{(\lambda_m - s^{(2lM)})^2} \cdot (\lambda_m - s^{(2lM)})^{\varepsilon} \cdot \frac{E_{m-1}^{(2(l+1)M)}}{(E_{m-1}^{(2lM)})^{2-\varepsilon}}. \end{aligned} \tag{6.24}$$

On the right-hand side of (6.24), $\tau^{(2(l+1)M)}/(\tau^{(2lM)})^{2-\varepsilon}$ and $(\lambda_m - s^{(2(l+1)M)})/(\lambda_m - s^{(2lM)})^2$ converge to some positive constants and $(\lambda_m - s^{(2lM)})^{\varepsilon}$ converges to 0 as $l \to \infty$. Furthermore, from Theorem 6.1, we see that $E_{m-1}^{(2(l+1)M)}/(E_{m-1}^{(2lM)})^{2-\varepsilon}$ also converges to 0. Therefore, the right-hand side of (6.24) converges to 0 as $l \to \infty$. This immediately leads to (6.14). □

We present a numerical example to demonstrate the accelerated convergence in the sequence of shifted LR transformations with the shift strategy proposed in Section 3. Numerical tests were carried out with IEEE double-precision arithmetic.

For test matrices, we prepare a TN matrix with $M = 5$

$$A^{(0)} = L^{(0)} L^{(1)} \ldots L^{(4)} R^{(0)},$$

$$L^{(n)} := \begin{pmatrix} 2 & & & \\ 1 & 2 & & \\ & \ddots & \ddots & \\ & & 1 & 2 \end{pmatrix} \in \mathbb{R}^{100\times 100}, \quad n = 0, 1, \ldots, 4,$$

$$R^{(0)} := \begin{pmatrix} 1 & 1 & & \\ & 1 & \ddots & \\ & & \ddots & 1 \\ & & & 1 \end{pmatrix} \in \mathbb{R}^{100\times 100}.$$

In the sequence of shifted LR transformations, the initial values are given as $Q_k^{(0)} = Q_k^{(1)} = \ldots = Q_k^{(4)} = 2$ for $k = 1, 2, \ldots, m$, and $E_k^{(0)} = 1$ for $k = 1, 2, \ldots, m-1$. With respect to the convergence history, we numerically compare the proposed shift strategy with $s^{(0)} = 0$ to the zero-shift strategy, namely, $s^{(n)} = 0$ for every n in the shifted LR transformations. Figure 1 shows the convergence of $E_{99}^{(lM)}$ to 0 as l increases in the shifted LR transformations. In Figure 1, the horizontal and vertical axes denote the values of l and $E_{99}^{(lM)}$, respectively, and the solid lines with the symbols "$*$" and "$\circ$" represent the values of $E_{99}^{(lM)}$ in the cases of the zero-shift strategy and in the proposed strategy, respectively. From Figure 1, we see that the proposed shift strategy enables us to accelerate the convergence of the shifted LR transformations in comparison with the linear convergence of the zero-shift LR transformations. Similarly to the convergence of $E_{99}^{(lM)}$, we observe that a higher-order convergence is achieved in the other variables in the shifted LR transformations with the proposed shift strategy.

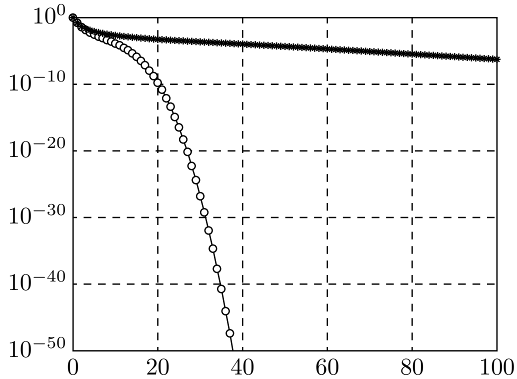

Figure 1. A graph of the values l of $n = lM$ (horizontal axis) and the values $E_{99}^{(n)}$ (vertical axis with logarithmic scale) in the sequence of shifted LR transformations. $*$: zero-shift strategy and $\circ$: proposed shift strategy.

7. Conclusion

In this paper, we briefly explained the shifted LR transformation for a TN banded Hessenberg matrix, which is based on the discrete hungry Toda equation. We then designed an efficient shift strategy for the shifted LR transformation using the idea of the Newton shift. The Newton shift usually produces a valid shift, which is smaller than the minimum eigenvalue of the updated target matrix. However, a simple application of the Newton shift results in only linear convergence. This is because all the updated target matrices are all similar and the Newton shift strategy always generates the same shift for all iterations. To develop a more efficient shift strategy, we proposed a method of computing the Newton shift not from the updated target matrix itself, but from the shifted one, which implicitly appears in the shifted LR transformations. We showed that the resulting shift strategy achieves a convergence rate of order $2 - \varepsilon$ for any $\varepsilon > 0$ for the variables appearing in the shifted LR transformations. We also numerically verified the convergence acceleration by comparing it with the zero-shift LR transformations.

References

[1] *F. L. Bauer*: qd-method with Newton shift. Technical Report 56. Computer Science Department, Stanford University, Stanford, 1967.

[2] *K. V. Fernando, B. N. Parlett*: Accurate singular values and differential qd algorithms. Numer. Math. *67* (1994), 191–229. zbl MR doi

[3] *A. Fukuda, E. Ishiwata, Y. Yamamoto, M. Iwasaki, Y. Nakamura*: Integrable discrete hungry systems and their related matrix eigenvalues. Ann. Mat. Pura Appl. *192* (2013), 423–445. zbl MR doi

[4] *A. Fukuda, Y. Yamamoto, M. Iwasaki, E. Ishiwata, Y. Nakamura*: Error analysis for matrix eigenvalue algorithm based on the discrete hungry Toda equation. Numer. Algorithms *61* (2012), 243–260. zbl MR doi

[5] *A. Fukuda, Y. Yamamoto, M. Iwasaki, E. Ishiwata, Y. Nakamura*: On a shifted LR transformation derived from the discrete hungry Toda equation. Monatsh. Math. *170* (2013), 11–26. zbl MR doi

[6] *G. H. Golub, C. F. Van Loan*: Matrix Computations. Johns Hopkins Studies in the Mathematical Sciences. The Johns Hopkins University Press, Baltimore, 2013. zbl MR

[7] *R. A. Horn, C. R. Johnson*: Matrix Analysis. Cambridge University Press, Cambridge, 1985. zbl MR doi

[8] *LAPACK Team*: LAPACK: Linear Algebra PACKage. Available at `http://www.netlib.org/lapack/`. sw

[9] *C.-K. Li, R. Mathias*: Interlacing inequalities for totally nonnegative matrices. Linear Algebra Appl. *341* (2002), 35–44. zbl MR doi

[10] *B. N. Parlett*: The new qd algorithms. Acta Numerica *4* (1995), 459–491. zbl MR doi

[11] *A. Pinkus*: Totally Positive Matrices. Cambridge Tracts in Mathematics 181. Cambridge University Press, Cambridge, 2009. zbl MR doi

[12] *C. Reinsch, F. L. Bauer*: Rational QR transformation with Newton shift for symmetric tridiagonal matrices. Numer. Math. *11* (1968), 264–272. zbl MR doi

[13] *H. Rutishauser*: Lectures on Numerical Mathematics. Birkhäuser, Boston, 1990. zbl MR doi

[14] *J.-Q. Sun, X.-B. Hu, H.-W. Tam*: Short note: An integrable numerical algorithm for computing eigenvalues of a specially structured matrix. Numer. Linear Algebra Appl. *18* (2011), 261–274. zbl MR doi

[15] *T. Tokihiro, A. Nagai, J. Satsuma*: Proof of solitonical nature of box and ball systems by means of inverse ultra-discretization. Inverse Probl. *15* (1999), 1639–1662. zbl MR doi

Authors' addresses: *Akiko Fukuda*, Department of Mathematical Sciences, Shibaura Institute of Technology, 307 Fukasaku, Minuma-ku, Saitama 337-8570, Japan, e-mail: `afukuda@shibaura-it.ac.jp`; *Yusaku Yamamoto*, Department of Communication Engineering and Informatics, The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan, e-mail: `yusaku.yamamoto@uec.ac.jp`; *Masashi Iwasaki*, Department of Life and Environmental Sciences, Kyoto Prefectural University, 1-5 Shimogamo Nakaragi-cho, Sakyo-ku, Kyoto 606-8522, Japan, e-mail: `imasa@kpu.ac.jp`; *Emiko Ishiwata*, Department of Applied Mathematics, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan, e-mail: `ishiwata@rs.tus.ac.jp`; *Yoshimasa Nakamura*, Graduate School of Informatics, Kyoto University, 36-1 Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan, e-mail: `ynaka@i.kyoto-u.ac.jp`.