A Novel Nonsymmetric K _-Lanczos Algorithm for the Generalized Nonsymmetric K _-Eigenvalue Problems

William R. Ferng

Department of Applied Mathematics
National Chiao Tung University
Hsin-Chu, Taiwan 30050, Republic of China

and

Kun-Yi Lin and Wen-Wei Lin

Institute of Applied Mathematics
National Tsing Hua University
Hsin-Chu, Taiwan 30043, Republic of China

ABSTRACT

   In this article, we present a novel algorithm, named nonsymmetric K_--Lanczos algorithm, for computing a few extreme eigenvalues of the generalized eigenvalue problem $Mx = \lambda Lx$, where the matrices M and L have the so-called K_+-structures. We demonstrate a K_--tridiagonalization procedure preserves the K_+-structures. An error bound for the extreme K_--Ritz value obtained from this new algorithm is presented. When compared with the class nonsymmetric Lanczos approach, this method has the same order of computational complexity and can be viewed as a special 2×2-block nonsymmetric Lanczos algorithm. Numerical experiments with randomly generated K_--matrices show that our algorithm converges faster and more accurate than the nonsymmetric Lanczos algorithm.   © *Elsevier Science Inc., 1997*

1. INTRODUCTION

Many modeling problems in scientific computing require the development of efficient numerical methods for solving the associated eigenvalue problems. In quantum chemistry, for example, the time-independent Hartree–Fock model [1, 2] leads to the following generalized eigenvalue problem

$$Mx = \lambda Lx, \tag{1.1}$$

where $M, L \in \mathbf{R}^{2n \times 2n}$, $M = M^T$, $L = L^T$, $\lambda \in \mathbf{C}$ is the eigenvalue of (1.1), and $x \in \mathbf{C}^{2n}$ is the corresponding eigenvector. Moreover, M is a so-called symmetric K_+-matrix and can be represented in the block structure

$$M = \begin{bmatrix} A & B \\ B & A \end{bmatrix} \tag{1.2}$$

with $A, B \in \mathbf{R}^{n \times n}$, $A = A^T$, $B = B^T$, and L is a symmetric K_--matrix

$$L = \begin{bmatrix} \Sigma & \Delta \\ -\Delta & -\Sigma \end{bmatrix} \tag{1.3}$$

with $\Sigma, \Delta \in \mathbf{R}^{n \times n}$, $\Sigma = \Sigma^T$, $\Delta = -\Delta^T$. If we define the matrix K to be

$$K = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}, \tag{1.4}$$

where I_n is the $n \times n$ identity matrix, then we have

$$(MK)^T = MK = KM, \tag{1.5}$$

and

$$(LK)^T = -LK = KL. \tag{1.6}$$

If matrix L is invertible, the generalized eigenvalue problem (1.1) can be transformed to the eigenvalue problem

$$L^{-1}Mx = \lambda x. \tag{1.7}$$

That is, the generalized eigenvalue problem (1.1) is equivalent to the eigenvalue problem (1.5). Let

$$N \equiv L^{-1}M; \tag{1.8}$$

one can verify that N is a $2n \times 2n$ K_--matrix [2].

In general, the matrices M and L can be very large ($n \approx 10^6$) and sparse, and only a few extreme eigenvalues are required. The nonsymmetric Lanczos algorithm [3] is one of the most widely used techniques for computing some extreme eigenvalues of a large sparse nonsymmetric matrix. In this approach, the matrix N is reduced to a tridiagonal matrix, which no longer has the K_--structure, using a general similarity transformation. Information about N's extreme eigenvalues tends to emerge long before the tridiagonalization process is complete. The Ritz values (eigenvalues of this tridiagonal submatrix) are used to approximate the extreme eigenvalues of N [4]. A convergence analysis for the nonsymmetric Lanczos method was recently presented by Ye in [5].

In [6], Flaschka considered the case when N is symmetric. The symmetric K_--algorithm and KQR algorithm were proposed for solving problem (1.1).

In this article, we propose a new algorithm; we name it nonsymmetric K_--Lanczos algorithm, for computing a few extreme eigenvalues of large sparse nonsymmetric K_--matrices for problem (1.1). This new method utilizes the special properties of the K_--structure, preserves the structure at each step, and reduces the K_--matrix to a K_--tridiagonal matrix. The extreme eigenvalues are approximated by the K_--Ritz values of the K_--tridiagonal matrix using nonsymmetric KQR or KQZ algorithm [7]. An error bound, which demonstrates the convergence behavior of this nonsymmetric K_--Lanczos algorithm, is developed and analyzed for the K_--Ritz values. This algorithm has the same order of computational complexity as the nonsymmetric Lanczos algorithm in terms of floating-point operations. A our numerical experiments with randomly generated nonsymmetric K_--matrices show that the new algorithm converges faster than the classic nonsymmetric Lanczos algorithm. With minor modifications, the algorithm can be adapted for K_+-matrices.

We organize this article as follows. In Section 2, we establish the existence of the nonsymmetric K_--tridiagonalization theorem. A new elimination procedure and the nonsymmetric K_--Lanczos algorithm is introduced in Section 3. Then we present an error bound for the extreme K_--Ritz values and analyze the convergence behavior in Section 4. Numerical experiments and some remarks are summarized in Section 5 followed by the conclusion in Section 6.

2. NONSYMMETRIC K_--TRIDIAGONALIZATION

When the nonsymmetric Lanczos method is applied to a K_--matrix, the algorithm treats this matrix as a general matrix and reduces it to a tridiagonal matrix which no longer has the K_--structure. In this section, we shall prove that under certain conditions, if $N \in \mathbf{R}^{2n \times 2n}$ is a K_--matrix, there exists a K_+-similarity transformation $X \in \mathbf{R}^{2n \times 2n}$ such that $X^{-1}NX$ is an unreduced K_--tridiagonal matrix. That is,

$$X^{-1}NX = T \equiv \begin{bmatrix} T_1 & T_2 \\ -T_2 & -T_1 \end{bmatrix},$$

where $T_1, T_2 \in \mathbf{R}^{n \times n}$ are tridiagonal, and

$$\left| (T_1)_{i, i+1} \right| \neq \left| (T_2)_{i, i+1} \right|,$$

$$\left| (T_1)_{i+1, i} \right| \neq \left| (T_2)_{i+1, i} \right|,$$

for $i = 1, 2, \ldots, n - 1$. Unlike the nonsymmetric Lanczos method, this similarity transformation preserves the K_--structure.

The following theorem can be proved immediately from theorems in [7].

THEOREM 2.1.   *Suppose $N \in \mathbf{R}^{2n \times 2n}$ is a K_--matrix, $K \in \mathbf{R}^{2n \times 2n}$ is as defined in* (1.2), *and the K_--Krylov matrix*

$$V \equiv V[N, x, n]$$

$$\equiv \left[x, Nx, \ldots, N^{n-1}x \right| - Kx, -KNx, \ldots, -KN^{n-1}x \right] \qquad (2.1)$$

for some $x \in \mathbf{R}^{2n}$, has full rank.

(1) *If $X \in \mathbf{R}^{2n \times 2n}$ is a nonsingular K_+-matrix such that $R \equiv X^{-1}V$ is a K_--upper triangular matrix, then $H \equiv X^{-1}NX$ is an unreduced K_--upper Hessenberg matrix.*

(2) *Let $X \in \mathbf{R}^{2n \times 2n}$ be a K_+-matrix with the first column x. If $H \equiv X^{-1}NX$ is an unreduced K_--upper Hessenberg matrix, then the K_--Krylov matrix $V[N, x, n] = XV[H, e_{1, 2n}, n]$, where $V[H, e_{1, 2n}, n]$ is a nonsingular K_--upper triangular matrix and $e_{1, 2n}$ is the first column of I_{2n}.*

(3) *Let $H \equiv X^{-1}NX$ and $\tilde{H} \equiv Y^{-1}NY$ be K_--upper Hessenberg matrices with X and Y both nonsingular K_+-matrices. If H is unreduced and the first*

columns of X and Y are linearly dependent, then $\tilde{H}$ is unreduced and $X^{-1}Y$ is K_+-upper triangular.

Next we show that a K_+-matrix can be factored into the product of a K_--lower triangular matrix and a K_--upper triangular matrix if the conditions are satisfied. This property is similar to the general LU-decomposition theorem but in special $K_\pm$-structure.

THEOREM 2.2 (Real K_+-LR Factorization). *Let $M \in \mathbf{R}^{2n \times 2n}$ be a K_+-matrix and be partitioned as follows:*

$$M = \begin{bmatrix} A & B \\ B & A \end{bmatrix},$$

with $A, B \in \mathbf{R}^{n \times n}$. If all the leading principal submatrices of $A + B$ and $A - B$ are nonsingular, then there exists a K_--lower triangular matrix L and a K_--upper triangular matrix R such that $M = LR$.

Proof. Let

$$Z = \frac{1}{\sqrt{2}} \begin{bmatrix} I_n & I_n \\ I_n & -I_n \end{bmatrix},$$

where I_n is the $n \times n$ identity matrix. Then Z is nonsingular and

$$Z^T M Z = \begin{bmatrix} A + B & 0 \\ 0 & A - B \end{bmatrix}.$$

Since all leading principal submatrices of $A + B$ and $A - B$ are nonsingular, $A + B$ and $A - B$ have the LR-decompositions, say, $A + B = L_1 R_1$ and $A - B = L_2 R_2$, where L_1, L_2 are lower triangular and R_1, R_2 are upper triangular. Then

$$Z^T M Z = \begin{bmatrix} L_1 R_1 & 0 \\ 0 & L_2 R_2 \end{bmatrix} = \begin{bmatrix} L_1 & 0 \\ 0 & L_2 \end{bmatrix} \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}.$$

It follows that

$$
M = Z^{-T} \begin{bmatrix} L_1 & 0 \\ 0 & L_2 \end{bmatrix} \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix} Z^{-1}
$$

$$
= \frac{1}{2} \begin{bmatrix} L_1 R_1 + L_2 R_2 & L_1 R_1 - L_2 R_2 \\ L_1 R_1 - L_2 R_2 & L_1 R_1 + L_2 R_2 \end{bmatrix}
$$

$$
\equiv LR,
$$

where

$$
L = \begin{bmatrix} \frac{1}{2}(L_1 + L_2) & \frac{1}{2}(L_2 - L_1) \\ -\frac{1}{2}(L_2 - L_1) & -\frac{1}{2}(L_1 + L_2) \end{bmatrix}
$$

and

$$
R = \begin{bmatrix} \frac{1}{2}(R_1 + R_2) & \frac{1}{2}(R_1 - R_2) \\ -\frac{1}{2}(R_1 - R_2) & -\frac{1}{2}(R_1 + R_2) \end{bmatrix}.
$$

It is easy to check that L is a K_--lower triangular matrix and R is a K_- upper triangular matrix. Now we prove the existence theorem for the nonsymmetric K_--Lanczos tridiagonalization.

THEOREM 2.3 (Existence Theorem).   *Let N be a K_--matrix. If $V[N, x, n]$ and $V[N^T, y, n]$ are nonsingular K_--Krylov matrices for some $x, y \in \mathbf{R}^{2n}$ such that $B \equiv V[N^T, y, n]^T V[N, x, n]$ has the real K_+-LR decomposition, $B = LR$, then there exists a nonsingular K_+-matrix $X \in \mathbf{R}^{2n}$ such that $T \equiv X^{-1}NX$ is an unreduced K_--tridiagonal matrix.*

Proof.   Let

$$
X = V[N, x, n]R^{-1}.
$$

Then X is a nonsingular K_+-matrix. Since R is K_--upper triangular, it follows from Theorem 1 that $H = X^{-1}NX$ is an unreduced K_--upper Hessenberg matrix. Also from the assumption $B = V[N^T, y, n]^T V[N, x, n] = LR$, we have $V[N^T, y, n] = X^{-T}L^T$. L is K_--lower triangular; hence L^T is

K_--upper triangular. Apply Theorem 1 once again; one obtains that $\tilde{H} = X^T N^T X^{-T}$ is unreduced K_--upper Hessenberg. This implies that $H = X^{-1}NX = \tilde{H}^T$ is unreduced K_--lower Hessenberg. Therefore

$$T \equiv X^{-1}NX$$

is an unreduced K_--tridiagonal matrix.   ∎

THEOREM 2.4.   *Let X and Y be nonsingular K_+-matrices. Suppose that*

(a) $T \equiv X^{-1}NX$ *and* $\tilde{T} \equiv Y^{-1}NY$ *are unreduced K_--tridiagonal matrices,*
(b) *the first columns of X and Y are linearly dependent, and*
(c) *the first rows of X^{-1} and Y^{-1} are linearly dependent;*

then $X^{-1}Y$ is a K_+-diagonal matrix.

Proof. Since T and $\tilde{T}$ are unreduced K_--tridiagonal matrices, they can be viewed as unreduced K_--upper Hessenberg matrices. By using Theorem 1, $X^{-1}Y$ is a K_+-upper triangular matrix. Apply the same argument once again. $T^T = X^T N^T X^{-T}$ and $\tilde{T}^T = Y^T Y^{-T}$ are unreduced K_--upper Hessenberg matrices; hence $Y^T X^{-T} \equiv (X^{-1}Y)^T$ is K_+-upper triangular. It implies that $X^{-1}Y$ is a K_+-diagonal matrix.   ∎

We have proved that for a given K_--matrix N and $2n$-vectors x and y, if the product of K_--Krylov matrices $V[N^T, y, n]^T V[N, x, n]$ has the real K_+-LR decomposition, then there exists a nonsingular K_+-matrix X such that $X^{-1}NX$ is unreduced K_--tridiagonal. Similar results can be developed if the K_--matrix N is replaced by a K_+-matrix M.

In the next section, we introduce a new elimination procedure and derive the nonsymmetric K_--Lanczos algorithm for the unreduced K_--tridiagonal reduction on nonsymmetric K_--matrices.

3. NONSYMMETRIC K_--LANCZOS ALGORITHM

The nonsymmetric Lanczos method for extreme eigenvalue problems is to reduce matrix $N \in \mathbf{R}^{2n \times 2n}$ to tridiagonal form using a general similarity

transformation X as [4]

$$X^{-1}NX = T = \begin{bmatrix} \alpha_1 & \gamma_1 & & \cdots & & 0 \\ \beta_1 & \alpha_1 & \ddots & & & \vdots \\ & \ddots & \ddots & \ddots & & \\ \vdots & & \ddots & \ddots & \ddots & \gamma_{2n-1} \\ 0 & \cdots & & & \beta_{2n-1} & \alpha_{2n} \end{bmatrix}.$$

With the column partitionings, one denotes

$$X = [x_1, \ldots, x_{2n}],$$

$$X^{-T} = Y = [y_1, \ldots, y_{2n}].$$

Upon comparing columns in $NX = XT$ and $N^T Y = YT^T$, one obtains that

$$Nx_j = \gamma_{j-1} x_{j-1} + \alpha_j x_j + \beta_j x_{j+1}$$

and

$$N^T y_j = \beta_{j-1} y_{j-1} + \alpha_j y_j + \gamma_j y_{y+1},$$

for $j = 1, \ldots, 2n - 1$, with $\gamma_0 x_0 = 0$ and $\beta_0 y_0 = 0$. These equations together with $Y^T X = I_{2n}$ imply the three-term recurrence formulae for computing α_j, β_j, and γ_j:

$$\alpha_j = y_j^T N x_j, \tag{3.1}$$

and

$$\beta_j x_{j+1} = r_j = (N - \alpha_j I) x_j - \gamma_{j-1} x_{j-1}, \tag{3.2}$$

$$\gamma_j y_{j+1} = p_j = (N - \alpha_j I)^T y_j - \beta_{j-1} y_{j-1}. \tag{3.3}$$

There is some flexibility in choosing β_j and γ_j. A canonical choice is to set $\beta_j = \|r_j\|_2$ and $\gamma_j = x_{j+1}^T p_j$.

In the previous section we proved that there exists a K_+-similarity transformation such that a K_--matrix can be reduced to an unreduced

K_--tridiagonal matrix. Suppose

$$X^{-1}NX = T = \begin{bmatrix} T_1 & T_2 \\ -T_2 & -T_1 \end{bmatrix},$$

with

$$T_1 = \begin{bmatrix} \alpha_1 & \gamma_2 & & \\ \beta_2 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \gamma_n \\ & & \beta_n & \alpha_n \end{bmatrix} \quad \text{and} \quad T_2 = \begin{bmatrix} \tilde{\alpha}_1 & \tilde{\gamma}_2 & & \\ \tilde{\beta}_2 & \tilde{\alpha}_2 & \ddots & \\ & \ddots & \ddots & \tilde{\gamma}_n \\ & & \tilde{\beta}_n & \tilde{\alpha}_n \end{bmatrix}.$$

One wishes to be able to derive a set of formulae that are similar to the three-term recurrence formulae (3.1)–(3.3). However, there are more scalar factors α_j, $\tilde{\alpha}_j$, β_j, $\tilde{\beta}_j$, γ_j, and $\tilde{\gamma}_j$ to be determined than the identities at hand. Because of this difficulty, we develop a new elimination procedure to annihilate the lower subdiagonals $\tilde{\beta}_j$'s of T_2 so that we are able to derive the five-term recurrence formulae and propose the novel K_--Lanczos algorithm. First we prove the following lemma.

LEMMA 3.1. *Suppose N is a 4×4 matrix with the following structure:*

$$N = \begin{bmatrix} A & B \\ -B & -A \end{bmatrix}, \tag{3.4}$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix},$$

and $|a_{21}| \neq |b_{21}|$. Then there exists a 4×4 nonsingular K_+-diagonal matrix D such that

$$DND^{-1} = \begin{bmatrix} \tilde{A} & \tilde{B} \\ -\tilde{B} & -\tilde{A} \end{bmatrix} \quad \text{with} \quad \tilde{B} = \begin{bmatrix} \tilde{b}_{11} & \tilde{b}_{12} \\ 0 & \tilde{b}_{22} \end{bmatrix}. \tag{3.5}$$

Proof. The entry a_{21} is used as a "pivot" element. Depending on zero or nonzero of a_{21}, we derive this lemma as follows.

Case I. If $a_{21} = 0$, let D be the K_+-permutation matrix

$$D = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right];$$

then it is easy to check that DND^{-1} has the structure defined in (3.5).

Case II. If $a_{21} \neq 0$, let

$$D = \left[\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & \dfrac{b_{21}}{a_{21}} & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \dfrac{b_{21}}{a_{21}} \end{array}\right];$$

then D is a K_+-diagonal matrix. Since $|a_{21}| \neq |b_{21}|$, D is nonsingular and

$$D^{-1} = \left[\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & \dfrac{-a_{21}b_{21}}{a_{21}^2 - b_{21}^2} & 0 & \dfrac{a_{21}^2}{a_{21}^2 - b_{21}^2} \\ \hline 1 & 0 & 0 & 0 \\ 0 & \dfrac{a_{21}^2}{a_{21}^2 - b_{21}^2} & 0 & \dfrac{-a_{21}b_{21}}{a_{21}^2 - b_{21}^2} \end{array}\right].$$

One can verify that DND^{-1} has the structure in (3.5)   ■

This elimination procedure can be generalized to the general $2n \times 2n$ K_--tridiagonal cases. Suppose N is an unreduced K_--tridiagonal matrix with structure as defined in (3.4) but with

$$A = \begin{bmatrix} a_1 & c_2 & & & \\ b_2 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & c_n \\ & & & b_n & a_n \end{bmatrix}, \quad B = \begin{bmatrix} \tilde{a}_1 & \tilde{c}_2 & & & \\ \tilde{b}_2 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \tilde{c}_n \\ & & & \tilde{b}_n & \tilde{a}_n \end{bmatrix}, \quad (3.6)$$

and $|b_j| \neq |\bar{b}_j|$ for all $j = 2, \ldots, n$. Without loss of generality, we assume $b_j \neq 0$, for all $j = 2, \ldots, n$. Now let

$$D_2 = \left[\begin{array}{cccccc|cccccc} 0 & 0 & & & & & 1 & 0 & & & & \\ 0 & \dfrac{\tilde{b}_2}{b_2} & & & & & 0 & 1 & & & & \\ & & 0 & & & & & & 1 & & & \\ & & & \ddots & & & & & & \ddots & & \\ & & & & 0 & & & & & & 1 & \\ 1 & 0 & & & & & 0 & 0 & & & & \\ 0 & 1 & & & & & 0 & \dfrac{\tilde{b}_2}{b_2} & & & & \\ & & 1 & & & & & & 0 & & & \\ & & & \ddots & & & & & & \ddots & & \\ & & & & 1 & & & & & & 0 & \end{array}\right]. \quad (3.7)$$

One can verify that D_2 is nonsingular and

$$D_2 N D_2^{-1} = N^{(2)} = \begin{bmatrix} A^{(2)} & B^{(2)} \\ -B^{(2)} & -A^{(2)} \end{bmatrix}$$

with

$$A^{(2)} = \left[\begin{array}{cccccc} -a_1 & & * & & & \\ * & & * & & * & \\ & \dfrac{b_2\left(\tilde{b}_2\tilde{b}_3 - b_2 b_3\right)}{b_2^2 - \tilde{b}_2^2} & -a_3 & -c_4 & & \\ & & -b_3 & -a_4 & -c_5 & \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & -c_n \\ & & & & & -b_n & -a_n \end{array}\right],$$

and

$$
B^{(2)} = \begin{bmatrix}
-\tilde{a}_1 & & * & & & & \\
0 & & * & & * & & \\
& \dfrac{b_2(\bar{b}_2\tilde{b}_3 - b_2\bar{b}_3)}{b_2^2 - \bar{b}_2^2} & & -\tilde{a}_3 & -\tilde{c}_4 & & \\
& & & -\bar{b}_3 & -\tilde{a}_4 & -\tilde{c}_5 & \\
& & & & \ddots & \ddots & \ddots \\
& & & & & \ddots & \ddots & -\tilde{c}_n \\
& & & & & & -\bar{b}_n & -\tilde{a}_n
\end{bmatrix}.
$$

Here $*$ denotes a component that has been modified. Note that the $(2,1)$ entry of B has been eliminated.

Next we verify that $N^{(2)}$ is still unreduced by showing $|\bar{b}_2\tilde{b}_3 - b_2\bar{b}_3| \neq |\bar{b}_2\tilde{b}_3 - b_2\bar{b}_3|$. Suppose, otherwise, one has $\bar{b}_2\tilde{b}_3 - b_2\bar{b}_3 = \pm(b_2\tilde{b}_3 - b_2\bar{b}_3)$. It implies $(b_2 \pm \bar{b}_2)(b_3 \mp \bar{b}_3) = 0$. However, this contradicts that the assumption $|b_2| \neq |\bar{b}_2|$ and $|b_3| \neq |\bar{b}_3|$. Thus, all corresponding lower subdiagonal entries of $A^{(2)}$ and $B^{(2)}$ are unequal in absolute value. This means $N^{(2)} = D_2 N D_2^{-1}$ is maintained to be unreduced K_--tridiagonal.

This procedure can be repeated for $i = 3, \dots, n$, by defining

$$
\begin{bmatrix}
0 & & & & 1 & & & \\
& \ddots & & & & \ddots & & \\
& & \dfrac{\bar{b}_i^{(i)}}{b_i^{(i)}} & & & & 1 & \\
& & & \ddots & & & & \ddots \\
& & & & 0 & & & & 1 \\
\hline
1 & & & & 0 & & & \\
& \ddots & & & & \ddots & & \\
& & 1 & & & & \dfrac{\bar{b}_i^{(i)}}{b_i^{(i)}} & \\
& & & \ddots & & & & \ddots \\
& & & & 1 & & & & 0
\end{bmatrix} \begin{matrix} \\ \\ \\ \\ i\text{th row} \\ \\ \\ \\ (n+i)\text{th row} \\ \end{matrix},
$$

where $b_i^{(i)}$ and $\bar{b}_i^{(i)}$ are the $(i, i-1)$ entries of $A^{(i)}$ and $B^{(i)}$, respectively. $A^{(i)}$ and $B^{(i)}$ denote the matrices after $(i-1)$ modifications. Note that whenever the "pivot" element $b_i^{(i)}$ equals zero, one can choose D_i to be a proper K_+-permutation matrix. By letting $D = D_n D_{n-1} \cdots D_2$, we proved the following theorem.

THEOREM 3.1. *Suppose*

$$N = \begin{bmatrix} A & B \\ -B & -A \end{bmatrix},$$

is a $2n \times 2n$ unreduced K_--tridiagonal matrix; then there exists a nonsingular K_+-matrix D such that the lower subdiagonals of B are annihilated. That is,

$$DND^{-1} = \begin{bmatrix} A^{(n)} & B^{(n)} \\ -B^{(n)} & -A^{(n)} \end{bmatrix},$$

where $A^{(n)}$ remains tridiagonal and $B^{(n)}$ is upper bidiagonal.

With this elimination procedure and the result of Theorem 3, we have the following theorem.

THEOREM 3.2 (K_--Lanczos Reduction Theorem). *Let N be a K_--matrix. If $V[N, x, n]$ and $V[N^T, y, n]$ are nonsingular K_--Krylov matrices for some $x, y \in \mathbf{R}^{2n}$ such that $V[N^T, y, n]^T V[N, x, n]$ has the real K_+-LR decomposition, then there exists a nonsingular K_+-matrix $X \in \mathbf{R}^{2n}$ such that*

$$X^{-1}NX = T = \begin{bmatrix} T_1 & T_2 \\ -T_2 & -T_1 \end{bmatrix}, \tag{3.8}$$

where T_1 is tridiagonal and T_2 is upper bidiagonal.

With this theorem, we derive a set of five-term recurrence formulae, hence the novel approach: the K_--Lanczos algorithm. These formulae are similar to the three-term formulae (3.1)–(3.3) for nonsymmetric Lanczos algorithm.

With column partitionings, we denote

$$X = [x_1, x_2, \ldots, x_n | Kx_1, Kx_2, \ldots, Kx_n], \qquad (3.9)$$

$$X^{-T} = Y = [y_1, y_2, \ldots, y_n | Ky_1, Ky_2, \ldots, Ky_n], \qquad (3.10)$$

and

$$
T_1 = \begin{bmatrix} \alpha_1 & \gamma_2 & & \\ \beta_2 & \alpha_2 & \cdot & \\ & \cdot & \cdot & \gamma_n \\ & & \beta_n & \alpha_n \end{bmatrix}, \qquad
T_2 = \begin{bmatrix} \tilde{\alpha}_1 & \tilde{\gamma}_2 & & \\ 0 & \tilde{\alpha}_2 & \cdot & \\ & \cdot & \cdot & \tilde{\gamma}_n \\ & & 0 & \tilde{\alpha}_n \end{bmatrix}. \qquad (3.11)
$$

Upon comparing columns in $NX = XT$ and $N^T Y = YT^T$, we obtain

$$Nx_j = \gamma_j x_{j-1} + \alpha_j x_j + \beta_{j+1} x_{j+1} - \tilde{\gamma}_j Kx_{j-1} - \tilde{\alpha}_j Kx_j, \qquad (3.12)$$

and

$$N^T y_j = \beta_j y_{j-1} + \alpha_j y_j + \gamma_{j+1} y_{j+1} + \tilde{\alpha}_j Ky_j + \tilde{\gamma}_{j+1} Ky_{y+1}, \qquad (3.13)$$

for $j = 1, \ldots, n - 1$, with $x_0 = y_0 = 0$. These identities together with $Y^T X = I_{2n}$ imply

$$\alpha_j = y_j^T Nx_j, \qquad (3.14)$$

$$\tilde{\alpha}_j = y_j^T NKx_j, \qquad (3.15)$$

and

$$\beta_{j+1} x_{j+1} = r_{j+1} = Nx_j - \gamma_j x_{j-1} - \alpha_j x_j + \tilde{\gamma}_j Kx_{j-1} + \tilde{\alpha}_j Kx_j, \qquad (3.16)$$

$$\left(\gamma_{j+1} I_{2n} + \tilde{\gamma}_{j+1} K \right) y_{j+1} = p_{j+1} = N^T y_j - \beta_j y_{j-1} - \alpha_j y_j - \tilde{\alpha}_j Ky_j. \qquad (3.17)$$

There is some flexibility in choosing the scalar factor β_{j+1}. An intuitive choice is to set $\beta_{j+1} = \|r_{j+1}\|_2$; then one can derive $\gamma_{j+1} = x_{j+1}^T p_{j+1}$, and $\tilde{\gamma}_{j+1} =$

$(Kx_{j+1})^T p_{j+1}$. If $|\gamma_{j+1}| \neq |\tilde{\gamma}_{j+1}|$, one obtains

$$y_{j+1} = \left(\gamma_{j+1} I_{2n} + \tilde{\gamma}_{j+1} K \right)^{-1} p_{j+1}$$

$$= \left(\frac{\gamma_{j+1}}{\gamma_{j+1}^2 - \tilde{\gamma}_{j+1}^2} I_{2n} - \frac{\tilde{\gamma}_{j+1}}{\gamma_{j+1}^2 - \tilde{\gamma}_{j+1}^2} K \right) p_{j+1}. \qquad (3.18)$$

We now summarize these mathematical formulations in the following, which we call nonsymmetric K_--Lanczos algorithm.

Algorithm: Nonsymmetric K_--Lanczos
    Given nonzero vectors x_1 and y_1 such that $x_1^T y_1 = 1$ and
        $\|x_1\|_2 = \|y_1\|_2 = 1$.
    Set $\beta_1 = \gamma_1 = 1$, $\tilde{\gamma}_1 = 0$.
    Set $x_0 = y_0 = 0$, and $p_1 = y_1$.
    Set $j = 1$.
    While $\beta_j \neq 0 \wedge |\gamma_j| \neq |\tilde{\gamma}_j|$
$$y_j = \left(\frac{\gamma_j}{\gamma_j^2 - \tilde{\gamma}_j^2} I_{2n} - \frac{\tilde{\gamma}_j}{\gamma_j^2 - \tilde{\gamma}_j^2} K \right) p_j;$$
        $\alpha_j = y_j^T N x_j;$
        $\tilde{\alpha}_j = y_j^T N K x_j;$
        $j = j + 1;$
        $r_j = N x_{j-1} - \gamma_{j-1} x_{j-2} - \alpha_{j-1} x_{j+1} \tilde{\gamma}_{j-1} K x_{j-2} + \tilde{\alpha}_{j-1} K x_{j-1};$
        $p_j = N^T y_{j-1} - \beta_{j-1} y_{j-2} - \alpha_{j-1} y_{j-1} - \tilde{\alpha}_{j-1} K y_{j-1};$
        $\beta_j = \|r_j\|_2;$
        $x_j = r_j / \beta_j;$
        $\gamma_j = x_j^T p_j;$
        $\tilde{\gamma}_j = (Kx_j) p_j;$
    End While
End Algorithm

Let

$$X_j = \left[x_1, x_2, \ldots, x_j \,|\, Kx_1, Kx_2, \ldots, Kx_j \right],$$

$$Y_j = \left[y_1, y_2, \ldots, y_j \,|\, Ky_1, Ky_2, \ldots, Ky_j \right],$$

and

$$T_j = \begin{bmatrix} (T_1)_j & (T_2)_j \\ -(T_2)_j & -(T_1)_j \end{bmatrix},$$

where $(T_1)_j$ and $(T_2)_j$ are the leading j-by-j principal submatrices of T_1 and T_2, respectively. It can be verified that

$$NX_j = X_j T_j + \beta_{j+1} x_{j+1} e^T_{j,2n} - \beta_{j+1} K x_{j+1} e^T_{2j,2n},$$

and

$$N^T Y_j = Y_j T_j^T + (\gamma_{j+1} + \tilde{\gamma}_{j+1}) x_{j+1} e^T_{j,2n} - (\gamma_{j+1} + \tilde{\gamma}_{j+1}) K x_{j+1} e^T_{2j,2n},$$

where $e_{j,2n}$ is the jth column of I_{2n}. Thus, whenever $\beta_j = \|r_j\|_2$, the columns of X_j define an invariant subspace for N. Termination in this regard is therefore a welcome event. However, if $|\gamma_j| = \tilde{\gamma}_j|$, then the iteration terminates without any invariant subspace information. This problem is one of the difficulties associated with Lanczos-type algorithms.

Eigenvalues of T_j are called the K_--Ritz values and used for the approximations of N's eigenvalues. Note that K is of the form defined in (1.2), and the computations Kx_j and Ky_j are free and can be viewed as permutations. It should not be counted as matrix-vector multiplications. Therefore, this nonsymmetric K_--Lanczos algorithm has the same order of computational complexity as the nonsymmetric Lanczos algorithm in terms of floating point operations.

4. ERROR BOUND ANALYSIS

An error bound for the K_--Ritz values obtained from the nonsymmetric K_--Lanczos algorithm is presented in this section. Hereinafter $\mathscr{P}^k$ is the set of polynomials of degree less than or equal to k. The next theorem establishes a relation between a 2×2-block tridiagonal matrix and its leading principal submatrices.

THEOREM 4.1. Let $\tilde{T}_n \in \mathbf{R}^{2n \times 2n}$ be a block tridiagonal matrix with 2×2 blocks and $\tilde{T}_m \in \mathbf{R}^{2m \times 2m}$ be a leading principal submatrix of $\tilde{T}_n$. Then

$$[e_{1,2n}, e_{2,2n}]^H f(\tilde{T}_n)[e_{1,2n}, e_{2,2n}]$$
$$= [e_{1,2m}, e_{2,2m}]^H f(\tilde{T}_m)[e_{1,2m}, e_{2,2m}], \tag{4.1}$$

for all $f \in \mathscr{P}^{2m-1}$, where $e_{i,2n}$ and $e_{i,2m}$ are the ith columns of I_{2n} and I_{2m}, respectively.

Proof. The proof follows from Theorem 3.3 in [5].

Suppose T_n is the K_--tridiagonal matrix obtained from applying n iterations of the nonsymmetric K_--Lanczos algorithm to a nonsymmetric K_--matrix N and T_m is a K_--principal submatrix of T_n. One can find permutations matrices $\Pi_{2n} \in \mathbf{R}^{2n \times 2n}$ and $\Pi_{2m} \in \mathbf{R}^{2m \times 2m}$ such that $\Pi_{2n}^T T_n \Pi_{2n}$ and $\Pi_{2m}^T T_m \Pi_{2m}$, are block tridiagonal matrices with 2×2 blocks. Following from the previous theorem, we have

$$[e_{1,2n}, e_{2,2n}]^H \Pi_{2n}^H f(T_n) \Pi_{2n}[e_{1,2n}, e_{2,2n}]$$

$$= [e_{1,2m}, e_{2,2m}]^H \Pi_{2m}^H f(T_m) \Pi_{2m}[e_{1,2m}, e_{2,2m}], \qquad (4.2)$$

for all $f \in \mathscr{P}^{2m-1}$. Since Π_{2n} and Π_{2m} are permutation matrices, (4.2) becomes

$$[e_{1,2n}, e_{n+1,2n}]^H f(T_n)[e_{1,2n}, e_{n+1,2n}]$$

$$= [e_{1,2m}, e_{m+1,2m}]^H f(T_m)[e_{1,2m}, e_{m+1,2m}], \qquad (4.3)$$

for all $f \in \mathscr{P}^{2m-1}$.

For simplicity, our error-bound analysis focuses on the case where T_n and T_m are both diagonalizable. Suppose $T_n = X^H \Lambda Y$ and $T_m = P^H \Theta Q$, where

$$\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n \mid -\lambda_1, \ldots, -\lambda_n) \equiv \begin{bmatrix} \Lambda_1 & 0 \\ 0 & -\Lambda_1 \end{bmatrix},$$

and

$$\Theta = \mathrm{diag}(\theta_1, \ldots, \theta_m \mid -\theta_1 \ldots, -\theta_m) \equiv \begin{bmatrix} \Theta_1 & 0 \\ 0 & -\Theta_1 \end{bmatrix}.$$

Substituting into Eq. (4.3), one has

$$[e_{1,2n}, e_{n+1,2n}]^H X^H f(\Lambda) Y[e_{1,2n}, e_{n+1,2n}]$$

$$= [e_{1,2m}, e_{m+1,2m}]^H P^H f(\Theta) Q[e_{1,2m}, e_{m+1,2m}], \qquad (4.4)$$

for all $f \in \mathscr{P}^{2m-1}$. This implies

$$
\begin{bmatrix} x_1^H \\ x_{n+1}^H \end{bmatrix} \begin{bmatrix} f(\Lambda_1) & 0 \\ 0 & f(-\Lambda_1) \end{bmatrix} [y_1, y_{n+1}]
$$

$$
= \begin{bmatrix} p_1^H \\ p_{n+1}^H \end{bmatrix} \begin{bmatrix} f(\Theta_1) & 0 \\ 0 & f(-\Theta_1) \end{bmatrix} [q_1, q_{n+1}]. \qquad (4.5)
$$

Here x_i, y_i, p_i, and q_i are the ith column of X, Y, P, and Q, respectively. Denote

$$
x_1 \equiv \begin{bmatrix} x_1^{(1)} \\ x_1^{(2)} \end{bmatrix},
$$

where $x_1^{(1)}$, $x_1^{(2)} \in \mathbf{R}^n$. Using the similar notation for x_{n+1}, y_1, y_{n+1}, p_1, p_{n+1}, q_1, and q_{n+1}, we obtain the identity from (4.5),

$$
\left(x_1^{(1)} + x_1^{(2)} \right)^H \left(f(\Lambda_1) + f(-\Lambda_1) \right) \left(y_1^{(1)} + y_1^{(2)} \right)
$$

$$
= \left(p_1^{(1)} + p_2^{(2)} \right)^H \left(f(\Theta_1) + f(-\Theta_1) \right) \left(q_1^{(1)} + q_1^{(2)} \right) \qquad (4.6)
$$

for all $f \in \mathscr{P}^{2m-1}$. Let $g(\Lambda_1) = f(\Lambda_1) + f(-\Lambda_1)$, $g \in \mathscr{P}^{2m-2}$. Then (4.6) can be rewritten as

$$
\sum_{i=1}^n g(\lambda_i)(\tilde{x}_{i,1} + \tilde{x}_{n+i,1})(y_{i,1} + y_{n+i,1})
$$

$$
= \sum_{i=1}^n g(\theta_i)(\bar{p}_{i,1} + \bar{p}_{n+i,1})(q_{i,1} + q_{n+i,1}). \qquad (4.7)
$$

Define

$$
\varepsilon_1^{(k)} = \inf_{p \in \mathscr{P}^k,\, p(\lambda_1)=1} \max_{x \in S} |p(x)|, \qquad (4.8)
$$

and

$$
\delta_1(S, \tilde{S}) = \max \left\{ |x - \theta_1| \prod_{\lambda \in S} \frac{|x - \lambda|}{|\lambda_1 - \lambda|} : x \in S \cup \tilde{S} \right\}, \qquad (4.9)
$$

where S and $\tilde{S}$ are two disjoint sets. An error bound for the extreme K_--Ritz values obtained from the nonsymmetric K_--Lanczos algorithm is derived in the following theorem.

THEOREM 4.2. *Suppose T_n and T_m are diagonalizable so that $\lambda(T_n) = \{\lambda_1, \ldots, \lambda_n \mid -\lambda_1, \ldots, -\lambda_n\}$ and $\lambda(T_m) = \{\theta_1, \ldots, \theta_n \mid -\theta_1, \ldots, -\theta_m\}$. Assume that λ_1 is the largest eigenvalue in magnitude of T_n and $|\lambda_1 - \theta_1| = \min_j |\lambda_1 - \theta_j|$. Let $\sigma_1 = \{\lambda_2, \ldots, \lambda_n\}$ and $\hat{\sigma}_1 = \{\theta_2, \ldots, \theta_m\}$. If $\sigma_1 \cup \hat{\sigma}_1 = S_1 \cup S_2$ with S_1 and S_2 disjoint, and $s = |S_2| \leqslant 2m - 3$, then*

$$
|\lambda_1 - \theta_1| \leqslant \varepsilon_1^{(2m-3-s)}(S_1)\,\delta_1(S_2, S_1)\frac{1}{|x_{11} + x_{n+1,1}|\,|y_{11} + y_{n+1,1}|}
$$

$$
\times \left(\sum_{i=2}^{n} |x_{i,1} + x_{n+i,1}|\,|y_{i,1} + y_{n+i,1}| \right.
$$

$$
\left. + \sum_{i=2}^{m} |p_{i,1} + p_{n+i,1}|\,|q_{i,1} + q_{m+i,1}| \right). \tag{4.10}
$$

Proof. Substituting $g(x) = (x - \theta_1)h(x)\prod_{\lambda \in S_2}(x - \lambda)$ for any $h \in \mathscr{P}^{2m-3-s}$ with $h(\lambda_1) = 1$ into (4.7), one obtains

$$
(\lambda_1 - \theta_1)(\bar{x}_{11} + \bar{x}_{n+1,1})(y_{11} + y_{n+1,1})h(\lambda_1)
$$

$$
= \sum_{\theta_i \in S_1} (\theta_i - \theta_1) \prod_{\lambda \in S_2} \frac{(\theta_i - \lambda)}{(\theta_1 - \lambda)} h(\theta_i)(\bar{p}_{i,1} + \bar{p}_{n+i,1})(q_{i,1} + q_{m+i,1})
$$

$$
- \sum_{\lambda_i \in S_1} (\lambda_i - \theta_1) \prod_{\epsilon S_2} \frac{(\lambda_i - \lambda)}{(\lambda_1 - \lambda)} h(\lambda_i)(\bar{x}_{i,1} + \bar{x}_{i,1})(y_{i,1} + y_{m+i,1}).
$$

The error bound (4.10) follows immediately.

In [5], Ye presented a convergence analysis for the nonsymmetric Lanczos algorithm. Error bounds for Ritz values and Ritz vectors were established. In [6], Flaschka discussed the error bound from applying Lanczos-type algorithm to symmetric K_--matrices. If we apply the nonsymmetric Lanczos algorithm on K_--matrix N to obtain a tridiagonal matrix T_m after m iterations and utilize the analysis in [5], we have the following error bound.

THEOREM 4.3.   *Suppose that N and T_m are diagonalizable with $\sigma(N) = \{\lambda_1, \ldots, \lambda_n, -\lambda_1, \ldots, -\lambda_n\}$ and $\sigma(T_m) = \{\theta_1, \ldots, \theta_m\}$. Assume that $|\lambda_1 - \theta_1| = \min_j |\lambda_1 - \theta_j|$, and let $\sigma_2 = \{\lambda_2, \ldots, \lambda_n, -\lambda_1, \ldots, -\lambda_n\}$, $\hat{\sigma}_2 = \{\theta_2, \ldots, \theta_m\}$. If $\sigma_2 \cup \hat{\sigma}_2 = \tilde{S}_1 \cup \tilde{S}_2$ with $\tilde{S}_1$ and $\tilde{S}_2$ disjoint, and $s = |\tilde{S}_2| \leqslant 2m - 2$, then*

$$|\lambda_1 - \theta_1| \leqslant \varepsilon_1^{(2m-2-s)}(\tilde{S}_1)\delta_1(\tilde{S}_2, \tilde{S}_1)$$

$$\times \frac{\left(\sum_{i=2}^{2n}|x_{i,1}|^2 + \sum_{i=2}^{m}|p_{i,1}|^2\right)^{1/2}}{|x_{1,1}|}$$

$$\times \frac{\left(\sum_{i=2}^{2n}|y_{i,1}|^2 + \sum_{i=2}^{m}|q_{i,1}|^2\right)^{1/2}}{|y_{1,1}|}. \qquad (4.11)$$

Following from the discussion in [5] we analyze the magnitude of $\varepsilon_1^{(k)}(S_1)$ as follows. Suppose S_1 lies inside an ellipse E and λ_1 lies on the major axis of E but outside S_1. Let $|\lambda_1| = d$ and the length of foci and semimajor axis of E are c and a, respectively. Then $0 < c \leqslant a \leqslant d$. Let $T_k(x)$ be the Chebyshev polynomial of degree k on $[-1, 1]$ (see [8, 9]). Then

$$\min_{p \in \mathscr{P}^k, \, p(\lambda_1)=1} \, \max_{x \in E}|p(x)| = \max_{x \in E}|p_k(x)| = \frac{T_k(a/c)}{T_k(d/c)},$$

where $p_k(x) = T_k((d-x)/c)/T_k(d/c)$ [8]. Hence

$$\varepsilon_1^{(k)}(S_1) \leqslant \frac{T_k(a/c)}{T_k(d/c)}. \qquad (4.12)$$

Since $1 \leqslant a/c \leqslant d/c$, we have $T_k(a/c) \leqslant T_k(d/c)$. Furthermore, if $a \ll d$, that is, λ_1 is well separated from the ellipse E, then $\varepsilon_1^{(k)}(S_1)$ is small. The bigger the difference between a/c and d/c, the smaller the bound of (4.12). Note that d/c is a measure of separation of λ_1 from E, and a/c is a measure of flatness of the ellipse E.

On the other hand, if $s = |S_2|$ is small, $\delta_1(S_2, S_1)$ is a bound number. If s is large and, in addition, $|x - \lambda| < |\lambda_1 - \lambda|$ for most $\lambda \in S_2$ and any $x \in S_1$, then $\delta_1(S_2, S_1)$ is a product of s numbers (see (4.9)), most of which are less than one. Hence it is a small number.

Thus, for an extreme eigenvalue λ_1, we can partition $\sigma_1 \cup \hat{\sigma}_1$ into a union of S_1 and S_2 so that s is small and S_1 lies in a flat ellipse, which is well separated from λ_1. Then Theorem 4.2 says that we can expect a good approximation bound for λ_1. By comparing Theorems 4.2 and 4.3, since the $|S_2|$ is less than $|\tilde{S}_2|$, (4.10) is a tighter error bound.

5. NUMERICAL EXPERIMENTS AND REMARKS

In this section, we use several numerical experiments to access the viability of the proposed nonsymmetric K_--Lanczos algorithms to extract the eigenpairs for the K_--eigenvalue problem (1.5). In the experiments, we focused on finding the eigenvalues with maximal absolute values and the corresponding eigenvectors. Based on the numerical results, we compare the convergence behavior and numerical efficiency of this novel structure-preserving algorithm with the conventional nonsymmetric Lanczos algorithm.

The results reported herein were obtained using Pro-Matlab 3.x. Random K_--matrices with known exact eigenvalues are generated. The spectral distributions of these random K_--matrices are manipulated to be clustered along the real axis and imaginary axis to match the typical phenomena arising in the time-independent Hartree–Fock model. The ratios of the largest eigenvalue in magnitude to the second largest, $|\lambda_1|/|\lambda_2|$, are all less than 1.02 in order to test the robustness of the algorithms. Many experiments have been conducted and we report here only a few of them in Table 1, which shows the number of iterations required for the K_--Ritz value and Ritz value to converge to the largest eigenvalue in magnitude for matrices of different sizes. Figures 1–3 illustrate the convergent behavior of these two algorithms for the first three testings. We note that the results in Table 1 are typical among all testings.

TABLE 1

COMPARISON OF NUMBER OF ITERATIONS REQUIRED TO CONVERGE TO THE
LARGEST EIGENVALUE IN MAGNITUDE

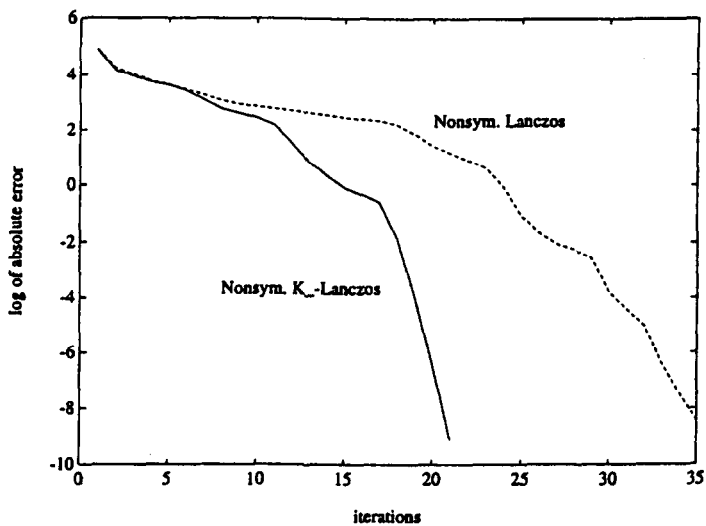| Size of matrix N	Nonsym. K_--Lanczos	Nonsym. Lanczos
500	21	35
1000	31	51
5000	82	113
10000	84	126
50000	87	131

WILLIAM R. FERNG ET AL

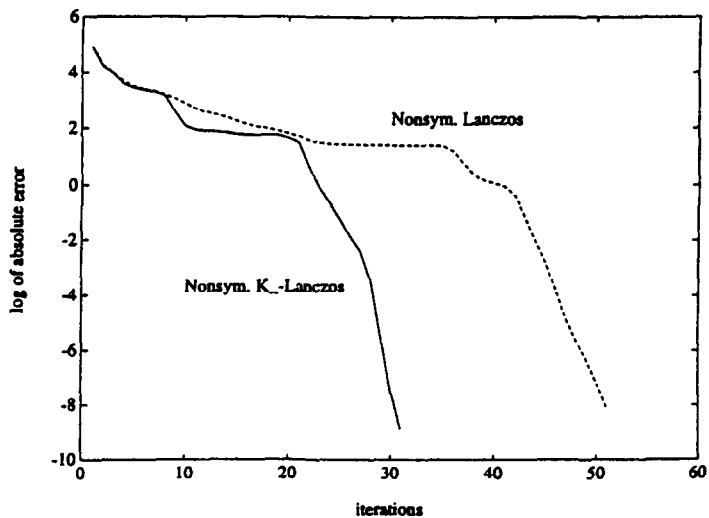FIG. 1. Comparison of convergence behavior for a random 500×500 K_- matrix.

FIG. 2. Comparison of convergence behavior for a random 1000×1000 K_- matrix.
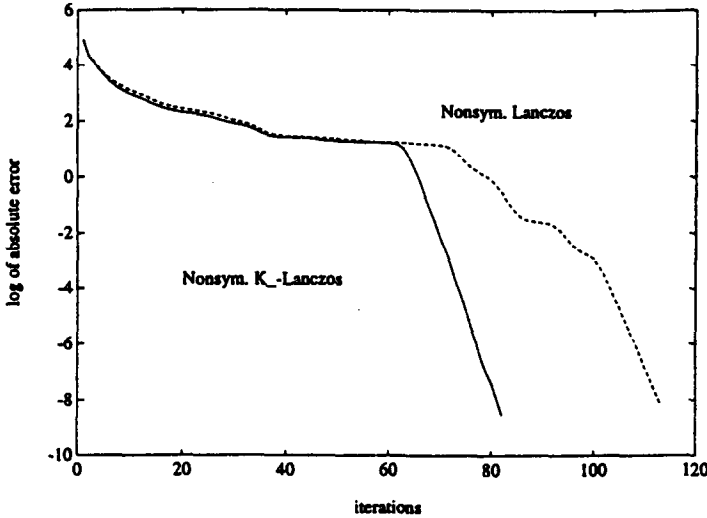
FIG. 3.   Comparison of convergence behavior for a random 5000×5000 K_--matrix.

On the basis of these results, we observe that the proposed nonsymmetric K_--Lanczos algorithm converges faster in terms of the number of iterations to converge than the classic nonsymmetric Lanczos approach in all experiments we conducted. Some remarks are in order.

1. After j iterations, nonsymmetric K_--Lanczos algorithm produces a $2j \times 2j$ k_--tridiagonal matrix, while nonsymmetric Lanczos algorithm produces a $j \times j$ tridiagonal matrix. The computational complexities for the nonsymmetric K_--Lanczos reduction and the nonsymmetric Lanczos reduction are approximately even.

2. The KQR algorithm [6] or KQZ algorithm [7] should be applied to compute the K_--Ritz values from the K_--tridiagonal matrix. On the other hand, the QR algorithm [10, 11, 12] or QZ algorithm [13] can be applied for computing the Ritz values from the tridiagonal matrix. The computational cost for obtaining the same number of K_--Ritz values via KQZ algorithm and Ritz values via QZ algorithm are equal [7].

3. If the extreme eigenvalue in magnitude λ_1 of a K_--matrix N is complex, it follows that λ_1, $\bar{\lambda}_1$, $-\lambda_1$, $-\bar{\lambda}_1$ are all extreme eigenvalues in magnitude of N. There will be four K_--Ritz values obtained at one time from the nonsymmetric K_--Lanczos algorithm to approximate these four eigenval-

ues at the same time. However, the Ritz values produced by the nonsymmetric Lanczos algorithm cannot attain this behavior.

4. If T_j is the $2j \times 2j$ K_--tridiagonal matrix produced by the nonsymmetric K_--Lanczos algorithm, there exists a permutation matrix Π_{2j} such that $\Pi_{2j} T_j \Pi_{2j}^{-1} = \tilde{T}_j$ is a 2×2-block tridiagonal matrix. Indeed, nonsymmetric K_--Lanczos algorithm can be viewed as a special 2×2-block nonsymmetric Lanczos algorithm. Hence better convergent rate can be expected.

5. Neither the nonsymmetric Lanczos algorithm nor the 2×2-block nonsymmetric Lanczos algorithm can maintain the K_--structure like the nonsymmetric K_--Lanczos algorithm does.

6. One may construct counterexamples to make the algorithm break down. However, the breakdown situation has not been observed in our numerical experiments with randomly generated K_--matrices.

6. CONCLUSION

In this article, we propose a novel method, the nonsymmetric K_--Lanczos algorithm, for solving the K_--eigenvalue problems, $Nx = \lambda x$, hence the generalized eigenvalue problems $Mx = \lambda Lx$. We proved the existence of K_--tridiagonalization theorem, developed a new elimination procedure, and derived an error bound for the extreme K_--Ritz value. Numerical experiments show that nonsymmetric K_--Lanczos algorithm converges faster than the classic nonsymmetric Lanczos algorithm in all randomly generated test K_--matrices. We remark that the error bound is not intended to provide a practical computable estimation of the number of iterations required, but rather to illustrate the convergence behavior of the nonsymmetric K_--Lanczo algorithm.

REFERENCES

1. J. Olsen, H. J. A. Jensen, and P. Jorgensen, Solution of the large matrix equations which occur in response theory, *J. Comput. Phys.* 74:265–282 (1973).
2. J. Olsen and P. Jorgensen, Linear and nonlinear response functions for an exact state and for an MCSCF state, *J. Chem. Phys.* 82:3235 (1985).
3. B. N. Parlett, D. R. Taylor, and Z. A. Liu, A look-ahead Lanczos-algorithm for unsymmetric matrices, *Math. Comput.* 44:105–124 (1985).
4. G. H. Golub and C. Van Loan, *Matrix Computations*, 2nd ed., Johns Hopkins Press, Baltimore, 1989.
5. Q. Ye, A convergence analysis for nonsymmetric Lanczos algorithms, *Math. Comput.* 56:677–691 (1991).

6. U. Flaschka, Eine variante des Lanczos-algorithmus fuer grosse, duenn besetzte symmetrische Matrizen mit Blockstruktur, Doctoral Dissertation, Bielefeld Universitaet, 1992.
7. U. Flaschka, W. W. Lin, and J. L. Wu, A KQZ algorithm for solving linear-response eigenvalue equations, *Linear Algebra Appl.* 165:93–123 (1992).
8. T. A. Manteuffel, The Tchebychev iteration for nonsymmetric linear systems, *Numer. Math.* 28:307–327 (1977).
9. T. A. Manteuffel, Adaptive procedure for estimating parameters for the nonsymmetric Tchebychev iteration, *Numer. Math.* 31:183–208 (1978).
10. J. G. F. Francis, The QR transformation: A unitary for the solution of the LR transformation, *Comput. J.* 4:332–334 (1961).
11. V. N. Kublanovskaya, On some algorithms for the solution of the complete eigenvalue problem, *USSR. Comp. Math. Phys.* 3:637–657 (1961).
12. H. Rutishaueser, Solution of eigenvalue problems with the LR transformation, *Nat. Bur. Stand. App. Math. Ser.* 49:47–81 (1958).
13. J. Olsen, H. J. A. Jensen, and P. Jorgensen, Solution of the large matrix equations which occur in response theory, *J. Comput. Phys.* 74:265–282 (1988).