Numer. Math. (1998) 80: 239–266

Numerische
Mathematik

Electronic Edition

# Generalized block Lanczos methods for large unsymmetric eigenproblems

**Zhongxiao Jia**

Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, P.R. China; e-mail: zxjia@gingko.dlut.edu.cn

Received July 7, 1994 / Revised version received March 1, 1997

**Summary.** Generalized block Lanczos methods for large unsymmetric eigenproblems are presented, which contain the block Arnoldi method, and the block Arnoldi algorithms are developed. The convergence of this class of methods is analyzed when the matrix $A$ is diagonalizable. Upper bounds for the distances between normalized eigenvectors and a block Krylov subspace are derived, and a priori theoretical error bounds for Ritz elements are established. Compared with generalized Lanczos methods, which contain Arnoldi's method, the convergence analysis shows that the block versions have two advantages: First, they may be efficient for computing clustered eigenvalues; second, they are able to solve multiple eigenproblems. However, a deep analysis exposes that the approximate eigenvectors or Ritz vectors obtained by general orthogonal projection methods including generalized block methods may fail to converge theoretically for a general unsymmetric matrix $A$ even if corresponding approximate eigenvalues or Ritz values do, since the convergence of Ritz vectors needs more sufficient conditions, which may be impossible to satisfy theoretically, than that of Ritz values does. The issues of how to restart and to solve multiple eigenproblems are addressed, and some numerical examples are reported to confirm the theoretical analysis.

*Mathematics Subject Classification (1991):* 65F15, 15A18, 41A10

## 1. Introduction

In a large number of disciplines of applications, we frequently want to compute a subset of the spectrum of a large unsymmetric matrix $A$ and possibly

the corresponding eigenvectors or invariant subspaces [3,9,31,41]. For this kind of problem, a major class of methods are orthogonal projection methods [43]. Arnoldi's method [2] is an orthogonal projection method, where the projection subspace is a Krylov subspace. The method was developed by Saad [39], who defines a class of generalized Lanczos methods which reduce to Arnoldi's method when the basis of Krylov subspace is orthonormally constructed by Arnoldi's process. Since then much efforts have been made to this method both in theory and algorithms, e.g., [17]–[29], [36]–[43] and [47].

Saad [39] was the first to make a convergence analysis for generalized Lanczos methods, and he established bounds for the distances between normalized eigenvectors and a Krylov subspace when the matrix $A$ only has real simple eigenvalues. His results show that these distances usually converge to zero first for the eigenvectors associated with outer part eigenvalues. In a later paper [40], Saad extended the convergence theory to the case where $A$ is diagonalizable and has complex eigenvalues, and he dealt with the most right outer eigenpair $\lambda_1, \varphi_1$. In [19], the convergence theory of this class of methods is analyzed for more eigenpairs when $A$ is diagonalizable and complex eigenvalues are present. The results indicate that these distances usually tend to zero first for the eigenvectors corresponding to eigenvalues with the largest and smallest real parts. However, all these papers have not considered the implications of these distances on the behavior of Ritz elements, while these are very important and essential in understanding how generalized Lanczos methods converge and in guiding how to develop related algorithms. Recently, the convergence theory of this class of methods has been investigated in detail in [20,21] for the case that the matrix $A$ is defective, where a priori theoretical error bounds for Ritz elements are established. The results show that the methods still favor the outer part eigenvalues and the associated eigenvectors of $A$ usually though they may converge quite slowly due to the defectiveness of $A$. However, a detailed analysis [25,26, 28] has exposed that the approximate eigenvectors or Ritz vectors obtained by generalized Lanczos methods cannot be guaranteed to converge theoretically for a general unsymmetric matrix $A$ even if corresponding approximate values or Ritz values do, so that Arnoldi type algorithms may converge very slowly or irregularly and even may fail to converge. Moreover, this possible nonconvergence may occur to a general orthogonal projection method [25, 26,28].

In applications, however, some of the wanted eigenvalues are clustered, multiple, and the eigenproblem of $A$ is very ill conditioned or $A$ is highly nonnormal [3,5,6,11,31]. When $A$ is highly nonnormal, the eigenproblem of $A$ may be hard to solve by some iterative algorithms like subspace iteration; see Chatelin [8]. In terms of the theoretical analysis [19–21,39,

41,43], iterative Arnoldi type algorithms either exhibit poor performance or fail in these cases. For clustered eigenvalues and ill-conditioned eigenproblems, more precisely, it means that the minimal steps $m$ of iterative Arnoldi algorithms per restart might be quite large in order to ensure the convergence, which makes the algorithm become very expensive and even impractical. The Chebyshev acceleration [17,41], rational Krylov sequences techniques [36]–[38], etc., can overcome this difficulty to some extent, but they are useless to solve multiple eigenproblems unless some deflation techniques are used. Here we should distinguish the convergence of a method from the stability of related algorithms. It is shown in [1,9,13] that Arnoldi type algorithms can be guaranteed to be absolutely stable in finite precision arithmetic using the Householder transformation, the Givens rotation and the Gram–Schmidt orthogonalization with iterative refinement and they are only conditionally stable using the modified Gram–Schmidt orthogonalization. However, we should keep in mind that the stability of an algorithm does not mean the convergence of it and vice versa.

To overcome, at least to some extent, these drawbacks, we describe a class of generalized *block* Lanczos methods, which reduce to the *block* Arnoldi method when a basis of the block Krylov subspace orthonormally constructed by the block Arnoldi process. We present a basic and an iterative block Arnoldi algorithms. Block versions are not only for improving numerical efficiency, but also suitable for exploiting vector and parallel computers.

The idea of generalized block Lanczos methods is a natural extension of the single vector versions presented in [39]. In particular, one of them, the block Arnoldi method, has appeared somewhere, e.g., [4,43] and in many papers, e.g., [44–46]. However, much less work has been done for theoretically understanding the block versions, and in fact no convergence theory has been established hitherto. Just as the symmetric block Lanczos method [16] and its single vector version as well as the block conjugate gradient type methods [33] and their single vector versions, although there are some similar points, there exist a number of essential difficuties in extending the convergence theory of generalized Lanczos methods to their block versions. Therefore, how to carry out a convergence analysis and reveal features of generalized block Lanczos methods, in particular the block Arnoldi method, is nontrivial and difficult. Concerning implementations, it is obvious that the block Arnoldi algorithms are much more complicated than the Arnoldi algorithms. Because of the lack of a necessary and strong theoretical background, the development of block algorithms have been empirical and quite limited in the literature, and some important and essential issues concerning the block versions have not been dealt with, e.g., how to solve multiple eigenproblems both in theory and in practice. Although it is believed intuitively in the literature that the the block versions can solve multiple eigenproblems

and are efficient when some of the wanted eigenvalues are clustered, these assertions have not yet been proved theoretically.

In Sect. 2 we present generalized block Lanczos methods. In Sect. 3 we establish a convergence theory of these methods when the matrix $A$ is diagonalizable, in which bounds for the distances between normalized eigenvectors and a block Krylov subspace are derived and a priori theoretical error bounds for Ritz elements are given. Using them we expose why this class of methods may work well for computing clustered eigenvalues and may solve multiple eigenproblems under some circumstances as well as why Ritz vectors cannot be guaranteed to converge theoretically for a general unsymmetric $A$ even if Ritz values do. In Sect. 4 we present the block Arnoldi algorithms. In Sect. 5 we discuss implementations and address the issues of how to restart and to solve multiple eigenproblems. In Sect. 6 we report some numerical examples to confirm our theoretical analysis, followed by some concluding remarks in Sect. 7.

## 2. Generalized block Lanczos methods

### *2.1. General orthogonal projection methods and notation*

Assume the matrix $A \in \mathbb{R}^{N\times N}$ to be large, and $U_m = (V_1, V_2, \ldots, V_m)$ a matrix consisting of $mp$ linearly independent column vectors, where $V_i$, $i = 1, 2, \ldots, m$ have $p$ columns, and let $\lambda_i, \varphi_i$, $i = 1, 2, \ldots, N$ be the eigenpairs of $A$ with $\|\varphi_i\| = 1$, and $\|\cdot\|$ the Euclidean norm throughout the paper unless stated otherwise. We shall sometimes write $\lambda_i, \varphi_i$ simply as $\lambda, \varphi$. In the context we denote by the superscript * the conjugate transpose of a matrix and a vector. An orthogonal projection method on the subspace $\mathrm{span}\{U_m\}$ is as follows:

Determine $\lambda^{(m)}, \varphi^{(m)}$ as solutions of

$$\begin{cases} \varphi^{(m)} \in \mathrm{span}\{U_m\}, \\ (A - \lambda^{(m)} I)\varphi^{(m)} \perp \mathrm{span}\{U_m\}, \end{cases} \tag{1}$$

and take them as approximations to $\lambda$, $\varphi$, where $\|\varphi^{(m)}\| = 1$.

From the first relation of (1), there exists a vector $y^{(m)}$ of dimension $mp$ such that

$$\varphi^{(m)} = U_m y^{(m)}. \tag{2}$$

It is easily seen from (1) and (2) that $\lambda^{(m)}$, $y^{(m)}$ satisfy the following eigenproblem:

$$(B_m^{-1} C_m - \lambda^{(m)} I) y^{(m)} = 0, \tag{3}$$

where $C_m = U_m^* A U_m$, $B_m = U_m^* U_m$. Here $B_m^{-1} C_m$ is called the matrix representation of the restriction of $A$ to $\mathrm{span}\{U_m\}$ in the basis $\{V_i\}_1^m$, and $\lambda^{(m)}, \varphi^{(m)}$ are the Ritz values and Ritz vectors of $A$ in $\mathrm{span}\{U_m\}$.

### *2.2. Generalized block Lanczos methods*

We treat now the case that

$$\mathrm{span}\{U_m\} = \mathcal{K}_m(V_1, A),$$

where $\mathcal{K}_m(V_1, A)$ is the block Krylov subspace spanned by $V_1, AV_1, \ldots, A^{m-1}V_1$.

Let us give a definition.

**Definition.** Any projection process (1) on the block Krylov subspace $\mathcal{K}_m(V_1, A)$ applied to an unsymmetric matrix $A$ is called a generalized block Lanczos method.

The block Arnoldi method we will discuss is a generalized block Lanczos method when a basis $\{V_i\}_1^m$ of $\mathcal{K}_m(V_1, A)$ is orthonormally constructed by the block Arnoldi process; see Sect. 4 for details. It can be seen from this definition that different choices of the basis will give rise to different generalized block Lanczos methods and thus different algorithms. But it can be easily verified that mathematically $\lambda^{(m)}, \varphi^{(m)}$ are uniquely determined, independently of bases of $\mathcal{K}_m(V_1, A)$.

Note that assume $\pi_m$ to be the orthogonal projector on $\mathcal{K}_m(V_1, A)$. Then it can be shown that (3) is equivalent to the eigenproblem

$$A_m\varphi^{(m)} = \lambda^{(m)}\varphi^{(m)}, \tag{4}$$

where $A_m = \pi_m A\pi_m$. (4) has $N - mp$ zero eigenvalues, whose corresponding eigenvectors are orthogonal to $\mathcal{K}_m(V_1, A)$, and the rest just satisfy (3), whose corresponding eigenvectors $\varphi^{(m)}$ belong to $\mathcal{K}_m(V_1, A)$.

## 3. Convergence analysis

Now let us study the convergence theory of generalized block Lanczos methods. Before this it is necessary to remind the reader that the theory to be established will in fact work for this class of methods, not only for the block Arnoldi method, this is, we have no restriction to choices of the basis of $\mathcal{K}_m(V_1, A)$. But we will only develop block Arnoldi algorithms in this paper.

It is clear that in the finite $N$-dimensional space generalized block Lanczos methods must converge when $mp = N$, because then $\pi_m = I$, $\lambda^{(m)} = \lambda$ and $\varphi^{(m)} = \varphi$.

We prove four points: First, $\lambda^{(m)}$, $\varphi^{(m)}$ may first become good approximations to the outer part $\lambda$ and the corresponding eigenvectors $\varphi$; second, compared with generalized Lanczos methods, the block versions may be efficient for computing clustered eigenvalues under some circumstances;

third, the block versions may be able to solve multiple eigenproblems provided that the block size $p$ is no less than the multiplicities possessed by the wanted eigenvalues; fourth, Ritz vectors may be much harder to converge than Ritz values, more precisely, compared with that of $\lambda^{(m)}$, the convergence of $\varphi^{(m)}$ needs more sufficient conditions which may not be satisfied theoretically.

Before discussions, we need the following result proved by Saad [40].

**Theorem 1.** *Let $\gamma_m = \|\pi_m A(I-\pi_m)\|$, and $\lambda$, $\varphi$ an eigenpair of $A$. Then*

$$\begin{cases} \|(A_m - \lambda I)\pi_m\varphi\| \le \gamma_m \|(I-\pi_m)\varphi\|, \\ \|(A_m - \lambda I)\varphi\| \le \sqrt{|\lambda|^2+\gamma_m^2}\,\|(I-\pi_m)\varphi\|. \end{cases} \tag{5}$$

It is easily seen from Theorem 1 that $\gamma_m \le \|A\|$. Thus, we can get a small residual provided that $\|(I-\pi_m)\varphi\|$ is small. Moreover, in terms of [15,49], if $\|(I-\pi_m)\varphi\|$ is very small, there exists some pair $\lambda^{(m)}, \varphi^{(m)}$ to well approximate $\lambda, \varphi$ provided that the eigenproblem of $A_m$ is not too ill conditioned.

Now let us estimate $\|(I-\pi_m)\varphi\|$. To this end, assume $A$ to be diagonalizable, and denote by $Q_k$ the set consisting of all polynomials of degree not exceeding $k$, by $\theta(u, \mathcal{K}_m(V_1, A))$ the acute angle between a nonzero vector $u$ and the block Krylov subspace $\mathcal{K}_m(V_1, A)$, defined by

$$\theta(u, \mathcal{K}_m(V_1, A)) = \arcsin \frac{\|(I-\pi_m)u\|}{\|u\|}, \tag{6}$$

and by $P_j$ the spectral projectors associated with $\lambda_j$, $j = 1, 2, \ldots, N$.

### *3.1. Bounds for $\|(I-\pi_m)\varphi\|$*

In order to prove Theorem 2, which is the key result to bound $\|(I-\pi_m)\varphi\|$ in the block versions, we need the following simple but important lemma.

**Lemma 1.** *Assume the matrix $A$ to be diagonalizable, and define $E_1 = \mathrm{span}\{V_1\}$ and $\hat{P}_i = \sum_{j=i}^{i+p-1} P_j$. Assume $V_1$ such that the vectors $\hat{P}_i v_1, \ldots, \hat{P}_i v_p$ are linearly independent, where $V_1 = (v_1, \ldots, v_p)$. Then there exists a unique vector $\hat{x}_j$ in $E_1$ such that*

$$\hat{P}_i \hat{x}_j = \varphi_j, j = i, \ldots, i+p-1. \tag{7}$$

*Proof.* By assumption, it is obvious that the vectors $\hat{P}_i v_1, \ldots, \hat{P}_i v_p$ constitute a basis of the subspace spanned by the eigenvectors $\varphi_j$ associated with $\lambda_j, j = i, \ldots, i+p-1$, from which it follows immediately that (7) holds. □

**Theorem 2.** *Assume that $A$ is diagonalizable, and the vectors $\hat{P}_i v_j$, $j = 1, \ldots, p$ are linearly independent. Let $\hat{x}_i$ be the vector defined by Lemma 1, $N_i = \{1, 2, \ldots, N\} \setminus \{i, \ldots, i+p-1\}$ and*

$$\hat{x}_i = \sum_{j=1}^{N} \alpha_j \varphi_j, \ \xi_i = \sum_{j=1}^{i-1} \mid \alpha_j \mid + \sum_{j=i+p}^{N} \mid \alpha_j \mid .$$

*Then*

$$\|(I - \pi_m)\varphi_i\| \leq \xi_i \epsilon_i^{(m)}, \tag{8}$$

*where*

$$\epsilon_i^{(m)} = \min_{p \in Q_{m-1}, p(\lambda_i)=1} \max_{j \in N_i} \mid p(\lambda_j) \mid . \tag{9}$$

*Proof.* In terms of Lemma 1, since $\hat{P}_i \hat{x}_i = \varphi_i$, we have from $\hat{x}_i = \sum_{j=1}^{N} \alpha_j \varphi_j$

$$\alpha_i = 1, \alpha_j = 0, j = i+1, \ldots, i+p-1.$$

Hence

$$\hat{x}_i = \sum_{j=1}^{i-1} \alpha_j \varphi_j + \varphi_i + \sum_{j=i+p}^{N} \alpha_j \varphi_j.$$

Let us consider an element $u \in \mathcal{K}_m(V_1, A)$ of form $u = q(A)\hat{x}_i$, where $q \in Q_{m-1}$. Then

$$u = \sum_{j=1}^{i-1} q(\lambda_j)\alpha_j \varphi_j + q(\lambda_i)\varphi_i + \sum_{j=i+p}^{N} q(\lambda_j)\alpha_j \varphi_j,$$

from which it follows that

$$\frac{u}{q(\lambda_i)} - \varphi_i = \frac{1}{q(\lambda_i)} \left( \sum_{j=1}^{i-1} q(\lambda_j)\alpha_j \varphi_j + \sum_{j=i+p}^{N} q(\lambda_j)\alpha_j \varphi_j \right).$$

Because

$$\frac{u}{q(\lambda_i)} \in \mathcal{K}_m(V_1, A)$$

and $q \in Q_{m-1}$ is arbitrary, it follows from definition that

$$\begin{aligned} \|(I - \pi_m)\varphi_i\| &= \min_{u \in \mathcal{K}_m(V_1, A)} \|u - \varphi_i\| \\ &\leq \frac{1}{q(\lambda_i)} \left( \sum_{j=1}^{i-1} \mid q(\lambda_j)\alpha_j \mid + \sum_{j=i+p}^{N} \mid q(\lambda_j)\alpha_j \mid \right) \\ &= \sum_{j=1}^{i-1} \mid q(\lambda_j)/q(\lambda_i) \mid\mid \alpha_j \mid + \sum_{j=i+p}^{N} \mid q(\lambda_j)/q(\lambda_i) \mid\mid \alpha_j \mid \\ &\leq \xi_i \max_{j \in N_i} \mid q(\lambda_j)/q(\lambda_i) \mid . \end{aligned}$$

Now let $p(x) = q(x)/q(\lambda_i)$. Then $p(\lambda_i) = 1$. Therefore, we obtain

$$\begin{aligned}\|(I - \pi_m)\varphi_i\| &\leq \xi_i \min_{p \in Q_{m-1}, p(\lambda_i)=1} \max_{j \in N_i} \mid p(\lambda_j) \mid \\ &= \xi_i \epsilon_i^{(m)}. \quad \square\end{aligned}$$

It thus follows from Theorem 2 that we only need to estimate $\epsilon_i^{(m)}$ once $\xi_i$ is not too large.

*Remark 1.* If $V_1$ is taken as a vector and $A$ has only simple eigenvalues, this theorem reduces to Proposition 2.1 in [39].

*Remark 2.* $\xi_i$ essentially reflects ill-conditioning of the eigenvector matrix of $A$ and the nonnormality of $A$. In general, it can be shown that i) the larger $\xi_i$ is, the more ill conditioned the eigenvalue problem of $A$ is and the more highly nonnormal $A$ is; ii) the smaller $\xi_i$ is, the better conditioned the eigenvalue problem of $A$ may be and the more normal $A$ may be. Therefore, for ill-conditioned eigenproblems and highly nonnormal $A$, the block versions may converge slowly if they do.

*Remark 3.* If $\lambda_i$ has multiplicity $d_i \leq p$, the block size, say

$$\lambda_i = \lambda_{i+1} = \cdots = \lambda_{i+d_i-1},$$

then from Lemma 1 and Theorem 2 there exist the vectors $\hat{x}_j$ such that $\|(I - \pi_m)\varphi_j\| \leq \xi_j \epsilon_j^{(m)}$ for $j = i, \ldots, i + d_i - 1$. Therefore, if $\epsilon_j^{(m)}$ and thus $\|(I - \pi_m)\varphi_j\|$ *approach zero* for $j = i, \ldots, i + d_i - 1$ as $m$ increases, then in terms of Theorem 1 there exists one pair $\lambda^{(m)}, \varphi^{(m)}$ to approximate $\lambda_j, \varphi_j$ for each $j$, $i \leq j \leq i + d_i - 1$. This means that if $d_i \leq p$, then there are $d_i$ Ritz values $\lambda^{(m)}$ to approximate the $d_i$ multiple eigenvalue $\lambda_i$ and the corresponding $d_i$ Ritz vectors $\varphi^{(m)}$ to approximate the $d_i$ independent eigenvectors $\varphi_j$, $j = i, \ldots, i + d_i - 1$ associated with the $d_i$ multiple eigenvalue $\lambda_i$. Note that $\varphi_j$, $j = i, \ldots, i + d_i - 1$ span the eigenspace associated with the $d_i$ multiple eigenvalue $\lambda_i$. Then $d_i$ Ritz vectors $\varphi^{(m)}$ span an approximate eigenspace associated with such $d_i$ multiple eigenvalue $\lambda_i$. Thus, generalized block Lanczos methods may determine a multiplicity up to the block size $p$ and solve multiple eigenproblems provided that the block size $p$ is no less than all multiplicities of the wanted eigenvalues.

Now we analyze how $\epsilon_i^{(m)}$ approaches zero as $m$ increases.

*Case 1.* The eigenvalues of $A$ are real.

**Theorem 3.** *Assume that $A$ is diagonalizable, the eigenvalues are real and are labeled in the decreasing order: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$, every multiplicity*

*of them does not exceed $p$ and $\lambda_{i-1} \neq \lambda_i$. Let*

$$\gamma_i = 1 + 2\frac{\lambda_i - \lambda_{i+p}}{\lambda_{i+p} - \lambda_N}, \kappa_i = \begin{cases} \prod_{\lambda_k \in \sigma_i} \frac{\lambda_k - \lambda_N}{\lambda_k - \lambda_i} & \text{if } i \neq 1, \\ 1 & \text{if } i = 1, \end{cases}$$

*where $\sigma_i$ is the set of distinct ones among the first $i-1$ eigenvalues, and $\mid \sigma_i \mid$ the cardinality of it. Then*

$$\epsilon_i^{(m)} \leq \kappa_i / T_{m-|\sigma_i|-1}(\gamma_i), \tag{10}$$

*where $T_k(x)$ is the first kind Chebyshev polynomial of degree k.*

*Proof.* Let $\tilde{Q}_{m-1}$ be the set of polynomials of the form $q(x) = l_i(x)r(x)$, where

$$l_i(x) = \begin{cases} \prod_{\lambda_k \in \sigma_i} \frac{x - \lambda_k}{\lambda_i - \lambda_k} & \text{if } i \neq 1, \\ 1 & \text{if } i = 1, \end{cases}$$

$r(x)$ is a polynomial of degree not exceeding $m- \mid \sigma_i \mid -1$ and satisfies $r(\lambda_i) = 1$. Obviously $q \in Q_{m-1}$ and $q(\lambda_i) = 1$, so we get from (9)

$$\epsilon_i^{(m)} \leq \min_{q \in \tilde{Q}_{m-1}} \max_{j \in N_i} \mid q(\lambda_j) \mid . \tag{11}$$

Since $q(\lambda_k) = 0$ for $\lambda_k \in \sigma_i$, the maximum in (11) must be achieved for certain $\lambda_j$, where $j \geq i + p$. Hence

$$\begin{aligned} \epsilon_i^m &\leq \min_{r \in Q_{m-|\sigma_i|-1}} \max_{i+p \leq j \leq N} \left| \prod_{\lambda_k \in \sigma_i} \frac{\lambda_j - \lambda_k}{\lambda_i - \lambda_k} \right| \mid r(\lambda_j) \mid \\ &\leq \kappa_i \min_{r \in Q_{m-|\sigma_i|-1}} \max_{i+p \leq j \leq N} \mid r(\lambda_j) \mid \\ &\leq \kappa_i \min_{r \in Q_{m-|\sigma_i|-1}} \max_{\lambda_N \leq x \leq \lambda_{i+p}} \mid r(x) \mid . \end{aligned} \tag{12}$$

It is well known [10] that the polynomial achieving the minimum in (12) is

$$r_o(x) = T_{m-|\sigma_i|-1}\left(1 + 2\frac{x - \lambda_{i+p}}{\lambda_{i+p} - \lambda_N}\right) / T_{m-|\sigma_i|-1}\left(1 + 2\frac{\lambda_i - \lambda_{i+p}}{\lambda_{i+p} - \lambda_N}\right),$$

from which it follows immediately that (10) holds. □

*Remark 1.* It is seen that the key of proof of Theorem 3 is the construction of the polynomial $l_i(x)$. Note that our construction is neither the same as that of [39] nor that of [42] for the symmetric block Lanczos method.

*Remark 2.* Theorem 3 is a generalization of Theorem 2.2 in [39]. In fact, when $V_1$ reduces to a vector, i.e., $p = 1$, and the matrix $A$ only has real simple eigenvalues, this theorem coincides with it.

*Remark 3.* Combining Theorem 3 with (6) and (8), (10) shows that $\|(I - \pi_m)\varphi_i\|$ may converge to zero as rapidly as $\kappa_i / T_{m-|\sigma_i|-1}(\gamma_i)$. As $m$ increases, we have

$$T_{m-|\sigma_i|-1}(\gamma_i) \approx \frac{1}{2}\left(\gamma_i + \sqrt{\gamma_i^2 - 1}\right)^{m-|\sigma_i|-1},$$

where $\gamma_i > 1$ depends on the gaps $(\lambda_i - \lambda_{i+p})$ and $(\lambda_{i+p} - \lambda_N)$. It follows from this that the distances $\|(I-\pi_m)\varphi_i\|$, in general, first converge to zero for the eigenvectors associated with right outer part eigenvalues. Furthermore, for the same $m$, the bigger the gap $\lambda_i - \lambda_{i+p}$ is, i.e., the bigger the block size $p$ is, the more rapidly they converge. This is behavior analogous to that of the symmetric block Lanczos method [16, 42].

*Remark 4.* Compared with Theorem 2.2 in [39], it can be seen that generalized block Lanczos methods may work much better than single vector versions, i.e., $p = 1$, when some of the wanted eigenvalues are clustered. The argument is that when $p = 1$, $\gamma_i$ is very near one if $\lambda_i$ and $\lambda_{i+1}$ are close, such that the right-hand side of (10) converges to zero very slowly and its left-hand side may approach zero very slowly, while the block versions can overcome this drawback once a choice of $p$ makes the gap between $\lambda_i$ and $\lambda_{i+p}$ big. Therefore, it can be expected that the block methods are efficient when $p$ is no less than the number of the custered eigenvalues to be computed.

*Case 2.* The eigenvalues of $A$ are complex.

Assume that the eigenvalues of $A$ are labeled in the decreasing order of their real parts:

$$\mathrm{Re}(\lambda_1) \geq \mathrm{Re}(\lambda_2) \geq \cdots \geq \mathrm{Re}(\lambda_N) \tag{13}$$

(If $A$ has eigenvalues with the same real parts but different imaginary parts, we first label the ones with larger imaginary parts; for a complex pair of eigenvalues, we first label the one with positive real part). For simplicity, assume $\lambda_{i+p-1} \neq \lambda_{i+p}$ and $\lambda_{i+p-1} \neq \bar{\lambda}_{i+p}$, the complex conjugate of $\lambda_{i+p}$. Then from the conjugation of eigenvalues and rule of labeling order (13), we can use an ellipse $E_i(d, c, a)$ with the real center $d$, foci $d + c, d - c$, major semiaxis $a$ and the symmetry with respect to the real axis to contain the set $\mathcal{R}_i = \{\lambda_{i+p}, \ldots, \lambda_N\}$ but $\{\lambda_1, \ldots, \lambda_{i+p-1}\}$. Note that the main axis of the ellipse must be either the real axis or parallel to the imaginary axis.

**Theorem 4.** *Assume that A is diagonalizable, and $E_i(d, c, a)$ as described above. Let*

$$\tilde{\kappa}_i = \begin{cases} \max_{i+p \leq j \leq N} \prod_{\lambda_k \in \sigma_i} \left| \frac{\lambda_k - \lambda_j}{\lambda_k - \lambda_i} \right| & \text{if } i \neq 1, \\ 1 & \text{if } i = 1, \end{cases}$$

*where $\sigma_i$ and $\mid \sigma_i \mid$ are in Theorem 3. Then*

$$\epsilon_i^{(m)} \leq \tilde{\kappa}_i T_{m-|\sigma_i|-1}\left(\frac{a}{c}\right) / \left| T_{m-|\sigma_i|-1}\left(\frac{\lambda_i - d}{c}\right)\right|, \tag{14}$$

*where $T_k(z)$ is the first kind complex Chebyshev polynomial of degree $k$.*

*Proof.* Constructing $\tilde{Q}_{m-1}$ analogous to the proof of Theorem 3, we get

$$\epsilon_i^{(m)} \leq \min_{q \in \tilde{Q}_{m-1}} \max_{j \in N_i} \mid q(\lambda_j) \mid, \tag{15}$$

which is just (11). Based on the same arguments as the proof of Theorem 3, we obtain

$$\begin{aligned} \epsilon_i^{(m)} &\leq \min_{r \in Q_{m-|\sigma_i|-1}} \max_{i+p \leq j \leq N} \left| \prod_{\lambda_k \in \sigma_i} \frac{\lambda_k - \lambda_j}{\lambda_k - \lambda_i} \right| \mid r(\lambda_j) \mid \\ &\leq \tilde{\kappa}_i \min_{r \in Q_{m-|\sigma_i|-1}} \max_{i+p \leq j \leq N} \mid r(\lambda_j) \mid \\ &\leq \tilde{\kappa}_i \min_{r \in Q_{m-|\sigma_i|-1}} \max_{z \in E_i(d,c,a)} \mid r(z) \mid . \end{aligned} \tag{16}$$

It is well known [14] that

$$r_o(z) = T_{m-|\sigma_i|-1}\left(\frac{z-d}{c}\right) / T_{m-|\sigma_i|-1}\left(\frac{\lambda_i - d}{c}\right)$$

is often an optimal polynomial achieving the minimum in (16) though it is not always the case; if it is not optimal, it is still an approximate optimal one. So, taking the above polynomial, we can get (14) by using a property of Chebyshev polynomials in the complex plane, e.g., [32]. □

Without loss of generality, assume that the real part of $(\lambda_i - d)/c$ is non-negative. Then in terms of (14) and $\mathrm{Re}(\lambda_i) > \mathrm{Re}(\lambda_{i+p}) \geq \ldots \geq \mathrm{Re}(\lambda_N)$, obviously the right-hand side of (14) converges to zero as $m$ increases, and furthermore it is asymptotically equivalent to

$$\left| \frac{a/c + \sqrt{(a/c)^2 - 1}}{(\lambda_i - d)/c + \sqrt{((\lambda_i - d)/c)^2 - 1}} \right|^{m-|\sigma_i|-1} \to 0.$$

Thus, the asymptotic rate of convergence is

$$\tau_{1i} = \frac{\mid (\lambda_i - d) + \sqrt{(\lambda_i - d)^2 - c^2} \mid}{a + \sqrt{a^2 - c^2}} \text{ if } a \text{ and } c \text{ real}, \tag{17}$$

or

$$\tau_{2i} = \frac{\mid (\lambda_i - d) + \sqrt{(\lambda_i - d)^2 + \mid c \mid^2} \mid}{\mid a \mid + \sqrt{\mid a \mid^2 - \mid c \mid^2}} \text{ if } a \text{ and } c \text{ purely imaginary.} \tag{18}$$

If $c = 0$, the ellipse described above degenerates to a circle with the center $d$ and radius $a$. We have the following result.

**Corollary 1.** *Let $D_i(d, a)$ be a circle centered $d$, the radius $a$, and $D_i(d, a)$ contain the set $\mathcal{R}_i$ but $\{\lambda_1, \ldots, \lambda_{i+p-1}\}$, and $\tilde{\kappa}_i$ in Theorem 4. Then*

$$\epsilon_i^{(m)} \leq \tilde{\kappa}_i \left| \frac{a}{\lambda_i - d} \right|^{m-|\sigma_i|-1}, \tag{19}$$

*where $\sigma_i, \mid \sigma_i \mid$ are in Theorem 3.*

*Proof.* Construct $\tilde{Q}_{m-1}$ analogous to that in the proof of Theorem 3. Then we get

$$\epsilon_i^{(m)} \leq \tilde{\kappa}_i \min_{r \in Q_{m-|\sigma_i|-1}} \max_{z \in D_i(d,a)} \mid r(z) \mid .$$

The polynomial achieving the minimum in the above relation is

$$r_o(z) = \left( \frac{z - d}{\lambda_i - d} \right)^{m-|\sigma_i|-1},$$

according to [48]. Therefore, (19) holds. □

*Remark 1.* When $\lambda_i, i = 1, 2, \ldots, N$ become real, the ellipse degenerates to an interval $[\lambda_N, \lambda_{i+p}]$. In this case, we have $d = (\lambda_{i+p} + \lambda_N)/2$, $c = a = (\lambda_{i+p} - \lambda_N)/2$. Thus,

$$\epsilon_i^{(m)} \leq \tilde{\kappa}_i T_{m-|\sigma_i|-1}\left(\frac{a}{c}\right) / \left| T_{m-|\sigma_i|-1}\left(\frac{\lambda_i - d}{c}\right) \right| = \kappa_i / T_{m-|\sigma_i|-1}(\gamma_i).$$

So Theorem 4 will reduce to Theorem 3.

*Remark 2.* It can be seen from Theorem 4 that these distances usually converge to zero first for the eigenvectors corresponding to the eigenvalues with largest real parts.

### *3.2. A priori theoretical error bounds for Ritz elements*

We have established upper bounds for the distance $\|(I - \pi_m)\varphi_i\|$, which mean that there exists a vector, $\pi_m \varphi_i$, in $\mathcal{K}_m(V_1, A)$ to approximate $\varphi_i$ as $m$ increases. We now study the implications of these distances on behavior of Ritz elements, which are very crucial in understanding how generalized block Lanczos methods converge and in developing related block algorithms.

Before discussions, we want to emphasize that the results in this section hold for a general orthogonal projection method once we replace $\mathcal{K}_m(V_1, A)$ by a given projection subspace $E$, $\pi_m$ by the corresponding orthogonal

projector onto $E$, the sentence like "as $m$ increases" by the one like "as $E$ changes" and the phase like "uniformly bounded in $m$" by the one like "uniformly bounded in $E$". Meanwhile, we would like to point out that the error bounds in this section is only a priori and not computable but they can indeed be used to explain convergence behavior of orthogonal projection methods and to guide how to develop related algorithms.

Concerning error bounds for eigenvalues, for simplicity, assume $U_m$ to be orthonormal. Let us set $U_m^*\varphi_i = \tilde{y}_i^{(m)}$, and define $\mathcal{H}_m = U_m^* A U_m$, which is just the matrix representation of the restriction of $A$ to $\mathcal{K}_m(V_1, A)$. Noticing that $\|\tilde{y}_i^{(m)}\| = \|U_m \tilde{y}_i^{(m)}\| = \|\pi_m \varphi_i\|$, the first inequality of (5) then translates into:

$$\frac{\|(\mathcal{H}_m - \lambda_i I)\tilde{y}_i^{(m)}\|}{\|\tilde{y}_i^{(m)}\|} \leq \gamma_m \frac{\|(I - \pi_m)\varphi_i\|}{\|\pi_m \varphi_i\|}. \tag{20}$$

**Theorem 5.** *Let $S_m^{-1}\mathcal{H}_m S_m = J_m$ be the Jordan form of $\mathcal{H}_m$, $\kappa(S_m) = \|S_m^{-1}\|\|S_m\|$, and assume $\|(I-\pi_m)\varphi_i\|$ to be small and $\lambda_i^{(m)}$ the eigenvalue of $\mathcal{H}_m$ with the* index $l_i$. *Then, there exists a Ritz value $\lambda_i^{(m)}$ such that*

$$|\ \lambda_i^{(m)} - \lambda_i\ | \leq 2(\gamma_m \kappa(S_m))^{1/l_i} \left( \frac{\|(I - \pi_m)\varphi_i\|}{\|\pi_m \varphi_i\|} \right)^{1/l_i}. \tag{21}$$

The proof is analogous to that of Theorem 3.7 in [21], which is based on [5] or [7].

Theorem 5 can be simplified when some $l_i = 1$, as described below.

**Theorem 6.** *Assume that $l_i = 1$ and the associated $\|(I-\pi_m)\varphi_i\|$ are small. Let $P_i^{(m)}$ be the spectral projectors associated with $\lambda_i^{(m)}$. Then*

$$|\ \lambda_i^{(m)} - \lambda_i\ | \leq \|P_i^{(m)}\| \gamma_m \frac{\|(I - \pi_m)\varphi_i\|}{\|\pi_m \varphi_i\|} + o\left( \frac{\|(I - \pi_m)\varphi_i\|}{\|\pi_m \varphi_i\|} \right). \tag{22}$$

*Proof*. In terms of the first inequality of (5), it can be easily shown from [49, p. 69] and [15, p. 344] that (22) holds. □

*Remark 1*. (22) indicates that the *sufficient* conditions for $\lambda_i^{(m)} \to \lambda_i$ are that $\|P_i^{(m)}\|$ is uniformly bounded in $m$ and $\|(I - \pi_m)\varphi_i\|$ tends to zero as $m$ increases.

*Remark 2*. In Theorems 5–6, the assumption that $\|(I-\pi_m)\varphi_i\|$ is small is not strong. In fact, it has been proved in [25, 26, 28] that a *necessary* condition for convergence of $\varphi_i^{(m)}$ and that of orthogonal projection methods is that $\|(I - \pi_m)\varphi_i\|$ tends to zero as $m$ increases.

We now give a priori theoretical error bounds for eigenvectors.

**Theorem 7.** *Assume that $A_m$ (i.e., $\mathcal{H}_m$) is diagonalizable and has $s$ distinct eigenvalues $\lambda_j^{(m)}$ in $\mathcal{K}_m(V_1, A)$. Let $P_j^{(m)}$ denote the spectral projectors associated with $\lambda_j^{(m)}$, $d_{i,m} = \min_{\lambda_j^{(m)} \neq \lambda_i^{(m)}} \mid \lambda_i - \lambda_j^{(m)} \mid$, and define the matrix*

$$\Phi_i^{(m)} = (P_1^{(m)}\varphi_i, \ldots, P_{i-1}^{(m)}\varphi_i, P_{i+1}^{(m)}\varphi_i, \ldots, P_s^{(m)}\varphi_i).$$

*Then*

$$\|(I - P_i^{(m)})\varphi_i\| \leq \left(1 + \frac{\inf_D \text{ diag. } \kappa(\Phi_i^{(m)} D)(1 + \|P_i^{(m)}\|)\gamma_m}{d_{i,m}}\right) \times \|(I - \pi_m)\varphi_i\|. \tag{23}$$

*Let $P_i^{(m)}\varphi_i / \|P_i^{(m)}\varphi_i\| = \varphi_i^{(m)}$. Then*

$$\sin\theta(\varphi_i, \varphi_i^{(m)}) \leq \left(1 + \frac{\inf_D \text{ diag. } \kappa(\Phi_i^{(m)} D)(1 + \|P_i^{(m)}\|)\gamma_m}{d_{i,m}}\right) \times \sin\theta(\varphi_i, \mathcal{K}_m(V_1, A)). \tag{24}$$

The proof is similar to that of Theorem 3.10 in [21]; also see [20].

According to the results of Sect. 3.1 and (23),(24), it is clear that the *sufficient* conditions for $\varphi_i^{(m)} \to \varphi_i$ are that both $\inf_D \text{ diag. } \kappa(\Phi_i^{(m)} D)$ and $\|P_i^{(m)}\|$ are uniformly bounded in $m$ and $\|(I - \pi_m)\varphi\|$ tends to zero as $m$ increases. On the one hand, if $\lambda_i^{(m)}$ is ill conditioned, then $\varphi_i^{(m)}$ may approximate $\varphi_i$ quite slowly even though $\|(I - \pi_m)\varphi_i\|$, i.e., $\sin\theta(\varphi_i, \mathcal{K}_m(V_1, A))$, tends to zero rapidly as $m$ increases. On the other hand, if $\inf_D \text{ diag. } \kappa(\Phi_i^{(m)} D)$ is very large or unbounded, it is known [49] that at least one other $\|P_j^{(m)}\|, j \neq i$ is very large or unbounded, that is, at least one other Ritz value $\lambda_j^{(m)} (\neq \lambda_i^{(m)})$ is very or arbitrarily ill conditioned, and vice versa. So the *sufficient* conditions for convergence of $\varphi_i^{(m)}$ are that *all* the Ritz values are not arbitrarily ill conditioned, while those for convergence of the corresponding $\lambda_i^{(m)}$ only require itself not to be so. However, for a general unsymmetric $A$, we cannot guarantee theoretically that $\inf_D \text{ diag. } \kappa(\Phi_i^{(m)} D)$ is uniformly bounded in $m$, even if assuming that the eigenproblem of $A$ is well conditioned and $\|P_i^{(m)}\|$ is uniformly bounded in $m$. This factor may be getting larger or smaller as $m$ increases. Therefore, values of the right-hand sides of (23) and (24) may oscillate and even may not

converge to zero even if $\|(I-\pi_m)\varphi_i\|$ does. Thus $\varphi_i^{(m)}$ may converge irregularly and even *may not* converge to $\varphi_i$ even though $\|(I-\pi_m)\varphi_i\|$ tends to zero and $\lambda_i^{(m)}$ converges to $\lambda_i$. Therefore, Ritz vectors may be much harder to converge than Ritz values, so that the block methods may not converge if we test convergence in terms of the residual norms $\|(A-\lambda_i^{(m)}I)\varphi_i^{(m)}\|$. However, this potential nonconvergence cannot happen to a symmetric $A$ because then both $\inf_D$ diag. $\kappa(\Phi_i^{(m)}D) = 1$ and $\|P_i^{(m)}\| = 1$ for a symmetric $A_m$. We refer to [25,26,28] for more details, and in [26,28] how a possible ill-conditioning of the eigenvalue problem of $\mathcal{H}_m$ and thus that of $A_m$ may occur is interpreted from a viewpoint of matrix theory for an unsymmetric $A$. A concrete example is constructed in [21] to illustrate that the eigenvalue problem of $\mathcal{H}_m$ and thus that of $A_m$ is arbitrarily ill conditioned even though that of $A$ is well conditioned.

In summary, we can draw a few conclusions: First, generalized block Lanczos methods will favor the right outer part eigenvalues and the associated eigenvectors of $A$ usually; second, they may work better than single vector versions when some of the wanted eigenvalues are clustered; third, they can solve multiple eigenproblems up to multiplicities equal to the block size $p$; fourth, Ritz vectors cannot be guaranteed to converge theoretically for a general unsymmetric matrix $A$ even if Ritz values do.

Finally, we point out that the convergence theory in this section can apply as well to the left outer part eigenvalues and the associated eigenvectors of $A$ provided that $A$ is replaced by $-A$. This shows that generalized block Lanczos methods may be suitable for computing both right and left outer part eigenvalues and the associated eigenvectors of $A$.

## 4. The block Arnoldi algorithms

First we present the following well known algorithm, which generates an orthonormal sequence $\{V_i\}_1^m$ in exact arithmetic.

**Algorithm 4.1.** The block Arnoldi process

1. Start: Choose a starting $N \times p$ orthonormal matrix $V_1$ and the steps $m$.
2. For $i = 1, \ldots, m$ do
   $Z = AV_i$;
   For $j = 1, \ldots, i$ do
   $H_{ji} = V_j^* Z$,
   $Z = Z - V_j H_{ji}$.
   Endfor
   $Z = V_{i+1}R_{i+1}$, the $QR$ decomposition [12].
   Endfor

If $p = 1$, the stability of the above algorithm has been analyzed in [1, 9, 13]. It is shown that the algorithm is absolutely stable in finite precision arithmetic for a general matrix $A$ using the Householder transformation, the Givens rotation and the Gram–Schmidt orthogonalization with iterative refinement, while it is only conditionally stable using the modified Gram–Schmidt orthogonalization. For $p > 1$, it seems enough to ensure the stability with one reorthogonalization at step 2 in practice.

The block Arnoldi process can be written in matrix form

$$AU_m = U_m\mathcal{H}_m + V_{m+1}R_{m+1}E_m^*, \tag{25}$$

where

$$U_m = (V_1, V_2, \ldots, V_m),$$

$$\mathcal{H}_m = \begin{pmatrix} H_{11} & H_{12} & H_{13} & \cdots & H_{1m} \\ R_2 & H_{22} & H_{23} & \cdots & H_{2m} \\ & R_3 & H_{33} & \cdots & \vdots \\ & & \ddots & \ddots & H_{m-1m} \\ & & & R_m & H_{mm} \end{pmatrix}$$

and $E_m$ is an $mp \times p$ zero matrix except for the last $p$ rows being an $p \times p$ identity matrix.

Since $V_{m+1}$ is orthogonal to $V_1, V_2, \ldots, V_m$, we have that

$$U_m^* A U_m = \mathcal{H}_m. \tag{26}$$

Obviously, in terms of Sect. 2, the block upper Hessenberg matrix $\mathcal{H}_m$ is the matrix representation of the restriction of $A$ to $\mathcal{K}_m(V_1, A)$ in the basis $\{V_i\}_1^m$, and thus Ritz values $\lambda^{(m)}$ of $A$ in $\mathcal{K}_m(V_1, A)$ are just the eigenvalues of $\mathcal{H}_m$ and Ritz vectors $\varphi^{(m)} = U_m y^{(m)}$, where $y^{(m)}$ are eigenvectors of $\mathcal{H}_m$ associated with eigenvalues $\lambda^{(m)}$.

In order to compute some eigenpairs $\lambda, \varphi$, the block Arnoldi process has to be combined with other methods, e.g., the $QR$ method, and the resulting method is called the block Arnoldi method and the corresponding algorithm is called the basic block Arnoldi algorithm.

It can be proved that if $\dim(\mathcal{K}_m(V_1, A)) = mp$ then all subdiagonal block entries $R_i, i = 2, \ldots, m$ of block upper Hessenberg matrix $\mathcal{H}_m$, which are upper triangular, are nonsingular and $\{V_i\}_1^m$ exactly constitutes an orthonormal basis of $\mathcal{K}_m(V_1, A)$; see [20]. However, the case $\dim(\mathcal{K}_m(V_1, A)) < mp$ can occur both in theory and in numerical computation. In such a case, it is obvious that there is at least one $Z$ at step 2 of Algorithm 4.1 to be column rank deficient such that the corresponding $R_{i+1}$ is singular. In order to solve this problem, just as in the symmetric case [16], we choose those columns of $V_{i+1}$ corresponding to zero diagonal elements of $R_{i+1}$ randomly, such that

they are orthogonal to all previous $V_j, j = 1, 2, \ldots, i$. By adopting such a strategy, Algorithm 4.1 will not break down unless $Z = 0$, and furthermore we can prove in this case that $\mathcal{K}_m(V_1, A)$ is a subspace of that spanned by the columns of $U_m$ [20].

As in the symmetric block Lanczos method [16], one can easily compute a residual norm by using the formula

$$\|(A - \lambda_i^{(m)} I)\varphi_i^{(m)}\| = \|R_{m+1} w_i^{(m)}\| \tag{27}$$

without explicitly forming $\varphi_i^{(m)}$, where $w_i^{(m)}$ is a vector consisting of the last $p$ elements of $y_i^{(m)}$, and $y_i^{(m)}$ is the eigenvector of $\mathcal{H}_m$ associated with the eigenvalue $\lambda_i^{(m)}$. This formula is very useful as a stopping criterion, and it follows immediately from (25) by postmultiplying two hand sides by $y_i^{(m)}$ and then subtracting them.

*Remark*. When $A$ is highly nonnormal, $\|A\|$ is large. In [9], the nonnormality of $A$ is taken into account to design a reliable and realistic stopping criterion for iterative eigensolvers. For the block Arnoldi method, it amounts to requiring that the method stops when

$$\frac{\|(A - \lambda_i^{(m)} I)\varphi_i^{(m)}\|}{\|A\|_{\mathrm{F}}} = \frac{\|R_{m+1} w_i^{(m)}\|}{\|A\|_{\mathrm{F}}} \leq c\epsilon_M, \tag{28}$$

where $\|A\|_{\mathrm{F}}$ is the Frobenius norm of $A$ and $c$ is some moderate constant, say 10, and $\epsilon_M$ is the machine precision.

In practice, due to the limitation of primary memory and amount of computations, we often have to restart the basic block Arnoldi algorithm.

**Algorithm 4.2.** The iterative block Arnoldi algorithm

1. Start: Choose $p, m$ and $V_1$, where $1 \leq p \leq N/2$, $2 \leq m \leq N/p$ and $V_1$ is an $N \times p$ orthonormal matrix.
2. Perform Algorithm 4.1 to construct $\mathcal{H}_m$ and $U_m$.
3. Compute the eigenvalues $\lambda_i^{(m)}$, $i = 1, 2, \ldots, mp$ of $\mathcal{H}_m$, and select certain $\lambda_1^{(m)}, \ldots, \lambda_r^{(m)}$ as approximations to the wanted $\lambda_1, \ldots, \lambda_r$, where $r \leq mp$.
4. Decide if $\lambda_i^{(m)}, \varphi_i^{(m)}$, $i = 1, 2, \ldots, r$ satisfy the desired accuracy using (27). If yes, then compute $\varphi_i^{(m)} = U_m y_i^{(m)}$ and stop; otherwise, return to step 2 with an updated $V_1$.

## 5. Implementations

In this section we discuss how to restart in Algorithm 4.2 and how to solve multiple eigenproblems.

### *5.1. Choices of initial $V_1$*

When performing the first iteration of Algorithm 4.2, we choose an $N \times p$ column orthonormal matrix $V_1$ randomly.

Much more attention has to be paid to step 4 of Algorithm 4.2. To this end, let us introduce the unitary matrix

$$U = \frac{1}{2}\begin{pmatrix} 1+\mathrm{i} & 1-\mathrm{i} \\ 1-\mathrm{i} & 1+\mathrm{i} \end{pmatrix},$$

where "i" is the imaginary unit.

Let $\lambda_j^{(m)} = \bar{\lambda}_{j+1}^{(m)}$ be a complex pair of eigenvalues of $\mathcal{H}_m$. Then it can be easily shown that the following equality holds:

$$(\varphi_j^{(m)}, \bar{\varphi}_j^{(m)})U^* = (\mathrm{Re}(\varphi_j^{(m)}) + \mathrm{Im}(\varphi_j^{(m)}), \mathrm{Re}(\varphi_j^{(m)}) - \mathrm{Im}(\varphi_j^{(m)})).$$

The right-hand side of the above relation only contains real matrices. Here $\bar{\varphi}_j^{(m)}$ is the complex conjugate of $\varphi_j^{(m)}$.

With the argument stated above, we may update $V_1$ as follows.

*Case 1.* $p = r$.

For real $\varphi_j^{(m)}$, $1 \le j \le r$, we take it as one column of $Z_1^{\mathrm{new}}$; for complex $\varphi_j^{(m)}, 1 \le j \le r$, we take $\mathrm{Re}(\varphi_j^{(m)}) + \mathrm{Im}(\varphi_j^{(m)})$ as one column of $Z_1^{\mathrm{new}}$; if the number of columns of $Z_1^{\mathrm{new}}$ is still less than $p$, then also take $\mathrm{Re}(\varphi_j^{(m)}) - \mathrm{Im}(\varphi_j^{(m)})$ as one column of $Z_1^{\mathrm{new}}$, and proceed analogously until $Z_1^{\mathrm{new}}$ has $p$ columns.

*Case 2.* $p > r$.

The first $r$ columns of $Z_1^{\mathrm{new}}$ may be chosen in the same way as that in case 1, and the rest $(p - r)$ columns are chosen randomly.

*Case 3.* $p < r$.

It is important to note that $A$ is real, so $\mathrm{Re}(\varphi_j^{(m)})$ is a combination of the complex pair $\varphi_j^{(m)}$ and $\bar{\varphi}_j^{(m)}$. We thus combine the real parts of $\varphi_j^{(m)}, j = 1, 2, \ldots, r$ randomly, and can in practice obtain $p$ linearly independent vectors which constitute $Z_1^{\mathrm{new}}$. Numerical experiments show that the iterative block Arnoldi algorithm with this choice works quite well.

If $p = 1$, for simplicity we still adopt Saad's strategy [39] though the implicitly restarted strategy of Sorensen [47] has appeared to be more promising.

Modifying $V_1$ in such a way has the advantage that Algorithm 4.1 avoids complex arithmetic. Note that it is easy to design Algorithm 4.2 so as to avoid forming $\varphi_j^{(m)}$ explicitly before convergence occurs.

### *5.2. Multiple eigenvalues*

Although, theoretically speaking, the block methods can solve multiple eigenproblems, neither we know the multiplicity of an eigenvalue of $A$ in advance, nor we can assert that the multiplicity of an eigenvalue is exactly equal to the block size $p$ even though we have computed a $p$ (numerically) multiple eigenvalue. Therefore, more cares should be taken on how to determine multiplicities of the wanted eigenvalues of $A$. In practice, given $p$, the block size, if we compute a $p$ numerically multiple eigenvalue, then this eigenvalue is at least $p$ multiple, based on the previous theoretical analysis. At this time, in comparisons with the symmetric block Lanczos method [16, 34], it is quite difficult to determine how big a multiplicity is. Theoretically, a simulation of the idea used in [16] is to use the iterative block Arnoldi algorithm, starting with a new $V_1$ orthogonal to all the *left* eigenvectors of $A$ associated with these converged Ritz vectors $\varphi^{(m)}$ corresponding to such numerically multiple eigenvalue, such that all the new converged eigenvectors $\varphi^{(m)}$ associated with this eigenvalue, if still appear, have no components in directions of the previous converged $\varphi^{(m)}$. Proceed this way until this eigenvalue does not appear any more in some repeated use of Algorithm 4.2, starting with a new $V_1$ orthogonal to all the left eigenvectors corresponding to all the converged $\varphi^{(m)}$ computed previously associated with the eigenvalue. At this moment, we have determined an approximate eigenspace associated with the multiple eigenvalue and thus found the multiplicity, which is the number of all the converged $\varphi^{(m)}$ associated with the eigenvalue. Unfortunately, an easy analysis similar to that of Arnoldi's method [35] exposes that the block Arnoldi method is ineffective for computing the left eigenvectors of $A$.

To this end, the following approach can be used thanks to a technique adapted from [20, 22]. The idea can be simply stated as follows:

Starting with $p_1, V_1^{(1)}$, if we have found an eigenvalue with the multiplicity $l_1$, then we apply Algorithm 4.2 with the block size $p_2$ and a new initial $V_1^{(2)}$, which is chosen randomly. Now we can compute an eigenvalue with the multiplicity $l_2$, which is numerically equal to the one computed with $p_1, V_1^{(1)}$. We then determine the rank of the matrix consisting of these $l_1 + l_2$ converged Ritz vectors with the numerically multiple converged Ritz values. Note here that when the eigenproblem of $A$ is not too ill conditioned, if some singular values of this matrix are of the same order as the maximum of residual norms of these $l_1 + l_2$ converged eigenpairs, then we consider them to be zero numerically. If the rank of the matrix is less than $l_1 + l_2$, then the multiplicity of this eigenvalue is just the rank of such a matrix. Otherwise, we repeat Algorithm 4.2 with $p_3, V_1^{(3)}$ and so on until the rank of the matrix consisting of these $l_1 + l_2 + \cdots + l_k$ converged Ritz

vectors with the numerically multiple converged Ritz values, starting with $p_1, V_1^{(1)}, p_2, V_1^{(2)}, \ldots, p_k, V_1^{(k)}$ respectively, is less than $l_1 + l_2 + \cdots + l_k$. Then the multiplicity of the eigenvalue has been determined, and equals the rank of this matrix. For this idea and its theoretical background, we refer to [20,22].

## 6. Numerical experiments

In this section we report several numerical examples to illustrate our theory and algorithms, all of them were performed on a DEC RISC workstation using MATLAB4.1 with the double precision $eps \approx 2.22 \times 10^{-16}$. The efficiency of Algorithm 4.2 may be measured in terms of the number of matrix by vector multiplications, indicated by $m.v$.

*Example 1.* This test problem is a common one [3,39]. Consider the following convection diffusion differential operator:

$$-\triangle u(x,y) + \rho u_x(x,y) = \lambda u(x,y)$$

on a square region $[0,1] \times [0,1]$ with the boundary condition $u(x,y) = 0$.

Taking $\rho = 1$ and discretizing with centered differences yield a block tridiagonal matrix $A(n) = \mathrm{tri}(-I, B(n), -I)$, where $B(n) = \mathrm{tri}(b, 4, a)$ is a tridiagonal matrix with $a = -1+1/2(n+1)$ and $b = -1-1/2(n+1)$, and $n$ is chosen the number of interior mesh points on each side of the square. $A(n)$ is of order $N = n^2$. The eigenvalues $\lambda_2$ and $\lambda_3$ are very clustered and they get closer and closer as $n$ increases.

Algorithm 4.2 was run on the $576 \times 576$ matrix $A(24)$ obtained by taking $n = 24$ mesh points to each side. We want to compute the four eigenvalues with the largest real parts and require that the algorithm stop as soon as all actual residual norms of the approximating eigenpairs are below $tol = 10^{-7}$. For different $p$, all initial block vectors $V_1$ were chosen randomly in a normal distribution.

Table 1 shows the results obtained for different $m$ and $p$, where $it$ denotes the number of iterations.

Note that the wanted eigenvalues $\lambda_2$ and $\lambda_3$ are very clustered. So, in terms of the previous theoretical analysis, the iterative Arnoldi algorithm, i.e., $p = 1$, may be inefficient for computing $\lambda_2, \lambda_3$. This was confirmed by Table 1. We saw that the iterative Arnoldi algorithm did not converge for $m \leq 60$ after enough iterations (restarts) and it only succeeded when the steps $m$ were quite large, e.g., $m = 70, 80$, while the block versions worked very well even for $mp \leq 60$ and $p > 1$. In fact, in terms of the theoretical analysis in Sect. 3, when the block size $p$ is such that $\lambda_i - \lambda_{i+p}$ is not small, Algorithm 4.2 may work well usually. Therefore, in order to compute

**Table 1.** Example 1, $\lambda_1 \approx 7.96806192, \lambda_2 \approx 7.92100825, \lambda_3 \approx 7.92099884, \lambda_4 \approx 7.87394517$, *n.c* denotes no convergence

| $m$ | $p$ | $it$ | $m.v$ | Max. Res. Norms |
|---|---|---|---|---|
| 50 | 1 | 2155 | 107750 | $n.c$ |
| 60 | 1 | 557 | 33420 | $n.c$ |
| 70 | 1 | 24 | 1680 | $1.5D-9$ |
| 80 | 1 | 7 | 560 | $1.1D-8$ |
| 15 | 2 | 65 | 1950 | $3.3D-8$ |
| 20 | 2 | 19 | 760 | $4.D-8$ |
| 25 | 2 | 11 | 550 | $5.4D-8$ |
| 30 | 2 | 6 | 360 | $5.5D-8$ |
| 7 | 3 | 87 | 1827 | $5.1D-8$ |
| 10 | 3 | 35 | 1050 | $9.D-8$ |
| 15 | 3 | 14 | 630 | $5.3D-8$ |
| 20 | 3 | 9 | 540 | $5.3D-8$ |
| 6 | 4 | 78 | 1872 | $9.5D-8$ |
| 10 | 4 | 26 | 1040 | $6.6D-8$ |
| 15 | 4 | 11 | 660 | $9.D-8$ |
| 20 | 4 | 7 | 560 | $1.6D-8$ |

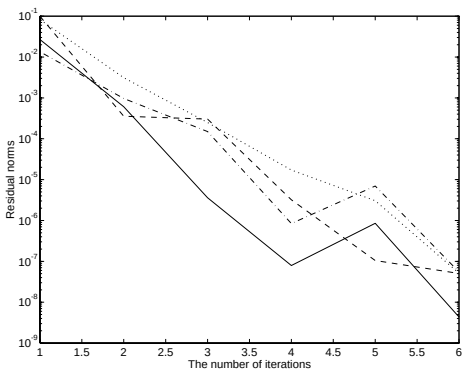


**Fig. 1.** Example 1, Algorithm 4.2 with $m = 30, p = 2$

$\lambda_i, i = 1, 2, 3, 4$, the block versions may be efficient provided $p \geq 2$, which was confirmed by Table 1. Figures 1–2 depict the convergence curves for $m = 30, p = 2$ and the iterative Arnoldi algorithm with $m = 70$, where the solid line denotes the residual norms of $\lambda_1^{(m)}, \varphi_1^{(m)}$, the dashdot those of $\lambda_2^{(m)}, \varphi_2^{(m)}$, the dashed line those of $\lambda_3^{(m)}, \varphi_3^{(m)}$ and the dotted line those of $\lambda_4^{(m)}, \varphi_4^{(m)}$.

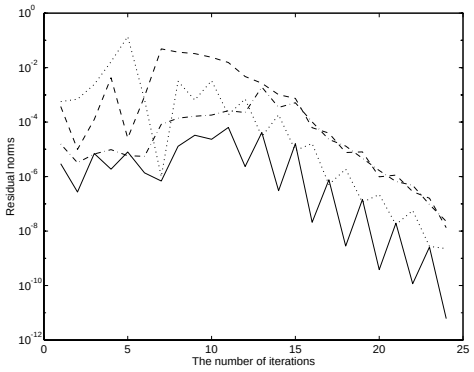


**Fig. 2.** Example 1, the iterative Arnoldi algorithm with $m = 70$

*Example 2*. This test problem is taken from the set of test matrices in $netlib$, and the matrix $A$ is of form

$$A = \begin{pmatrix} 0 & 1 & & & \\ N-1 & 0 & 2 & & \\ & \ddots & \ddots & \ddots & \\ & & 2 & 0 & N-1 \\ & & & 1 & 0 \end{pmatrix}.$$

$A$ has zero diagonal entries and known eigenvalues, and is singular if $N$ is odd. The eigenvalues are plus and minus the numbers $N-1, N-3$, $N-5, \ldots, (1 \text{ or } 0)$. The eigenvalue problem of $A$ is highly ill conditioned when $N = 500$. In fact, it can be verified that the eigenvector matrix of $A$ is too ill conditioned to compute its condition number. It is thus expected that it may be difficult to find a few eigenpairs of $A$.

We perform Algorithm 4.2 on the $500 \times 500$ matrix $A$, and require that the algorithm stop as soon as all actual residual norms of the approximating eigenpairs are below $tol = 10^{-8}$. We want to compute three largest eigenvalues $499, 497, 495$. Table 2 shows the results obtained.

Observe from Table 2 that the iterative block Arnoldi algorithm performed much better than the iterative Arnoldi algorithm. When $mp$ was about equal, that is, flops in Algorithm 4.1 were roughly same (but Algorithm 4.1 is faster on parallel computers for bigger $p$!), Algorithm 4.2 with $p > 1$ worked much better than that with $p = 1$, i.e., the iterative Arnoldi algorithm. Figures 3–4 depicts the typical convergence curves of Algorithm 4.2 with $m = 25, p = 2$ and the iterative Arnoldi algorithm with $m = 50$, where the solid line denotes the residual norms of $\lambda_1^{(m)}, \varphi_1^{(m)}$, the dashdot line those of $\lambda_2^{(m)}, \varphi_2^{(m)}$ and the dashed line those of $\lambda_3^{(m)}, \varphi_3^{(m)}$. Note that

**Table 2.** Example 2, Algorithm 4.2 for various $m$ and $p$

| $m$ | $p$ | $it$ | $m.v$ | Max. Res. Norms |
|---|---|---|---|---|
| 50 | 1 | 542 | 27100 | $7.6D-9$ |
| 60 | 1 | 656 | 39360 | $1.D-8$ |
| 80 | 1 | 384 | 30720 | $9.4D-9$ |
| 100 | 1 | 320 | 32000 | $2.7D-9$ |
| 25 | 2 | 77 | 3850 | $7.5D-9$ |
| 30 | 2 | 90 | 5400 | $9.4D-9$ |
| 40 | 2 | 57 | 4560 | $8.8D-9$ |
| 50 | 2 | 51 | 5100 | $9.4D-9$ |
| 16 | 3 | 84 | 4032 | $6.9D-9$ |
| 20 | 3 | 56 | 3360 | $2.4D-9$ |
| 26 | 3 | 67 | 5226 | $9.6D-9$ |
| 33 | 3 | 46 | 4554 | $7.2D-9$ |

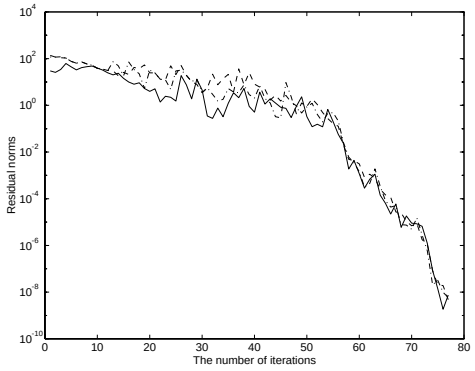


**Fig. 3.** Example 2, Algorithm 4.2 with $m = 25, p = 2$

Algorithm 4.2 sometimes converged irregularly, and for the same $p$ a bigger $mp$ may use more iterations than a smaller $mp$, e.g., $m = 30, p = 2$ versus $m = 25, p = 2$. This is not suprising and coincides with the previous convergence analysis.

*Example 3*. This is a multiple eigenproblem. $A$ is a $400 \times 400$ unsymmetric matrix and is constructed as follows:

$$A = \mathrm{diag}\left(\begin{bmatrix} 1 & 0.2 \\ -3.2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0.2 \\ -3.2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0.2 \\ -3.2 & 1 \end{bmatrix}, \begin{bmatrix} a_i & b_i/4 \\ -4b_i & a_i \end{bmatrix}\right),$$

where $a_i, b_i$, $i = 1, 2, \ldots, 197$ were generated randomly in a uniform distribution.

Algorithm 4.2 was run on $A$. We want to find the four right outer part eigenvalues of $A$ and determine their multiplicities. Algorithm 4.2 stopped

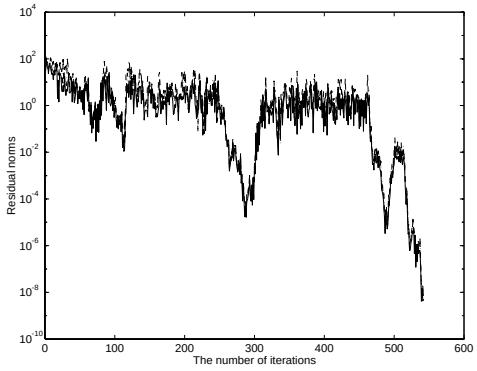


**Fig. 4.** Example 2, the iterative Arnoldi algorithm with $m = 50$

**Table 3.** Example 3, $\lambda_1 = \lambda_3 = 1 + 0.8i, \lambda_2 = \lambda_4 = 1 - 0.8i$

| $m$ | $p$ | $it$ | $m.v$ | Max. Res. Norms |
|---|---|---|---|---|
| 20 | 2 | 37 | 1480 | $6.1D-9$ |
| 25 | 2 | 21 | 1050 | $5.3D-9$ |
| 10 | 3 | 28 | 840 | $9.8D-9$ |
| 15 | 3 | 24 | 1080 | $8.1D-9$ |
| 10 | 4 | 21 | 840 | $8.7D-9$ |
| 5 | 4 | 12 | 720 | $2.2D-9$ |
| 8 | 5 | 28 | 1120 | $9.3D-9$ |
| 10 | 5 | 20 | 1000 | $9.9D-9$ |

as soon as all actual residual norms dropped below $10^{-8}$. Table 3 shows the results obtained for different $m$ and $p$.

Obviously, based on the previous theoretical analysis, it is seen from Table 3 that the multiplicities of $\lambda_1, \lambda_2$ are at least two since the converged $\lambda_1^{(m)}, \lambda_3^{(m)}$ and $\lambda_2^{(m)}, \lambda_4^{(m)}$ were numerically equal for each pair $m, p$.

For $m = 20, p = 2$ and $m = 25, p = 2$, the singular values of the matrix consisting of all the converged $\varphi^{(m)}$ associated with the numerically multiple $\lambda_1^{(m)}$ were

$$1.64893690475688,\ 1.08637886946588$$
$$0.31741315411112,\ 0.00000002494237.$$

Of them there was only one singular value less than $10^{-7}$, which was very close to the prescribed tolerance $tol$ and thus was considered to be zero numerically, according to [20, 22]. Therefore, $\lambda_1$ is three multiple. As $\lambda_2 = \bar{\lambda}_1$ and $A$ is real, $\lambda_2$ is three multiple too.

For $m = 10, p = 3$ and $m = 15, p = 3$, the singular values of the matrix consisting of all the converged $\varphi^{(m)}$ associated with the numerically

mutlple $\lambda_1^{(m)}$ were

$$1.85586943393065,\ 0.56282254375647$$
$$0.48885522237306,\ 0.00000002666527.$$

Of them there was also only one singular value less than $10^{-7}$, which was considered to be zero numerically. So $\lambda_1$ is three multiple, and $\lambda_2$ is also three multiple as $\lambda_2 = \bar{\lambda}_1$ and $A$ is real.

For $p = 4, 5$, it can be verified similarly that each corresponding converged Ritz vector matrix associated with numerically multiple $\lambda_1^{(m)}$ had only one singular value less than $10^{-7}$. It means that $\lambda_1$ and also $\lambda_2$ are three multiple.

## 7. Concluding remarks

We have presented generalized block Lanczos methods and investigated their convergence theory. The theoretical analysis show that generalized block Lanczos methods may be efficient when some of the wanted eigenvalues are clustered and may solve multiple eigenproblems under some circumstances. Based on the theoretical analysis, we have developed block Arnoldi algorithms that, among others, can sucessfully solve multiple eigenproblems practically. Numerical experiments have confirmed our theory. However, a deep analysis shows that Ritz vectors obtained by general orthogonal projection methods including the block methods may fail to converge for a general unsymmetric matrix $A$ even if corresponding Ritz values do. As a result, Algorithm 4.2 might converge slowly or irregularly and even fail to converge since in such a case combinations of current bad Ritz vectors would result in a very bad updated initial block vector $V_1$ in the next restart, such that Algorithm 4.2 may not work well. Numerical experiments in [27] have confirmed this phenomenon, where Algorithm 4.2 could exhibit very poor numerical behavior. Thus, we should seek new strategies that can theoretically guarantee convergence of eigenvectors when Ritz values do. Recently, a refined has been proposed in [27] that replaces Ritz vectors by the vectors minimizing residual norms formed with Ritz values over the block Krylov subspace. Numerical examples there have shown that the refined iterative block Arnoldi algorithms are often considerably more efficient than their counterparts. Besides, other block variants are available, e.g., [23]. For a systematic convergence theory of refined orthogonal projection methods and development of refined orthogonal projection type algorithms, see [28–30].

*Acknowledgements.* This paper is based on part of [20], and I thank Professor L. Elsner for his careful reading of the paper and many helpful discussions. Meanwhile, I thank the

anonymous referee for his/her valuable comments. Research work was supported by the Graduiertenkolleg Mathematik at the University of Bielefeld, Germany, the China State Major Key Project for Basic Researches, the National Natural Science Foundation of China, the Doctoral Program of the China State Educational Commission, the Foundation of Returning Scholars of China and the Natural Science Foundation of Liaoning Province, China.

## References

1. Arioli, M., Fassino, C. (1994): Roundoff error analysis of algorithms based on Krylov subspace methods. Technical Report, Cerfacs, Toulouse
2. Arnoldi, W.E. (1951): The principle of minimized iteration in the solution of the matrix eigenvalue problem. Quart. Appl. Math. **9**, 17–29
3. Bai, Z., Barrett, R., Day, D., Demmel, J., Dongarra, J. (1996): Test matrix collection (non-hermitian eigenvalue problems). Techical Report, Math. Dept., University of Kentucky
4. Chatelin, F. (1983): Spectral Approximation of Linear Operators. Academic Press, New York
5. Chatelin, F. (1986): Ill-conditioned eigenproblems. In: Cullum, J., Willoughby, R.A., eds., Large Scale Eigenvalue Problem, pp. 267–292
6. Chatelin, F., Godet-Thobie, S. (1991): Stability analysis in aeronautical Industries. In: Durand, M., Dabaghi, F.E., eds., High Performance Computing II, pp. 415–422
7. Chatelin, F. (1993): Eigenvalues of Matrices. Wiley, Chichester
8. Chatelin, F. (1994): Influence of nonnormality on the reliability of iterative methods in linear algebra. ILAS 94, Rotterdam
9. Chatelin, F., Fraysse, V. (1996): Lectures on Finite Precision Computations. SIAM, Philadephia
10. Cheney, E.W. (1966): Introduction to Approximation Theory. McGraw-Hill, New York
11. Cullum, J., Willoughby, W.R. (1986): A practical procedure for computing eigenvalues of large sparse nonsymmetric matrices. In: Cullum, J., Willoughby, R.A., eds., Large Scale Eigenvalue Problem, pp. 199–240
12. Daniel, G.W., Gragg, W.B., Kaufmann, L., Stewart, G.W. (1976): Reorthogonalization and stable algorithms for updating the Gram-Schmidt $QR$ factorization. Math. Comp. **30**, 772–795
13. Drkosova, J., Greenbaum, A., Rozloznik, M and Strakes, Z. (1995): Numerical stability of GMRES. BIT **35**, 309–330
14. Fischer, B., Freund, R.W. (1991): Chebyshev polynomials are not always optimal. J. Approx. Theory **65**, 261–272
15. Golub, G.H., Van Loan, C.F. (1989): Matrix Computations. The Johns Hopkins University Press, 2nd Edition, Baltimore
16. Golub, G.H., Underwood, R.R. (1977): The block Lanczos method for computing eigenvalues. In: Rice, J.R., ed., Mathematical Software **3**, pp. 361–377. Academic Press, New York
17. Ho, D. (1990): Tchebychev acceleration technique for large scale nonsymmetric matrices. Numer. Math. **56**, 721–734
18. Ho, D., Chatelin, F., Patau, J.C. (1988): Arnoldi-Tchebychev procedure for large nonsymmetric matrices. Math. Mod. Numer. Anal. **24**, 53–65
19. Jia, Z., Xiong, X. (1990): The convergence of generalized Lanczos methods for solving eigenproblems of large unsymmetric matrices. J. Dalian Univ. of Tech.(in Chinese) **30**, 1–7

20. Jia, Z. (1994): Some numerical methods for large unsymmetric eigenproblems. Ph.D. thesis, Math. Dept., University of Bielefeld, Germany
21. Jia, Z. (1995): The convergence of generalized Lanczos methods for large unsymmetric eigenproblems. SIAM J. Matrix Anal. Appl. **16**, 843–862
22. Jia, Z. (1995): Arnoldi type algorithms for large unsymmetric multiple eigenproblems. Preprint 95-026, Math. Dept., University of Bielefeld, Germany. J. Comput. Math. (to appear)
23. Jia, Z. (1995): A block incomplete orthogonalization method for large nonsymmetric eigenproblems. BIT **35**, 516–539
24. Jia, Z., Elsner, L. (1995): Improving eigenvectors in Arnoldi's method. Preprint 95-077, Math. Dept., University of Bielefeld, Germany
25. Jia, Z. (1997): Refined iterative algorithms based on Arnoldi's process for large unsymmetric eigenproblems. Linear Algebra Appl. **259**: 1–23
26. Jia, Z. (1998): A class of new projection methods for non-Hermitian matrix eigenproblems. Proceedings of the Third China-Japan Seminar on Numerical Mathematics, Shi, Z., ed., Academic Press, Beijing
27. Jia, Z. (1998): A refined iterative algorithm based on the block Arnoldi process for large unsymmetric eigenproblems. Linear Algebra Appl. **270**: 171–189
28. Jia, Z. (1996): Some natures and properties of refined orthogonal projection methods for large eigenproblems. Technical Report, Dept. of Applied Mathematics, Dalian University of Technology, China, submitted
29. Jia, Z. (1997): A theoretical comparison of classical and refined orthogonal projection methods. Technical Report, Dept. of Applied Mathematics, Dalian University of Technology, submitted
30. Jia, Z. (1997): A refined subspace iteration algorithm for large sparse eigenproblems. Technical Report, Dept. of Applied Mathematics, Dalian University of Technology, submitted
31. Kerner, W. (1989): Large-scale complex eigenvalue problems. J. Comp. Phys. **53**, 1–85
32. Manteuffel, T.A. (1977): The Tchebychev iteration for nonsymmetric linear system. Numer. Math. **28**, 307–327
33. O'Leary, D.P. (1980): The block conjugate gradient algorithm and the related methods. Linear Algebra Appl. **29**, 293–322
34. Parlett, B.N. (1980): The Symmetric Eigenvalue Problem. Prentice Hall, Englewood Cliffs, NJ
35. Ruhe, A. (1983): The two sided Arnoldi algorithm for nonsymmetric eigenvalue problem. In: Kågström, B., Ruhe, A., eds., Matrix Pencils, Lecture Notes in Mathematics **973**, pp. 104–120. Springer, Berlin
36. Ruhe, A. (1984): Rational Krylov sequence method for eigenvalue computation. Linear Algebra Appl. **58**, 391–405
37. Ruhe, A. (1994): Rational Krylov algorithms for nonsymmetric eigenvalue problems, II: Matrix pairs. Linear Algebra Appl. **197/198**, 283–295
38. Ruhe, A. (1994): The rational Krylov algorithm for nonsymmetric eigenvalue problems, III: Complex shifts for real matrices. BIT **34**, 165–176
39. Saad, Y. (1980): Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices. Linear Algebra Appl. **34**, 269–295
40. Saad, Y. (1983): Projection methods for solving large sparse eigenvalue problems. In: Kågström, B., Ruhe, A. eds., Matrix Pencils, Lecture Notes in Mathematics **973**, pp. 121–144. Springer, Berlin
41. Saad, Y. (1984): Chebyshev acceleration techniques for solving nonsymmetric eigenvalue problems. Math. Comp. **42**, 567–588

42. Saad, Y. (1980): On the rates of convergence of the Lanczos and the block-Lanczos methods. SIAM J. Numer. Anal. **17**, 687–706
43. Saad, Y. (1992): Numerical Methods for Large Eigenvalue Problems. Manchester University Press
44. Sadkane, M. (1993): A block Arnoldi-Chebyshev method for computing the leading eigenpairs of large sparse unsymmetric matrices Numer. Math. **64**, 181–194
45. Sadkane, M. (1993): Block-Arnoldi and Davidson methods for unsymmetric large eigenvalue problems. Numer. Math. **64**, 195–212
46. Scott, J.A. (1995): An Arnoldi code for computing selected eigenvalues of sparse real unsymmetric matrices. ACM Trans. Math. Softw. **21**, 432–475
47. Sorensen, D.C. (1992): Implicit application of polynomials filters in a $k$-step Arnoldi method. SIAM J. Matrix Anal. Appl. **13**, 357–385
48. Varga, R.S. (1957): A comparison of the successive overrelaxation method and semi-iterative method using Chebyshev polynomials. J. Soc. Indust. Appl. Math. **5**, 39–46
49. Wilkinson, J.H. (1965): The Algebraic Eigenvalue Problem. Clarendon Press, Oxford