University of Maryland College Park

Institute for Advanced Computer Studies TR–95–123
Department of Computer Science TR–3576

Arnoldi versus Nonsymmetric Lanczos Algorithms for Solving Nonsymmetric Matrix Eigenvalue Problems

Jane K. Cullum*

ABSTRACT

We obtain several results which may be useful in determining the convergence behavior of eigenvalue algorithms based upo n Arnoldi and nonsymmetric Lanczos recursions. We derive a relationship between nonsymmetric Lanczos eigenvalue procedures and Arnoldi eigenvalue procedures. We demonstrate that the Arnoldi recursions preserve a property which characterizes normal matrices, and that if we could determine the appropriate starting vectors, we could mimic the nonsymmetric Lanczos eigenvalue convergence on a general diagonalizable matrix by its convergence on related normal matrices. Using a unitary equivalence for each of these Krylov subspace methods, we define sets of test problems where we can easily vary certain spectral properties of the matrices. We use these and other test problems to examine the behavior of an Arnoldi and of a nonsymmetric Lanczos procedure.

*Mathematical Sciences Department, IBM Research Division, T.J. Watson Research Center, Yorktown Heights, NY 10598, USA, and Department of Computer Science, Institute for Advanced Computer Studies, and Institute for Systems Research, University of Maryland, College Park, MD 20742, NSF grant GER-9450081

1. Introduction

Lanczos recursions can be used to transform a general matrix eigenvalue problem

$$Ax = \lambda x \tag{1}$$

into a family of easier matrix eigenvalue problems whose solutions can be used to obtain approximations to eigenvalues and eigenvectors of A. In theory each easier problem is a matrix representation of projections of the given eigenvalue problem onto certain Krylov subspaces.

We consider two types of recursions for nonsymmetric A. The first variant, the Arnoldi recursion, is a direct analog of the real symmetric Lanczos recursion applied to a nonsymmetric matrix. For a given matrix and vector pair, $\{A, v_1\}$, the Arnoldi recursion simultaneously generates orthonormal bases for the Krylov subspaces, $\mathcal{K}_k(A, v_1)$, associated with that pair, and Hessenberg matrices H_k which are matrix representations of the orthogonal projections of A onto the Krylov subspaces. Theoretically, each step of the Arnoldi recursion is well-defined. However, at each step in the basic Arnoldi recursion all of the previously-generated Arnoldi basis vectors must be kept in storage.

Nonsymmetric Lanczos variants consist of two recursions. For a given matrix and vector triplet, $\{A, v_1, w_1\}$, the nonsymmetric Lanczos recursion simultaneously generates bi-orthogonal bases for the Krylov subspaces $\mathcal{K}_k(A, v_1)$ and $\mathcal{K}_k(A^T, w_1)$, and tridiagonal matrices T_k which are matrix representations of the bi-orthogonal projections of A onto these Krylov subspaces. These recursions are an implementation of a two-sided Gram-Schmidt orthogonalization. Therefore, there is no guarantee that they will not break down. However, in contrast with the Arnoldi procedure, at any step in a nonsymmetric Lanczos recursion only a few of the most recently-generated Lanczos vectors must be kept in storage. When A is real and symmetric and $w_1 = v_1$, then the nonsymmetric Lanczos recursions reduce to the real symmetric Lanczos recursion,

We would like to be able to answer the question: 'What spectral properties of A control the convergence of each of these methods'? We have not answered this question but describe several results which might be useful in such studies. See for example, the related work [2, 3, 4, 5, 6, 7, 8, 27, 28, 29].

In section 2 we outline briefly the Arnoldi and the nonsymmetric Lanczos eigenvalue procedures we are considering. In section 3 we exhibit a certain relationship between these two methods. We prove that given any matrix A and any application of a nonsymmetric Lanczos procedure to A, there exists a matrix B with the same eigenvalues as A such that the eigenvalue approximations and unnormalized residual norm estimates generated by applying an Arnoldi method to B are identical to those obtained from the nonsymmetric Lanczos computation on A. From this we can conclude, at least in exact arithmetic, that any type of eigenvalue convergence observed using the nonsymmetric Lanczos procedure can also be observed on some other problem using the Arnoldi procedure. Therefore, in this global sense, one of these procedures is not better than the other.

In section 4 we consider our nonsymmetric Lanczos procedure in more detail, indicating how we implement it in finite precision arithmetic. As in the symmetric case, convergence of eigenvalue approximations occurs in conjunction with losses in the biorthogonality of the Lanczos vectors. Using a symmetrized version of the nonsymmetric Lanczos procedure, we derive a variant of a theorem in Bai [1] connecting losses in biorthogonality to convergence of eigenvalue approximations, relaxing his assumptions of exact local biorthogonality and normalization.

In section 5 we derive a simple unitary invariance for each of these methods. In section 6 we use this invariance to identify a set of test matrices for Krylov subspace methods and express them in terms of normal matrices. In section 7 we demonstrate that Arnoldi methods preserve the Hermitian–skew Hermitian decomposition which characterizes normal matrices. We also demonstrate that if we were able to select the starting vectors in a nonsymmetric Lanczos procedure appropriately, we could simulate any eigenvalue convergence using only normal test matrices.

In section 8 we consider the Grcar test matrix [29] and several variants of it based upon the test matrices discussed in section 6, in an attempt to gain some insight into the behavior of both the Arnoldi and the nonsymmetric Lanczos procedure. The results of these tests suggest that to characterize the behavior of these methods on nonnormal problems it is not sufficient to know the singular values of the eigenvector matrix. They also suggest a potential source of numerical difficulties for both types of methods.

In [9] we consider similar questions for the problem $Ax = b$. We use the following notation.

1.1. Notation.

$A = (a_{ij})$, $1 \leq i, j \leq n$, $n \times n$ real or complex matrix
$A^T = (a_{ji})$, $1 \leq i, j \leq n$, transpose of A
$A^H = (\bar{a}_{ji})$, $1 \leq i, j \leq n$, complex conjugate transpose of A
$D = \text{diag } \{d_1, \ldots, d_n\}$, $n \times n$ diagonal matrix
$\lambda_j(A)$, $1 \leq j \leq n$, eigenvalues of A
$w(A) = \{\lambda_j(A), 1 \leq j \leq n\}$
$\sigma_j(A)$, $1 \leq j \leq n$, singular values of A where $\sigma_1 \geq \ldots \geq \sigma_n$
$\Sigma = \text{diag } \{\sigma_1, \ldots, \sigma_n\}$
$\mathcal{K}_j(A, b) = \text{span}\{b, Ab, \ldots, A^{j-1}b\}$, jth Krylov subspace generated by A and b
$\kappa(A) = \sigma_{\max}(A)/\sigma_{\min}(A)$, condition number of A
$\|A\|_2 = \sigma_{\max}(A)$, $\|x\|_2 = \sqrt{\sum_{j=1}^n x_j^2}$
v_j, jth vector in any sequence of vectors, $V_j = \{v_1, \ldots, v_j\}$
$r_j = -Az_j + \mu_j z_j$, jth residual vector for $\{\mu_j, z_j\}$
R^m, $m-$dimensional Euclidean space
e_j, jth coordinate vector in R^m where m is specified in the context
$\hat{e}_j$, jth coordinate vector in R^{m+1} where m is specified in the context
I_j, $j \times j$ identity matrix

2. Arnoldi and Nonsymmetric Lanczos Eigenvalue Procedures

In this section we review briefly Arnoldi and nonsymmetric Lanczos, Krylov subspace eigenvalue procedures. We consider an Arnoldi method and two nonsymmetric Lanczos methods. Consider Equation(1) where A is a $n \times n$ nonsymmetric matrix. A may be real or complex.

2.1. Arnoldi Methods

The Arnoldi method is based upon the Arnoldi recursion [26].

Arnoldi Recursion:

1. Given v_1 with $\|v_1\| = 1$, for $j = 2, 3, \ldots$ compute: $v_{j+1} = Av_j$
2. For each j and for $i = 1, \ldots, j$ compute:

$$h_{ij} = v_i^H v_{j+1}, \;\; v_{j+1} = v_{j+1} - h_{ij} v_i$$

3. For each j compute:

$$h_{j+1,j} = \|v_{j+1}\|, \text{ and } v_{j+1} = v_{j+1}/h_{j+1,j}.$$

Theoretically, $V_j = \{v_1, \ldots, v_k\}$ is an orthonormal basis of the Krylov subspaces $\mathcal{K}_j(A, v_1)$ and the Hessenberg Arnoldi matrices $H_j \equiv (h_{ik})$ are matrix representations of A onto $\mathcal{K}_j(A, v_1)$ with respect to the V_j. The preceding implementation is a modified Gram-Schmidt orthogonalization of the vectors $\{v_1, Av_1, A^2v_1, \ldots\}$. Other implementations exist [31]. In matrix form these recursions become

$$AV_j = V_j H_j + h_{j+1,j} v_{j+1} e_j^T \text{ where } H_j = (h_{ik}), \; 1 \leq i, k \leq j. \tag{2}$$

Basic Arnoldi Eigenvalue Procedure:

1. Given v_1 use the Arnoldi recursion to generate Hessenberg matrices H_j for $j = 1, \ldots, m$.
2. For some $j \leq m$ compute eigenvalues $H_j u = \mu u$. Compute convergence error estimates $u_\mu(m)$.
3. If desired eigenvalues are not converged, increase m and repeat steps 1 and 2.

Typically, v_1 would be chosen randomly. As with any iterative method there is an assumption that the starting vector has a projection on each part of the eigenspace to be computed.

2.2. Nonsymmetric Lanczos Methods

The corresponding nonsymmetric Lanczos variants generate two sets of Lanczos vectors, and the Lanczos matrices are tridiagonal. We consider a nonsymmetric variant which generates nonsymmetric tridiagonal matrices, and a symmetric variant which generates complex but symmetric Lanczos matrices.

The nonsymmetric variant is used in theorem 3.1 where we derive a relationship between it and an Arnoldi method. The complex symmetric variant is used in theorem 4.1 where we obtain a theorem relating losses in biorthogonality of the Lanczos vectors to convergence of eigenvalue approximations. We note however, that we could obtain a version of theorem 3.1 which uses the complex symmetric variant but the statement of the theorem would have to be modified slightly and the notation would be more complicated.

Nonsymmetric Lanczos Recursion (Nonsymmetric Variant):

1. Given v_1 and w_1 with $\|w_1\| = \|v_1\| = 1$, set $v_0 = w_0 = 0$, and $\rho_1 = 1$, $\xi_1 = 1$, and $\beta_1 = 0$. For each $j = 1, \ldots, m$ compute:

$$v_{j+1} = Av_j \text{ and } w_{j+1} = A^T w_j$$

2. For each $j = 1, \ldots, m$ compute:

$$\alpha_j = w_j^T v_{j+1} / w_j^T v_j,$$

$$p_j = v_{j+1} - \alpha_j v_j - \beta_j v_{j-1}$$

$$s_j = w_{j+1} - \alpha_j w_j - (\beta_j \rho_j / \xi_j) w_{j-1}$$

$$\rho_{j+1} = \|p_j\|, \quad v_{j+1} = p_j / \rho_{j+1}$$

$$\xi_{j+1} = \|s_j\|, \quad w_{j+1} = s_j / \xi_{j+1}$$

$$\beta_{j+1} = \xi_{j+1} w_{j+1}^T v_{j+1} / w_j^T v_j,$$

In this variant the Lanczos vectors are scaled to have unit norm. In the complex symmetric variant the Lanczos vectors are scaled so that $w_j^T v_j = 1$ for each j. There is no agreement as to which variant is preferable in practice. In both variants the coefficients in the recursions are chosen to make the Lanczos vectors $V_m = \{v_1, \ldots, v_m\}$ and $W_m = \{w_1, \ldots, w_m\}$ biorthogonal. Theoretically, for each m, V_m is a basis for the Krylov subspace $\mathcal{K}_m(A, v_1)$, and W_k is a basis for the subspace $\mathcal{K}_m(A^T, w_1)$.

Nonsymmetric Lanczos Recursion (Complex Symmetric Variant):

1. Given v_1 and w_1 with $w_1^T v_1 = 1$, set $v_0 = w_0 = 0$, and $\beta_1 = 0$. For each $j = 1, \ldots, m$ compute;

$$v_{j+1} = Av_j \text{ and } w_{j+1} = A^T w_j$$

2. For $j = 1, \ldots, m$ compute:

$$\alpha_j = w_j^T v_{j+1},$$

$$p_j = v_{j+1} - \alpha_j v_j - \beta_j v_{j-1}$$

$$s_j = w_{j+1} - \alpha_j w_j - -\beta_j w_{j-1}$$

$$\beta_{j+1} = \sqrt{p_j^T s_j},\ v_{j+1} = p_j/\beta_{j+1},\ w_{j+1} = s_j/\beta_{j+1}.$$

In practice, the complex symmetric variant is implemented using a modified two–sided Gram Schmidt orthogonalization.

For each variant we can define Lanczos tridiagonal matrices

$$T_k = \begin{pmatrix} \alpha_1 & \beta_2 & & & \\ \gamma_2 & \alpha_2 & \beta_3 & & \\ & \gamma_3 & \ddots & \ddots & \\ & & \ddots & \alpha_{k-1} & \beta_k \\ & & & \gamma_k & \alpha_k \end{pmatrix}, \tag{3}$$

In the complex symmetric variant $\gamma_k = \beta_k$ and in the nonsymmetric variant $\gamma_k = \rho_k$. Theoretically, each T_k is the matrix representation of a bi-orthogonal projection of A onto the Krylov subspaces $\mathcal{K}_k(A, v_1)$ and $\mathcal{K}_k(A^T, w_1)$. We have the following basic procedure for either variant.

Basic Nonsymmetric Lanczos Eigenvalue Procedure:

1. Given v_1 and w_1 use the nonsymmetric Lanczos recursion to generate tridiagonal Lanczos matrices T_k for $k = 1, \ldots, m$.
2. For some $k \leq m$ compute eigenvalues, $T_k u = \mu u$. Select some subset of the eigenvalues of T_k as approximations to eigenvalues of A. For all relevant μ compute convergence error estimates $u_\mu(m)$.
3. If the desired eigenvalues are not converged, increase m and repeat steps 1 and 2.

Typically, v_1 is chosen randomly and $w_1 \equiv v_1$. If A is a real, normal matrix, and v_1 has reasonable projections on the desired right eigenvectors of A, then setting $w_1 = v_1$ may be an optimal choice in terms of the mismatch theorem [25]. .

Lemma 2.1. *Let A be a real, normal matrix with n distinct eigenvalues. Let v_1 have a significant projection on each unit right eigenvector of A. Then v_1 has a significant projection on each unit left eigenvector of A.*

Proof. Let r_j, l_j denote unit right and left eigenvectors, $1 \leq j \leq n$. Since A is normal, each $l_j = \bar{r_j}$. Let $v_1 = \sum_{k=1}^n \gamma_k r_k$. If λ_j is real, then r_j is real and $l_j = r_j$. If λ_j is complex, then $\lambda_{j+1} = \bar{\lambda_j}$ is also an eigenvalue, and the projection of v_1 on l_j equals its projection on $r_{j+1} = \bar{r_j}$ and its projection on l_{j+1} equals its projection on r_j.

In matrix form the nonsymmetric Lanczos recursions can be written as

$$\begin{aligned} AV_m &= V_m T_m + \gamma_{m+1} v_{m+1} e_m^T \\ A^T W_m &= W_m \tilde{T}_m + \omega_{m+1} w_{m+1} e_m^T \end{aligned} . \tag{4}$$

In the complex symmetric variant, $\tilde{T}_m = T_m$ and $\gamma_{m+1} = \beta_{m+1} = \omega_{m+1}$. In the nonsymmetric variant, $\tilde{T}_m = \Phi_m^{-1} T_m \Phi_m$ where $\Phi_m = \mathrm{diag}(\phi_1, \ldots, \phi_m)$ with $\phi_1 = 1$ and each $\phi_j = \phi_{j-1}\rho_j/\xi_j$.

In each Krylov subspace method there is a right Ritz vector, $z_j^k = V_k u_j^k / \|V_k u_j^k\|$ corresponding to any eigenvalue approximation μ_j^k. In the nonsymmetric Lanczos procedures we also obtain a left Ritz vector. In the complex symmetric variant the left vector has the form $y_j^k = W_k u_j^k / \|W_k u_j^k\|$. ¿From recursions(4), for any k and each $j \leq k$, we have, the normalized right residual norm error estimates .

$$\|r_j^R\| \equiv \|Az_j^k - \mu_j^k z_j^k\| \leq (|\beta_{k+1} u_j^k(k)| \|v_{k+1}\| + \|F_k\|)/\|V_k u_k^j\|. \tag{5}$$

For the nonsymmetric Lanczos procedure we also have the normalized left residual norm estimate.

$$\|r_j^L\| \equiv \|A^T y_j^k - \mu_j^k y_j^k\| \leq (|\beta_{k+1} u_j^k(k)| \|w_{k+1}\| + \|G_k\|)/\|W_k u_k^j\|. \tag{6}$$

F_j and G_j represent the errors introduced into the recursions by the finite precision arithmetic. Existing error estimates for any μ_j^k, when it is considered as an estimate of some eigenvalue λ of A, require estimates of both a right and a left normalized residual norm [25], and the condition of that eigenvalue, cond(λ).

$$\|\mu_j^k - \lambda\| \leq \text{cond}(\lambda)\max(\|r_j^R\|, \|r_j^L\|). \tag{7}$$

In practice we do not know the condition of the eigenvalues, and do not have an estimate of the error matrices F_j and G_j. If $w_1 = v_1$ in the nonsymmetric Lanczos recursions and A is real and symmetric, then both the Arnoldi and the nonsymmetric Lanczos recursions reduce to the real symmetric Lanczos recursions.

Each step of the nonsymmetric Lanczos recursions requires matrix-vector multiplications by both A and A^T. The Arnoldi recursions use only A. However, the computation of the $(k+1)^{st}$ Arnoldi vector requires all k preceding Arnoldi vectors. In contrast, the computation of the $(k+1)^{st}$ Lanczos vector requires only the two most recently-generated left and right Lanczos vectors. Therefore, if there is no re-bi-orthogonalization of the Lanczos vectors, the storage requirements of a basic nonsymmetric Lanczos procedure are at most some small multiple of the order of A.

Since the nonsymmetric Lanczos recursion is an implementation of a two-sided Gram Schmidt bi-orthogonalization of two sets of vectors, it is possible for it to break down even if each set V_k and W_k is linearly independent. Serious breakdown occurs if for some j, $w_j^T v_j = 0$ but $w_j \neq 0$ and $v_j \neq 0$. If this occurs then the recursions cannot be continued. If $w_j = 0$ or $v_j = 0$, then this means an invariant subspace for either A^T or A has been found.

Exact breakdown is highly improbable, near breakdowns may cause numerical instabilities. To avoid such problems, various look-ahead strategies have been proposed, see e.g. [25, 18]. The discussions in this paper are equally applicable to the look-ahead variants of these methods. If look-ahead steps are performed, then the scalar coefficients in Equations(4) become matrices, and the Lanczos vectors become block biorthogonal.

Assumption 1.1: In any statement or theorem about the nonsymmetric Lanczos procedures we will always assume that no breakdown has occurred and that all quantities are well-defined.

3. A Relationship Between a Nonsymmetric Lanczos Method and an Arnoldi Method

In this section we use B and C to denote two different matrices. We use a superscript A to denote quantities associated with an Arnoldi computation, and use a superscript L to denote quantities associated with nonsymmetric Lanczos computations.

We want to examine the behavior of the Arnoldi and the nonsymmetric Lanczos procedures as we vary the spectral properties of a matrix. How different are these two methods? Can we obtain a general relationship between these two methods? In exact arithmetic, we prove that given any matrix B and any application of a nonsymmetric Lanczos procedure to B, there exists a matrix C of the same size as B and with the same eigenvalues as B such that the eigenvalue approximations, the matrix residual norm, and the unnormalized residual norm estimates generated by applying an Arnoldi procedure to C are identical to those obtained from the nonsymmetric Lanczos computation on B. ¿From this we can conclude, at least in exact arithmetic, that any type of eigenvalue convergence observed using a nonsymmetric Lanczos procedure can also be observed using the Arnoldi procedure on some other problem with the same eigenvalues . In section 8 we will consider the convergence of these two procedures when they are applied to the same test problems. In [9] we derive an analog of this theorem for the $Ax = b$ problem,

In Theorem 3.1 we assume that B has n distinct eigenvalues, that the starting vectors have projections on each of the right and the left eigenvectors of B, and that there is no breakdown in the nonsymmetric Lanczos recursions. If there were fewer than n distinct eigenvalues, the recursions would terminate for some $m < n$. In this case Theorem 3.1 would still be valid when reworded in terms of m and the distinct eigenvalues of B.

Theorem 3.1. *(Exact arithmetic). Let B be any $n \times n$ matrix with n distinct eigenvalues. Let v_1 be any random vector such that it has a projection on each of the right and the left eigenvectors of B. Apply the nonsymmetric variant of the nonsymmetric Lanczos eigenvalue procedure with $w_1 = v_1$. Let V_k^L and T_k^L denote respectively, Lanczos vectors and matrices generated by applying recursions(4) to $\{B, v_1, v_1\}$. Then there exists a $n \times n$ matrix C with the same eigenvalues as B and a starting vector v_1^A such that for $1 \le k \le n$, the Arnoldi eigenvalue procedure applied to $\{C, v_1^A\}$, yields*

$$\begin{aligned} &H_k^A = T_k^L \text{ and} \\ &\|BV_k^L - V_k^L T_k^L\| = \|CV_k^A - V_k^A H_k^A\| = |\rho_{k+1}|. \end{aligned} \tag{8}$$

Furthermore, for any $T_k^L u_j^k = \mu_j^k u_j^k$

$$\|r_j^R(B,L)\| \equiv \|BV_k^L u_j^k - \mu_j^k V_k^L u_j^k\| = \|r_j^R(C,A)\| \equiv \|CV_k^A u_j^k - \mu_j^k V_k^A u_j^k\| = |\rho_{k+1} u_j^k(k)| \tag{9}$$

Proof. Let

$$T_n^L = Q_n U_n Q_n{}^H \tag{10}$$

be any Schur decomposition of T_n.

Multiplying both sides of equation(10) on the left by Q_n^H we obtain

$$U_n Q_n^H = Q_n^H T_n^L \ . \tag{11}$$

Define $V_n^A \equiv Q_n^H$, and $C \equiv U_n$. The columns of V_n^A are Arnoldi vectors corresponding to C and to the first column of Q_n^H. For each $k \le n$,

$$\|CV_k^A - V_k^A T_k^L\| = \|\rho_{k+1} v_{k+1}^A e_k^T\| = |\rho_{k+1}|. \tag{12}$$

Since $\|v_{k+1}^L\| = 1$, we have that

$$\|BV_k^L - V_k^L T_k^L\| = \|\rho_{k+1} v_{k+1}^L e_k^T\| = |\rho_{k+1}| \tag{13}$$

yielding equality of the norms of the residual matrices, see Equation(8).

Let $T_k^L u_j^k = \mu_j^k u_j^k$, and define $\tilde{u}$ equal to the $k+1$ vector whose first k components consist of u and whose $k+1$ component is zero. The corresponding residual for the Arnoldi right Ritz vector satisfies

$$r_k^R(C,A) = CV_k^A u_j^k - \mu_j^k V_k^A u_j^k = V_{k+1}^A z_k^A = V_{k+1}^A z_k^L = \rho_{k+1} u_j^k(k) v_{k+1}^A \tag{14}$$

where

$$z_k^L = z_k^A = -(T_k^L)^e u_j^k + \mu_j^k \tilde{u}_j^k \text{ with } (T_k^L)^e = \begin{pmatrix} T_k \\ \beta_{k+1} e_k^T \end{pmatrix} , \tag{15}$$

and z_k^L, z_k^A are the associated Lanczos and Arnoldi quasi-residual vectors. The corresponding unnormalized residual for the Lanczos right Ritz vector satisfies

$$r_k^R(B,L) = BV_k^L u_j^k - \mu_j^k V_k^L u_j^k = V_{k+1}^L z_k^L = \rho_{k+1} v_{k+1}^L u_j^k(k) \tag{16}$$

From the orthonormality of the columns of V_k^A, we have that

$$\|r_k^R(C,A)\| = \|z_k^A\| = \|z_k^L\| = |\rho_{k+1}u_j^k(k)|. \tag{17}$$

By construction $\|v_{k+1}^L\| = 1$. Therefore,

$$\|r_k^R(B,L)\| = \|z_k^L\| = |\rho_{k+1}u_j^k(k)| \tag{18}$$

4. Finite Precision Arithmetic and Nonsymmetric Lanczos Procedures

Typically, in finite precision the Lanczos vectors do not remain biorthogonal, and the basic procedure must be modified. We use modifications analogous to our modifications for the real symmetric Lanczos procedure [14]. We require the following assumptions.

Assumption 4.1:. *Lanczos Phenomenon.* For large enough m, all of the desired eigenvalues of A will appear in $\omega(T_m)$.

Assumption 4.2: Any *spurious* eigenvalues appearing in the spectra of any Lanczos matrix T_m are caused by losses in the biorthogonality of the Lanczos vectors and represent reappearances of converged eigenvalue approximations.

Ritz vectors are not computed during the eigenvalue computations so that the storage requirements for our eigenvalue computations are very small. Once the eigenvalues have been computed accurately, an appropriate size Lanczos matrix $m(\mu)$ can be determined for each relevant eigenvalue approximation μ and used to compute corresponding eigenvector approximations. These eigenvector computations require regeneration of the Lanczos vectors.

4.1. Spurious Eigenvalues

The success of any Lanczos procedure which does not use reorthogonalization depends upon a procedure for identifying the *spurious* eigenvalues which appear when biorthogonality is lost. For the real symmetric case, Paige [22] proposed that error estimates be used to make this identification. An eigenvalue of some T_m would be accepted as *good* and an approximation to an eigenvalue of A only if its corresponding error estimate was sufficiently small. There are two problems associated with that approach. First, any *spurious* eigenvalue which is close to a converged *good* eigenvalue will typically have an error estimate of the same order of magnitude as the estimate for a converged eigenvalue approximation. Thus, the procedure would indicate that two such eigenvalues were either distinct or should be combined, resulting in losses in achievable accuracy. Second, if error estimates are used to determine convergence, the user only sees the eigenvalues whose estimates meet the convergence tolerance, and none of the other Lanczos eigenvalues which are not copies of converged approximations but whose error estimates do not meet the convergence tolerance. The accuracy of these other approximations varies and they cannot be identified using a convergence tolerance. Presumably *spurious* eigenvalues which are not close to converged eigenvalues would not have small error estimates and would be excluded on that basis, along with these other *good* eigenvalues. Error estimates cannot distinquish between these two types of Lanczos eigenvalues. Alternatively, references [24, 30] track the convergence of approximations as the size of the Lanczos matrix is increased and only accept converged approximations. That approach suffers from the same difficulties as an approach which uses error estimates.

We use a completely different approach. No convergence tolerances are used. Our identification test is a simple extension of the test used in our real symmetric Lanczos procedures. This extension is discussed in detail in [11]. The argument requires only the symmetry of the Lanczos tridiagonal matrices and is valid in finite precision arithmetic, as long as the error terms in the Lanczos recursions remain small. There is no proof that our test identifies all of the extra eigenvalues but it seems to work well in practice. This test is implemented using inverse iteration on the indicated submatrix.

C-W Identification Test: A simple eigenvalue μ of some T_k which is also numerically an eigenvalue of the submatrix obtained by deleting the first row and column of T_k is labelled *spurious* and discarded.

All of the remaining simple and numerically-multiple Lanczos eigenvalues are labelled *good*, and used as approximations to eigenvalues of A.

Error estimates for relevant simple *good* eigenvalues μ are computed by computing eigenvectors u of the Lanczos matrix, T_k, and then computing $|\beta_{k+1}u(k)|$. Convergence is declared when all relevant estimates are small. In the nonsymmetric case, the accuracy of the error estimates as measures of convergence depends upon the error propagation, the condition of the eigenvalues being computed, and the sizes of various vector norms.

4.2. Losses in Biorthogonality Imply Convergence

The stability of our algorithm depends upon the error propagation. In Theorem 4.1, we derive an apparently stronger version of a theorem in [1] relating the loss of biorthogonality of the Lanczos vectors to convergence of Ritz vector approximations. Bai [1] considered a different nonsymmetric variant of the nonsymmetric Lanczos eigenvalue procedure and obtained extensions of theorems in Paige [20, 21] to the error terms F_j and G_j in equations (19). In his proof, Bai assumed exact local biorthogonality and normalizability. We prove a similar result for the complex symmetric variant which requires only ϵ-biorthogonality and normalizability.

Theorem 4.1. *(Loss of Bi-orthogonality implies Convergence): Let*

$$\begin{aligned} AV_j &= V_jT_j + \beta_{j+1}v_{j+1}e_j^T + F_j \\ A^TW_j &= W_jT_j + \beta_{j+1}w_{j+1}e_j^T + G_j \end{aligned} \tag{19}$$

$$W_j^TV_j = I_j + \Delta_j + L_j + U_j \tag{20}$$

where

$$\begin{aligned} \Delta_j &= \text{diag } (\delta_1, \ldots, \delta_j) \text{ with } \delta_j \equiv (w_j^Tv_j - 1) \\ L_j &= \text{strictly lower triangular part of } W_j^TV_j \\ U_j &= \text{strictly upper triangular part of } W_j^TV_j \end{aligned}$$

and V_j and W_j are $n \times j$, T_j is $j \times j$, and F_j and G_j are $n \times j$ matrices.

Define

$$K_j - K_j^T \equiv W_j^TF_j - F_j^TW_j + V_j^TG_j - G_j^TV_j + 2(\Delta_jT_j - T_j\Delta_j) \tag{21}$$

where K_j is strictly upper triangular.

$$\epsilon_i^V = v_{i+1}^Tw_i \quad , \quad \epsilon_i^W = w_{i+1}^Tv_i, \text{ for } 1 \le i \le j \ , \tag{22}$$

$$D_j = \text{ diag } (d_1, \ldots, d_j) \text{ with } d_i = \beta_i[\epsilon_{i-1}^W - \epsilon_{i-1}^V] + \beta_{i+1}[\epsilon_i^V - \epsilon_i^W] \ . \tag{23}$$

For any $T_ju_i = \mu_iu_i$ define corresponding unnormalized right and left Ritz vectors

$$z_i^R = V_ju_i \text{ and } z_i^L = W_ju_i \quad . \tag{24}$$

Then for any $1 \le i \le j$,

$$\left(z_i^L\right)^T v_{j+1} + \left(z_i^R\right)^T w_{j+1} = \frac{-u_i^TD_ju_i + u_i^TK_ju_i + (\epsilon_j^V + \epsilon_j^W)u_i^2(j)}{\beta_{j+1}u_i(j)}. \tag{25}$$

Therefore if $\|F_j\|, \|G_j\|, \|\Delta_j\|$, and $\|D_j\|$ are small, $\|W_j\|, \|V_j\|$, and $\|T_j\|$ are not too big, and either $|(z_i^L)^Tv_{j+1}|$ or $|(z_i^R)^Tw_{j+1}|$ is large, then $|\beta_{j+1}u_i(j)|$ must be small.

We need the following lemmas.

Lemma 4.1. *Let T, L and U be as defined in Theorem 4.1. Let $C_1 = TL - LT$ and $C_2 = TU - UT$. Then C_1 is lower triangular, C_2 is upper triangular, and*

$$\begin{aligned} C_1(i,i) &= \beta_{i+1}\epsilon_i^W - \beta_i\epsilon_{i-1}^W \\ C_2(i,i) &= \beta_i\epsilon_{i-1}^V - \beta_{i+1}\epsilon_i^V \quad . \end{aligned} \tag{26}$$

Proof. Consider C_1. The proof for C_2 is similar. C_1 is lower Hessenberg. Straight-forward applications of the facts $T_{ik} = 0$ unless $k = i-1, i, i+1$ and $L_{ik} = 0$ when $k \geq i$, yield $C_1(i, i+1) = 0$ for all $1 \leq i \leq j$, and the first half of Equation (26).

Lemma 4.2. *For each j, $W_j^T r_v e_j^T - \epsilon_j^V e_j e_j^T$ is strictly upper triangular.*

Proof. The proof follows directly from the definition of ϵ_j^V.

Proof. (Theorem 4.1): ¿From the nonsymmetric Lanczos recursions (19), we obtain

$$W_j^T A V_j - \left(W_j^T V_j\right) T_j = \left(W_j^T r_v\right) e_j^T + W_j^T F_j \ . \tag{27}$$

$$V_j^T A^T W_j - \left(V_j^T W_j\right) T_j = \left(V_j^T r_w\right) e_j^T + V_j^T G_j \ . \tag{28}$$

where we set $r_v = \beta_{j+1} v_{j+1}$ and $r_w = \beta_{j+1} w_{j+1}$. Taking the transposes of both sides of equations (27, 28) we obtain

$$V_j^T A^T W_j - T\left(V_j^T W_j\right) = e_j\left(r_v^T W_j\right) + F_j^T W_j \tag{29}$$

$$W_j^T A V_j - T\left(W_j^T V_j\right) = e_j\left(r_w^T V_j\right) + G_j^T V_j \quad . \tag{30}$$

Subtracting Equation (30) from Equation (27) we obtain

$$T_j\left(W_j^T V_j\right) - \left(W_j^T V_j\right) T_j = \left(W_j^T r_v\right) e_j^T - e_j\left(r_w^T V_j\right) + W_j^T F_j - G_j^T V_j \ . \tag{31}$$

Subtracting Equation(28) from Equation (29) we obtain

$$\left(V_j^T W_j\right) T_j - T_j\left(V_j^T W_j\right) = e_j\left(r_v^T W_j\right) - \left(V_j^T r_w\right) e_j^T + F_j^T W_j - V_j^T G_j \ . \tag{32}$$

Subtracting Equation (32) from Equation (31) we obtain

$$\begin{aligned} \left[W_j^T r_v e_j^T - e_j r_v^T W_j\right] + \left[V_j^T r_w e_j^T - e_j r_w^T V_j\right] &+ \left[W_j^T F_j - F_j^T W_j\right] + \left[V_j^T G_j - G_j^T V_j\right] \\ = \left[T_j\left(W_j^T V_j\right) - \left(V_j^T W_j\right) T_J\right] &+ \left[T_j\left(V_j^T W_j\right) - \left(W_j^T V_j\right) T_j\right] \ . \end{aligned} \tag{33}$$

The matrices within each pair of square brackets are skew symmetric. ¿From Equation (20) and the symmetry of T_j we have that

$$\begin{aligned} T_j\left(W_j^T V_j\right) - \left(V_j^T W_j\right) T_j + T_j\left(V_j^T W_j\right) &- \left(W_j^T V_j\right) T_j = \\ 2\left(T_j \Delta_j - \Delta_j T_j\right) + \left(T_j L_j - L_j T_j\right) &+ \left(T_j U_j - U_j T_j\right) \\ + \left(T_j L_j^T - L_j^T T_j\right) &+ \left(T_j U_j^T - U_j^T T_j\right) \ . \end{aligned} \tag{34}$$

¿From Lemma 4.1 we have that

$$\tilde{L}_j \equiv \left(T_j L_j - L_j T_j\right) + \left(T_j U_j^T - U_j^T T_j\right) \tag{35}$$

is lower triangular and

$$\tilde{U}_j \equiv \left(T_j L_j^T - L_j^T T_j\right) + \left(T_j U_j - U_j T_j\right) \tag{36}$$

is upper triangular. Rewrite Equation (33) as

$$\left[W_j^T r_v e_j^T - e_j r_v^T W_j\right] + \left[V_j^T r_w e_j^T - e_j r_w^T V_j\right] = \tilde{L}_j + \tilde{U}_j + K_j - K_j^T \ . \tag{37}$$

Equate the strictly upper triangular parts of the right and of the left hand sides of Equation (37)

$$\left[W_j^T r_v e_j^T - \epsilon_j^V e_j e_j^T\right] + \left[V_j^T r_w e_j^T - \epsilon_j^W e_j e_j^T\right] = \tilde{U}_j - D_j + K_j \ . \tag{38}$$

Let $T_j u_i = \mu_i u_i$ with $\|u_i\| = 1$. If we apply u_i^T to the left side of Equation (38) and u_i to the right side, we obtain

$$\begin{aligned}(\beta_{j+1} u_i(j)) \left[(z_i^L)^T v_{j+1} + (z_i^R)^T w_{j+1}\right] \quad &= \\ u_i^T \tilde{U}_i u_i - u_i^T D_j u_i \quad &+ \quad u_i^T K_j u_i + \left[\epsilon_j^W + \epsilon_j^V\right] u_i^2(j) \ .\end{aligned} \tag{39}$$

¿From Equation (36), $u_i^T \tilde{U}_j u_i = 0$ and therefore Equation (39) reduces to Equation (25). Furthermore, $u_i^T D_j u_i = E_j$ where $\|E_j\| \leq 4\left(\max_i |\beta_{i+1}|\right)\left(\max_i |\epsilon_i^{V,W}|\right)$. Therefore, if $\|F_j\|, \|G_j\|, \|\Delta_j\|$, and the local near-biorthogonality terms $\epsilon_i^V, \epsilon_i^W$, $1 \leq i \leq j$, are small, and $\|W_j\|, \|V_j\|, \|T_j\|$ are not too large, then the right hand side of Equation (39) is small. Therefore, if for some i, either $|(z_i^L)^T v_{j+1}|$ or $|(z_i^R)^T w_{j+1}|$, is large, then $|\beta_{j+1} u_i(j)|$ must be small. Furthermore, if $\|F_j\|$ and $\|G_j\|$ are small, $\|v_{j+1}\|$ and $\|w_{j+1}\|$ are not large, and $\|V_j u_i\|$ and $\|W_j u_i\|$ are not too small, Equation (19) implies that the residual norms $\|A z_i^R - \mu_i z_i^R\| / \|z_i^R\|$ and $\|A^T z_i^L - \mu_i z_i^L\| / \|z_i^L\|$ are small, so that convergence, at least in the sense of the residual norms, has occurred.

5. Invariant Properties of Arnoldi and Nonsymmetric Lanczos Methods

To be able to fully utilize Krylov subspace iterative methods in practical problems we need a better understanding of their convergence behavior. In order to study the numerical behavior of these algorithms we need test matrices where we can systematically vary spectral properties. In this section we list several theorems which identify one such possible class of matrices.

We have the following unitary invariance for the real symmetric Lanczos methods [11]. Theorem 5.1 is stated for real C but the complex Hermitian analog follows easily. In this section we use C and $\tilde{C}$ to denote two matrices which are unitarily similar.

Theorem 5.1. *(Exact Arithmetic): Let C and $\tilde{C}$ be similar real symmetric matrices. Let U be a unitary matrix such that $\tilde{C} = U^T C U$. For $k = 1, \ldots, K$, let T_k^C , $T_k^{\tilde{C}}$ and V_k^C, $V_k^{\tilde{C}}$ denote respectively, Lanczos matrices and Lanczos vectors, obtained by applying the real symmetric Lanczos recursions to C and to $\tilde{C}$. If the starting vector for C is v_1 and for $\tilde{C}$ is $v_1^{\tilde{C}} = U^T v_1$, then for $k = 1, \ldots, K$, $T_k^C = T_k^{\tilde{C}}$, $V_k^{\tilde{C}} = U^T V_k^C$, and the two computations yield identical eigenvalue approximations, and identical residual and error norms for the corresponding Ritz vectors.*

Corollary 5.1. *if A is real and symmetric and the arithmetic is exact, to understand the convergence behavior of a real symmetric Lanczos procedure, it is sufficient to consider diagonal test matrices.*

We consider extensions of Theorem 5.1 to nonsymmetric matrices. We will use these extensions to obtain a family of test matrices. First consider an Arnoldi method.

Theorem 5.2. *(Exact Arithmetic) Let U be any unitary matrix. Let C be a nonsymmetric matrix and define $\tilde{C} \equiv U^H C U$. For $k = 1, \ldots, K$,, let H_k^C , $H_k^{\tilde{C}}$ and V_k^C, $V_k^{\tilde{C}}$ denote respectively, Arnoldi matrices and Arnoldi vectors, obtained by applying the Arnoldi recursions in exact arithmetic to $\{C, v_1^C\}$ and to $\{\tilde{C}, \tilde{v}_1^{\tilde{C}}\}$ with $\tilde{v}_1 = U^H v_1^C$. Then for each k, $H_k^C = H_k^{\tilde{C}}$, $V_k^C = U V_k^{\tilde{C}}$, and the two computations yield identical eigenvalue approximations, and identical residual and error norms for the corresponding Ritz vectors.*

Proof. For each k and in exact arithmetic

$$CV_k^C = V_k H_k^C + h_{k+1,k} v_{k+1}^C e_k^T \text{ and } \tilde{C}(U^H V_k^C) = (U^H V_k^C) H_k^C + h_{k+1,k} U^H v_{k+1}^C e_k^T. \tag{40}$$

Therefore, for each k, $V_k^{\tilde{C}} = U^H V_k^C$ and $H_k^{\tilde{C}} = H_k^C$.

Since C and $\tilde{C}$ are similar, they have the same eigenvalues and if $Cx = \lambda x$ with $\|x\| = 1$, then $\tilde{C} U^H x = \lambda U^H x$. Moreover, for any $H_k^C u = \mu u$, the corresponding Arnoldi right Ritz vectors satisfy $z_\mu^{\tilde{C}} = U^H z_\mu^C$ where we set $\|z_\mu^C\| = 1$. Let λ be the eigenvalue of C closest to μ and let x be a corresponding unit right eigenvector. Then the error and the residual norms for these eigenvector approximations satisfy the following.

$$\|e_k^C\| = \|z_k^C - x\| = \|z_k^{\tilde{C}} - U^H x\| = \|e_k^{\tilde{C}}\| \text{ and } \|r_k^C\| = \| - C z_k^C + \mu z_k^C\| = \| - \tilde{C} z_k^{\tilde{C}} + \mu z_k^{\tilde{C}}\| = \|r_k^{\tilde{C}}\|. \tag{41}$$

Theorem 5.2 tells us that applications of the Arnoldi procedure to any $\{C, v_1\}$ and to the corresponding $\{U^H C U, U^H v_1\}$ where U is any unitary matrix will yield the same eigenvalue and Ritz vector convergence behavior. This is not unexpected since the corresponding starting vectors have the same size projections on corresponding right eigenvectors. We note that the matrices were not assumed to be diagonalizable.

In Theorem 5.3 we use the complex variant of the nonsymmetric Lanczos recursions but the basic arguments are independent of the particular scaling used.

Theorem 5.3. *(Exact Arithmetic) Let U be any unitary matrix. Let C be a nonsymmetric matrix and define $\tilde{C} \equiv U^H C U$. For $k = 1, \ldots, K$, let T_k^C , $T_k^{\tilde{C}}$ and V_k^C, W_k^C, $V_k^{\tilde{C}}$, $W_k^{\tilde{C}}$ denote respectively, Lanczos matrices and Lanczos vectors, obtained by applying the nonsymmetric Lanczos recursions in equation(4) to C and to $\tilde{C}$ with $v_1^{\tilde{C}} = U^H v_1^C$ and $w_1^{\tilde{C}} = U^T w_1^C$. Then for each k, $T_k^C, = T_k^{\tilde{C}}$, $V_k^C = U V_k^{\tilde{C}}$, and $W_k^C = \bar{U} W_k^{\tilde{C}}$, where U^T is the simple transpose of U. The two computations yield identical eigenvector approximations, and identical residual and error norms for the corresponding right and left Ritz vectors.*

Proof. For each k and in exact arithmetic,

$$\begin{aligned} CV_k^C &= V_k^C T_k^C + \rho_{k+1} v_{k+1}^C e_k^T \\ C^T W_k^C &= W_k^C T_k^C + \rho_{k+1} w_{k+1}^C e_k^T \end{aligned} \tag{42}$$

Similarly,

$$\begin{aligned} \tilde{C}(U^H V_k^C) &= (U^H V_k^C) T_k^C + \rho_{k+1} U^H v_{k+1}^C e_k^T \\ \tilde{C}^T (U^T W_k^C) &= (U^T W_k^C) T_k^C + \rho_{k+1} U^T w_{k+1}^C e_k^T \end{aligned} \tag{43}$$

Therefore, for each k, $T_k^C = T_k^{\tilde{C}}$, $V_k^{\tilde{C}} = U^H V_k^C$ and $W_k^{\tilde{C}} = U^T W_k^C$.

Since C and $\tilde{C}$ are similar, they have the same eigenvalues, and if $Cx = \lambda x$ and $C^T y = \lambda y$, then $\tilde{C} U^H x = \lambda U^H x$ and $\tilde{C}^T U^T y = \lambda U^T y$. Moreover, for any $T_k u = \mu u$, the corresponding right and left Ritz vectors satisfy $z_\mu^{R^{\tilde{C}}} = U^H z_\mu^{R^C}$ and $z_\mu^{L^{\tilde{C}}} = U^T z_\mu^{L^C}$ where we set $\|z_\mu^{R^C}\| = \|z_\mu^{L^C}\| = 1$. Let λ be the eigenvalue of C closest to μ, and let x and y be corresponding unit right and left eigenvectors. The error and residual norms satisfy the following.

$$\begin{aligned} &\|e_k^{R^C}\| = \|z_k^{R^C} - x\| = \|z_k^{R^{\tilde{C}}} - U^H x\| = \|e_k^{R^{\tilde{C}}}\| \\ &\|e_k^{L^C}\| = \|z_k^{L^C} - y\| = \|z_k^{L^{\tilde{C}}} - U^H y\| = \|e_k^{L^{\tilde{C}}}\| \\ &\|r_k^{R^C}\| = \| - C z_k^{R^C} + \mu z_k^{R^C}\| = \| - \tilde{C} z_k^{R^{\tilde{C}}} + \mu z_k^{R^{\tilde{C}}}\| = \|r_k^{R^{\tilde{C}}}\| \\ &\|r_k^{L^C}\| = \| - C^T z_k^{L^C} + \mu z_k^{L^C}\| = \| - \tilde{C}^T z_k^{L^{\tilde{C}}} + \mu z_k^{L^{\tilde{C}}}\| = \|r_k^{L^{\tilde{C}}}\| \\ &. \end{aligned} \tag{44}$$

Theorem 5.3 tells us that applications of the nonsymmetric Lanczos procedure to any $\{C, v_1, w_1\}$ and to any $\{U^H C U, U^H v_1, U^T w_1\}$ where U is a unitary matrix will yield the same eigenvalue and Ritz vector convergence behavior. As in Theorem 5.2, equal size projections on each of the left and the right eigenvectors are maintained.

6. Test Matrices

From Theorems 5.2 and 5.3, we obtain the well-known fact that we can study the behavior of each of these Krylov methods on normal matrices by considering only complex diagonal matrices. For general matrices we have the following theorems. Theorem 5.4 provides a form for test matrices.

6.1. A Family of Test Matrices

Theorem 6.1. *(Exact Arithmetic). For either an Arnoldi or a nonsymmetric Lanczos eigenvalue procedure, all possible sequences of eigenvalue approximations and all corresponding sequences of Ritz vector residual and error norms can be generated by considering matrices of the form*

$$A = \Sigma V^H J V \Sigma^{-1} \tag{45}$$

where Σ is a diagonal matrix with positive diagonal entries, V is a unitary matrix and J is the Jordan canonical form of A. If A is diagonalizable, then J is a diagonal matrix Λ of the eigenvalues of A.

Proof. Let C be any $n \times n$ matrix then there exists an invertible matrix X and a Jordan matrix J such that

$$C = XJX^{-1}\ . \tag{46}$$

Let $X = U\Sigma V^H$ be a singular value decomposition of X. Then $C = U\Sigma V^H J V \Sigma^{-1} U^H$. Define $\tilde{C} \equiv \Sigma V^H J V \Sigma^{-1}$. By Theorems 5.2 and 5.3, any one of the two methods will generate the same eigenvalue approximations and Ritz vector residual and error norms when applied to both C and $\tilde{C}$, when the starting vectors are chosen according to Theorems 5.2 and 5.3. If C is diagonalizable, then X is a matrix of eigenvectors of C and J is a diagonal matrix of corresponding eigenvalues of C.

If A is real and diagonalizable we can replace complex V and Λ in equation(45) by a real orthogonal matrix and a real block diagonal matrix with 1×1 and 2×2 blocks, [10]. We can use equations(45) to specify various eigenvalue distributions and eigenvector spaces. In this paper we focus on diagonalizable test matrices. In [10] where we study the convergence of iterative procedures for $Ax = b$, we also consider defective matrices.

6.2. Some Properties of These Test Matrices

Theorem 6.1 states that it is sufficient to consider eigenvector matrices of the form ΣV^H where Σ is a positive diagonal matrix and V is a unitary matrix. Since for any unitary V and diagonal Λ, $B = V^H \Lambda V$ is normal, the following lemma is a simple restatement of theorem 6.1 in terms of normal matrices.

Lemma 6.1. *Let A be diagonalizable and defined by equation(45) for some choice of Σ , unitary V, and Λ, then A is a positive diagonal similarity transformation of the normal matrix*

$$B = V^H \Lambda V. \tag{47}$$

If all of the eigenvalues of A are real, then A is a positive diagonal similarity transformation of a Hermitian matrix.

We observe that the matrices in equation(45) are invariant under any scaling of the Σ matrices. In our test problems we could for example choose a scale γ such that we

$$\min \left(\|\gamma\Sigma - I\|_2^2 + \|\gamma^{-1}\Sigma^{-1} - I\|_2^2 \right) . \tag{48}$$

If $\sigma_1 > 1$ and $\sigma_n < 1$, we can approximate this minimization by letting $\gamma^4 {\sigma_1}^2 = 1/{\sigma_n}^2$. This suggests a scale of $\gamma = 1/\sqrt{\sigma_1 \sigma_n}$ for which the scaled singular values and the inverses of the scaled singular values lie in the same interval. We used this scale in each example of the form equation(45) which we considered.

7. Arnoldi versus Nonsymmetric Lanczos Methods

7.1. A Property of Arnoldi Methods

The following lemma may indicate that an Arnoldi method should perform well when applied to a matrix which is nearly normal. Theorem 7.1 gives a characterization of any normal matrix A in terms of its Hermitian-skew Hermitian decomposition. For a proof, see for example, [10]. We use λ^C to denote an eigenvalue of C.

Theorem 7.1. *Let A be any $n \times n$ matrix. Define $M^A \equiv (A + A^H)/2$ and $N^A \equiv (A - A^H)/2$. A is normal if and only if $M^A N^A = N^A M^A$. Equivalently, A is normal if and only if each eigenvector of A is also an eigenvector of M^A and of N^A, and each eigenvalue of A is of the form $\lambda^A = \lambda^{M^A} + i\lambda^{N^A}$ for some orderings of the eigenvalues of A, M^A, and N^A.*

Theorem 7.1 states that if A is normal we can compute its eigenvectors and eigenvalues by Hermitian eigenvalue computations on M^A and iN^A and and Hermitian eigenvector computation on either M^A or N^A. In other words, the Hermitian-skew Hermitian decomposition is the natural decomposition when the matrix is normal. Lemma 7.1 states that for any A, and at each stage of an Arnoldi process, the Hermitian-skew Hermitian decomposition of A is preserved in the Arnoldi orthogonal projection matrices H_k. Therefore, at each stage we are implicitly generating orthogonal projection matrices for M^A and N^A.

Lemma 7.1. *(Exact Arithmetic): Let A be any $n \times n$ matrix. Apply the Arnoldi recursions to A, generating orthogonal projection matrices $H_k = V_k^H A V_k$, $k = 1, \ldots, m$. Then*

$$M^{H_k} = V_k^H M^A V_k \text{ and } N^{H_k} = V_k^H N^A V_k \tag{49}$$

are respectively, orthogonal projection matrices for M^A and N^A, on the Krylov subspaces.

7.2. A Property of Nonsymmetric Lanczos Methods

For any A defined by equation(45) and the corresponding normal matrix B defined in lemma 6.1, we derive relationships between the Lanczos matrices, the residual norms and the error norms of corresponding Ritz vectors generated by applying the nonsymmetric Lanczos procedure to both A and B when the starting vectors satisfy the relationships specified in this theorem.

This theorem says that to understand the eigenvalue convergence of the nonsymmetric Lanczos procedure when it is applied to any diagonalizable matrix, it is sufficient to understand the behavior of this procedure when it is applied to normal matrices with certain well-chosen starting vectors. However, the corresponding residual and error norms for the nonnormal and the normal problem are not equal. Moreover, this is not a practical result since to generate the correct starting vectors we would need to know Σ. We note that one cannot prove a similar result for the Arnoldi procedure. Moreover, if the eigenvalue convergence is monitored by estimates of the residual norms for the Ritz vectors, we would not see this invariance. The proof is analogous to the proof of Theorem 5.3.

Theorem 7.2. *(Exact arithmetic): Let $A = \Sigma B \Sigma^{-1}$ where B is a normal matrix. For $k = 1, \ldots, K$, let T_k^A , T_k^B and V_k^A, W_k^A, V_k^B, and W_k^B denote respectively, Lanczos matrices and Lanczos vectors, obtained by applying the complex symmetric variant of the nonsymmetric Lanczos recursions to $\{B, v_1, w_1\}$ and to $\{A, \Sigma v_1, \Sigma^{-1} w_1\}$. Then for each k, $T_k^A, = T_k^B$, $V_k^A = \Sigma V_k^B$, and $W_k^A = \Sigma^{-1} W_k^B$.*

For any $T_k u = \mu u$, the corresponding right and left Ritz vectors satisfy $z_\mu^{RA} = \Sigma z_\mu^{RB}$ and $z_\mu^{LA} = \Sigma^{-1} z_\mu^{LB}$. Let λ be the eigenvalue of B closest to μ, and let x and y be corresponding unit right and left eigenvectors, then the right error norms satisfy

$$\|e_k^{RB}\| = \|z_k^{RB} - x\| = \|\Sigma^{-1}(z_k^{RA} - \Sigma x)\| \neq \|e_k^{RA}\|. \tag{50}$$

Similar relationships exist for the right and the left residual norms and the left error norm.

8. Numerical comparisons, Arnoldi versus Nonsymmetric Lanczos

We are interested in studying the convergence behavior of Arnoldi and nonsymmetric Lanczos eigenvalue procedures as we vary the spectral properties of the matrix A. These methods are invariant under shifts. In exact arithmetic either method when applied to A and $A + \sigma I$ for any shift σ will generate the same vectors, and matrices which differ only by the specified shift. Therefore , without loss of generality, we can restrict ourselves to test matrices with eigenvalues whose real parts are all positive.

We have written a suite of MATLAB [19] codes which allow the user to generate and regenerate test matrices of the form given in equation(45). The user can also call either a $Ax = b$ routine or a basic real or complex Arnoldi eigenvalue routine. The matlab routines will write a test matrix to a file which can then be read into the codes for the nonsymmetric Lanczos procedure which are written in Fortran. Output from any of these computations can be plotted using MATLAB routines written specifically for such tests. There are also codes for generating pseudospectra of test matrices and contour plots of the pseudospectra. These codes are described in [10].

We considered 9 different test problems. We applied a complex Arnoldi method with reorthogonalization and our complex symmetric variant of the nonsymmetric Lanczos method with no re-biorthogonalization to each problem and computed eigenvalue approximations and true errors for various size Arnoldi and Lanczos matrices. In each test both methods used the same randomly-generated starting vector. In the nonsymmetric Lanczos tests we set $w_1 = v_1$.

Ideally we would like to be able to characterize the convergence behavior in terms of the nonnormality of the test matrix. To be able to make comparisons across different test matrices, we would need to eliminate the effects of using different starting vectors on different problems, and we have not done that. These tests however indicate, at least for these test matrices, that in order to completely characterize the convergence behavior of the eigenvalue approximations generated by either method, we need not only the singular values of an eigenvector matrix but also knowledge of the relative sizes of entries in the associated right singular vector matrix. They also indicate that small but nonzero entries in strategic positions in the singular vector matrix may interact with the singular values to yield an eigenvector matrix whose columns are even more poorly conditioned than the singular values of the eigenvector matrix suggest. We will develop these ideas in another paper.

8.1. Test Problems Used

We consider 9 different examples. For examples $1-8$, $n = 48$, and all have the same eigenvalue distribution. The first two examples are the Grcar matrix $n = 48$, [29], and its transpose. We denote this matrix by $Gr48true$. This matrix is real. The subdiagonal entries are -1. The diagonal entries and the entries on the first 3 superdiagonals are 1. All other entries are 0. Each eigenvalue is simple. The eigenvalue distribution is depicted by o's in Fig. 1.

We used the MATLAB function EIG to compute right eigenvectors X_G of $Gr48true$ and eigenvalues Λ_G. We then used the MATLAB function SVD to compute a singular value decomposition $X_G = U_X \Sigma_X V_X^H$ of X_G. The singular values Σ_X of X_G vary from 5.3×10^{-8} to 4.5. We scaled Σ_X using $s \equiv 1/\sqrt{\sigma_n \sigma_1}$ and set $\Sigma_s = s\Sigma_X$.

Examples 3 and 4 are normal matrices.

$$\mathrm{Gr48id2} \equiv V_X^H \Lambda_G V_X \text{ and } \mathrm{Gr48id3} \equiv V_R^H \Lambda_G V_R, \tag{51}$$

where V_R was generated randomly using a uniform distribution and orthonormalized using the MATLAB function ORTH.

Examples 5 and 6 were nonnormal matrices.

$$\mathrm{Gr482} \equiv \Sigma_s V_X^H \Lambda_G V_X {\Sigma_s}^{-1} \text{ and } \mathrm{Gr483} \equiv \Sigma_s V_R^H \Lambda_G V_R {\Sigma_s}^{-1}, \tag{52}$$

where V_R is the unitary matrix used in examples 3 and 4.

Examples 7 and 8 were also nonnormal.

$$\text{Gr48a2r2b02} \equiv \Sigma_a V_X^H \Lambda_G V_X {\Sigma_a}^{-1} \text{ and } \text{Gr48a2r2b03} \equiv \Sigma_a V_R^H \Lambda_G V_R {\Sigma_a}^{-1}, \tag{53}$$

where V_R is the unitary matrix used in examples 4 and 6. Σ_a was obtained by computing the 48 points in the interval $[10^{-4}, 1]$ whose logarithms of their square roots are equally spaced in the interval $[10^{-2}, 1]$ and then scaling these numbers by the reciprocal of the square root of the product of the smallest and the largest values. The scaled values in Σ_a varied from 10^{-2} to 10^2.

Example 9 is of size $n = 32$. This matrix is bidiagonal with the entries $1/\sqrt{k}$ on and above the main diagonal.

In Figs 1-4 and 7-14, we use o's to denote the true eigenvalues, x's to denote the Arnoldi approximations, and a combined + and x to denote the Lanczos approximations. In Figs 5-6, we use x to denote the errors at the smallest size considered, a combined + and x at the next largest size, and + to denote the errors at the largest size considered. In Figs 15-16 we plot the sizes of entries in certain matrices. In each of these pictures the 'origin' is the $(1, 1)$ position in the matrix.

8.2. Observations

We make several comments based upon the tests we have run. However, more testing is needed before any definite conclusions can be made. We believe that both Arnoldi and nonsymmetric Lanczos methods can be effective for computing eigenvalues (and eigenvectors) of many nonsymmetric matrices. We observe that, for examples 1-8, the approximations to the left tails of the eigenvalue distribution converge prior to convergence of the leading edge.

Example 1 is well-conditioned in the sense that its singular values vary from .922 to 3.23. This matrix is however very nonnormal [29]. In Figs. 1-2, for $m = 24$ and $m = 32$, we plot the true eigenvalues, the complex Arnoldi approximations, and the nonsymmetric Lanczos approximations . We observe that in both figures the nonsymmetric Lanczos approximations are closer to the actual eigenvalue curve than the Arnoldi approximations. At $m = 32$ the Lanczos approximations have correctly identified the left trailing eigenvalues. The Arnoldi estimates are still on the wrong side of the origin.

We ran the same tests using the real Arnoldi method and a real randomly-generated starting vector in both the Arnoldi and the nonsymmetric Lanczos method. See Figs. 3-4. These plots are very similar to Figs. 1-2. In corresponding tests on the transpose of $Gr48true$, the order in which the eigenvalues converged changed slightly but the overall plots were very similar to the plots for $Gr48true$.

In Figs 5-6 we plot true errors for the eigenvalue approximations generated by each method for different size Arnoldi and Lanczos matrices. The x-axis corresponds to the number of an eigenvalue when the eigenvalues $\lambda_j, 1 \leq j \leq n$, are ordered from algebraically-smallest to algebraically-largest real parts. With this ordering, the numbering moves from left to right along the tails of the eigenvalue distribtuion for the first 15 complex pairs of eigenvalues. It then moves to the center of the curved front edge of the distribution, moving upwards and downward from that center through λ_{36}. $\lambda_{37,38}$ complete the left arms of the distribution, afterwhich the numbering resumes along the front edge, continuing up into the top(bottom) cusp. Comparing Figs. 5 and 6 at $m = 40$, we observe that the Lanczos errors for the left portion of the spectrum are smaller than the corresponding Arnoldi errors at that size. By $m = 48$ the nonsymmetric Lanczos procedure has an approximation to each eigenvalue. Good accuracy is obtained by $m = 60$.

Examples 3 and 4 are normal. Fig. 7 for example 3, and Fig. 8 for example 4, both for $m = 24$, indicate that the Arnoldi method has identified the shape of the spectrum. The nonsymmetric Lanczos iterates are not constrained by the normality and are more scattered than the Arnoldi iterates.

Examples 5 and 6 are not normal. The condition number of example 5 is 3.5 with the singular values ranging from .92 to 3.23. The singular values of the eigenvector matrix range from 5.28×10^{-8} to 4.5. The condition number of example 6 is 2.4×10^{14}. The singular values of example 6 range from 1.1×10^{-7} to 2.5×10^{7}. The singular values of the eigenvector matrix range from 5.29×10^{-8} to 4.6. If we selected the starting vectors according to Theorems 5.2-5.3, in exact arithmetic, the convergence of example 5 should be the same as that observed for example 1. Although we did not modify the starting vectors appropriately, the observed convergence is very similar. See Fig. 9. We would like to be able to make a statement about differences corresponding to V_X versus V_R. A precise statement would require us to take starting vectors with equal projections on corresponding eigenspaces. We did not do that. In fact we used the same starting vector on both examples and methods. If however we compare Figs 9 and 10, for $m = 32$, we observe better convergence for the nonsymmetric Lanczos procedure when $V = V_R$ versus $V = V_X$, along the front edge of the distribution.

In examples 7 and 8 we attempted to construct problems with normality *between* that of $Gr48true$ and $Gr48id$. The norm of example 7 is 2.82. The singular values of example 7 range from 1.16 to 2.82. The singular values of the MATLAB eigenvector matrix for example 7 range from 4×10^{-4} to 4. The norm of example 8 is 4.2×10^3. The singular values of example 8 range from 7.8×10^{-4} to 4.2×10^3. The condition number is 5.4×10^6. The singular values of the MATLAB eigenvector matrix for example 8 range from 4×10^{-4} to 4.

In Fig. 11 at $m = 32$ we observe that both the Arnoldi and the nonsymmetric Lanczos approximations have correctly identified the left tails of the distribution. There is however still a bulge on the front edge in the Arnoldi plot but it is smaller than in Fig. 9. In Fig. 12 corresponding to $V = V_R$, the bulge is gone. The differences between Figs. 11 and 12, and between Figs. 9 and 10, led us to examine the structure of the V_X^H matrix. We observed that the numerical shape of this matrix, in terms of the sizes of the nonzero entries in this matrix, enhances the deterimental effects of the small singular values on the resulting eigenvector matrix. For examples $5 - 8$ these enhancements affect portions along the eigenvalue distribution near the cusps. This structure also allows the matrix to 'conceal' its ill-conditioning. We can illustrate this type of effect more clearly by considering example 9 where even the singular values alone are sufficient to prevent finite precisions implementations of either procedure from working properly.

Example 9 is well-conditioned and all of the eigenvalues are real, ranging from .176 to 1.0. The condition number is 154. The singular values of example 9 range from 10^{-2} to 1.55. The singular values of the eigenvector matrix, however, range from 1.2×10^{-24} to 5.5. First consider Fig. 13 with $m = 24$. Both methods have computed the 3 right-most eigenvalues correctly but the other approximations form rings around the true eigenvalues. If we increase the size to $m = 32$, we observe in Fig. 14, that the 8 rightmost eigenvalues are approximated by both procedures but the other approximations for both methods form rings around the rest of the spectrum even though the Arnoldi vectors are fully orthogonal and $m = n$.

For this example only 16 singular values are larger than 10^{-10}. Therefore, we might expect the eigenvector matrix to behave as though its rank is less than 32, and it does. Fig. 15 is a contour plot of the base 10 logarithms of the absolute values of the entries of the right singular vector matrix. This matrix was computed by first using MATLAB to compute an eigenvector matrix for example 9 and then using MATLAB to compute the singular value decomposition of example 9. We observe that the upper left hand portion of this plot corresponds to small entries. Fig. 16 is a contour plot of the base 10 logarithms of the absolute values of the entries of the corresponding canonical eigenvector matrix. The figure indicates that the numerical rank of this matrix is less than 10. Thus, we can expect to be able to compute at most 10 of the eigenvalues of this example. The 'shape' of the singular vector matrix has combined with the singular values to make things even worse. Since the well-separated eigenvalues converge first, the actual convergence is from the right to the left.

We also note that w.r.t several of the examples, without scaling Σ, premature termination of the Arnoldi recursions occurred even with limited reorthogonalization. Therefore, for each test we used the scaling and at each step of the Arnoldi procedure incorporated total reorthogonalization of each intermediate Arnoldi vector w.r.t all preceding Arnoldi vectors. However, there was no re-biorthogonalization in the Lanczos method.

9. Summary

We obtained several results which may be useful in determining the convergence behavior of eigenvalue algorithms based upon Arnoldi and nonsymmetric Lanczos recursions. We derived a relationship between nonsymmetric Lanczos eigenvalue procedures and Arnoldi eigenvalue procedures. We demonstrated that the Arnoldi recursions preserve a property which characterizes normal matrices, and that if we could determine the appropriate starting vectors, we could mimic the nonsymmetric Lanczos eigenvalue convergence on a general diagonalizable matrix by its convergence on related normal matrices. Using a unitary equivalence for each of these Krylov subspace methods, we defined sets of test problems where we can easily vary certain spectral properties of the matrices. We used these and other test problems to examine the behavior of an Arnoldi and of a nonsymmetric Lanczos procedure. The results of these tests suggest that to completely characterize the behavior of these methods on nonnormal problems it is not sufficient to know the singular values of the eigenvector matrix. They also suggest a potential source of numerical difficulties for both types of methods. In addition, these tests suggest that a nonsymmetric Lanczos procedure may more readily identify key portions of the spectrum of a highly nonnormal matrix than an Arnoldi method. However, each step of a nonsymmetric Lanczos procedure requires a matrix-vector multiply by both A and by A^T. Moreover, the eigenvalue approximations generated by the Arnoldi method typically exhibit less scatter than those generated by a Lanczos method.

There are many open questions regarding the behavior of either or both of these types of methods, and even questions about the design of tests for comparisons both between methods and for a given method. In [9] we examine similar questions in the context of the problem $Ax = b$.

References

[1] Z. Bai, *Error analysis of the Lanczos algorithm for nonsymmetric eigenvalue problem*, Math. Comp., 62(205) (1994), pp. 209–226.

[2] M. Bennani and T. Braconnier, *Comparative behaviour of eigensolvers on highly nonnormal matrices.* CERFACS Technical Report TR-PA-94-23 (1994). CERFACS, Toulouse, France.

[3] T. Braconnier, *Importance of the orthogonalization scheme used for eigensolvers applied to nonnormal matrices*, CERFACS Technical Report TR-PA-94-20, (1994). CERFACS, Toulouse, France.

[4] T. Braconnier, F. Chatelin and V. Fraysse, *The influence of large nonnormality on the quality of convergence of iterative methods in linear algebra*, CERFACS Technical Report TR-PA-94-07 (1994). CERFACS, Toulouse, France.

[5] F. Chatelin, V. Fraysse, and T. Braconnier, *Computations in the neighbourhood of algebraic singularities*, Numer. Funct. Anal. and Optimiz., 16(3&4) (1995), pp. 287–302.

[6] F. Chaitin-Chatelin, *Is nonnormality a serious difficulty?*, CERFACS Technical Report TR-PA-94-18, 1994. CERFACS, Toulouse, France.

[7] F. Chaitin–Chatelin, *Finite precision computations, reliability of numerical software*, CERFACS Technical Report TR-PA-94-05 (1994). CERFACS, Toulouse, France.

[8] F. Chatelin and V. Fraysse, numerical illustrations by T. Braconnier, *Qualitative computing, elements for a theory for finitie precision computation*, CERFACS Technical Report TR-PA-93-12 (1993). CERFACS, Toulouse, France.

[9] J. Cullum, *GMRES versus QMR/BICG methods for solving $Ax = b$*, to appear IBM Research Report, December 1995, IBM Research, Yorktown Heights, NY.

[10] J. Cullum, *Testing iterative methods for nonsymmetric matrices*, to appear, IBM Research Report, January 1996, IBM Research, Yorktown Heights, NY.

[11] J. Cullum, *Lanczos algorithms for large scale symmetric and nonsymmetric matrix eigenvalues problems*, Proceedings of the Cornelius Lanczos International Centenary Conference, eds. J. David Brown, Moody T. Chu, Donald C. Ellison, and Robert J. Plemmons, SIAM, Philadelphia, PA (1994) pp. 11–31.

[12] J. Cullum, W. Kerner, and R. Willoughby, *A generalized nonsymmetric Lanczos procedure*, Comput. Phys. Comm., 53 (1989), pp. 19–48.

[13] J. Cullum and R. Willoughby, *A practical procedure for computing eigenvalues of large sparse nonsymmetric matrices*, in Large Scale Eigenvalue Problems, J. Cullum and R. Willoughby (eds.), North-Holland, 1986, pp. 193–240.

[14] J. Cullum and R. A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations*, Vol. 1, Theory, Progress in Scientific Computing Vol. 3, eds. S. Abarbanel et al., Birkhaüser, Basel, Switzerland (1985).

[15] V. Druskin and L. Knizhnerman, *Krylov subspace approximation of eigenpairs and matrix functions in exact and computer arithmetic*, Schlumberger-Doll Research Note, October 20, 1992, Schlumberger-Doll, Ridgefield, CT; to appear in Numerical Linear Algebra.

[16] R. W. Freund, M. H. Gutknecht, and N. M. Nachtigal, *An implementation of the look-ahead Lanczos algorithm for non-Hermitian matrices*, SIAM J on Scientific and Statistical Computing, 14,(1993) pp. 137–58.

[17] R. W. Freund, G. H. Golub, Noel Nachtigal, *Iterative solution of linear systems*, Acta Numerica., 1 (1992), pp.57–100.

[18] R. Freund and N. Nachtigal, *QMR: a quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math., 8 (1992), pp. 43–71.

[19] C. Moler et al., *MATLAB User's Guide*, MathWorks, Inc., 24 Prime Park Way, Natick, MA (1992).

[20] C. C. Paige, *Accuracy and effectiveness of the Lanczos algorithm for the symmetric eigenproblem*, Lin. Alg. Appl. 34 (1980), pp. 235–258.

[21] C. C. Paige, *Error analysis of the Lanczos algorithm for tridiagonalizing a symmetric matrix*, J. Inst. Math. Appl. 18 (1976), pp. 341–349.

[22] C. C. Paige, *The computation of eigenvalues and eigenvectors of very large sparse matrices*, Ph.D. thesis, London University, Institute of Computer Science, London, England, 1971.

[23] B. N. Parlett, *Reduction to tridiagonal form and minimal realizations*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 567–593.

[24] B. N. Parlett and J. K. Reid, *Tracking the progress of the Lanczos algorithm for large symmetric eigenproblems*, IMA J. Num. Anal., 1 (1981), pp. 135–155.

[25] B. N. Parlett, D. R. Taylor, and Z.A. Liu, *A look-ahead Lanczos algorithm for unsymmetric matrices*, Math. Comp., 44 (1985), pp. 105–124.

[26] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, Manchester University Press, Manchester, UK, 1992.

[27] L. N. Trefethen, *Spectra and Psuedospectra: The Behavior of Non–Normal Matrices and Operators*, book in preparation.

[28] L. N. Trefethen, A. E. Trefethen, S. C. Reddy, and T. A. Driscoll, *Hydrodynamic stability without eigenvalues*, Science, 261 (1993), 578–584.

[29] L. N. Trefethen, *Pseudospectra of matrices*, in D. F. Griffiths and G. A. Watson, eds., Numerical Analysis 1991, Longman Scientific and Technical, Harlow, UK, 1992.

[30] J. M. van Kats and H. A. van der Vorst, *Automatic monitoring of Lanczos-schemes for symmetric or skew-symmetric generalized eigenproblems*, Report TR-7, Academisch Computer Centrum, Utrecht, The Netherlands, 1977.

[31] H. F. Walker, *Implementations of the GMRES method using Householder transformations*, SIAM J. Sci. Stat. Comput. 9(1) (1988), pp. 152–163.

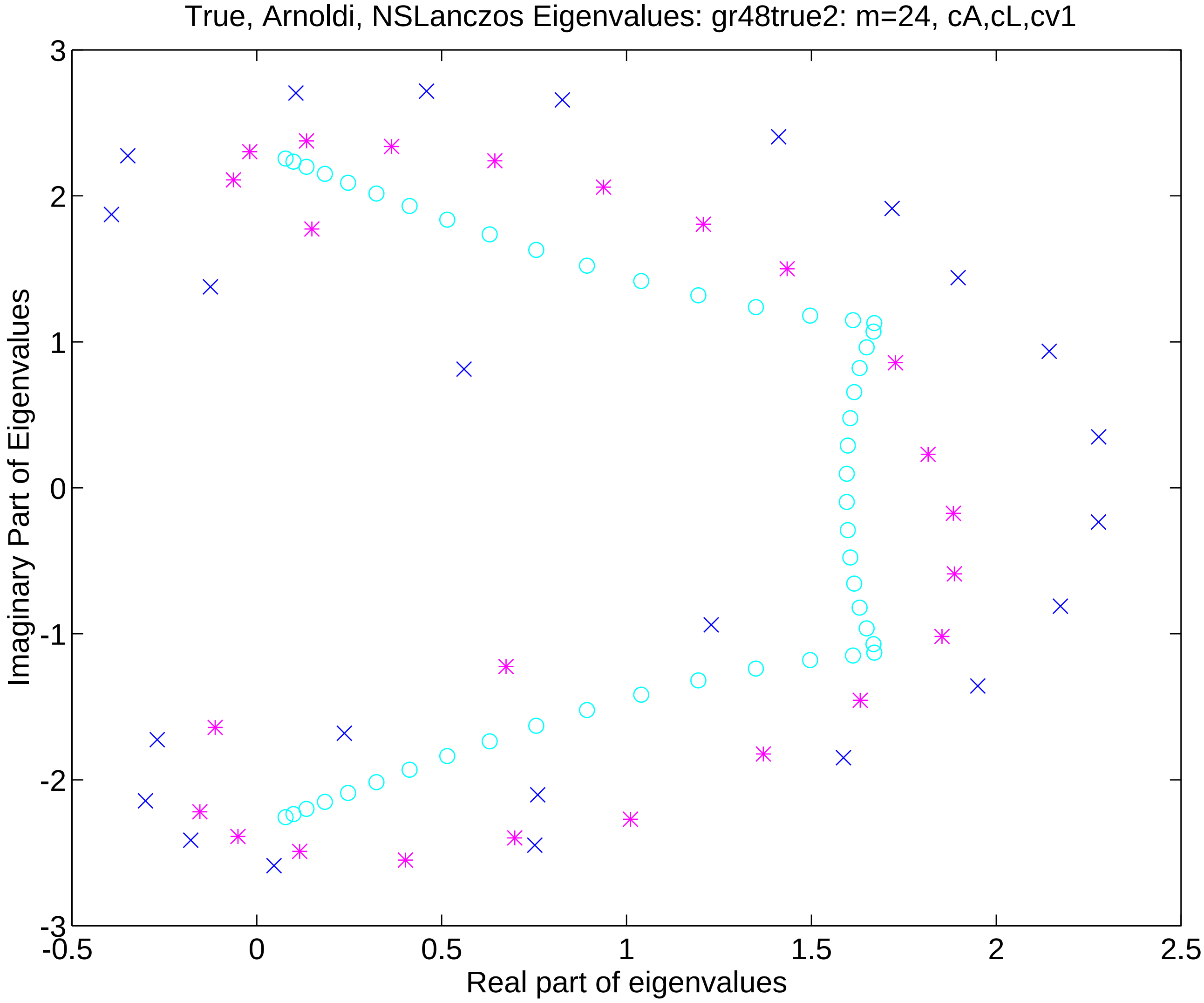

FIG. 1. *Example 1: Complex Arnoldi, NS Lanczos: True Eigenvalues and Estimates, m=24*

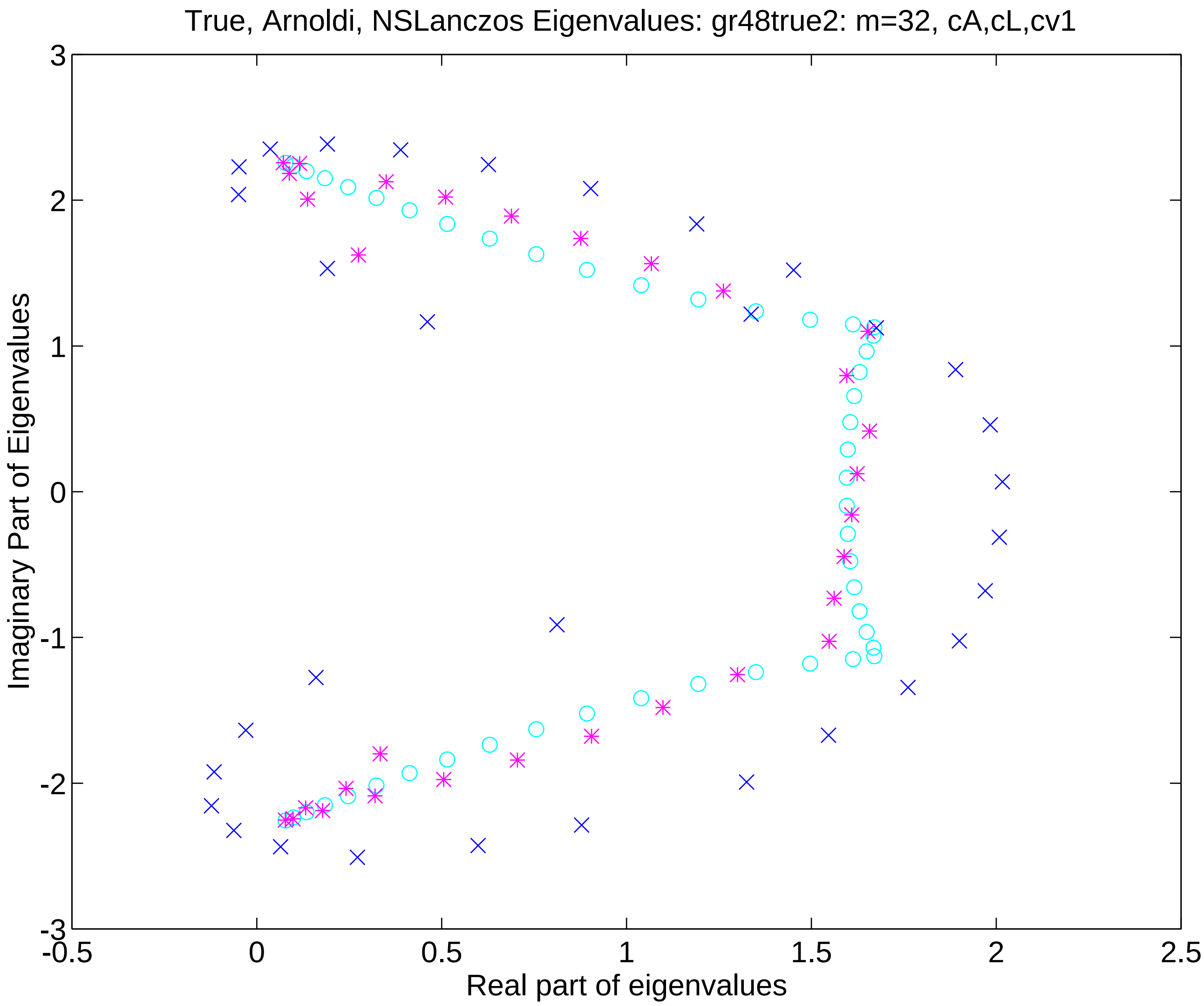

FIG. 2. *Example 1: Complex Arnoldi, NS Lanczos: True Eigenvalues and Estimates, m=32*

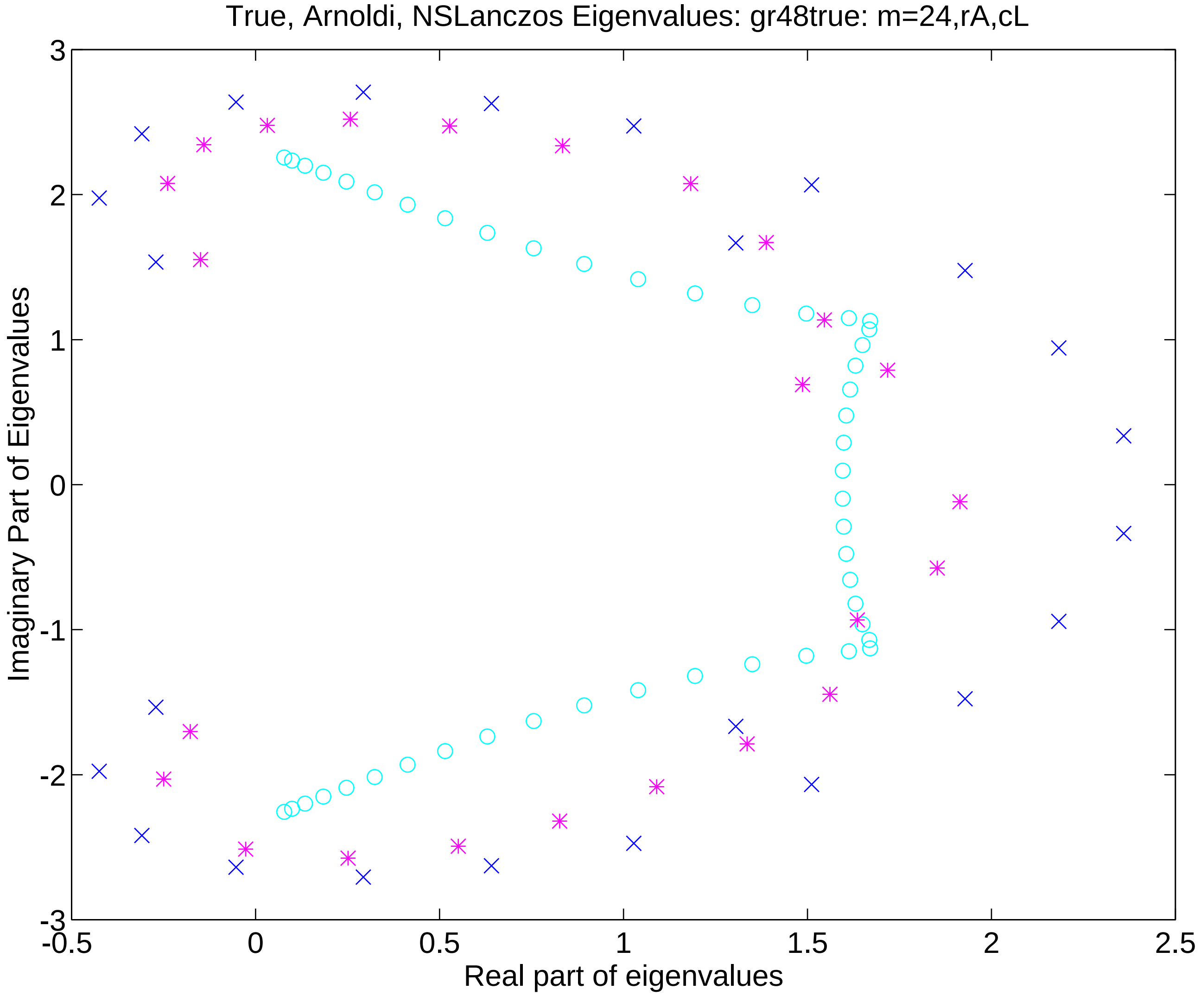

FIG. 3. *Example 1: Real Arnoldi, NS Lanczos: True Eigenvalues and Estimates, m=24*

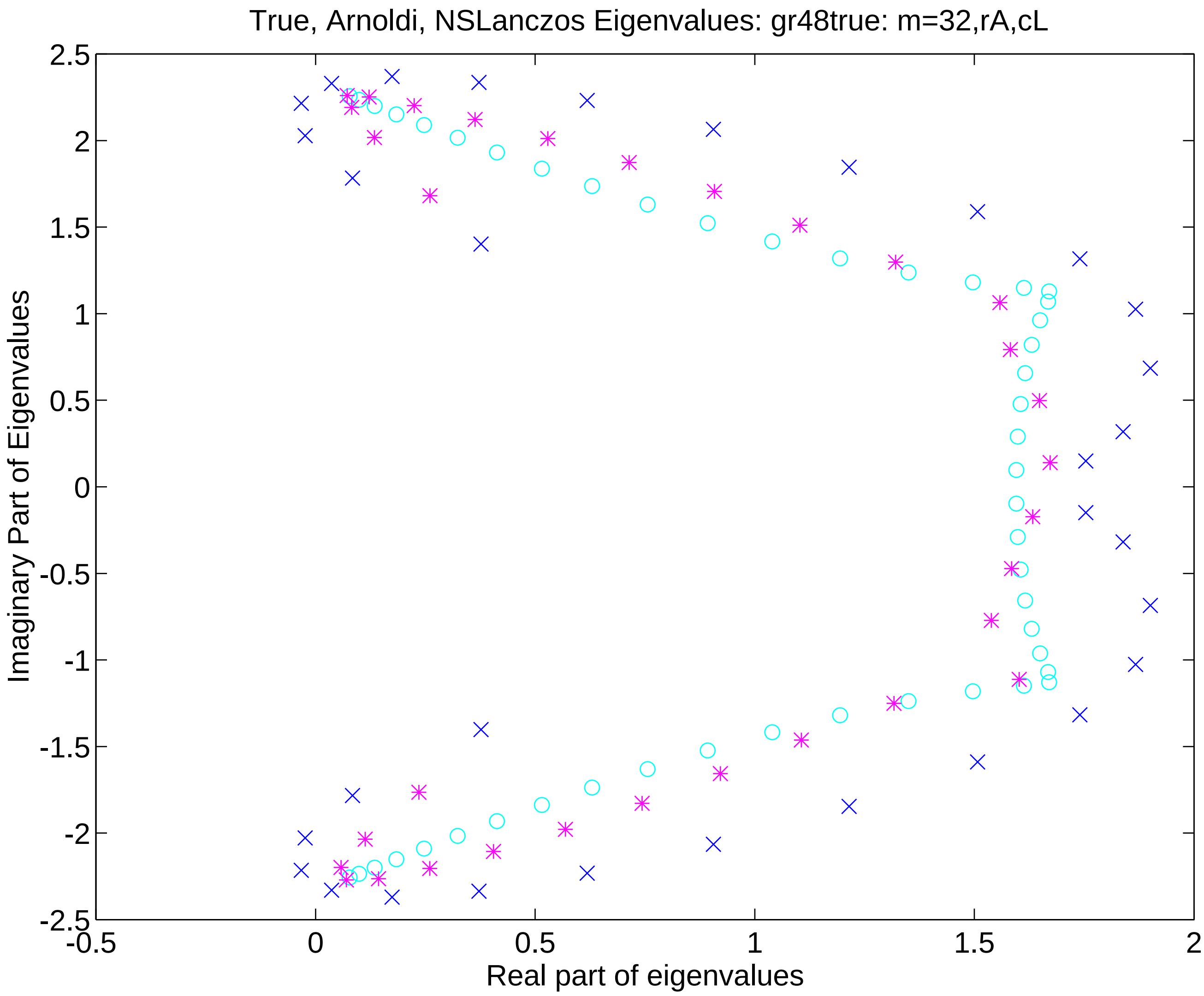

FIG. 4. *Example 1: Real Arnoldi, NS Lanczos: True Eigenvalues and Estimates, m=32*

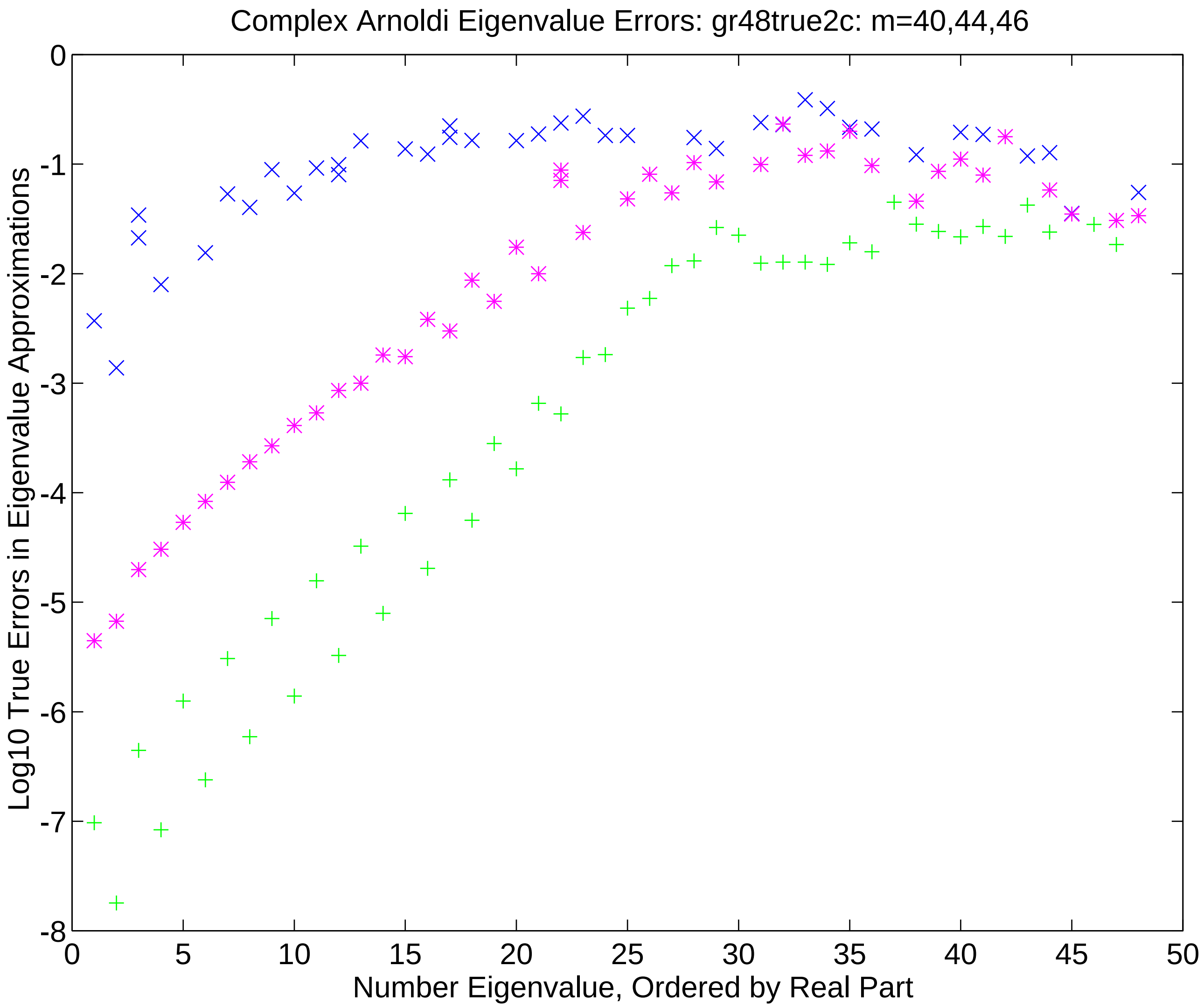

FIG. 5. *Example 1: Complex Arnoldi: Logarithms True Errors versus Eigenvalue Number*

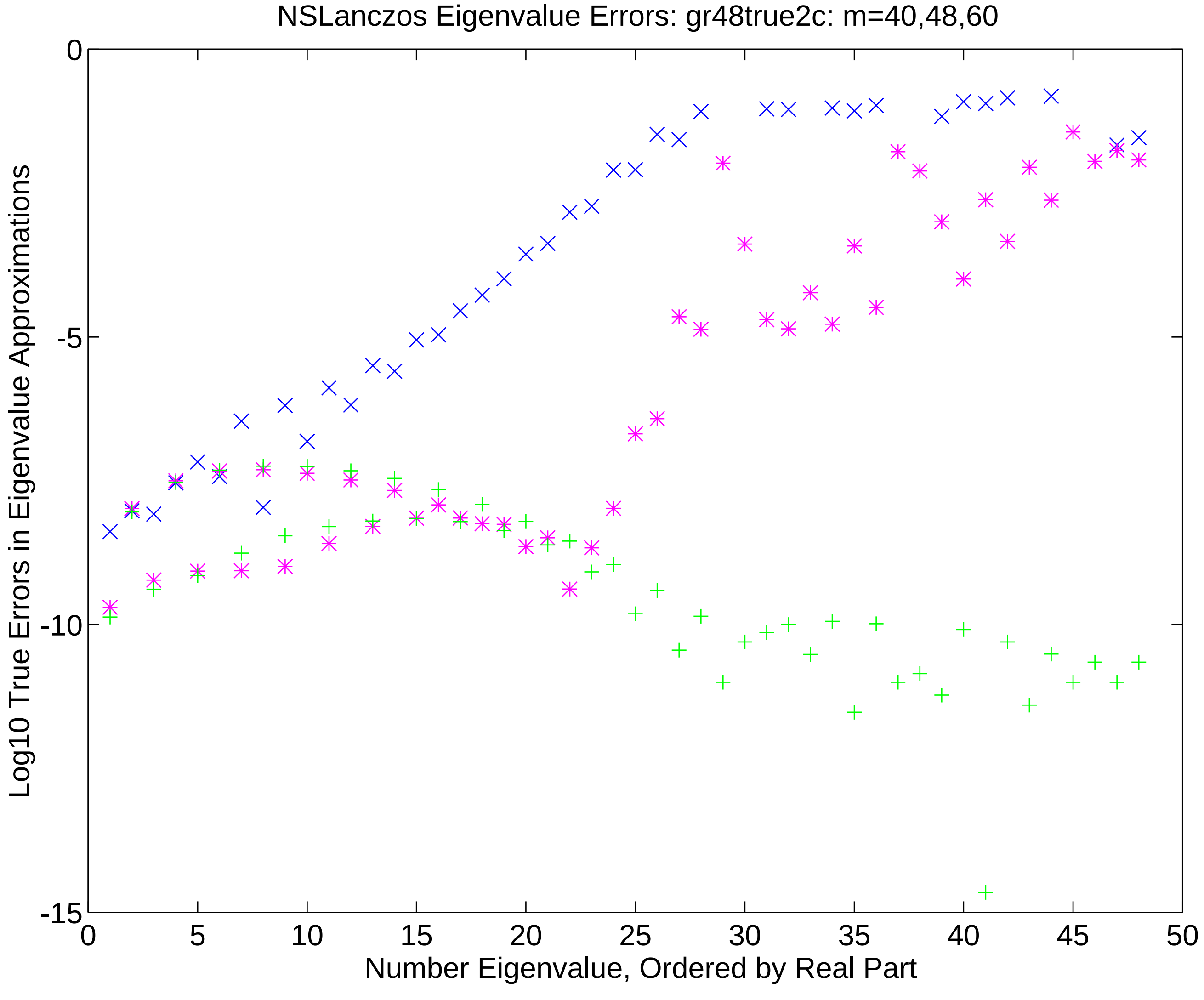

FIG. 6. *Example 1: NS Lanczosi: Logarithms True Errors versus Eigenvalue Number*

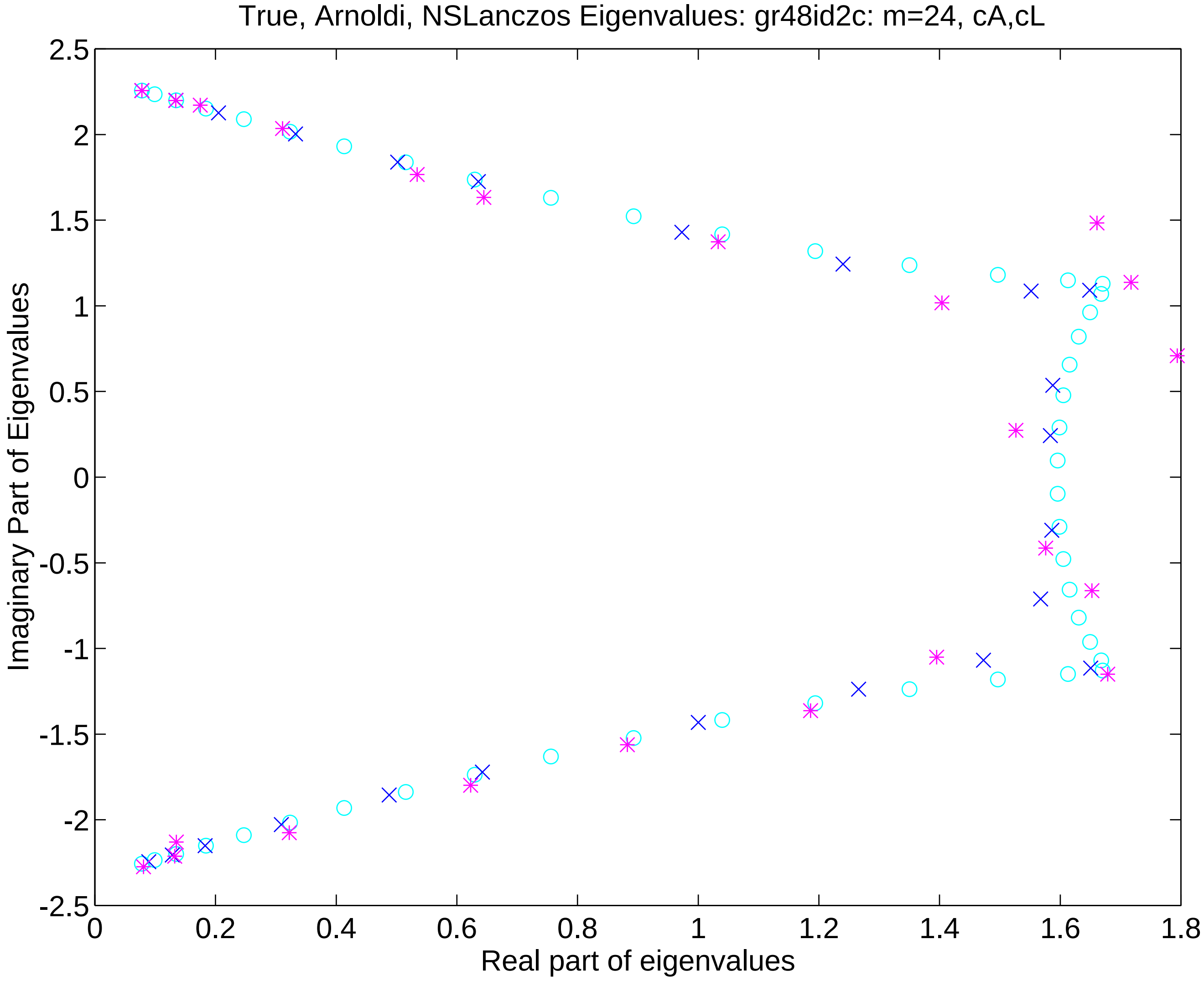

FIG. 7. *Example 3: Complex Arnoldi, NS Lanczos: True Eigenvalues and Estimates, m=24*

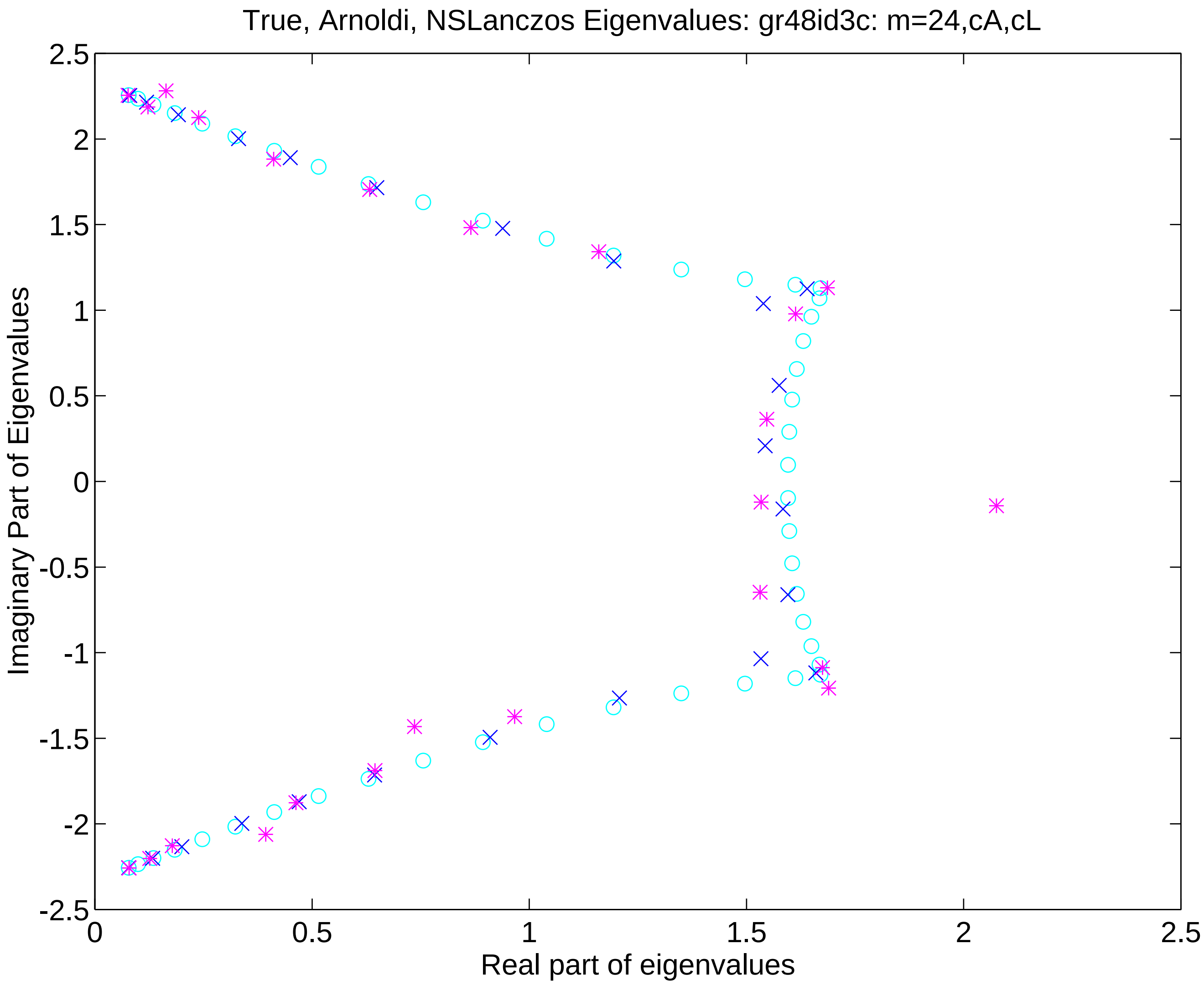

FIG. 8. *Example 4: Complex Arnoldi, NS Lanczos: True Eigenvalues and Estimates, m=24*

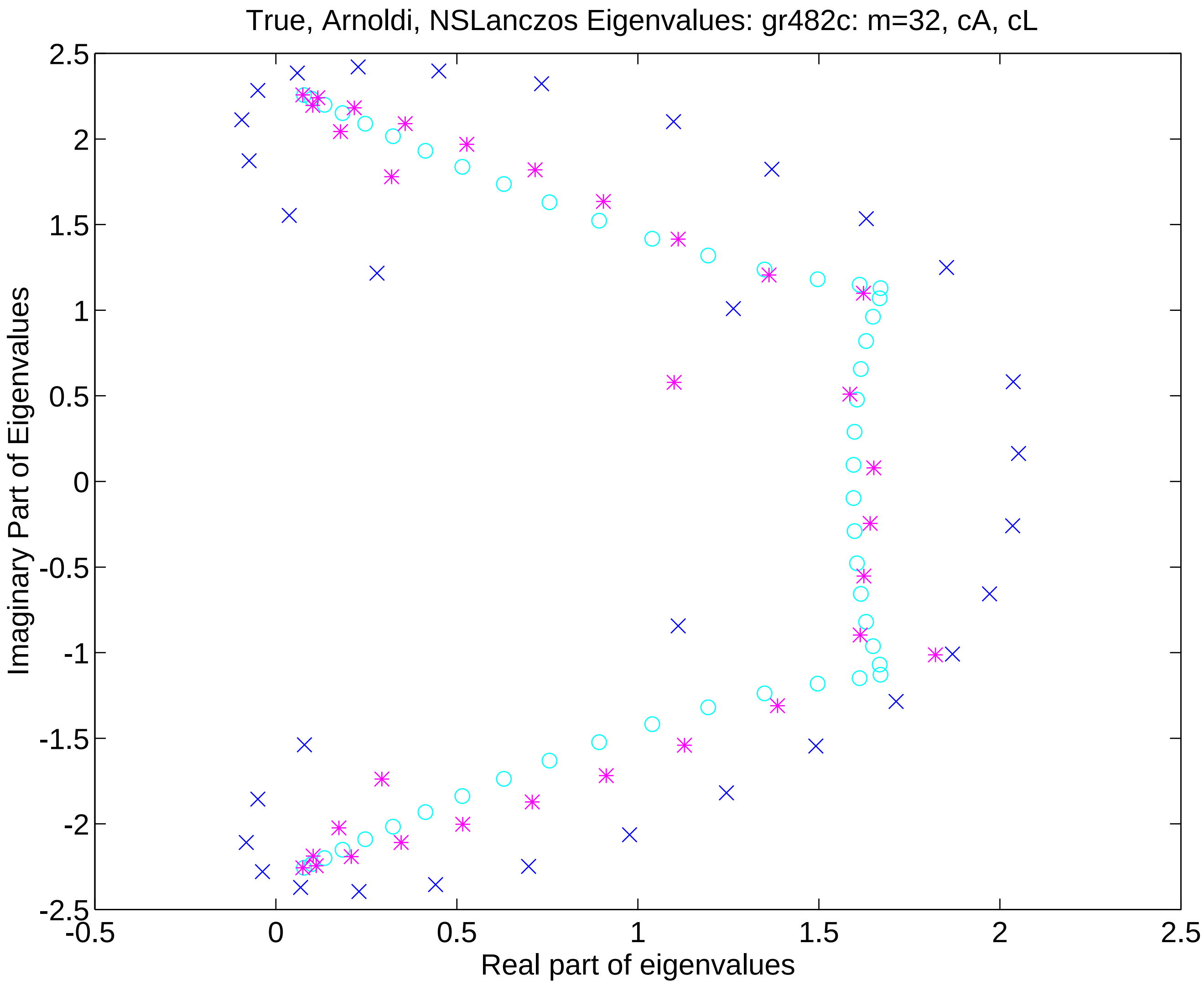

FIG. 9. *Example 5: Complex Arnoldi, NS Lanczos: True Eigenvalues and Estimates, m=32*

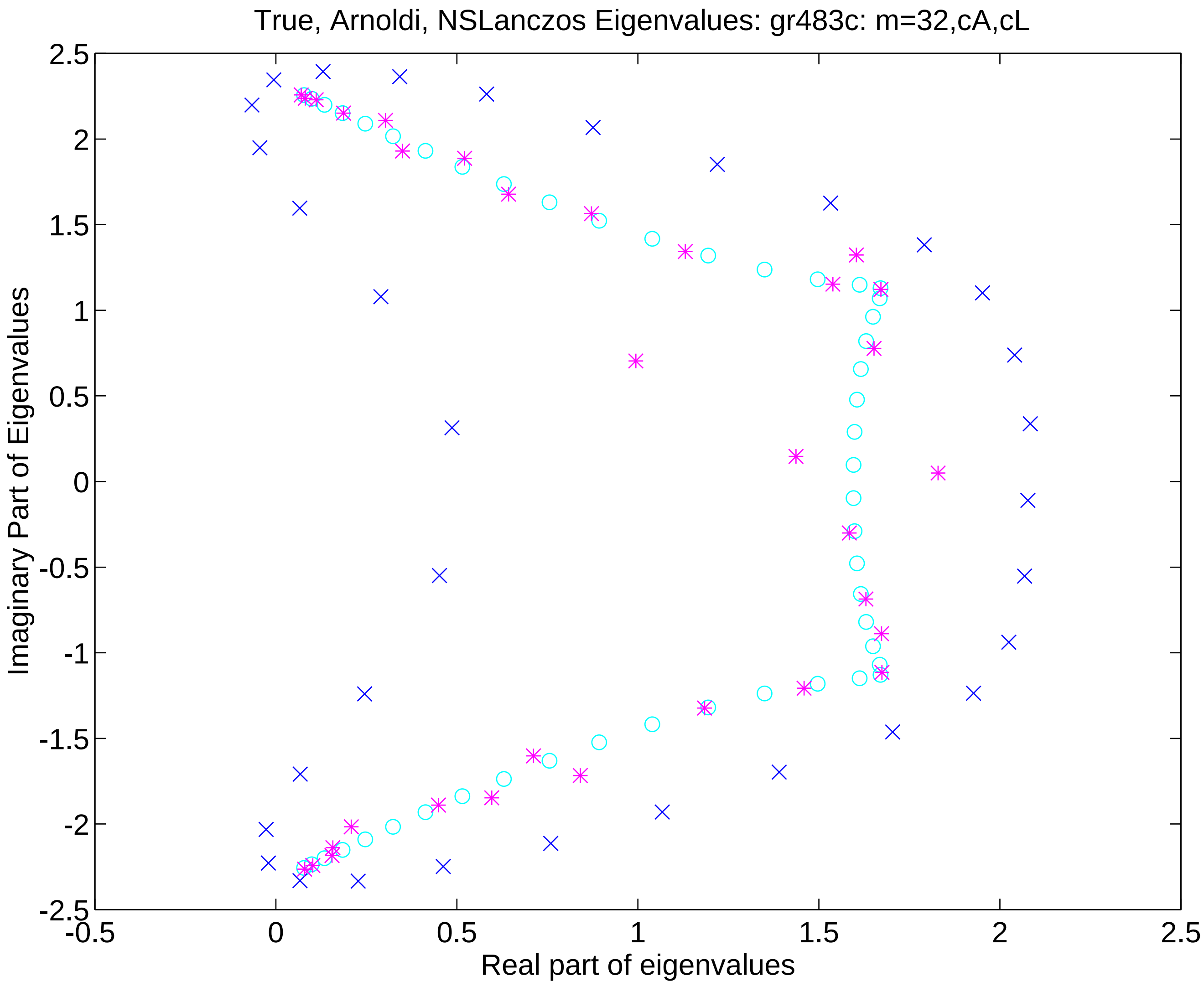

FIG. 10. *Example 6: Complex Arnoldi, NS Lanczos: True Eigenvalues and Estimates, m=32*

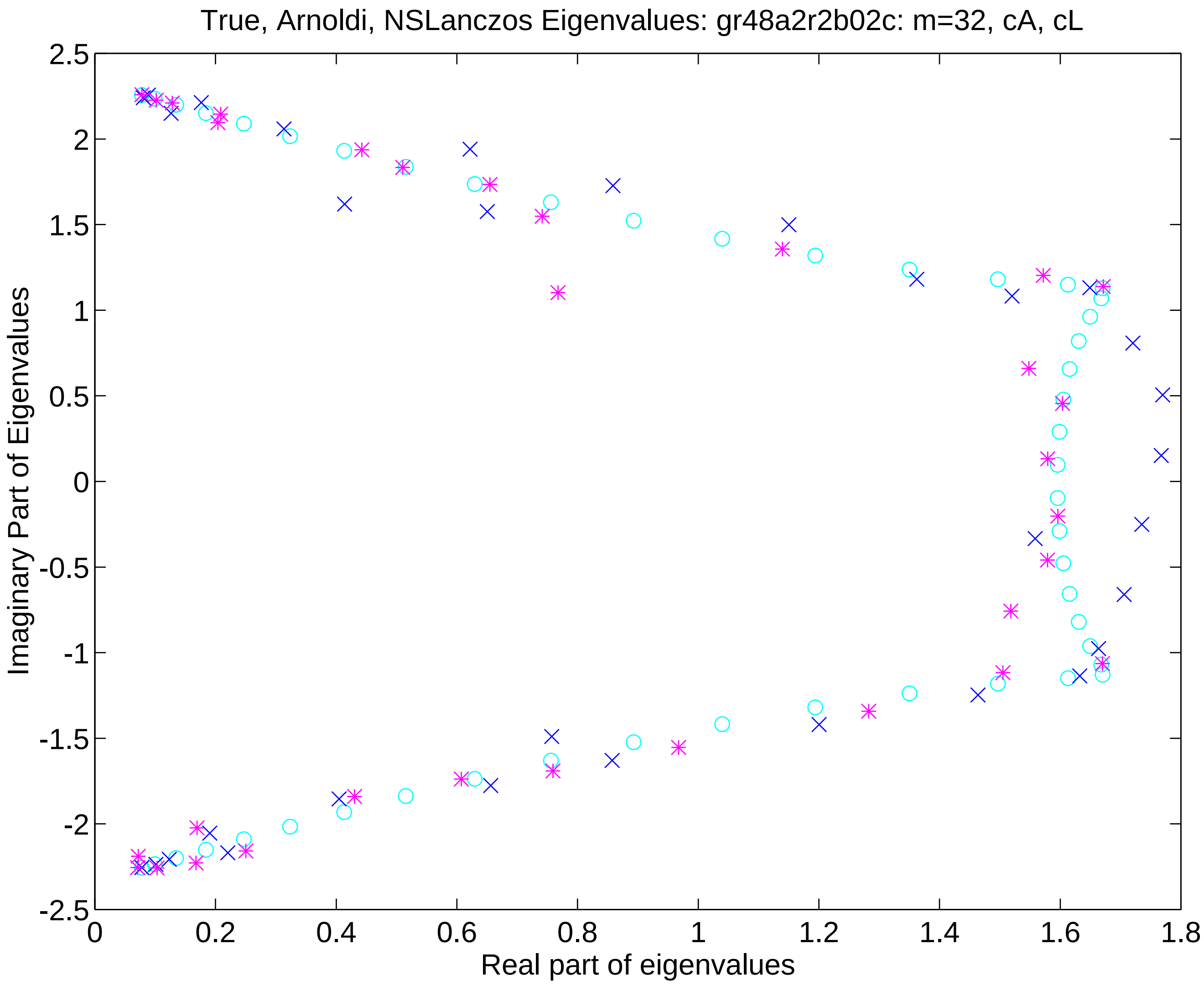

Fig. 11. *Example 7: Complex Arnoldi, NS Lanczos: True Eigenvalues and Estimates, m=32*

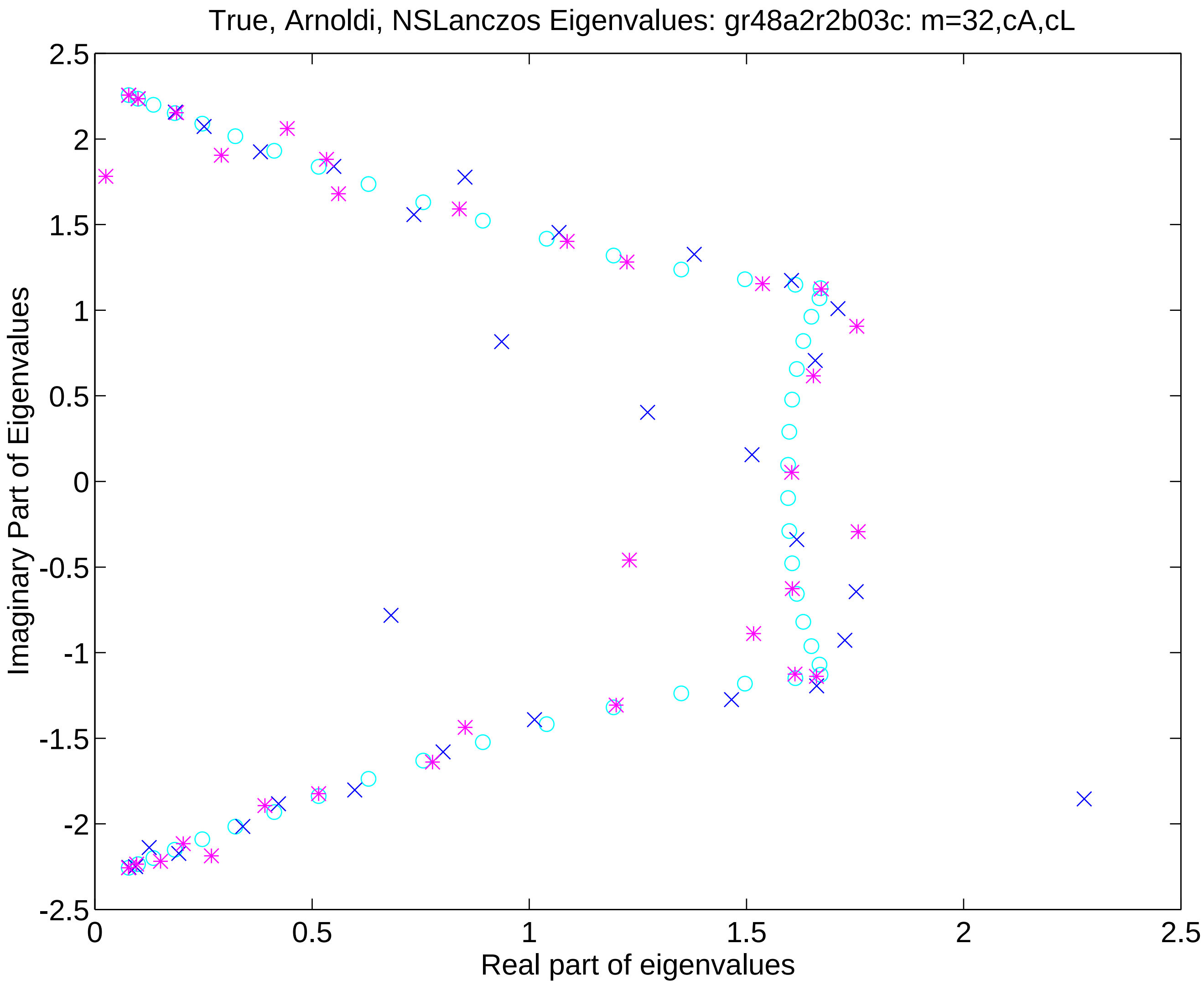

FIG. 12. *Example 8: Complex Arnoldi, NS Lanczos: True Eigenvalues and Estimates, m=32*

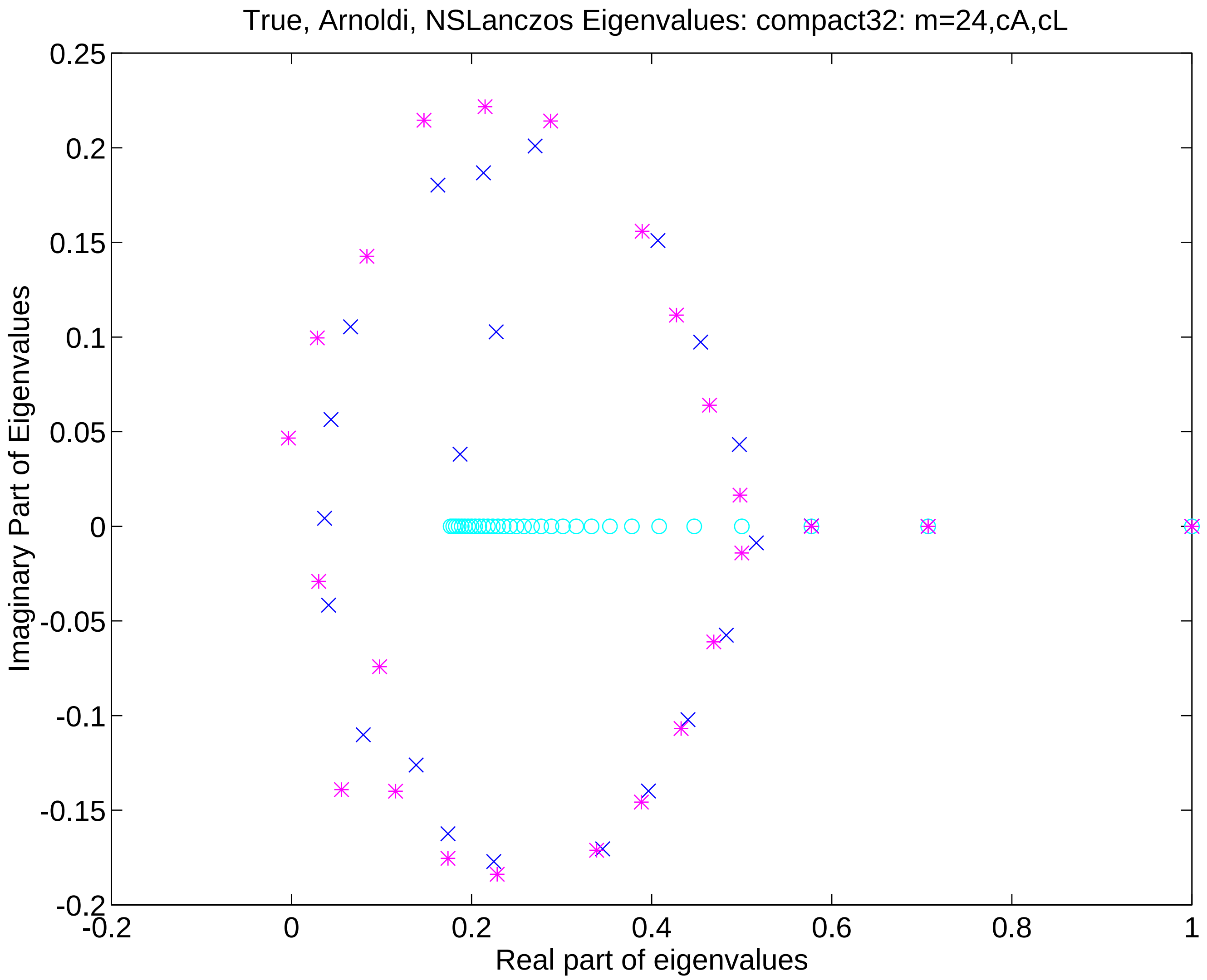

FIG. 13. *Example 9: Complex Arnoldi, NS Lanczos: True Eigenvalues and Estimates, m=24*

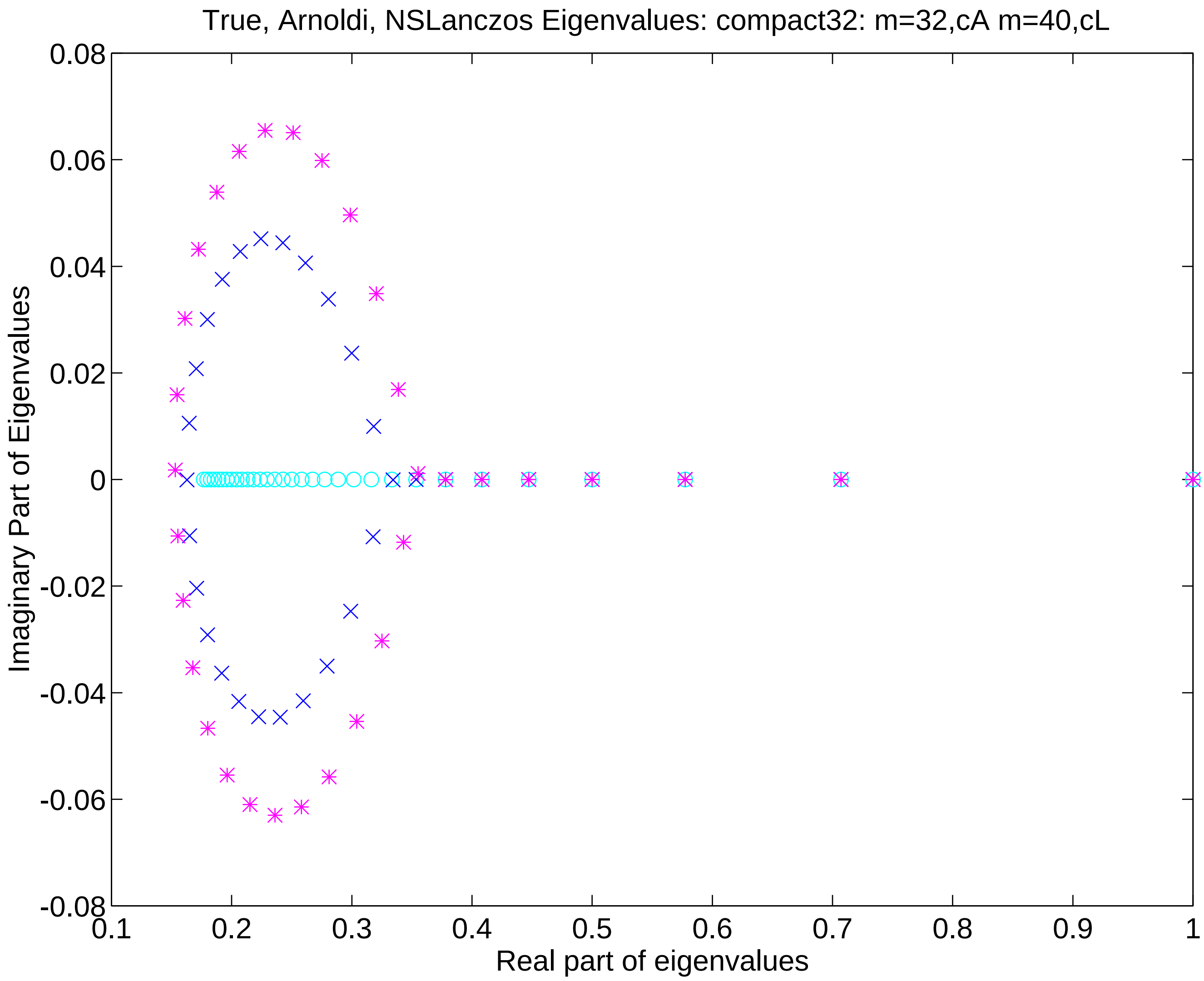

FIG. 14. *Example 9: Complex Arnoldi, m=32, NS Lanczos, m=40: True Eigenvalues and Estimates*

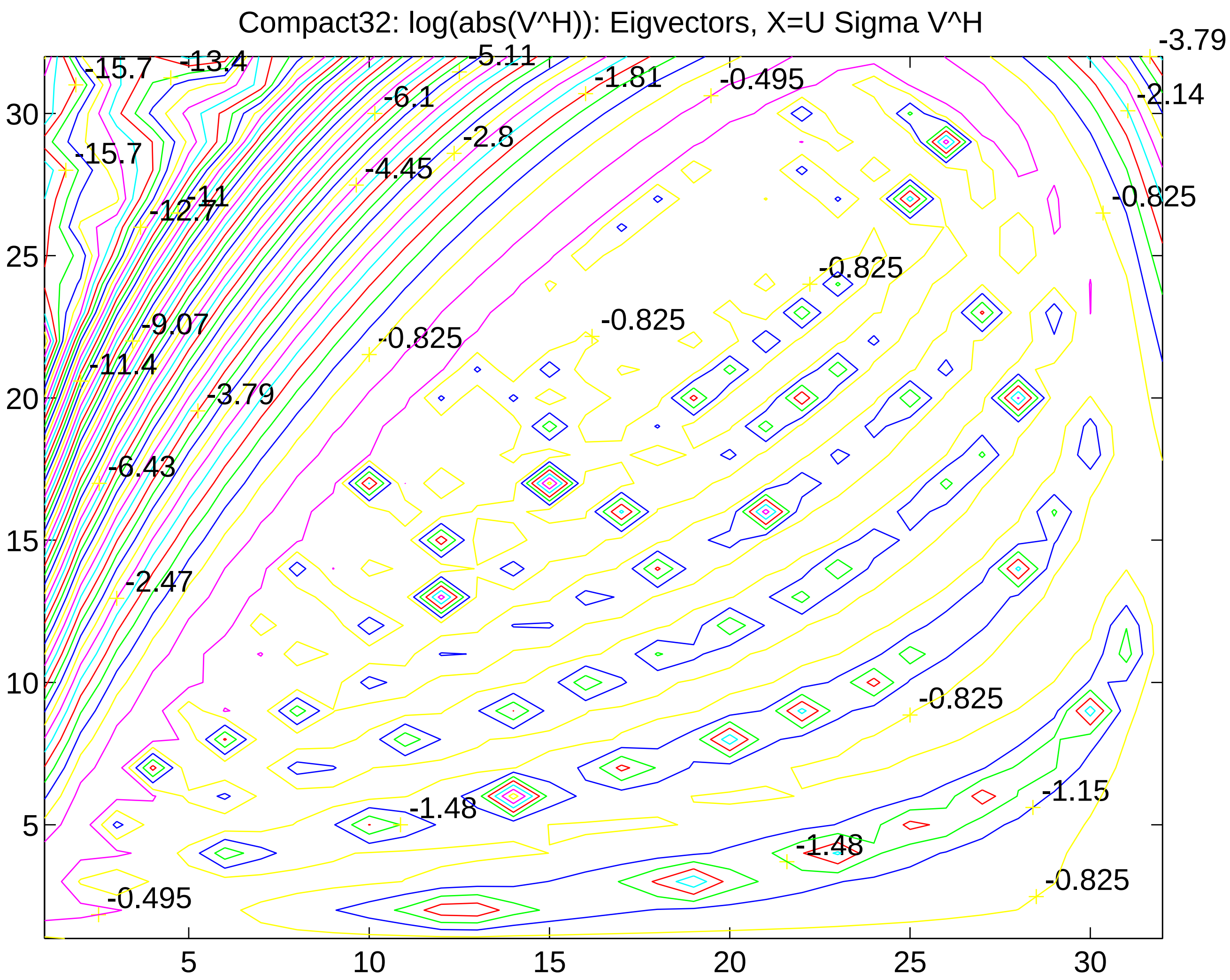

FIG. 15. *Example 9: Contours, log(abs(V^H)-matrix, Eigenvector Matrix, $X = U\Sigma_a V^H$.*

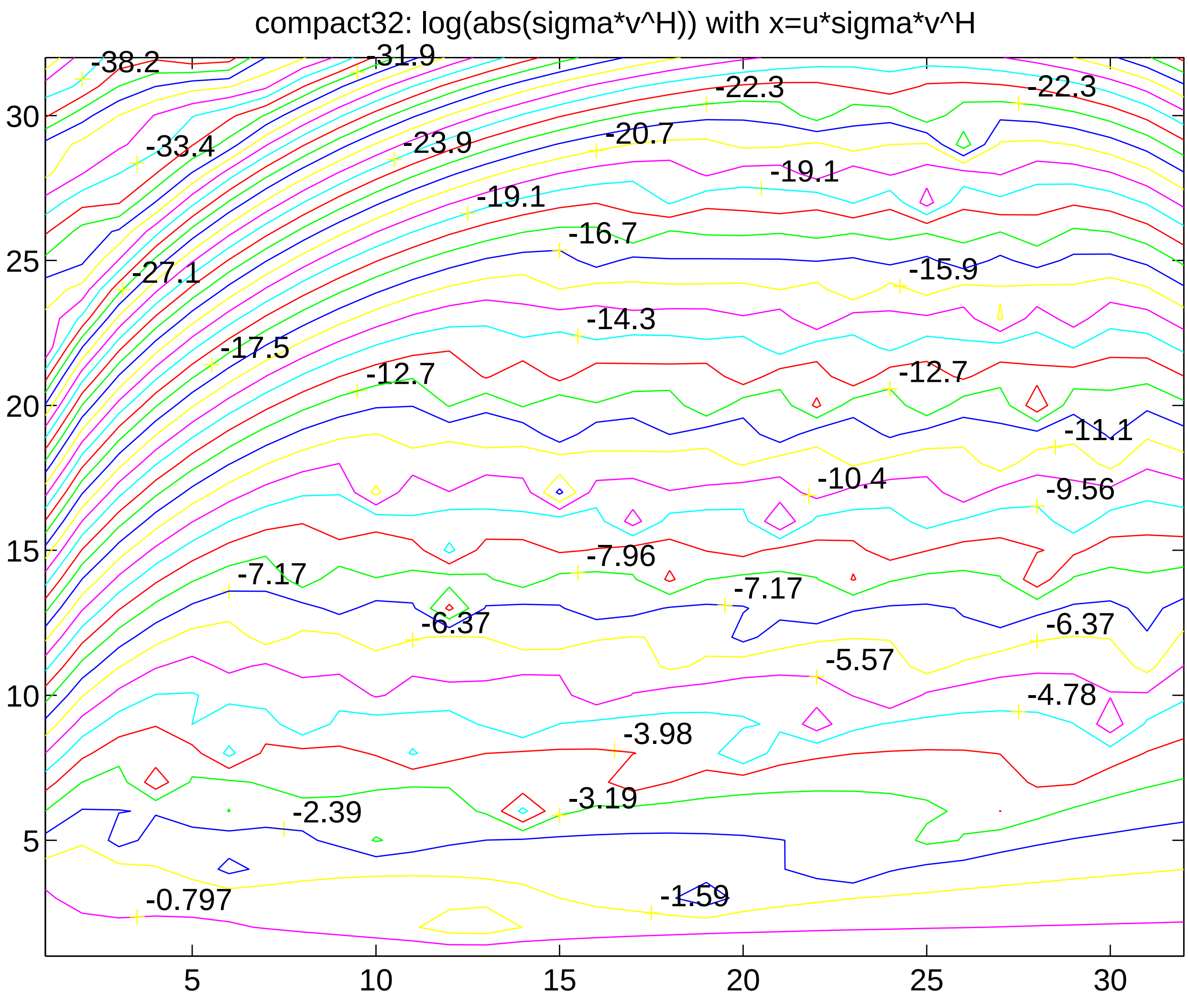

FIG. 16. *Example 9: Contours, log(abs(X)), X = Eigenvector Matrix*