Deflation in Krylov Subspace Methods
and
Distance to Uncontrollability

Daniel Kressner[*]

May 4, 2006

Abstract

The task of extracting from a Krylov decomposition the approximation to an eigenpair that yields the smallest backward error can be phrased as finding the smallest perturbation which makes an associated matrix pair uncontrollable. Exploiting this relationship, we propose a new deflation criterion, which potentially admits earlier deflations than standard deflation criteria. Along these lines, a new deflation procedure for shift-and-invert Krylov methods is developed. Numerical experiments demonstrate the merits and limitations of this approach.

AMS subject classifications: 65F15, 65F50, 93B05.

1 Introduction

Lanczos and Arnoldi methods are among the most popular numerical algorithms for computing a few eigenvalues and eigenvectors of a large, possibly sparse matrix. Both methods proceed by generating a sequence of Krylov subspaces hopefully containing approximations to a desired subset of eigenvectors. The purpose of deflation is to find, extract and lock such approximations from these subspaces. A crucial decision to be met in this process is when to consider an approximate eigenvalue/eigenvector pair $(\hat{\lambda}, \hat{x})$ of a matrix A converged. Assuming $\|\hat{x}^H \hat{x}\|_2 = 1$, the most common convergence criterion is

$$\|A\hat{x} - \hat{\lambda}\hat{x}\|_2 \leq \mathsf{tol}, \tag{1}$$

with some small, user-defined tolerance $\mathsf{tol} > 0$. Defining the residual $r = A\hat{x} - \hat{\lambda}\hat{x}$ and the backward error $\triangle A = -r\hat{x}^H$, this implies that $(\hat{\lambda}, \hat{x})$ is an exact eigenpair of the perturbed matrix $A + \triangle A$ with $\|\triangle A\|_2 = \mathsf{tol}$. In other words, $(\hat{\lambda}, \hat{x})$ has a small backward error, yielding a small forward error unless the corresponding eigenpair is ill-conditioned [7].

There are several strategies for finding an approximate pair $(\hat{\lambda}, \hat{x})$ in a given subspace $\mathcal{U}$, see, e.g., [2, 11] for an overview. The Rayleigh-Ritz method (also called Galerkin method) chooses $(\hat{\lambda}, \hat{x})$ to satisfy the conditions

$$\hat{x} \in \mathcal{U}, \qquad A\hat{x} - \hat{\lambda}\hat{x} \perp \mathcal{U}.$$

If the columns of U form an orthonormal basis for $\mathcal{U}$, these conditions amount to requiring $\hat{\lambda}$ to be an eigenvalue of $B = U^T A U$, and $\hat{x} = Uw$ with w being the corresponding eigenvector of B. In such a setting, $\hat{\lambda}$ and $\hat{x}$ are called Ritz value and Ritz vector, respectively. The Rayleigh-Ritz method works provably well for the extraction of eigenpairs belonging to well-separated exterior eigenvalues of Hermitian matrices, but may fail for other situations. In particular for non-Hermitian matrices, the Ritz vector $\hat{x}$ may not converge to an eigenvector even if $\hat{\lambda}$ converges to an eigenvalue, see, e.g., [17].

A strategy that avoids this drawback has been developed and analyzed by Jia [12] as well as Jia and Stewart [13]. Considering a Ritz value $\hat{\lambda}$, the vector $\hat{x}$ is chosen to minimize the norm of the residual:

$$\|A\hat{x} - \hat{\lambda}\hat{x}\|_2 = \min_{\substack{y \in \mathcal{U} \\ \|y\|_2 = 1}} \|Ay - \hat{\lambda}y\|_2. \tag{2}$$

Under this choice, $\hat{x}$ is called a refined Ritz vector. Optionally, the Ritz value can be replaced by the Rayleigh quotient $\hat{x}^T A \hat{x}$, which minimizes the residual norm (1) for fixed $\hat{x}$ and is likely to yield a more accurate eigenvalue approximation [17]. This procedure can be seen as a two-step optimization process; first (1) is minimized with respect to $\hat{x}$ for fixed $\hat{\lambda}$ and afterwards with respect to $\hat{\lambda}$ for fixed $\hat{x}$.

In this paper, we propose an approach which minimizes the residual norm (1) with respect to $\hat{\lambda}$ and $\hat{x}$ at the same time:

$$\|A\hat{x} - \hat{\lambda}\hat{x}\|_2 = \min_{\substack{\sigma \in \mathbb{C}, \ y \in \mathcal{U} \\ \|y\|_2 = 1}} \|Ay - \sigma y\|_2. \tag{3}$$

For general subspaces $\mathcal{U}$, this optimization problem is unlikely to be numerically tractable. However, if $\mathcal{U}$ is a Krylov subspace, it will be shown that (3) is equivalent to computing the distance to uncontrollability of a small-sized linear time-invariant system. Combined with recent advances of numerical methods for computing such distances, see [5, 8, 9], this relation makes (3) tractable for Krylov subspaces of modest dimension. Although the close connection between Krylov subspace methods and the notion of uncontrollability is well-known, see [3, 15], we are not aware of work turning this relation into (premature) deflation techniques. In the context of the QR algorithm, a similar relation has motivated the development of the highly successful aggressive early deflation technique [4].

2 Krylov decompositions and uncontrollability

Let $A \in \mathbb{C}^{n \times n}$ and let the columns of $[U, u] \in \mathbb{C}^{n \times (k+1)}$ form an orthonormal basis for a Krylov subspace $\mathcal{K}_{k+1}(A, u_0) = \mathrm{span}\{u_0, Au_0, \ldots, A^k u_0\}$ with some starting vector $u_0 \in \mathbb{C}^n$. Then there exists an orthogonal *Krylov decomposition* of the form

$$AU = UB + ub^H, \tag{4}$$

where $B \in \mathbb{C}^{k \times k}$ and $b \in \mathbb{C}^k$. Vice versa, a decomposition of the form (4) implies under a mild extra condition that the columns of U span a Krylov subspace $\mathcal{K}_k(A, u_0)$ for some u_0 [17]. The concept of Krylov decompositions was proposed by Stewart [18] in the context of the Krylov-Schur algorithm. We note that in the most general definition, U is only assumed to have linearly independent columns; this level of generality, however, is not suitable for our purpose.

Typical algorithms for producing (4) lead to decompositions that possess some particular structure. For example, the (implicitly restarted) Arnoldi method produces a matrix B which is upper Hessenberg and a vector b which is a scalar multiple of the kth unit vector e_k. In this case, (4) is called an *Arnoldi decomposition*. Krylov decompositions are not significantly more general than Arnoldi decompositions but admit elegant implicit restarting and deflation techniques.

What follows is not confined to a particular decomposition or algorithm. We just assume that a decomposition of the form (4) has been produced and the aim is now to extract the best approximation $(\hat{\lambda}, \hat{x})$ from $\mathcal{U} = \text{span}(U)$ to an eigenpair of A in the sense of (3). Inserting (4) into (3) yields

$$
\begin{aligned}
\|A\hat{x} - \hat{\lambda}\hat{x}\|_2 &= \min_{\substack{\sigma \in \mathbb{C}, \ w \in \mathbb{C}^k \\ \|w\|_2 = 1}} \|(A - \sigma I)(Uw)\|_2 \\
&= \min_{\substack{\sigma \in \mathbb{C}, \ w \in \mathbb{C}^k \\ \|x\| = 1}} \left\| \begin{bmatrix} U^H(A - \sigma I)U \\ u^H(A - \sigma I)U \end{bmatrix} w \right\|_2 \\
&= \min_{\substack{\sigma \in \mathbb{C}, \ w \in \mathbb{C}^k \\ \|w\|_2 = 1}} \left\| \begin{bmatrix} B - \sigma I \\ b^H \end{bmatrix} w \right\|_2 \\
&= \min_{\sigma \in \mathbb{C}} \sigma_{\min}([B^H - \sigma I, \ b]),
\end{aligned}
\tag{5}
$$

where $\sigma_{\min}(\cdot)$ denotes the minimal singular value of a matrix. The equality (5) can be used to relate deflations in Krylov decompositions to the concept of (un)controllability.

A matrix pair $(C, D) \in \mathbb{C}^{k \times k} \times \mathbb{C}^{k \times m}$ is called *controllable* if the associated linear time-invariant system $\dot{z}(t) = Cz(t) + Du(t)$ is controllable, see, e.g., [10]. Equivalently, by a result of Kalman [14], the pair (C, D) is controllable if and only if the matrix $[C - \sigma I, \ D]$ has full rank for all $\sigma \in \mathbb{C}$. Hence, from (5) it can be readily seen that a Krylov decomposition yields an exact eigenpair $(\hat{\lambda}, \hat{x})$ if and only if the pair (B^H, b) is uncontrollable, see also [15].

A reliable way to check controllability numerically is to compute the distance $\delta(C, D)$ of a given matrix pair (C, D) to the nearest uncontrollable matrix pair,

$$
\delta(C, D) = \inf\{\|[E, F]\|_2 : \ (C + E, D + F) \text{ is not controllable}, \ [E, F] \in \mathbb{C}^{k \times (k+m)}\}.
$$

By the following result, this multi-parameter optimization problem is reduced to a minimization problem in one (complex) parameter.

Theorem 1 (Eising [6]) *The distance $\delta(C, D)$ of a matrix pair $(C, D) \in \mathbb{C}^{k \times k} \times \mathbb{C}^{k \times m}$ to uncontrollability satisfies*

$$
\delta(C, D) = \inf_{\sigma \in \mathbb{C}} \sigma_{\min}([C - \sigma I, \ D]).
\tag{6}
$$

Combined with (5), this theorem shows that a minimizer $(\hat{\lambda}, \hat{x})$ of (3) satisfies

$$
\|A\hat{x} - \hat{\lambda}\hat{x}\|_2 = \delta(B^H, b).
$$

Moreover, a suitable $\hat{\lambda}$ is given by a minimizer σ of (6) and $\hat{x}$ can be set to $U\hat{w}$, where $\hat{w}$ is a left singular vector belonging to the smallest singular value of $[B^H - \hat{\lambda}I, \ b]$. To summarize, we obtain the following algorithm for extracting the optimal eigenpair approximation from a Krylov decomposition.

Algorithm 1 *Given a Krylov decomposition of the form (4), this algorithm produces the optimal eigenpair approximation $(\hat{\lambda}, \hat{x})$.*

1. *Compute $\hat{\lambda}$ such that $\delta(B^H, b) = \sigma_{\min}([B^H - \hat{\lambda}I, \, b])$.*

2. *Compute a left singular vector $\hat{w}$ belonging to the smallest singular value of $[B^H - \hat{\lambda}I, \, b]$.*

3. *Set $\hat{x} = U\hat{w}$.*

The computationally most expensive part of Algorithm 1 is Step 1. The algorithm presented in [9] requires $\mathcal{O}(k^5)$ floating point operations (flops) in the worst case and $\mathcal{O}(k^4)$ flops on average. This makes the computational cost of Algorithm 1 justifiable as long as k is moderate. For example for $k = \mathcal{O}(n^{1/3})$, the average cost is of the same order as the cost of orthogonalizing a new vector with respect to the Krylov basis. It should be noted that the algorithms presented in [8, 9] run into numerical difficulties if the distance to uncontrollability gets too small, say smaller than 10^{-8}. On the other hand, the value of Algorithm 1 for such tiny values of $\delta(B^H, b)$ is possibly limited. The following basic result indicates that Algorithm 1 amounts to the refined Ritz vector strategy as $\delta(B^H, b) \to 0$, provided that the eigenvalues of B are not ill-conditioned.

Lemma 2 *Assume that the eigenvalues of $C \in \mathbb{C}^{k \times k}$, denoted by $\Lambda(C)$, are mutually distinct and let $D \in \mathbb{C}^{k \times m}$. Then*

$$\min_{\sigma \in \Lambda(C)} \sigma_{\min}([C - \sigma I, \, D]) \leq (1 + \kappa)\delta(C, D) + \mathcal{O}(\delta(C, D)^2),$$

where κ is the maximal eigenvalue condition number of C.

Proof. If $\hat{\lambda}$ is a minimizer for $\sigma_{\min}([C - \sigma I, \, D])$ then there is a perturbation $[E, F]$ of norm $\delta(C, D)$ such that $[C - \hat{\lambda}I, \, D] + [E, F]$ has rank less than n. In particular, this implies that $\hat{\lambda}$ is an eigenvalue of $C + E$. By first order perturbation theory, there is an eigenvalue λ of C such that $|\hat{\lambda} - \lambda| \leq \kappa\delta(C, D) + \mathcal{O}(\delta(C, D)^2)$, see, e.g., [7]. Weyl's theorem implies

$$\sigma_{\min}([C - \lambda I, \, D]) = \sigma_{\min}([C - \hat{\lambda}I, \, D]) \leq \delta(C, D) + |\hat{\lambda} - \lambda|,$$

which concludes the proof. □

Once Algorithm 1 has found a pair $(\hat{\lambda}, \hat{x})$ satisfying the convergence criterion (1), we can lock this pair by performing a change of basis $UQ = [\hat{x}, \, U_2]$, where Q is an appropriate $k \times k$ unitary matrix. Since the Rayleigh quotient minimizes the residual for fixed $\hat{x}$, we may assume without loss of generality that $\hat{\lambda} = \hat{x}^H A\hat{x}$. This yields a transformed Krylov decomposition of the form

$$A[\hat{x}, \, U_2] = [\hat{x}, \, U_2] \begin{bmatrix} \hat{\lambda} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} + u[b_1^H, \, b_2^H], \tag{7}$$

where $[b_1^H, \, b_2^H] = b^H Q$ and $\begin{bmatrix} \hat{\lambda} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = Q^H B Q$. Since $\left\| \begin{bmatrix} B_{21} \\ b_1^H \end{bmatrix} \right\|_2 = \|A\hat{x} - \hat{\lambda}\hat{x}\|_2 \leq \mathsf{tol}$, we may safely set B_{21} and b_1 to zero. This decouples (7) and results into the reduced Krylov decomposition

$$\tilde{A}U_2 = U_2 B_{22} + ub_2^H,$$

with the projected matrix $\tilde{A} = (I - \hat{x}\hat{x}^T)A(I - \hat{x}\hat{x}^T)$. We can apply Algorithm 1 again to find another potential candidate for an eigenpair approximation. Some care must be taken since neglecting B_{21} and b_1 can have a (mild) influence on the attainable residuals in the further process of the Krylov subspace method, see [16] for more details.

3 Shift-and-invert Krylov decompositions

For computing interior eigenvalues by Krylov subspace methods, it is preferable to replace the matrix A by the matrix $(A - \tau I)^{-1}$, where $\tau \in \mathbb{C}$ is some target value. This leads to so called shift-and-invert methods, corresponding to Krylov decompositions of the form

$$(A - \tau I)^{-1} U = UB + u b^H. \tag{8}$$

Here, again under a mild extra condition, the columns of $U \in \mathbb{C}^{n \times k}$ form an orthonormal basis for a Krylov subspace $\mathcal{K}_k((A - \tau I)^{-1}, u_0)$. Usage of shift-and-invert methods has the advantage that eigenvalues close to τ are usually much faster approximated, but comes at the expense of having to solve a sequence of linear systems in the course of computing (8).

We could directly apply the results from the previous section to extract an eigenpair approximation $(\hat{\lambda}, \hat{x})$ from (8) which minimizes the residual with respect to $(A - \tau I)^{-1}$. However, a small residual of the form $(A - \tau I)^{-1} \hat{x} - \hat{\lambda} \hat{x}$ does not necessarily correspond to a small residual with respect to the original matrix A. To find a pair $(\hat{\lambda}, \hat{x})$ which minimizes the residual (and consequently the backward error) with respect to A, we assume B to be nonsingular and rewrite (8) as

$$AU = U(B^{-1} + \tau I) - (A - \tau I) u b^H B^{-1}.$$

This yields

$$\|A\hat{x} - \hat{\lambda}\hat{x}\|_2 = \min_{\substack{\sigma \in \mathbb{C}, \, w \in \mathbb{C}^k \\ \|w\|_2 = 1}} \|(A - \sigma I)(Uw)\|_2 \tag{9}$$

$$= \min_{\substack{\sigma \in \mathbb{C}, \, w \in \mathbb{C}^k \\ \|w\| = 1}} \left\| \begin{bmatrix} U^H(A - \sigma I)U \\ (I - UU^H)(A - \sigma I)U \end{bmatrix} x \right\|_2$$

$$= \min_{\substack{\sigma \in \mathbb{C}, \, x \in \mathbb{C}^k \\ \|x\| = 1}} \left\| \begin{bmatrix} B^{-1} + \tau I - U^H A u b^H B^{-1} - \sigma I \\ -((I - UU^H)A - \tau I) u b^H B^{-1} \end{bmatrix} x \right\|_2 \tag{10}$$

$$= \min_{\sigma \in \mathbb{C}} \sigma_{\min}([\tilde{B}^H - \sigma I, \, \tilde{b}]) \tag{11}$$

with

$$\tilde{B} = (I - U^H A u b^H)B^{-1} + \tau I, \quad \tilde{b} = \|(I - UU^H)Au - \tau u\| B^{-H} b. \tag{12}$$

Here we have used to fact that the two-norm is invariant under orthogonal transformations, which allows us to compress the vector $-((I - UU^H)A - \tau I)u$ in (10) to a multiple of the first unit vector by using, e.g., a Householder reflector [7].

Again, finding an pair $(\hat{\lambda}, \hat{x})$ satisfying (11) amounts to computing the distance to uncontrollability for a small-sized matrix pair. Algorithm 1 applies likewise with the quantities B and b replaced by $\tilde{B}$ and $\tilde{b}$, respectively. If k, the dimension of the Krylov subspace, is small compared to n then the most expensive steps when forming $\tilde{B}$ and $\tilde{b}$ consist of the matrix-vector multiplication $c = Au$ and the $\mathcal{O}(nk)$ computation of $U^H c$. In many situations, it can be expected that the cost of the (repeated) application of $(A - \tau I)^{-1}$ greatly outweighs these expenses.

To save the costly computation of the distance to uncontrollability, we can restrict the domain of (11) to the Ritz values $\Lambda(\tilde{B})$:

$$\|A\hat{x} - \hat{\lambda}\hat{x}\|_2 = \min_{\sigma \in \Lambda(\tilde{B})} \sigma_{\min}([\tilde{B}^H - \sigma I, \, \tilde{b}]). \tag{13}$$

In view of Lemma 2, this choice is certainly justified for sufficiently small residuals and leads to a shift-and-invert variant of refined Ritz vectors.

Algorithm 2 *Given a shift-and-invert Krylov decomposition of the form (8), this algorithm produces an eigenpair approximation $(\hat{\lambda}, \hat{x})$ satisfying (13).*

    1. Compute $\tilde{B}$ and $\tilde{b}$ according to (12).

    2. Find $\hat{\sigma}$ in the set of all eigenvalues of $\tilde{B}$ which minimizes $\sigma_{\min}([\tilde{B}^H - \hat{\sigma}I, \tilde{b}])$.

    3. Compute a left singular vector $\hat{w}$ belonging to the smallest singular value of $[\tilde{B}^H - \hat{\sigma}I, \tilde{b}]$.

    4. Set $\hat{x} = U\hat{w}$ and $\hat{\lambda} = \hat{x}^H A\hat{x}$.

4   Numerical experiments

We have performed some preliminary numerical experiments to assess the value of the new deflation strategies in comparison with the existing Rayleigh-Ritz and refined Ritz strategies. All numerical experiments were performed in MATLAB 6.5.

Being a popular test example, let us first consider the discretization of a convection-diffusion equation

$$-\triangle u(x,y) + \partial u(x,y)/\partial x = \lambda u(x,y)$$

on the unit square with Dirichlet boundary condition $u|_\partial = 0$. Taking centered differences yields an $n^2 \times n^2$ block tridiagonal matrix A with

$$A = \begin{bmatrix} D & -I_n & & \\ -I_n & D & \ddots & \\ & \ddots & \ddots & -I_n \\ & & -I_n & D \end{bmatrix}, \quad D = \begin{bmatrix} 4 & d & & \\ c & 4 & \ddots & \\ & \ddots & \ddots & d \\ & & c & 4 \end{bmatrix}, \tag{14}$$

where $c = -1 - 1/2(n+1)$ and $d = -1 + 1/2(n+1)$. For the following experiments, we have used $n = 30$ as in [12].

Example 1 To assess Algorithm 1, we have applied the standard Arnoldi method with random starting vector to successively generate Krylov decompositions of the form (4) for $k = 1, \ldots, 20$. Figure 1 shows the residuals of the optimal eigenpair approximations obtained by the Rayleigh-Ritz method, by the refined Ritz method, and by Algorithm 1. It can be seen that the Rayleigh-Ritz method leads to significantly larger residuals, which agrees with the results reported in [12]. On the other hand, there is no visible difference between refined Ritz method and Algorithm 1. In fact, the obtained norms of the residuals have 4 significant decimal digits in common. Although such a close agreement seems to be incidental, several experiments with other matrices from the Matrix Market collection [1] did not a reveal a single case where the use of Algorithm 1 led to a mentionable difference.

Example 2 To assess the different possibilities for extracting approximate eigenpairs from shift-and-invert Krylov decompositions, we have modified Example 1, replacing A by $(A - \tau I)^{-1}$ with $\tau \in \{4, 10\}$. Figure 2 shows the residuals of the optimal eigenpair approximations obtained by the Rayleigh-Ritz method, by the refined Ritz method, and by Algorithm 2. For
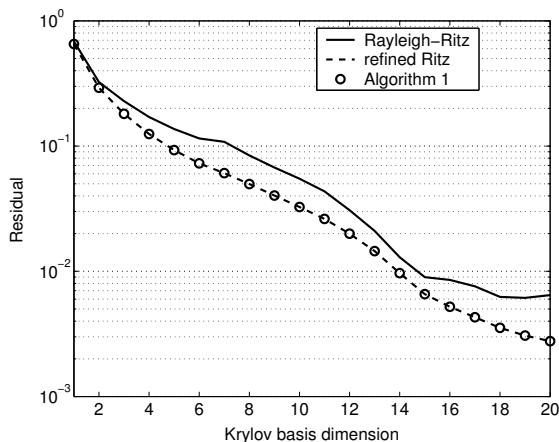
Figure 1: Optimal eigenpair approximations obtained from the Arnoldi method applied to the matrix A defined in (14).

$\tau = 10$, a target value far away from all eigenvalues of A, there is only a slight difference between the refined Ritz method and Algorithm 2. However, for $\tau = 4$, a target value very close to a large cluster of eigenvalues, Algorithm 2 yields considerably smaller residuals. This effect can be amplified by choosing τ even closer to an eigenvalue of A. Similarly as before, we found no significant difference when applying the full, distance-to-uncontrollability minimization (11) in place of (13).

To demonstrate that the observations from Example 2 are valid for other settings, we repeated the numerical experiments for the Tolosa matrix (TOLS340 with $\tau = -0.25 + 26\imath$), the Olmstead model (OLM1000 with $\tau = -5$), the reaction-diffusion Brusselator model (RDB1250 with $\tau = 0$), as well as the Tubular reactor model (TUB1000 with $\tau = -30+1.6\imath$); all from the Matrix Market collection. In all cases, the target value has been chosen so that it is close to at least one eigenvalue. In all experiments, Algorithm 2 yields significantly lower residuals.

5 Conclusions

We have investigated the relationship between finding the distance to uncontrollability for linear time-invariant systems and finding good eigenpair approximations in Krylov subspaces. This has resulted in an extraction strategy which provably delivers the eigenpair approximation with minimal residual for both standard and shift-and-invert Krylov subspace methods. Numerical experiments indicate that this new strategy is not significantly better than the refined Ritz method for standard Krylov subspace methods. In view of the increased computational effort, the use of this method without further modification is therefore not recommended. For shift-and-invert Krylov subspaces, however, Rayleigh-Ritz and refined Ritz methods are demonstrated to be inferior to a new strategy (Algorithm 2). The detailed analysis and use of this strategy in modern variants of the (implicitly restarted) Arnoldi algorithm are subject to further investigation. Also, theoretical and numerical comparisons to the extraction strategies presented in [11] need to be performed.
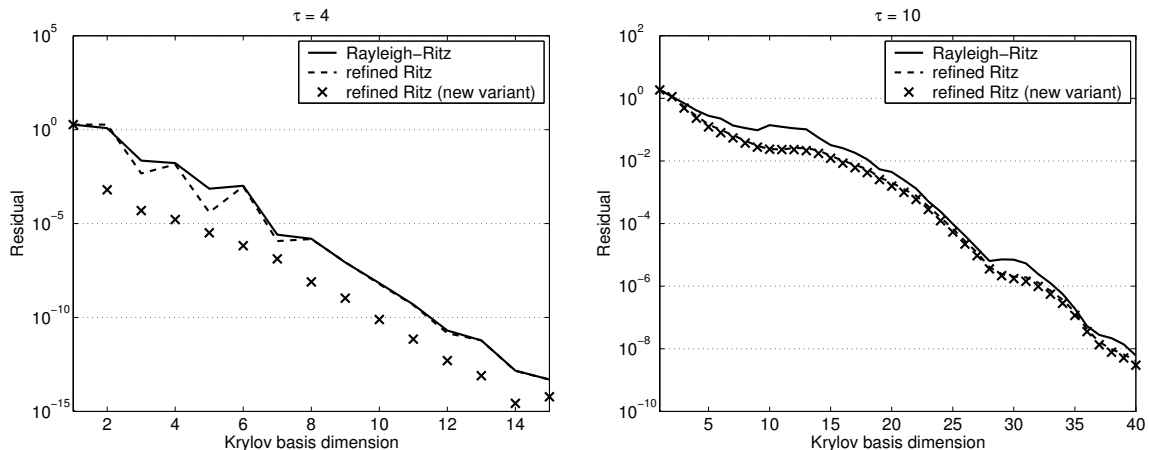
Figure 2: Optimal eigenpair approximations obtained from the shift-and-invert Arnoldi method with $\tau = 4$ and $\tau = 10$ applied for the matrix A defined in (14).

6   Acknowledgments

References

[1] Z. Bai, D. Day, J. W. Demmel, and J. J. Dongarra. A test matrix collection for non-Hermitian eigenvalue problems (release 1.0). Technical Report CS-97-355, Department of Computer Science, University of Tennessee, Knoxville, TN, USA, March 1997. Also available online from http://math.nist.gov/MatrixMarket.

[2] Z. Bai, J. W. Demmel, J. J. Dongarra, A. Ruhe, and H. van der Vorst, editors. *Templates for the Solution of Algebraic Eigenvalue Problems*. Software, Environments, and Tools. SIAM, Philadelphia, PA, 2000.

[3] D. L. Boley and G. H. Golub. The nonsymmetric Lanczos algorithm and controllability. *Systems Control Lett.*, 16(2):97–105, 1991.

[4] K. Braman, R. Byers, and R. Mathias. The multishift QR algorithm. II. Aggressive early deflation. *SIAM J. Matrix Anal. Appl.*, 23(4):948–973, 2002.

[5] J. V. Burke, A. S. Lewis, and M. L. Overton. Pseudospectral components and the distance to uncontrollability. *SIAM J. Matrix Anal. Appl*, 26(2):350–361, 2004/05.

[6] R. Eising. Between controllable and uncontrollable. *Systems Control Lett.*, 4(5):263–264, 1984.
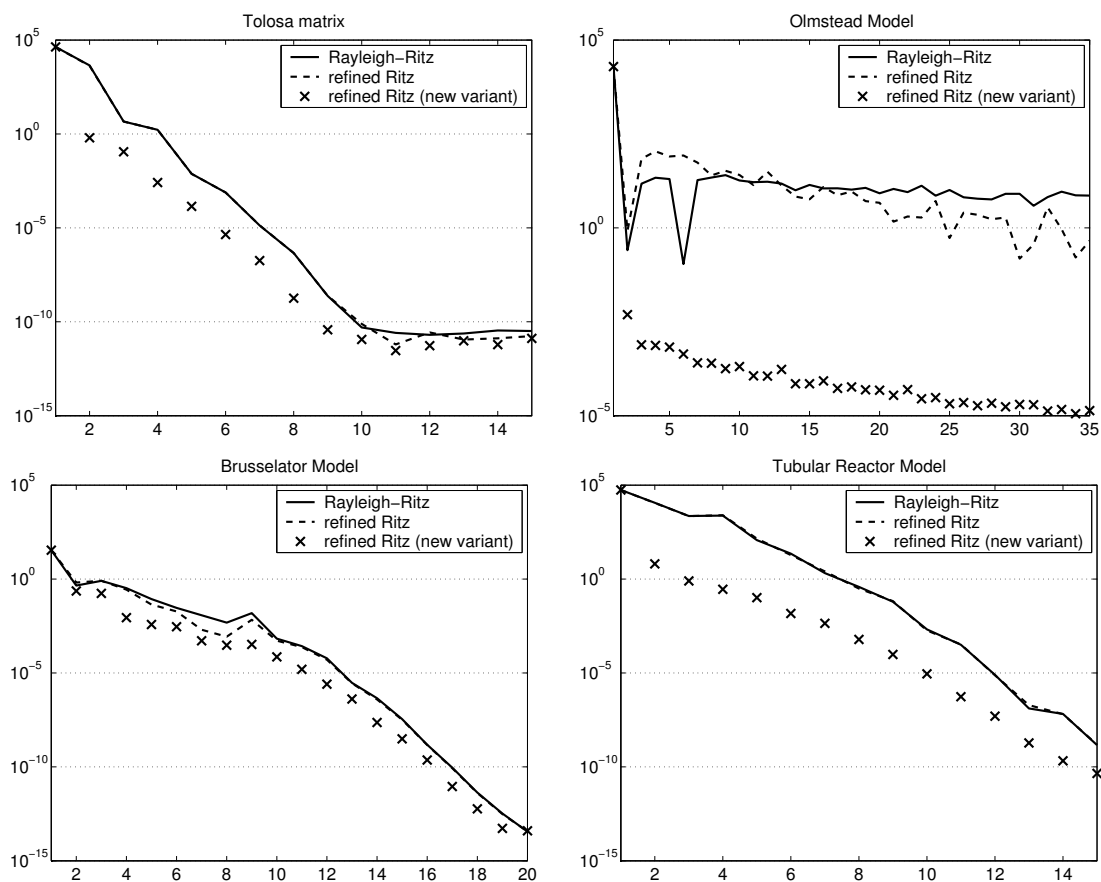
Figure 3: Optimal eigenpair approximations obtained from the shift-and-invert Arnoldi method for various examples from the Matrix Market collection.

[7] G. H. Golub and C. F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[8] M. Gu. New methods for estimating the distance to uncontrollability. *SIAM J. Matrix Anal. Appl.*, 21(3):989–1003, 2000.

[9] M. Gu, E. Mengi, M. L. Overton, J. Xia, and J. Zhu. Fast methods for estimating the distance to uncontrollability, 2005. To appear in SIAM J. Matrix Anal. Appl.

[10] D. Hinrichsen and A. J. Pritchard. *Mathematical Systems Theory I*, volume 48 of *Texts in Applied Mathematics.* Springer-Verlag, 2005.

[11] M. E. Hochstenbach and G. L. G. Sleijpen. Harmonic and refined Rayleigh-Ritz for the polynomial eigenvalue problem. Preprint, Dept. of Math., Case Western Reserve University, September 2004.

[12] Z. Jia. A refined iterative algorithm based on the block Arnoldi process for large unsymmetric eigenproblems. *Linear Algebra Appl.*, 270:171–189, 1997.

[13] Z. Jia and G. W. Stewart. An analysis of the Rayleigh-Ritz method for approximating eigenspaces. *Math. Comput.*, 70(234):637–647, 2001.

[14] R. E. Kalman. Mathematical description of linear systems. *SIAM J. Control Optim.*, 1:152–192, 1963.

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

[16] A. Stathopoulos. Locking issues for finding a large number of eigenvectors of Hermitian matrices. Tech Report WM-CS-2005-09, Computer Science, The College of William & Mary, July 2005.

[17] G. W. Stewart. *Matrix Algorithms. Vol. II.* SIAM, Philadelphia, PA, 2001. Eigensystems.

[18] G. W. Stewart. A Krylov-Schur algorithm for large eigenproblems. *SIAM J. Matrix Anal. Appl.*, 23(3):601–614, 2001/02.