University of Maryland College Park

Institute for Advanced Computer Studies TR–96–2
Department of Computer Science TR–3587

Iterative methods for solving $Ax = b$
GMRES/FOM versus QMR/BiCG

Jane K. Cullum*

ABSTRACT

We study the convergence of GMRES/FOM and QMR/BiCG methods for solving nonsymmetric $Ax = b$. We prove that given the results of a BiCG computation on $Ax = b$, we can obtain a matrix B with the same eigenvalues as A and a vector c such that the residual norms generated by a FOM computation on $Bx = c$ are identical to those generated by the BiCG computations. Using a unitary equivalence for each of these methods, we obtain test problems where we can easily vary certain spectral properties of the matrices. We use these test problems to study the effects of nonnormality on the convergence of GMRES and QMR, to study the effects of eigenvalue outliers on the convergence of QMR, and to compare the convergence of restarted GMRES and QMR across a family of normal and nonnormal problems. Our GMRES tests on nonnormal test matrices indicate that nonnormality can have unexpected effects upon the residual norm convergence, giving misleading indications of superior convergence when the error norms for GMRES are not significantly different from those for QMR. Our QMR tests indicate that the convergence of the QMR residual and error norms is influenced predominantly by small and large eigenvalue outliers and by the character, real, complex, or nearly real, of the outliers and the other eigenvalues. In our comparison tests QMR outperformed GMRES(10) and GMRES(20) on both the normal and nonnormal test matrices.

*Mathematical Sciences Department, IBM Research Division, T.J. Watson Research Center, Yorktown Heights, NY 10598, USA. This work was supported by NSF grant GER-9450081 while the author was a Visiting Professor in the Department of Computer Science, Institute for Advanced Computer Studies, and Institute for Systems Research, University of Maryland, College Park, MD 20742

1. Introduction

Computer simulation is an integral part of the analysis of physical phenomena and of the successful design and implementation of many systems and structures. Simulations require realistic mathematical models and practical numerical algorithms. Model complexity may be limited by the types of algorithms which are available. Many computer simulations require the repeated solution of large systems of linear equations.

$$Ax = b. \tag{1}$$

Often these systems are too large for numerical methods which explicitly modify the problem, and it becomes necessary to use iterative methods which obtain information about the problem through a sequence of matrix-vector computations.

Much of the current research on iterative methods for solving nonsymmetric equations(1) focuses on two sets of Krylov subspace methods and their variants [10]. Each set is based upon recursions which map the matrix A into a family of projection matrices which are then used to obtain approximations to a solution of Equation(1). The first set is based upon the Arnoldi recursion and includes the generalized minimal residual method (GMRES), the full orthogonal method (FOM), and their variants such as restarted GMRES(q) [17]. The second set of methods is based upon nonsymmetric Lanczos recursions, and includes the bi-conjugate gradient method (BiCG), the quasi-minimal residual method (QMR), and their variants such as BiCGSTAB(q) [9, 12, 22, 13, 19].

In this paper we study the convergence behavior of these methods using the MATLAB [14] templates discussed in [2]. In section 2 we review briefly the GMRES/FOM and the QMR/BiCG methods. In section 3 we obtain a relationship between BiCG and FOM, and a weaker relationship between QMR and GMRES. In section 4 we review unitary equivalences for each of these methods and how these equivalences can be used to obtain test matrices where certain spectral properties of the matrices can be varied [7]. In section 5 we address the question of consistency when testing a given method across changes in the nonnormality of the problem or in the eigenvalue distribution. In section 6 we use these test matrices in numerical experiments designed to provide some insight into the effects of nonnormality on the convergence of QMR/GMRES, on the effects of outliers on the convergence of QMR, and on the debate over which method is best.

In [8] we considered similar questions for the eigenvalue problem $Ax = \lambda x$. 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, $V_j = \{v_1, \ldots, v_j\}$
$x_j^G =$, jth iterate for method G
$\delta_j^G =$, $x_j - x_{true}$ for method G
$r_j = -Ax_j + b = -A\delta_j + r_0$, jth residual vector corresponding to x_j and δ_j

R^m, $m-$dimensional Euclidean space
e_j, jth coordinate vector in R^m for some specified m
$\hat{e}_j$, jth coordinate vector in R^{m+1} for some specified m
I_j, $j \times j$ identity matrix

2. Krylov Subspace Methods: GMRES/FOM and QMR/BiCG

In this section we review briefly the GMRES/FOM and QMR/BiCG methods for solving Equation(1). Without loss of generality we can consider the equivalent system

$$A\delta = r_0 \text{ where } r_0 = -Ax_0 + b \text{ and } \delta = x_{true} - x_0. \tag{2}$$

For simplicity we assume A is a real matrix.

2.1. Arnoldi Methods: GMRES/FOM/GMRES(q)

The GMRES/FOM/GMRES(q) methods are based upon the Arnoldi recursion [17].

Arnoldi Recursion:

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

$$h_{ik} = v_i^H v_{k+1}, \;\; v_{k+1} = v_{k+1} - h_{ik} v_i.$$

3. For each k compute:

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

Theoretically, $V_k = \{v_1, \ldots, v_k\}$ is an orthonormal basis of the Krylov subspace $\mathcal{K}_k(A, v_1)$ and the Hessenberg Arnoldi matrix $H_k \equiv (h_{ij})$ is a matrix representation of A on $\mathcal{K}_k(A, v_1)$ with respect to the V_k. The preceding implementation is a modified Gram-Schmidt orthogonalization. Other implementations exist [23]. In matrix form these recursions become

$$AV_k = V_k H_k + h_{k+1,k} v_{k+1} e_k^T \text{ where } H_k = (h_{ij}), \;\; 1 \leq i, j \leq k. \tag{3}$$

2.2. GMRES/FOM

In GMRES/FOM/GMRES(q), $v_1 \equiv r_0/\|r_0\|$ in (3). Theoretically, GMRES and FOM select their iterates, δ_k^G, δ_k^F, from the same Krylov subspaces but subject to different constraints on the corresponding residual vectors, r_k^G and r_k^F. At the k^{th} iteration in GMRES the iterate is selected so that the norm of the corresponding residual vector, r_k^G, is minimized over the k^{th} Krylov subspace corresponding to A and r_0 [16]. For each k,

$$\|r_k^G\| = \min_{\delta \epsilon \mathcal{K}_k(A, r_0)} \|r_0 - A\delta\| = \|r_0 - A\delta_k^G\|. \tag{4}$$

Saad [16] proved that Equation(4) is equivalent to the Petrov-Galerkin condition

$$r_k^G \perp AK_k(A, r^0). \tag{5}$$

We can rewrite equation (3) as

$$AV_k = V_{k+1} H_k^{(e)} \text{ where } H_k^{(e)} = \begin{pmatrix} H_k \\ 0 \cdots 0 \;\; h_{k+1,k} \end{pmatrix} \tag{6}$$

is a $k+1$ by k matrix.

Using equation(6), constraint (4), and the orthonormality of the Arnoldi vectors, it is not difficult to prove that the k^{th} GMRES iterate

$$\delta_k^G = V_k y_k^G \text{ where } \|r_k^G\| = \min_y \| - H_k^{(e)} y + \|r_0\| \hat{e}_1 \|. \tag{7}$$

At the k^{th} iteration in FOM the iterate is selected so that the residual vector, r_k^F, is orthogonal to the k^{th} Krylov subspace corresponding to A and r_0, [18]. For each k,

$$r_k^F \perp \mathcal{K}_k(A, r^0). \tag{8}$$

Using equation(3), the constraint (8), and the orthonormality of the Arnoldi vectors, it is not difficult to prove that the k^{th} FOM iterate

$$\delta_k^F = V_k y_k^F \text{ where } H_k y_k^F = \|r_0\| e_1. \tag{9}$$

Therefore, the k^{th} FOM iterate is obtained by solving a system of equations defined by the Galerkin iteration matrices H_k, and the k^{th} GMRES iterate is obtained by solving a least squares problem defined by H_k^e. We cannot guarantee that the Galerkin equations are well-conditioned or even nonsingular. If H_k is singular for some k, then the k^{th} FOM iterate does not exist but in this situation the k^{th} GMRES iterate is identical to the $(k-1)^{st}$ GMRES iterate [4].

For any $1 \le q \le n$, GMRES(q) is GMRES restarted with the current residual vector, after each q iterations [17].

2.3. Nonsymmetric Lanczos Methods: QMR/BiCG/BiCGSTAB(q)

QMR/BiCG and BiCGSTAB(q) are based upon nonsymmetric Lanczos recursions. Nonsymmetric Lanczos recursions generate two sets of Lanczos vectors, and tridiagonal matrices. We consider a variant used by [12].

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 $j = 1, \ldots, k$ compute:

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

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

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

$$p_j = v_{j+1} - \alpha_j v_j - \beta_j v_{j-1}$$
$$s_j = w_{j+1} - \alpha_j w_j - (\beta_j \rho_j / \xi_j) w_{j-1}$$
$$\rho_{j+1} = \|p_j\|, \quad v_{j+1} = p_j / \rho_{j+1}$$
$$\xi_{j+1} = \|s_j\|, \quad w_{j+1} = s_j / \xi_{j+1}$$
$$\beta_{j+1} = \xi_{j+1} w_{j+1}^T v_{j+1} / w_j^T v_j.$$

The Lanczos vectors are scaled to have unit norm. The coefficients in the recursions are chosen to make the Lanczos vectors $V_k = \{v_1, \ldots, v_k\}$ and $W_k = \{w_1, \ldots, w_k\}$ bi-orthogonal. Theoretically, for each k, V_k is a basis for the Krylov subspace $\mathcal{K}_k(A, v_1)$, and W_k is a basis for the subspace $\mathcal{K}_k(A^T, w_1)$.

In matrix form the nonsymmetric Lanczos recursions can be written as

$$
\begin{aligned}
AV_k &= V_k T_k + \gamma_{k+1} v_{k+1} e_k^T \\
A^T W_k &= W_k \tilde{T}_k + \omega_{k+1} w_{k+1} e_k^T
\end{aligned} \quad . \tag{10}
$$

where $\tilde{T}_k = \Phi_k^{-1} T_k \Phi_k$, $\Phi_k = \mathrm{diag}(\phi_1, \ldots, \phi_k)$, $\phi_1 = 1$, $\phi_k = \phi_{k-1}\rho_k/\xi_k$, and

$$
T_k = \begin{pmatrix}
\alpha_1 & \beta_2 & & & \\
\rho_2 & \alpha_2 & \beta_3 & & \\
& \rho_3 & \ddots & \ddots & \\
& & \ddots & \alpha_{k-1} & \beta_k \\
& & & \rho_k & \alpha_k
\end{pmatrix}, \tag{11}
$$

Theoretically, T_k is the matrix representation of the bi-orthogonal projection of A onto the Krylov subspaces $\mathcal{K}_k(A, r_0)$ and $\mathcal{K}_k(A^T, \tilde{r}_0)$ with respect to the bases generated by the nonsymmetric Lanczos recursions.

2.4. QMR/BiCG/BiCGSTAB(q)

In the QMR, BiCG, and BiCGSTAB(q) in [2], $v_1 \equiv r_0/\|r_0\| = w_1 = \tilde{r}_0/\|\tilde{r}_0\|$. Theoretically, at each iteration, QMR and BiCG select their iterates, δ_k^Q, δ_k^B, from the same Krylov subspaces , $\mathcal{K}_k(A, r_0)$, as GMRES and FOM but subject to different constraints on the residual vectors.

At the k^{th} iteration in QMR the iterate is chosen so that the norm of an associated quasi-residual vector, which we denote by z_k^Q, is minimized. The construction parallels the construction for GMRES. We can rewrite equation (10) as

$$
AV_k = V_{k+1} T_k^{(e)}. \text{ where } T_k^e \equiv \begin{pmatrix} T_k \\ 0 \ldots 0 \ \rho_{k+1} \end{pmatrix}. \tag{12}
$$

For any k, we define quasi-residual vectors z as any $k+1$ vector for which there is a k-vector y with

$$
z = -T_k^{(e)} y + \|r_0\| \hat{e}_1. \tag{13}
$$

The k^{th} QMR iterate $\delta_k^Q = V_k y_k^Q$ is chosen so that the norm of the associated quasi-residual vector is minimized [12].

$$
\delta_k^Q = V_k y_k^Q, \text{ where } \|z_k^Q\| \equiv \| - T_k^{(e)} y_k^Q + \|r_0\| \hat{e}_1\| = \min_y \| - T_k^{(e)} y + \|r_0\| \hat{e}_1\|. \tag{14}
$$

The k^{th} BiCG iterate δ_k^B is chosen so that the k^{th} residual vector, r_k^B, is orthogonal to the k^{th} Krylov subspace associated with A^T and $\tilde{r}_0$. Specifically,

$$
r_k^B \perp \mathcal{K}_k(A^T, \tilde{r}_0). \tag{15}
$$

Using equation(12), the constraint equation(15), and the bi-orthogonality of the Lanczos vectors, it is not difficult to prove that the k^{th} BiCG iterate

$$
\delta_k^B = V_k y_k^B \text{ where } T_k y_k^B = \|r_0\| e_1. \tag{16}
$$

Therefore, the k^{th} BiCG iterate is obtained by solving a system of equations defined by the Galerkin iteration matrices T_k, and the k^{th} QMR iterate is obtained by solving a least squares problem involving T_k^e. We cannot guarantee that the Galerkin equations are well-conditioned or even nonsingular. However, analogous to GMRES/FOM, if T_k is singular for some k, then the k^{th} BiCG iterate does not exist and in this situation the k^{th} QMR iterate is identical to the $(k-1)^{st}$ QMR iterate [12]. BiCGSTAB(q) are variants of BiCG which do not require A^T and which incorporate GMRES polynomials to accelerate the convergence. For details on their construction see [22, 19].

2.5. Comments on Both Types of Methods

In this paper we use superscripts Q, B, G, F to denote respectively, quantities associated with a QMR, a BiCG, a GMRES or a FOM computation.

For a given problem, QMR/BiCG and GMRES/FOM select their iterates from the same Krylov subspaces but subject to different constraints on the residuals of the iterates. The methods also differ in their storage requirements and the amount of work required at each iteration. QMR and BiCG require matrix-vector multiplications by A and by A^T but have minimal storage requirements, in addition to the storage required by the A and A^T computations. GMRES and FOM require only A but require access to all of the previously-generated Arnoldi vectors at each iteration. BiCGSTAB(q) also requires only A but since it is based upon BiCG and restarted GMRES, it also has small storage requirements. For details see [22, 19].

Theoretically, the recursions underlying GMRES/FOM cannot break down. Theoretically, however, the recursions underlying QMR/BiCG/BiCGSTAB(q) will break down if at some stage in the underlying nonsymmetric Lanczos recursions, a bi-orthogonality term $w^T v = 0$ for some $w \neq 0$ and $v \neq 0$. Within these schemes this corresponds to $\tilde{r}_k^T r_k = 0$ for some k but $\tilde{r}_k \neq 0$ and $r_k \neq 0$. If this occurs then the procedures cannot be continued. Exact breakdown is unlikely, near breakdowns can cause numerical instabilities. To avoid such problems, various look-ahead strategies have been proposed, see for example, [15, 11]. The discussions in this paper are equally applicable to the look-ahead variants of these methods. We do not include look-ahead capabilities in our tests. In addition the Galerkin methods, BiCG, FOM, and BiCGSTAB(q) can experience problems if the iteration matrices, T_k or H_k are nearly singular or singular.

Since the columns of the Lanczos vectors V_{k+1}^B are not orthonormal, the norms of the true residuals in QMR are not identical to those of the quasi-residuals. However, in exact arithmetic, for each k, [12, 11]

$$\|r_k^Q\| \leq \sqrt{k+1}\ \|z_k^Q\|. \tag{17}$$

In practice we observe that $\|r_k^Q\|$ and $\|z_k^Q\|$ appear to track each other until the limits of convergence are achieved. See section 5, figures 1-2 for examples of this behavior. We have the following lemma which we will need in section 3. See for example, [6].

Lemma 2.1. *(Exact Arithmetic). Apply either QMR/BiCG or GMRES/FOM to equation(1). For QMR/BiCG let $z^{Q,B} \equiv -T_k^e y + \|r_0\|\hat{e}_1$ for some k-vector y denote a corresponding quasi-residual vector. Similarly, for GMRES/FOM define $z^{G,F}$ using H_k^e. Define $x \equiv x_0 + \delta^{Q,B,G,F}$ where $\delta^{Q,B,G,F} \equiv V_k^{Q,B,G.F} y$. Then the corresponding residuals*

$$r \equiv -Ax^{Q,B,G,F} + b = -A\delta^{Q,B,G,F} + r_0 = V_{k+1}^{Q,B,G.F} z^{Q,B,G,F}. \tag{18}$$

Recursions (3,10) can be used to obtain error estimates for the iterates generated. Assuming no breakdowns at each iteration of QMR/BiCG, we obtain the following estimate of the norm of residuals at iteration k.

$$\|r_k^{Q,B}\| \equiv \|-A\delta_k^{Q.B} + r_0\| = \|\rho_{k+1} y_k^{Q,B}(k) v_{k+1}^B\| \tag{19}$$

where $y(k)$ denotes the k^{th} component of the vector y. We have a similar recursion for GMRES/FOM involving $h_{k+1,k}$. By construction for each k, $\|v_{k+1}^B\| = 1$. Therefore, we can estimate the norms of the residuals in GMRES/FOM or QMR/BiCG without computing the true residuals. In finite precision arithmetic, this estimate will still be valid as long as the errors in the recursions are sufficiently small [1]. Typically, each of these methods is implemented using recursive factorizations of either the Hessenberg or the tridiagonal matrices, and the preceding estimates of the norm of the residual are obtained recursively and used to track convergence.

At each iteration in each of the tests we discuss in section 6, we computed these estimates and the true residual norms. We used the true residual norms to track the convergence. Each residual norm plot in section 6 is a plot of true residual norms. We make the following assumption.

Assumption 1.1: In any statement or theorem about the QMR/BiCG/BiCGSTAB(q) methods, we will always assume that no breakdown has occurred and that all quantities are well-defined. Furthermore, in any statement about FOM we will always assume that all quantities are well-defined.

3. A Relationship Between BiCG and FOM

We obtain a relationship between BiCG and FOM residual norm plots and a weaker relationship between QMR quasi-residual and GMRES residual norm plots. In exact arithmetic we prove that given a BiCG residual norm plot we can construct an identical FOM residual norm plot for a different problem with the same eigenvalues. Furthermore, given a QMR quasi-residual norm plot, we can construct an identical GMRES residual norm plot for a different problem with the same eigenvalues. From this we can conclude, at least in exact arithmetic, that any type of residual norm convergence observed using BiCG can also be observed using FOM on some other problem with the same eigenvalues.

In this section we use C and $\tilde{C}$ to denote two different matrices. In Theorem 3.1 we assume that C 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, and the arithmetic were exact, the methods would terminate for some $m < n$, and Theorem 3.1 would have to be expressed in terms of m.

THEOREM 3.1. *: (Exact Arithmetic): Let C be a real $n \times n$ matrix with distinct eigenvalues. Apply BiCG/QMR to $Cx = c$. For $k = 1, \ldots, n$,*

$$CV_k^B = V_k^B (T_k^B)^e \tag{20}$$

where V_k^B and $(T_k^B)^e$ denote respectively the Lanczos vectors and matrices generated with $v_1^B = r_0/\|r_0\|$. Then there exists a $n \times n$ matrix $\tilde{C}$ with the same eigenvalues as C and a vector $\tilde{c}$ such that an application of the FOM/GMRES methods to $\tilde{C}x = \tilde{c}$, yields

$$\|r_k^F(\tilde{C})\| = \|z_k^F(\tilde{C})\| = \|z_k^B(C)\| = \|r_k^B(C)\| \text{ for } 1 \le k \le n, \tag{21}$$

where $z_k^B(C)$ and $z_k^F(\tilde{C})$ are the corresponding quasi-residual vectors. Furthermore, if $z_k^Q(C)$ and $z_k^G(\tilde{C})$ are the corresponding quasi-residual vectors for QMR and GMRES, then

$$\|r_k^G(\tilde{C})\| = \|z_k^G(\tilde{C})\| = \|z_k^Q(C)\|. \tag{22}$$

Proof. Let

$$T_n^B = Q_n U_n Q_n^H \tag{23}$$

be any Schur decomposition of T_n^B. Define $V_n^F \equiv Q_n^H$, then

$$U_n V_n^F = V_n^F T_n^B. \tag{24}$$

Define $\tilde{C} = U_n$ and $\tilde{c}$ equal to the first column of V_n^F scaled by $\|r_0\|$. Then V_n^F are FOM vectors corresponding to

$$\tilde{C}x = \tilde{c}, \text{ with } H_k^F(\tilde{C}) = T_k^B(C). \tag{25}$$

Furthermore, from Lemma 2.1, and equation(25),

$$r_k^F(\tilde{C}) = V_{k+1}^F z_k^F(\tilde{C}) = V_{k+1}^F z_k^B(C) = -\rho_{k+1} y_k^B(k) v_{k+1}^F \tag{26}$$

where y_k^B is the solution of the associated Galerkin equations

$$T_k^B y = \|r_0\| e_1 \text{ and } z_k^B(C) = z_k^F(\tilde{C}) = -(T_k^B)^e y_k^B + \|r_0\| \hat{e}_1 \tag{27}$$

are the associated BCG and FOM quasi-residual vectors.

From equations(26,27) and the orthonormality of the columns of V_k^F we have that

$$\|r_k^F(\tilde{C})\| = \|z_k^F(\tilde{C})\| = \|z_k^B(C)\| = |\rho_{k+1} y_k^B(k)|. \tag{28}$$

Moreover, from recursion(12) and Lemma 2.1 we have that

$$r_k^B(C) = V_{k+1}^B z_k^B(C) = -\rho_{k+1} v_{k+1}^B y_k^B(k) \tag{29}$$

By construction $\|v_{k+1}^B\| = 1$. Therefore, from equations(28,29)

$$\|r_k^B(C)\| = \|z_k^B(C)\| = |\rho_{k+1} y_k^B(k)|. \tag{30}$$

Combining equations(28,30) we obtain equation(21).

A similar argument yields equation(22). For each k we have that

$$r_k^G(\tilde{C}) = V_{k+1}^F z_k^G(\tilde{C}) \text{ and } z_k^G(\tilde{C}) = z_k^Q(C) \tag{31}$$

where the associated quasi-residual vectors $z_k^G(\tilde{C}) = z_k^Q(C)$ is obtained from the solution $y_k^G = y_k^Q$ of the associated least squares problem

$$\min_y \| - (T_k^B)^e y + \|r_0\| e_1\| \text{ by setting } z_k^Q(C) = z_k^G(\tilde{C}) = -(T_k^B)^e y_k^Q + \|r_0\|\hat{e}_1. \tag{32}$$

Therefore, from equation(31),

$$\|r_k^G(\tilde{C})\| = \|z_k^G(\tilde{C})\| = \|z_k^Q(C)\|. \tag{33}$$

Theorem 3.1 tells us that any portions of the corresponding BiCG/FOM residual norms plots are identical. If the errors in the BiCG recursions are sufficiently small, the BiCG residual norms are not too small, and the BiCG Galerkin equations are well-behaved, then we can obtain a finite precision variant of Theorem 3.1. In finite precision arithmetic the eigenvalues of $\tilde{C}_n$ will not usually equal the eigenvalues of C.

4. Unitary Invariant Properties of QMR/BiCG and of GMRES/FOM

We want to study the convergence behavior of QMR/BiCG and GMRES/FOM as we vary spectral properties of the test matrices. In practice users typically monitor estimates of the residual norms to identify convergence and may use true residual norms to verify convergence. Error norms are not available in practice. We obtain the following theorems in [7] which identify classes of matrices with identical residual and error norms. From this we obtain a set of test matrices. Details are in [7]. In this section we use C to denote a matrix and $\tilde{C}$ to denote the matrix $U^H C U$ where U is a unitary matrix.

THEOREM 4.1. *(Exact Arithmetic)(GMRES/FOM): 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, FOM matrices and vectors obtained by applying GMRES/FOM to $Cx = c$ and to $\tilde{C}x = \tilde{c}$ where $\tilde{c} = U^H c$ and $r_0^{\tilde{C}} = U^H r_0^C$. Then for each k, $H_k^C = H_k^{\tilde{C}}$, $V_k^C = U V_k^{\tilde{C}}$. Furthermore, if all of the iterates are defined for each method, corresponding iterates satisfy*

$$x_k^{F,G}(C) = U x_k^{F,G}(\tilde{C}) \text{ and } r_k^{F,G}(C) = U r_k^{F,G}(\tilde{C}). \tag{34}$$

Therefore, for each method, the corresponding computations yield identical residual and error norms,

$$\|r_k^{G,F}(C)\| = \|r_k^{G,F}(\tilde{C})\| \text{ and } \|e_k^{G,F}(C)\| = \|e_k^{G,F}(\tilde{C})\|. \tag{35}$$

THEOREM 4.2. *(Exact Arithmetic) (QMR/BiCG): 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, BiCG matrices and vectors, obtained by applying QMR/BiCG to $Cx = c$ and to $\tilde{C}y = \tilde{c}$ with $\tilde{c} = U^H c$, $r_0^{\tilde{C}} = U^H r_0^C$, and $\tilde{r}_0^{\tilde{C}} = U^T \tilde{r}_0^C$. 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 $\bar{U}$ is the complex conjugate of U. Furthermore if all of the iterates are well-defined for each method, corresponding iterates satisfy*

$$x_k^{B,Q}(C) = U x_k^{B,Q}(\tilde{C}) \text{ and } r_k^{B,Q}(C) = U r_k^{B,Q}(\tilde{C}). \tag{36}$$

Therefore, for each method, the corresponding two computations yield identical residual and error norms,

$$\|r_k^{B,Q}(C)\| = \|r_k^{B,Q}(\tilde{C})\| \text{ and } \|e_k^{B,Q}(C)\| = \|e_k^{B,Q}(\tilde{C})\|. \tag{37}$$

Theorems 4.1 and 4.2 demonstrate that as measured by the behavior of both the residual norms and the corresponding error norms, the GMRES/FOM methods and the QMR/BiCG methods for solving $Ax = b$ are invariant under unitary similarity transformations of A and corresponding unitary transformations of b, r_0, and $\tilde{r}_0$. Theorems 4.1 and 4.2 apply to complex or real versions of GMRES/FOM and QMR/BiCG. If A is normal these theorems imply that in exact arithmetic we can study the behavior of any of these procedures on normal matrices by studying their behavior on diagonal matrices. For arbitrary A we can use the preceding unitary equivalences to obtain the following theorem, [7].

THEOREM 4.3. *(Exact Arithmetic). For either GMRES, FOM, QMR or BiCG, all possible residual norm plots and all possible error norm plots can be generated by considering matrices of the form*

$$A = \Sigma V^H J V \Sigma^{-1} \tag{38}$$

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. If A is real and diagonalizable, we can replace complex V and Λ in equation(38) by a real orthogonal matrix and a real block diagonal matrix with 1×1 and 2×2 blocks.

Therefore, to study the behavior of any of these Krylov subspace methods on diagonalizable matrices it is sufficient to consider matrices with eigenvector matrices of the form ΣV^H where Σ is a positive diagonal matrix and V is a unitary matrix. We can use equations(38) to specify various eigenvalue distributions and eigenvector spaces. For any unitary V and diagonal Λ, $B = V^H \Lambda V$ is normal so that we have the following lemma.

LEMMA 4.1. *Let A be diagonalizable and defined by equation(38) for some choice of Σ , unitary V, and Λ, then A is a positive diagonal similarity transformation of the normal matrix*

$$B = V^H \Lambda V. \tag{39}$$

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

5. Numerical Comparisons

We are interested in studying the convergence behavior of GMRES/FOM, QMR/BiCG, BiCGSTAB, GMRES(10) and GMRES(20) as we vary the spectral properties of the test matrices. We have constructed MATLAB [14] codes which allow the user to generate and regenerate test matrices of the form given in equation(40). The user can also call any of the MATLAB Template codes [2] for solving equation(1). See [7] for more details. We note that in QMR/BiCG in [2], $\tilde{r}_0 = r_0$.

If all of the iterates are well-defined, the lower envelopes of the residual norm plots generated by applying GMRES and FOM to any problem track each other [6]. Whenever the residual norms in a FOM plot begin to increase the residual norms in the corresponding GMRES plot must flatten and neither residual norm can improve until the FOM residual norm plot finishes its excursion. Furthermore, whenever there is a strongly monotone decreasing segment in a GMRES residual norm plot there must be a strongly monotone decreasing segment in the corresponding FOM plot. Therefore, we do not need to run tests on FOM. A similar relationship exists for the BiCG residual norms and the corresponding QMR quasi-residual norms [6]. In practice we observe that until the finite precision errors become dominant, the true QMR residual norms and the QMR quasi-residual norms track each other. Serious deviations between these two norms could be used as a heuristic for estimating when the finite precision limits of convergence have been exceeded. Figures 1-2 illustrate this behavior on three nonnormal problems. In such situations we would not need to run BiCG tests.

In section 6 we examine the behavior of GMRES(10), GMRES(20), QMR, and BiCGSTAB. We restrict our tests to diagonalizable test matrices, only because including general Jordan forms would add to the number of cases to consider. Each test matrix is of the form

$$A_i = \Sigma_i V_i^H \Lambda_b V_i \Sigma_i^{-1} \tag{40}$$

where Λ_b is real and block diagonal with 1×1 and 2×2 blocks and, V_i is a real orthogonal matrix.

Many different factors may influence the convergence rates, the choice of Σ, of the eigenvalue distribution, of V, and of b or x_{true}. In addition we should consider possible effects from reordering the eigenvalues in Λ_b or the singular values in Σ. It is conceivable that if we constructed an example where the most dependent parts of the eigenspace corresponded to very small outlier eigenvalues, that the convergence of one or more of these procedures on that example could differ from its convergence on an example where the outliers corresponded to well-behaved eigenvectors. If we modify b in accordance with the permutation, then permuting either the singular values or the eigenvalues corresponds to some permutation of the columns of V [7]. In section 6 we fix V, generating it randomly, using the random number generator RAND and the ORTH routine in MATLAB [14], and we vary Σ and Λ_b.

In each test it is necessary to select either x_{true} or b. It is not sufficient to either simply fix b or fix x_{true} for all tests. We need to make choices which isolate the property we are varying, and are consistent across the set of problems being considered. We want to be able to look at each test problem individually and be able to compare results of tests across certain subsets of test problems.

In particular we want to trace the convergence of a method as we vary the nonnormality of the test problem while keeping the eigenvalue distribution fixed. We also want to trace the convergence of a method as we vary the eigenvalue distribution while keeping the eigenvectors fixed.

Theorems 4.1 and 4.2 indicate that if the test matrices differed by unitary similarity transformations, then we should select b so that the sizes of the projections of these vectors on corresponding right eigenvectors are preserved across the problems. Our test matrices do not differ by unitary similarity transformations. However, it still seems appropriate to maintain the sizes of the projections of b on the right eigenvectors across sets of test problems. Moreover, if A models some physical system and if the model is correct, then any x_{true} should live comfortably in the eigenspace of A. Therefore, for the nonnormality tests we fix the eigenvalue distribution and set $x_{true}^i = X^i \gamma$ where γ is generated randomly, $X^i = \Sigma_i V_i^H D^i$, and D^i is chosen as in Lemma 5.1 below. Lemma 5.1 states that this choice preserves the sizes of the projections of the right-hand sides on corresponding right eigenvectors, if the eigenvalue matrix is fixed.

In tests where we vary the eigenvalue distribution we cannot use equation(41). For those tests we fix the eigenvector matrix and set $b = X\gamma$ where γ is generated randomly and X is the specified eigenvector matrix, scaled as in Lemma 5.1.

We order the eigenvalues by algebraically-smallest real part to algebraically-largest real part, constructing any 2×2 blocks in Λ_b in the form

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

where $\lambda_j = a + bi$, $\lambda_{j+1} = a - bi$, and $b > 0$.

LEMMA 5.1. *Let A_i, $1 \leq i \leq q$, be real $n \times n$ matrices defined by equation(40). Let $Y^i \equiv \Sigma_i V_i^H$. Define $D^i = diag\{d_1^i, \ldots, d_n^i\}$ as follows. If λ_j^i is real, set d_j^i equal to the reciprocal of the norm of the j^{th} column of Y^i. If λ_j^i and λ_{j+1}^i are a conjugate pair of eigenvalues, set $d_j^i = d_{j+1}^i$ equal to the reciprocal of the square root of the sum of the norms of the j^{th} and the $(j+1)^{th}$ columns of Y^i. Define $X^i \equiv Y^i D^i$. Let γ be generated randomly and define*

$$x_{true}^i = X^i \gamma. \tag{41}$$

Then for all A_i, x_{true}^i is the same combination of corresponding unit right eigenvectors, and for any subset of matrices with $\Lambda_b^i = \Lambda_b$, both x_{true}^i and the corresponding right-hand sides

$$b^i = X^i \Lambda_b \gamma \tag{42}$$

are the same combinations of corresponding unit right eigenvectors.

Proof. For any i, let X_j^i denote the j^{th} column of X^i. If λ_j^i is real, then X_j^i is a unit right eigenvector of A_i. If λ_j^i and λ_{j+1}^i are a conjugate pair, define $z^R \equiv X_j^i$ and $z^L \equiv X_{j+1}^i$. Then $z_j^i \equiv z^R + \mathrm{i} z^L$ and $z_{j+1}^i \equiv z^R - \mathrm{i} z^L$ are unit right eigenvectors for λ_j^i and λ_{j+1}^i . Let Z^i denote the matrix whose columns are z_j^i, $1 \leq j \leq n$. It is straightforward to prove that $x_{true}^i = X^i\gamma = Z^i\bar{\gamma}$ where $\bar{\gamma}_j = \gamma_j$ if λ_j^i is real, and $\bar{\gamma}_j = .5(\gamma_j - \mathrm{i}\gamma_{j+1})$ and $\bar{\gamma}_{j+1} = .5(\gamma_j + \mathrm{i}\gamma_{j+1})$ if λ_j^i and λ_{j+1}^i are a conjugate pair. Therefore, $b^i = X^i \Lambda_b^i \gamma = Z^i \Lambda_d^i \bar{\gamma}$ where Λ_d^i is the diagonal matrix of eigenvalues corresponding to Λ_b^i.

Equation(38) is invariant under a scalar scaling of Σ. We scale each Σ so that the diagonal elements and the inverses of the diagonal elements are in the same subinterval of the real axis. We also order the singular values from largest to smallest.

6. Numerical Experiments and Observations

We ran a series of numerical experiments in attempts to gain some insight into the following issues.

1. How does nonnormality affect the convergence rates of GMRES and QMR as measured by the residual norms and the error norms?

2. How does the convergence rate of QMR vary with changes in the eigenvalue distribution? In particular, what role do outliers play in the convergence?

3. Which method is best for a given problem? Can we identify a method which is superior to the other methods in terms of both its residual norm and error norm convergence?

Our tests were limited and can only suggest possibly interesting types of behavior to explore theoretically and numerically. Answers to Question 2 may provide some insight into characterizations of effective preconditioners.

In section 6.1 we define the eigenvalue and the singular value distributions used in the test problems.

In section 6.2 we compare the observed behavior of QMR and GMRES across several nonnormal problems with variations in both the eigenvalue distribution and the singular value distribution of the eigenvector matrix in attempts to determine the sensitivity of these methods to nonnormality.

In section 6.3 we compare the behavior of QMR across a family of normal matrices and a corresponding family of nonnormal matrices where we systematically adjusted the eigenvalue distribution to determine the effects of outliers and of real versus complex eigenvalues on the convergence rate.

In section 6.4 we make some comments about the comparative behavior of GMRES(10), GMRES(20), QMR, and BiCGSTAB across a variety of normal and nonnormal test problems.

6.1. Test Matrices Used

Each test matrix is defined by equation(40) for different choices of Σ and Λ_b. V is generated randomly and fixed. All test problems are of size $n = 48$. In [8] we observed that small but nonzero entries in strategic positions in V can affect the convergence of eigenvalue computations based upon either the Arnoldi or nonsymmetric Lanczos recursions. In section 6.3 we observe, as others have observed, that the convergence rate of QMR appears to be correlated with the ability of the procedure to recognize small outliers [21]. If V has a nefarious structure which correlates with small outliers, we might expect similar effects upon the convergence of the QMR/GMRES procedures for $Ax = b$. We do not, however, consider such effects in this paper. Also we do not use any form of preconditioning.

In each test case all of the eigenvalues have positive real parts. The eigenvalue distributions considered were as follows. In each test, the corresponding real block diagonal form was used.

1. Λ_{gr48}: All of the eigenvalues are complex. There are no outliers. Λ_{gr48} is the eigenvalue distribution of the Gcrar matrix, $n = 48$ [20]. For a picture of this distribution see [8].

2. Λ_{asb2}: The eigenvalues are real and complex. There are both small and large complex pairs of outliers, two pairs with magnitude 10^{-4} and 10^{-2}, and three pairs with magnitudes 25, 55, and 100. The real parts of the other eigenvalues are in $[1, 10]$ and the imaginary parts are in $[-10, 10]$, all generated randomly in several boxes in the plane.

3. Λ_{bsb2}: The eigenvalues are real and complex. There is a small complex pair of outliers of magnitude 10^{-2}. There are 3 complex pairs of large outliers. This distribution is obtained from Λ_{asb2} by replacing the small complex pairs of eigenvalues of magnitude 10^{-4} and 10^{-2} by almost purely imaginary pairs of magnitude 10^{-2} and .96.

4. Λ_{dsb2}: The eigenvalues are real and complex. There are no small outliers. There are 3 large complex pairs of outliers. This distribution is obtained from Λ_{asb2} by replacing the small complex pairs of eigenvalues of magnitude 10^{-4} and 10^{-2} by almost purely imaginary pairs of magnitude .99 and .96.

5. Λ_{esb2}: The eigenvalues are real and complex. There are two complex pairs of small outliers of magnitude 10^{-4} and 10^{-2}. There are no large outliers. This distribution is obtained from Λ_{asb2} by replacing the large complex pairs by complex pairs of the same size as the rest of the eigenvalues in the distribution.

6. Λ_{fsb2}: The eigenvalues are real and complex. There are no small or large outliers. This distribution is obtained from Λ_{asb2} by replacing the small complex pairs of eigenvalues of magnitude 10^{-4} and 10^{-2} by almost purely imaginary pairs of magnitude .99 and .96, and replacing the large complex pairs by complex pairs of the same size as the rest of the eigenvalues in the distribution.

7. Λ_{msb2}: The eigenvalues are real and complex. The small outliers are nearly real, the big outliers are complex pairs. This distribution is obtained from Λ_{asb2} by replacing the small complex pairs of magnitude 10^{-4} and 10^{-2} by nearly real complex pairs with the same magnitudes.

8. Λ_{nsb2}: The eigenvalues are real and complex. Both the small and the large outliers are nearly real. This distribution is obtained from Λ_{asb2} by replacing the small complex pairs of magnitude 10^{-4} and 10^{-2} by nearly real complex pairs with the same magnitudes, and replacing the large complex pairs by nearly real complex pairs.

9. Λ_{osb2}: The eigenvalues are all nearly real. There are small outliers of magnitude 10^{-4} and 10^{-2}, and large outliers of magnitudes $21, 23$ and 28. This distribution is obtained from Λ_{asb2} by replacing all of the complex pairs by nearly real complex pairs obtained by multiplying all of the nonzero imaginary parts by 10^{-12}.

10. Λ_{psb2}: The eigenvalues are all nearly real. There are no small outliers. There are large outliers with magnitudes $21, 23$ and 28. This distribution is obtained from Λ_{asb2} by replacing all of the complex pairs by nearly real complex pairs obtained by multiplying all of the nonzero imaginary parts by 10^{-12}, and replacing the small eigenvalues of magnitude 10^{-4} and 10^{-2} by eigenvalues of magnitude .26 and .265.

11. Λ_{qsb2}: The eigenvalues are all nearly real. There are no small or large outliers. This distribution is obtained from Λ_{asb2} by replacing all of the complex pairs by nearly real complex pairs obtained by multiplying all nonzero imaginary parts by 10^{-12}, replacing the small eigenvalues of magnitude 10^{-4} and 10^{-2} by eigenvalues of magnitude .26 and .265. and reducing the large nearly real outliers to $2.1, 2.3$, and 2.8.

The three singular value distributions considered were as follows.

1. Σ_{I_n}: The identity matrix. Used to generate normal problems.
2. Σ_{a2r2b0}: Obtained by computing the equally spaced points t_i , $i = 1, \ldots, 48$, in the interval $[-2, 0]$ and setting the singular values $\sigma_i = 10^{-t_i^2}$.
3. Σ_{gr48}: Obtained from the Gcrar matrix of size 48 by applying the MATLAB function EIG to this matrix and then applying the MATLAB function SVD to the eigenvector matrix obtained from EIG.

In each test we track the true residual norms, the estimates of the residual norms, and the true error norms. True residual norms are used to determine convergence, and unless specifically indicated otherwise, each residual norm plot contains the true residual norms. Each procedure is terminated when $||r_k||/||r_0|| \leq 10{-}12$.

6.2. Effects of Nonnormality on Convergence

We compare the convergence of QMR and GMRES residual and error norms for $\Sigma_{I_n}, \Sigma_{a2r2b0}$, and Σ_{gr48}, and Λ_b corresponding to

$$\Lambda_{gr48}, \Lambda_{asb2}, \Lambda_{fsb2}, \Lambda_{msb2}, \Lambda_{osb2}, \text{ and } \Lambda_{qsb2}. \tag{43}$$

In each figure in this section we plot at most the first 150 iterations of each test.

In Figures $3 - 14$ we plot the residual norms and the corresponding error norms obtained by applying GMRES and QMR to each of the 18 problems obtained by combining each of the above eigenvalue distributions with each of the three singular value distributions. In each of these figures the eigenvalue distribution is fixed and we have applied GMRES and QMR to the test problems defined by that eigenvalue distribution, the prespecified V and γ, and each of the three Σ distributions, Σ_{I_n}, Σ_{a2r2b0}, and Σ_{gr48}. In each of these tests we define x_{true} according to Equations(41). Other tests using a different distribution for generating V and γ yielded similar results.

We make several observations. First, the residual norm plots in figures $(2j-1)$ for $2 \leq j \leq 7$, indicate that for each of the two nonnormal problems corresponding to each eigenvalue distribution, GMRES is converging significantly more rapidly than QMR. These indications, however, are incorrect, as is clearly illustrated in the corresponding error norm plots, figures $2j$, $2 \leq j \leq 7$. The initial portions of the error curves for corresponding GMRES and QMR plots do not exhibit any large deviations which might be expected from the large deviations between the corresponding residual norm plots.

Second, in each of these figures the GMRES residual norm plot for the normal problem is essentially an upper envelope to the GMRES plots for the two nonnormal problems whereas the QMR residual norm plot for the normal problem is either a lower envelope to the QMR plots for the two nonnormal problems or intersects with those curves. Theoretically, GMRES and QMR are identical when A is real and symmetric, and we observe that in the tests on normal problems with nearly real eigenvalues, the corresponding GMRES and QMR residual norm curves track each other, at least initially.

Third, in these tests, as we increase the nonnormality, the GMRES residual norm curves decrease more rapidly. These misleading decreases in the GMRES residual norm plots for the nonnormal problems caused premature termination of the GMRES procedure on several nonnormal test problems. See for example, figures 9-12.

Fourth, for problems without outliers, see figures 3-4,7-8,13-14, the differences in the number of iterations required for QMR residual norm convergence as we vary the nonnormality, are small. In the problems with outliers, figures 5-6,9-10,11-12, these differences are larger but the corresponding differences in the QMR error norm plots are still reasonably small for most but not all iterations. The effects of outliers on GMRES can be seen most clearly in the error norm plots. Whenever small outliers are present, all of the significant error reductions in the GMRES error norms occur on the last few iterations.

In these tests we also observe significant changes in the number of iterations required by QMR as we vary both the outliers and whether or not the eigenvalues and the outliers are real or complex. The number of QMR iterations required appears to be more sensitive to changes in outliers than to the changes in nonnormality which we use. For additional discussions of the effects of nonnormality on iterative methods for equation(1), see for example [3, 5, 20].

6.3. Outliers and Their Effect on the Convergence of QMR

We want to study the effect of different eigenvalue distributions upon the convergence rate of QMR. We consider both normal and nonnormal test matrices. For each set of tests in this section we fix the eigenvector matrix X as defined in Lemma 5.1. Beginning with a general eigenvalue distribution with both real and complex eigenvalues and both small and large complex pairs of outlier eigenvalues, we systematically modify the eigenvalue distribution to attempt to identify properties of the distribution which have the most significant effects upon the observed convergence rate. Since we are varying the eigenvalue distribution, setting $x^i_{true} = X^i\gamma$ does not preserve the sizes of the projections of b^i on corresponding right eigenvectors. In fact that choice of x^i_{true} yields b^i with small projections on any right eigenvectors corresponding to small eigenvalues, and this might cause QMR to take longer to converge.

Therefore, for each set of test problems in this section, we set $b^i \equiv X\gamma$ for all i. This choice fixes the size of the projection of b^i on each unit right eigenvector of A^i but can yield distorted x^i_{true}. However, this choice avoids any artificial effects due to changes in the starting residual as we go from one problem to another.

We consider two sets of eigenvalue distributions. The first set

$$\Lambda_{asb2}, \Lambda_{bsb2}, \Lambda_{dsb2}, \Lambda_{esb2}, \text{ and } \Lambda_{fsb2}, \tag{44}$$

consists of eigenvalue distributions which contain both real and complex pairs of eigenvalues, and are obtained from Λ_{asb2} by modifying the outliers. Λ_{asb2} contains two small complex pairs of outliers with magnitudes 10^{-4} and 10^{-2} and three large complex pairs of outliers with magnitudes 25, 55, and 100. Λ_{fsb2} has no outliers. In figures 15-16 we consider normal problems with these eigenvalue distributions. In figures 19-20 we consider nonnormal problems with these eigenvalue distributions.

The second set

$$\Lambda_{asb2}, \Lambda_{msb2}, \Lambda_{nsb2}, \Lambda_{osb2}, \Lambda_{psb2}, \text{ and } \Lambda_{qsb2}, \tag{45}$$

consists of eigenvalue distributions which contain complex, real or nearly real pairs of eigenvalues, and are obtained from Λ_{asb2} by modifying the outliers and the imaginary parts of the other complex pairs of eigenvalues. In figures 17-18 we consider normal problems with these eigenvalue distributions. In figures 21-22 we consider nonnormal problems with these eigenvalue distributions.

Figures 15-22 clearly illustrate the significant decreases in numbers of iterations achieved by systematically simplifying the eigenvalue distribution, independent of the nonnormality. We observe that for both the normal and the nonnormal problems and for both sets of eigenvalues, QMR takes much longer to converge when the problem has outliers . In particular note the reductions achieved by modifying Λ_{asb2} which is a mixed real and complex distribution with both small and large complex pairs of outliers to Λ_{bsb2} where the magnitudes of the two smallest complex pairs of outliers of Λ_{asb2} have been increased or to Λ_{esb2} where the magnitudes of the large complex outliers have been reduced. Also note the significant reductions obtained by modifying Λ_{asb2} to Λ_{dsb2} which has the large complex pairs of outliers but no small outliers and then to Λ_{fsb2} with no large or small outliers.

The tests on the second sequence of eigenvalue distributions, indicate continuous significant reductions in the numbers of iterations required as we simplify the mixed complex and real distribution Λ_{asb2} to the nearly real distribution with no outliers Λ_{qsb2}. Comparing figures 15-16 and 19-20 with figures 17-18 and 21-22 we also observe significant improvements in the convergence for each of the nearly real distributions over the corresponding mixed real and complex distributions.

In particular note the reductions obtained by systematically changing the distribution for Λ_{asb2} to Λ_{msb2} where the small complex outliers have been made nearly real and then to Λ_{nsb2} where both the small and the large outliers in Λ_{asb2} have been made nearly real. Also note the additional significant reductions achieved by further modifying Λ_{nsb2} to Λ_{osb2}, by making all of the eigenvalues nearly real, to Λ_{psb2}, where all of the eigenvalues are nearly real and there are no small outliers, and then to Λ_{qsb2}, where all of the eigenvalues are nearly and there are no outliers.

For a preconditioning to be successful it should modify the problem to elminate significant outliers, large or small, and make the modified eigenvalue distribution as nearly real as possible.

In several cases the QMR iterations continued beyond the finite precision limits of convergence. We could have avoided that problem if we had incorporated the heuristic check on the relative magnitudes of the quasi-residuals and true residuals which we mentioned in section 5.

6.4. Comparisons of Methods

We want to compare the performance of these methods. For large problems it is not feasible to use GMRES, so in this section we consider restarted GMRES. Moreover, in almost all cases, the lower envelopes of the QMR and BiCG residual norms are very close together [6]. Because of these similarities we do not include BiCG norms in any of the plots in this section. In this section we consider only GMRES(10), GMRES(20), QMR, and BiCGSTAB.

We applied each of these methods to the test problems obtained using the eigenvalue distributions listed in section 5 and the two singular value distributions Σ_{I_n} and Σ_{gr48}. In figures 23-32 we use the tests on the following eigenvalue distributions

$$\Lambda_{bsb2}, \Lambda_{dsb2}, \Lambda_{fsb2}, \Lambda_{osb2}, \Lambda_{psb2}, \text{ and } \Lambda_{qsb2}, \tag{46}$$

to summarize these tests. In figures 23-26 we consider GMRES(10) and GMRES(20) in the normal cases, $\Sigma = \Sigma_{I_n}$. In figures 27-28 we consider GMRES(20) in the nonnormal cases, $\Sigma = \Sigma_{gr48}$.

Number of Iterations to Convergence, $\|r_k\|/\|r_0\| \leq 10^{-12}$

Method / Problem	GMRES (10)	GMRES (20)	QMR	BICGSTAB
psb2id2	91	70	35	33
psb2gr2	87	59	35	43
qsb2id2	56	48	31	23
qsb2gr2	56	47	31	28

True Error Norm at Convergence

Method / Problem	GMRES (10)	GMRES (20)	QMR	BICGSTAB
psb2id2	2×10^{-12}	5×10^{-13}	1×10^{-13}	8×10^{-14}
psb2gr2	2×10^{-11}	3×10^{-10}	3×10^{-13}	6×10^{-13}
qsb2id2	4×10^{-12}	9×10^{-13}	1×10^{-13}	9×10^{-12}
qsb2gr2	4×10^{-12}	2×10^{-10}	2×10^{-13}	3×10^{-12}

Table 6.1. Comparison Across Methods: Residual and Error Norm Convergence. Eigenvalue Distributions: Λ_{psb2} and Λ_{qsb2}. Singular Value Distributions: $\Sigma_{I_n} = id$ and $\Sigma_{\{gr48\}} = gr$.

In the tests run, GMRES(10) converged only when $\Lambda = \Lambda_{psb2}$ or Λ_{qsb2}. In both of these distributions the eigenvalues are real or nearly real and there are no small outliers. Convergence occurs in both the normal and the nonnormal cases. See figures 23-24 where we have plotted the results of applying GMRES(10) to the normal problems corresponding to the six eigenvalue distributions in equation(46). In the corresponding plots for the nonnormal cases, the residual and error norm curves for the four nonconvergent eigenvalue distributions are flatter and closer together.

GMRES(20) achieves some measure of convergence for all of the distributions in equation(46) except for Λ_{bsb2}, the only distribution with small outliers which are complex pairs. For Λ_{osb2},,which contains two pairs of nearly real outliers of magnitude 10^{-4} and 10^{-2}, convergence occurs only in the normal case. See figures 25-26, 27-28.

For the two nearly real eigenvalue distributions with no small outliers, Λ_{psb2} and Λ_{qsb2} the differences in the convergence rates of the GMRES(20) residual and error norm plots between the normal and nonnormal cases are not large. For the other distributions there are significant differences, and the error reductions in the nonnormal cases are less than in the normal cases.

In figures 29-32 we consider similar tests on BiCGSTAB. We observe that the convergence is less predictable for this method. BiCGSTAB converges well for the two eigenvalue distributions which are nearly real and have no small outliers, Λ_{psb2} and Λ_{qsb2}. See Table 6.1. However, it has trouble with $\Lambda_{dsb2}, \Lambda_{fsb2}$, and Λ_{osb2} in the normal case, and with Λ_{dsb2} and Λ_{osb2} in the nonnormal case. The difficulties with Λ_{dsb2} and Λ_{fsb2} may occur because in real arithmetic BiCGSTAB cannot handle complex eigenvalues very well. It fits linear GMRES polynomials at each iteration. These tests should be rerun using BiCGSTAB(2) which fits second order GMRES polynomials which can model pairs of complex eigenvalues in real arithmetic [13]. BiCGSTAB(2) was not available in the MATLAB Templates [2].

Finally we consider QMR on each of these test problems. See figures 33-36. On these tests, to within error norms of size 10^{-9} there were no significant differences between the convergence in the normal and the nonnormal cases. The decrease in the number of iterations required as we simplify the eigenvalue distribution is depicted clearly in these figures.

If we compare the number of iterations required by GMRES(20), BiCGSTAB, and QMR on corresponding test problems, we see that with the exception of the nearly real distributions. Λ_{psb2} and Λ_{qsb2} with no small outliers, that QMR converged more quickly and more reliably than the other two procedures. For Λ_{qsb2}, BiCGSTAB converged significantly more quickly than QMR in the normal case and was 10% faster in the nonnormal case. For Λ_{psb2}, BiCGSTAB converged slightly more quickly than QMR in the normal case. See Table 6.1.

In these tests as well as in the tests in section 6.3, the number of iterations required for convergence of the QMR residual norms seems to be more strongly influenced by the choice of the eigenvalue distribution than by the choice of singular value distribution. We note however that changing Σ is only one way to alter the nonnormality of a test problem. In these tests we have not considered general variations in V. We only generated V randomly, using either a uniform or a normal distribution. It may be possible to construct other unitary V such that Σ and V interact to make the problem more difficult than the Σ matrix alone might indicate. Moreover, we note again that these tests should be rerun using BiCGSTAB(2) which was not available in the MATLAB Templates [2].

7. Summary

We derived a relationship between residual norms generated by the BiCG method and residual norms generated by the FOM method. This relationship states that any residual norm plots seen in a BiCG computation can also be seen in a FOM computation on a different problem but with the same eigenvalues.

Using unitary equivalences for each of the methods GMRES/FOM/QMR/BiCG, we defined sets of test problems where we can easily vary certain spectral properties of the test matrices. Within the context of these test matrices, we discussed the question of consistency across test problems. We used test problems of this type to explore three questions. First, we considered the effect of nonnormality on the convergence of GMRES and QMR for different eigenvalue distributions. Second, we traced the behavior of QMR as we varied the eigenvalue distribution in test problems with the eigenvector matrix fixed. Finally, we compared the convergence of QMR, GMRES(10), GMRES(20), and BiCGSTAB over a set of normal and nonnormal test problems with different eigenvalue distributions.

Our GMRES tests on nonnormal test matrices indicate that nonnormality can have unexpected effects upon the residual norm convergence, giving misleading indications of superior convergence over QMR when the error norms for GMRES are not significantly different from those for QMR. This behavior can lead to premature termination of the GMRES procedure.

Our tests on QMR indicate that the convergence of the QMR residual and error norms is strongly controlled by both small and large eigenvalue outliers and by the character, real, complex or nearly real, of the outliers and the other eigenvalues. This behavior is not evident in our GMRES residual norm plots but is exhibited clearly in the corresponding GMRES error norm plots. In our comparison tests, we considered complex, mixed, and nearly real eigenvalue distributions with and without small and large outliers. Overall QMR was more robust and reliable and outperformed GMRES(10) and GMRES(20) on both the normal and nonnormal test matrices. Comparisons with BiCGSTAB should be made only after similar tests are made using BiCGSTAB(2).

References

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

[2] Richard Barrett et al. Templates for the Solution of Linear Systems: Building Blocks for Iterative MethodsSIAM Publications, 1994.

[3] 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.

[4] P. N. Brown, *A theoretical comparison of the Arnoldi and GMRES algorithms*, SIAM Jour. Sci. Stat. Comput., 20 (1991), pp. 58–78.

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

[6] J. Cullum and Anne Greenbaum, *Residual relationships within three pairs of iterative algorithms for solving $Ax = b$*, IBM Research Report RC 18672, January 1993, IBM T.J. Watson Research Center, Yorktown Heights, New York. SIAM J. Matrix Anal. Appl. (1996).

[7] J. Cullum, *Testing iterative methods for nonsymmetric matrices*, IBM Research Report, RC , January 1996, IBM Research, T.J.Watson Research Center, Yorktown Heights, NY 10598.

[8] J. Cullum, *Arnoldi versus nonsymmetric Lanczos algorithms for solving matrix eigenvalue problems*, IBM Research Report, RC 20303, December 4, 1995, IBM Research, T.J.Watson Research Center, Yorktown Heights, NY 10598. Submitted to BIT.

[9] R. Fletcher, *Conjugate gradient methods for indefinite systems*, in Numerical Analysis Dundee 1975, G. A. Watson, ed., Lecture Notes in Mathematics 506, Springer, Berlin, 1976, pp. 73–89..

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

[11] 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.

[12] R. W. Freund and N. M. Nachtigal, *QMR: a quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math., 60 (1991), pp. 315–339.

[13] M. H. Gutknecht. *Variants of BiCGSTAB for matrices with complex spectrum*, SIAM J on Scientific and Statistical Computing, 14,(1993) pp. 1020–1033.

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

[15] 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.

[16] Y. Saad and M. H. Schultz, *GMRES: A generalized minimum residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.

[17] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publishers, Cambridge, Mass., 1996.

[18] Y. Saad, *Krylov subspace methods for solving unsymmetric linear systems*, Math. Comp., 37 (1981), pp. 105–126.

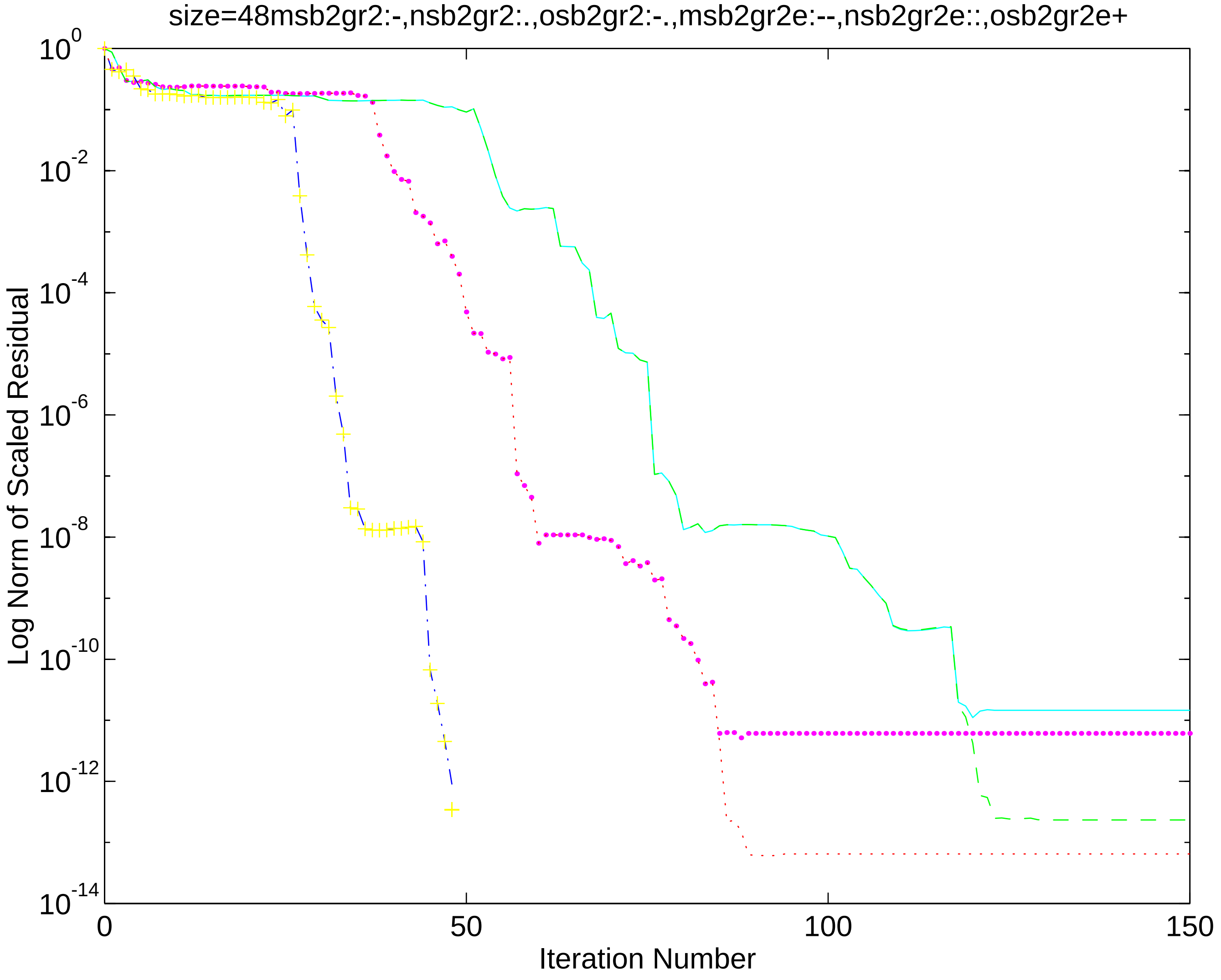

FIG. 1. *QMR: Residual Norms and Estimates of Residual Norms versus Iteration Number. Eigenvalue Distributions:* Λ_{msb2}*,* Λ_{nsb2}*,* Λ_{osb2}*. Singular Value Distribution:* Σ_{gr48}*. Residual Norms = solid, big dots, dash-dots. Estimates of Residual Norms = dash, little dots, +.*

[19] G.L.G. Sleijpen, H. A. van der Vorst, D.R. Fokkema, *BiCGSTAB(l) and other hybrid Bi-CG methods* , Numerical Algorithms, 7 (1994), pp. 75–109.

[20] 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, pp. 234–266.

[21] H. A. van der Vorst and C. Vuik, *The superlinear convergence behavior of GMRES*, J. Comp. Appl. Math., 48 (1993), pp. 327–341.

[22] H. A. van der Vorst, *A fast and smoothly converging variant of BiCG for the solutions of nonsymmetric linear systems*, SIAM J. Sci. Stat. Comp., 13 (1992), pp. 631-644.

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

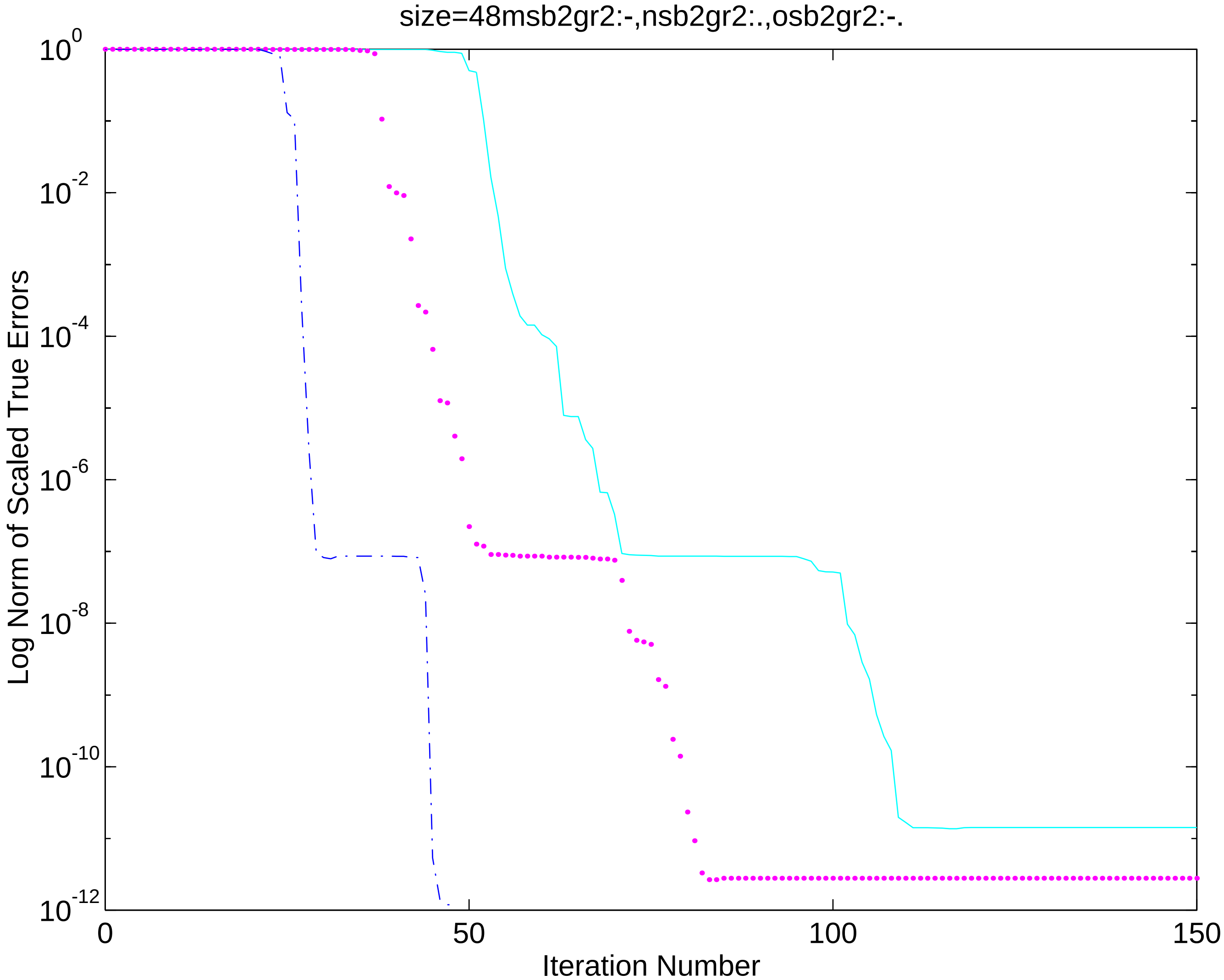

FIG. 2. *QMR: Error Norms versus Iteration Number. Eigenvalue Distributions:* Λ_{msb2}*,* Λ_{nsb2}*,* Λ_{osb2}*. Singular Value Distribution:* Σ_{gr48}*.*

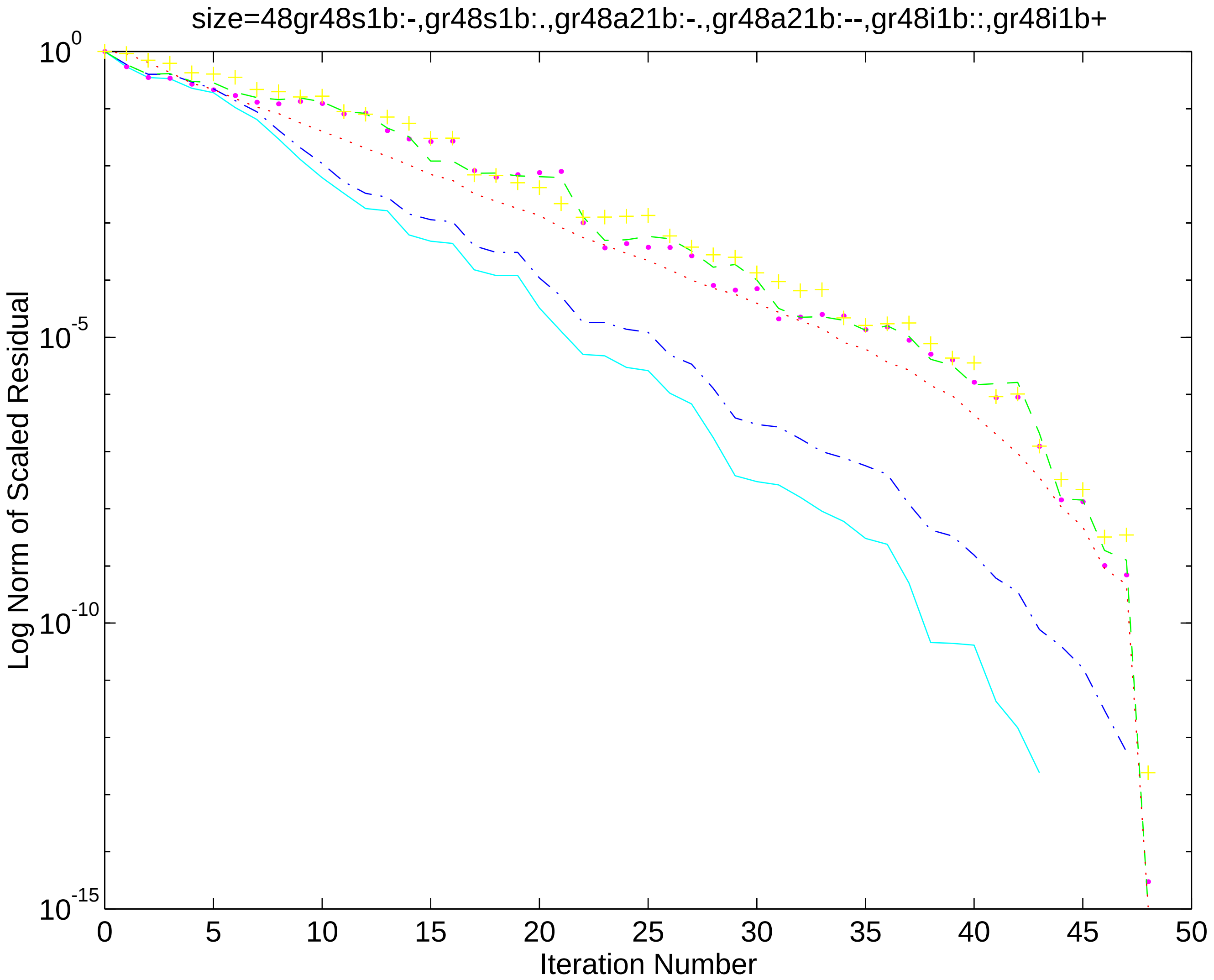

FIG. 3. *GMRES vs. QMR: Residual Norms versus Iteration Number. Eigenvalue Distribution:* Λ_{gr48}*. Singular Values Distributions:* $\Sigma_{gr48} = gr$*,* $\Sigma_{a2r2b0} = a2$*,* $\Sigma_{I_n} = id$*, Residual Norms: GMRES: solid, dot-dash, small dots. QMR: big dots, dashes, +.*

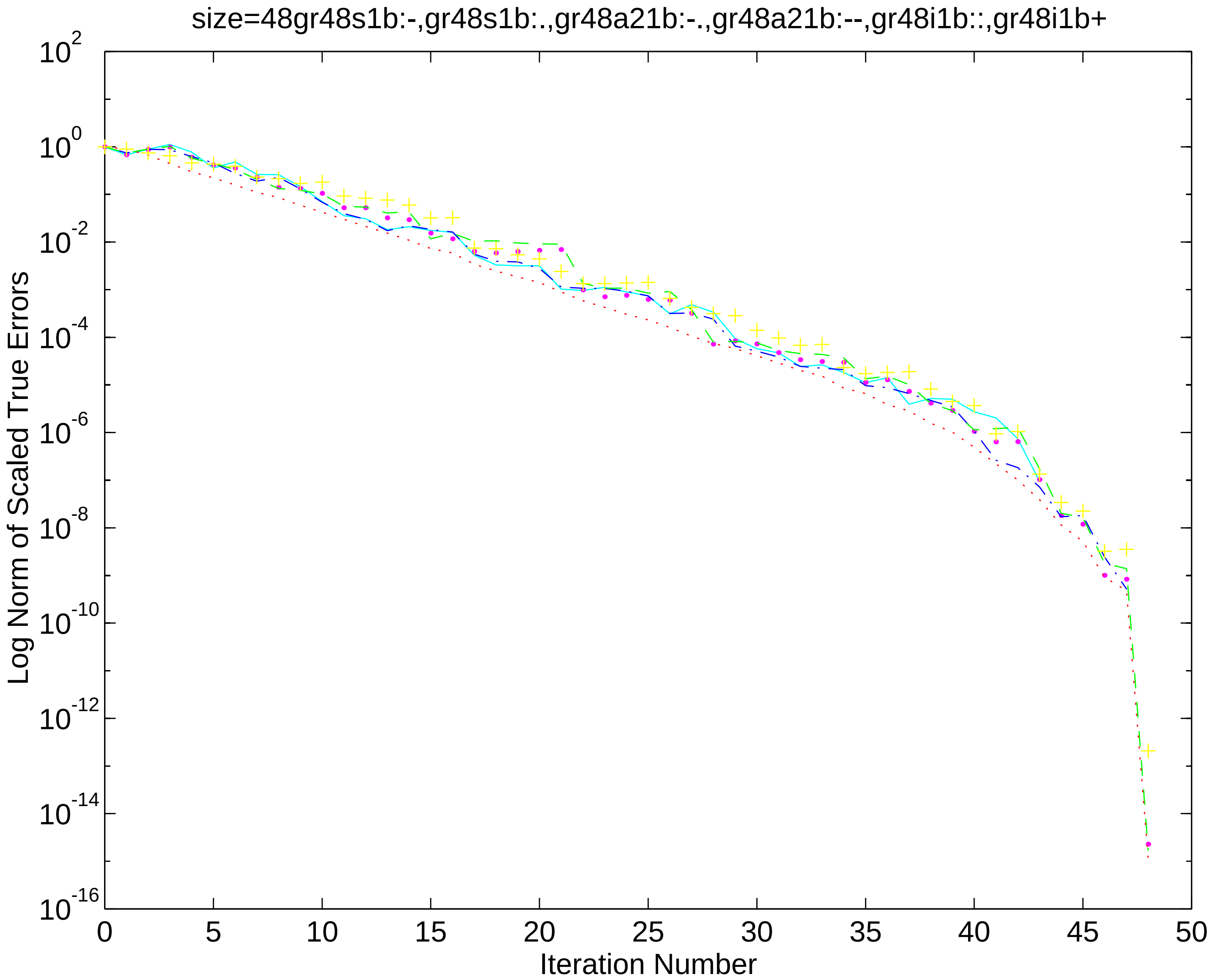

Fig. 4. *GMRES vs. QMR: Error Norms versus Iteration Number. Eigenvalue Distribution:* Λ_{gr48}. *Singular Values Distributions:* $\Sigma_{gr48} = gr$, $\Sigma_{a2r2b0} = a2$, $\Sigma_{I_n} = id$, *Error Norms: GMRES: solid, dot-dash, small dots. QMR: big dots, dashes, +.*

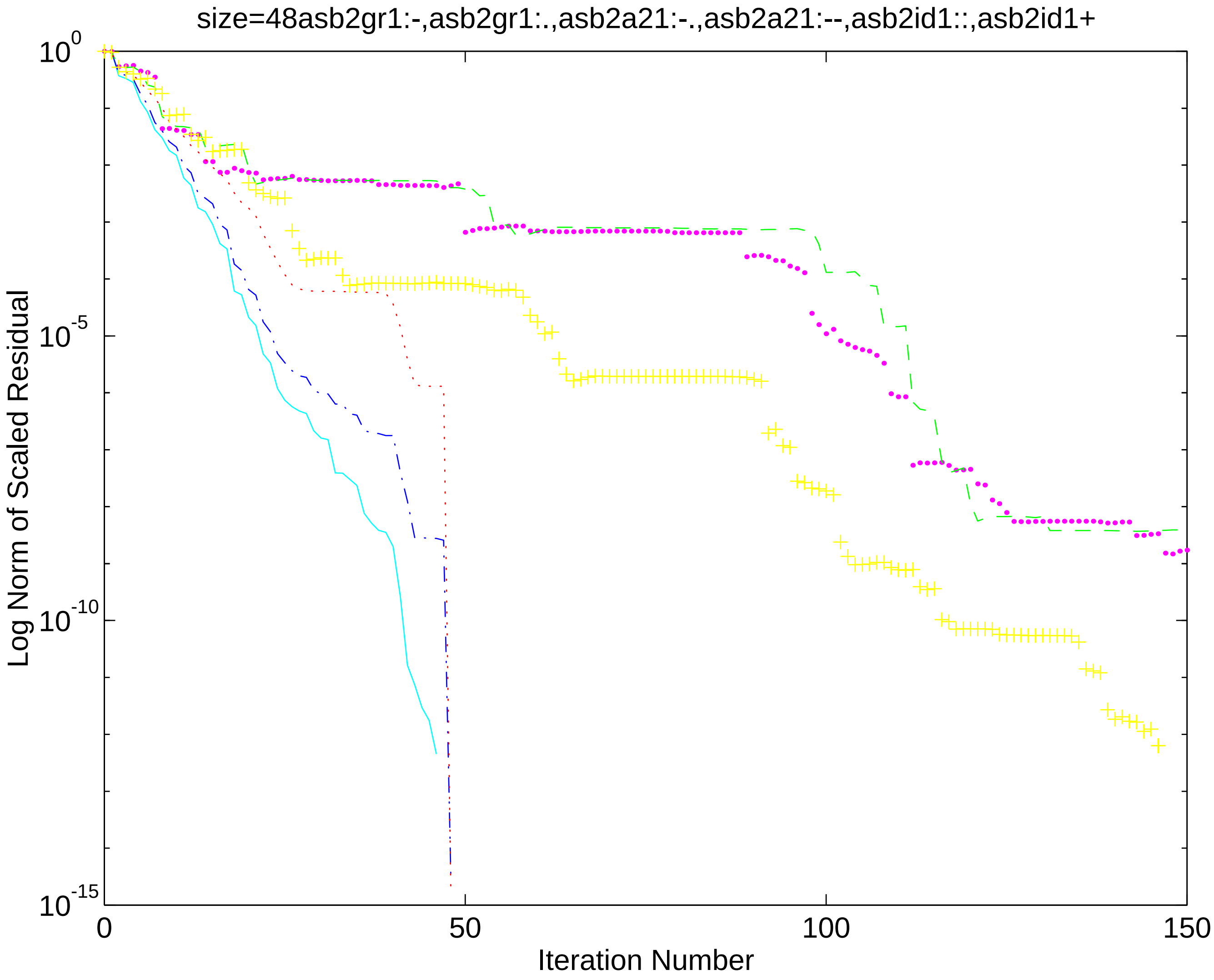

FIG. 5. *GMRES vs. QMR: Residual Norms versus Iteration Number. Eigenvalue Distribution:* Λ_{asb2}. *Singular Values Distributions:* $\Sigma_{gr48} = gr$, $\Sigma_{a2r2b0} = a2$, $\Sigma_{I_n} = id$, *Residual Norms: GMRES: solid, dot-dash, small dots. QMR: big dots, dashes, +.*

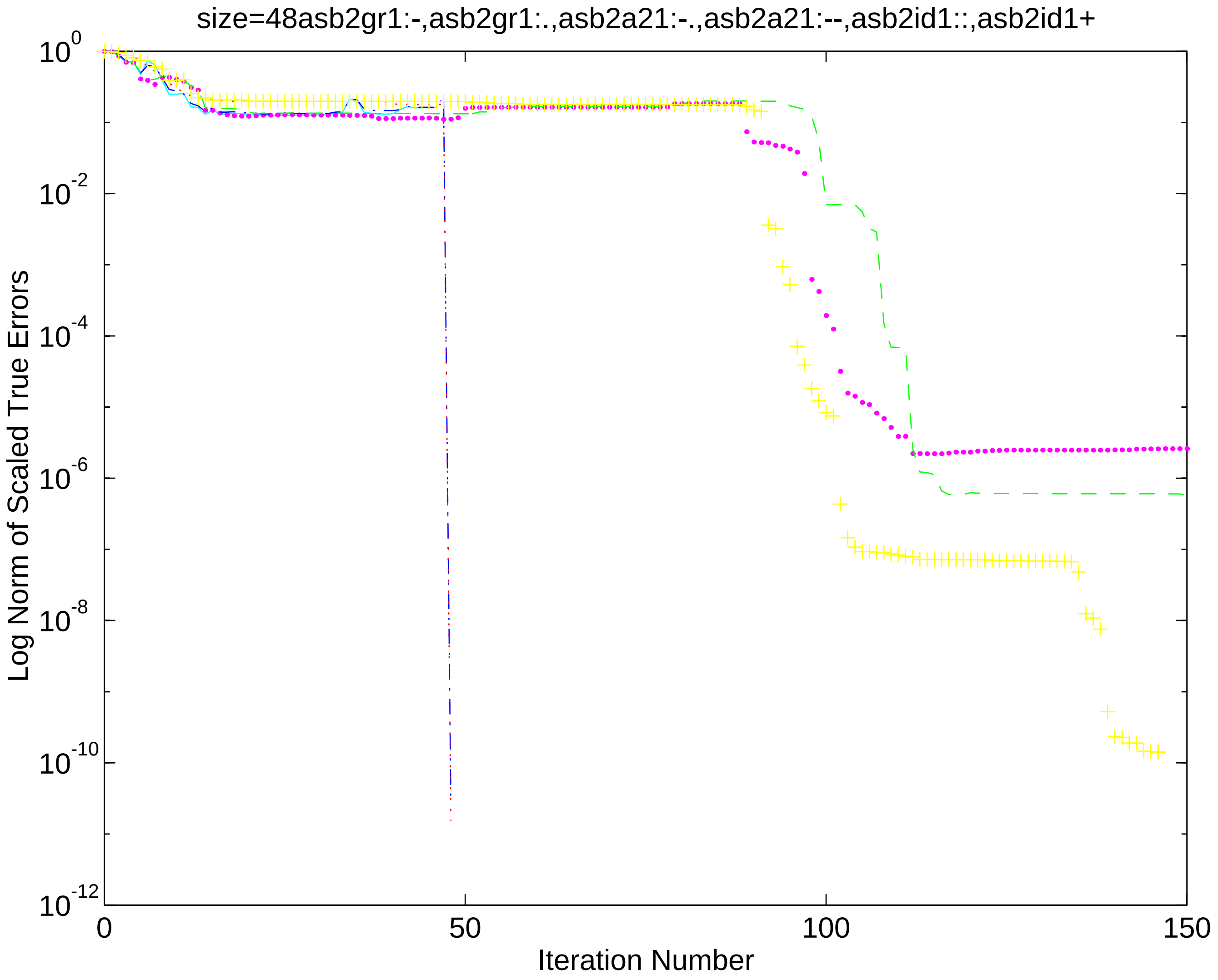

FIG. 6. *GMRES vs. QMR: Error Norms versus Iteration Number. Eigenvalue Distribution: Λ_{asb2}. Singular Values Distributions: $\Sigma_{gr48} = gr$, $\Sigma_{a2r2b0} = a2$, $\Sigma_{I_n} = id$, Error Norms: GMRES: solid, dot-dash, small dots. QMR: big dots, dashes, +.*

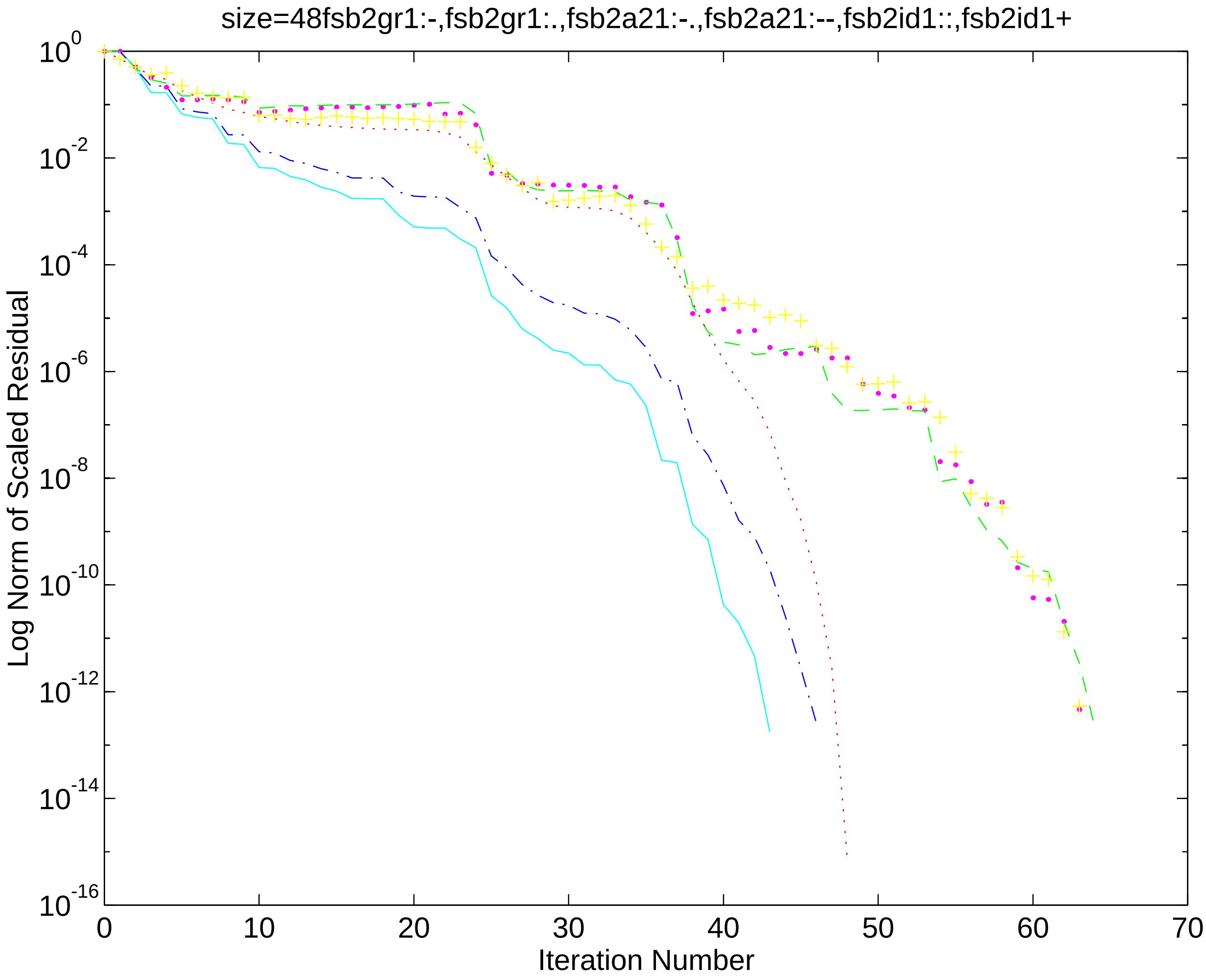

FIG. 7. *GMRES vs. QMR: Residual Norms versus Iteration Number. Eigenvalue Distribution:* Λ_{fsb2}*. Singular Values Distributions:* $\Sigma_{gr48} = gr$*,* $\Sigma_{a2r2b0} = a2$*,* $\Sigma_{I_n} = id$*, Residual Norms: GMRES: solid, dot-dash, small dots. QMR: big dots, dashes, +.*

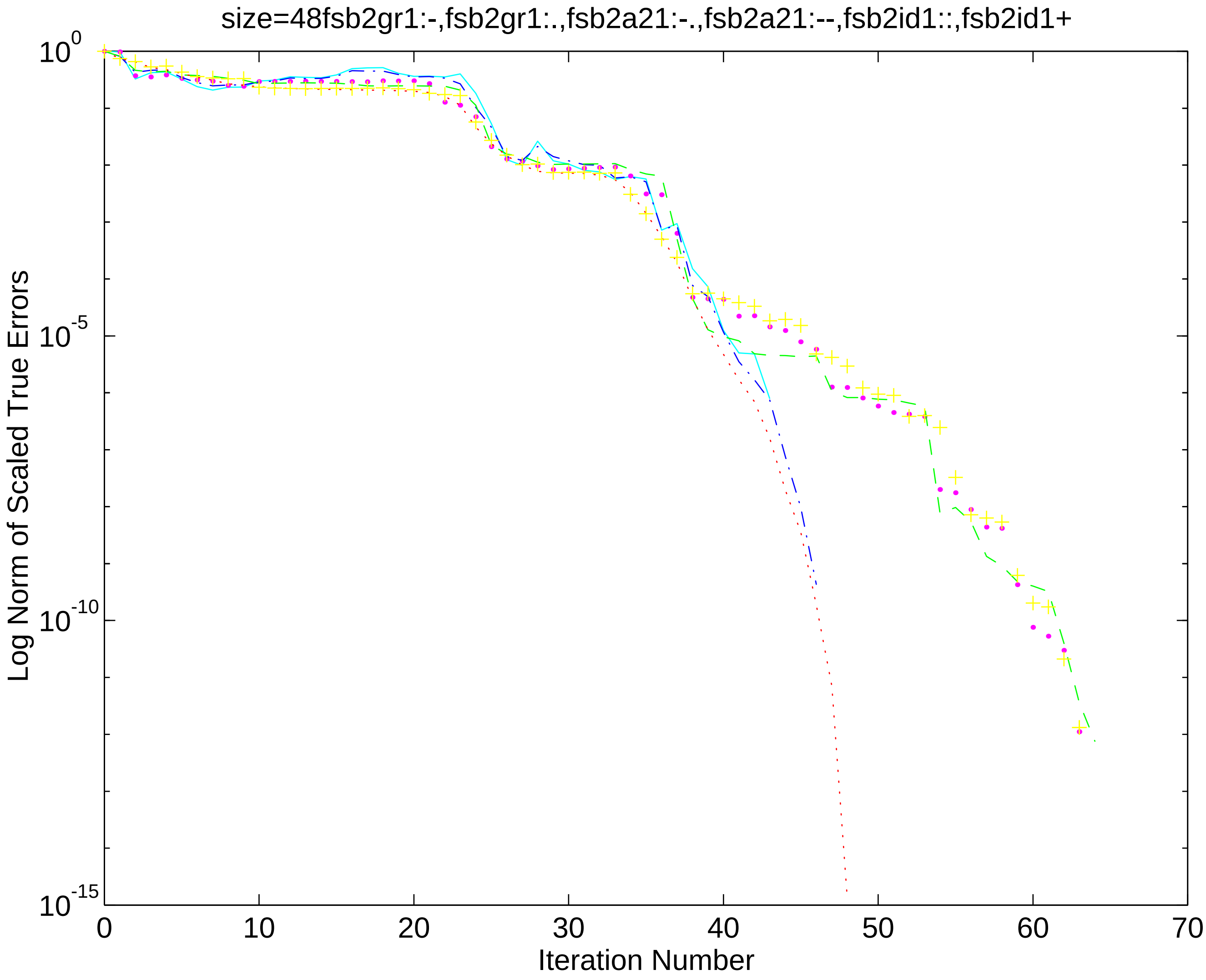

FIG. 8. *GMRES vs. QMR: Error Norms versus Iteration Number. Eigenvalue Distribution:* Λ_{fsb2}. *Singular Values Distributions:* $\Sigma_{gr48} = gr$, $\Sigma_{a2r2b0} = a2$, $\Sigma_{I_n} = id$, *Error Norms: GMRES: solid, dot-dash, small dots. QMR: big dots, dashes, +.*

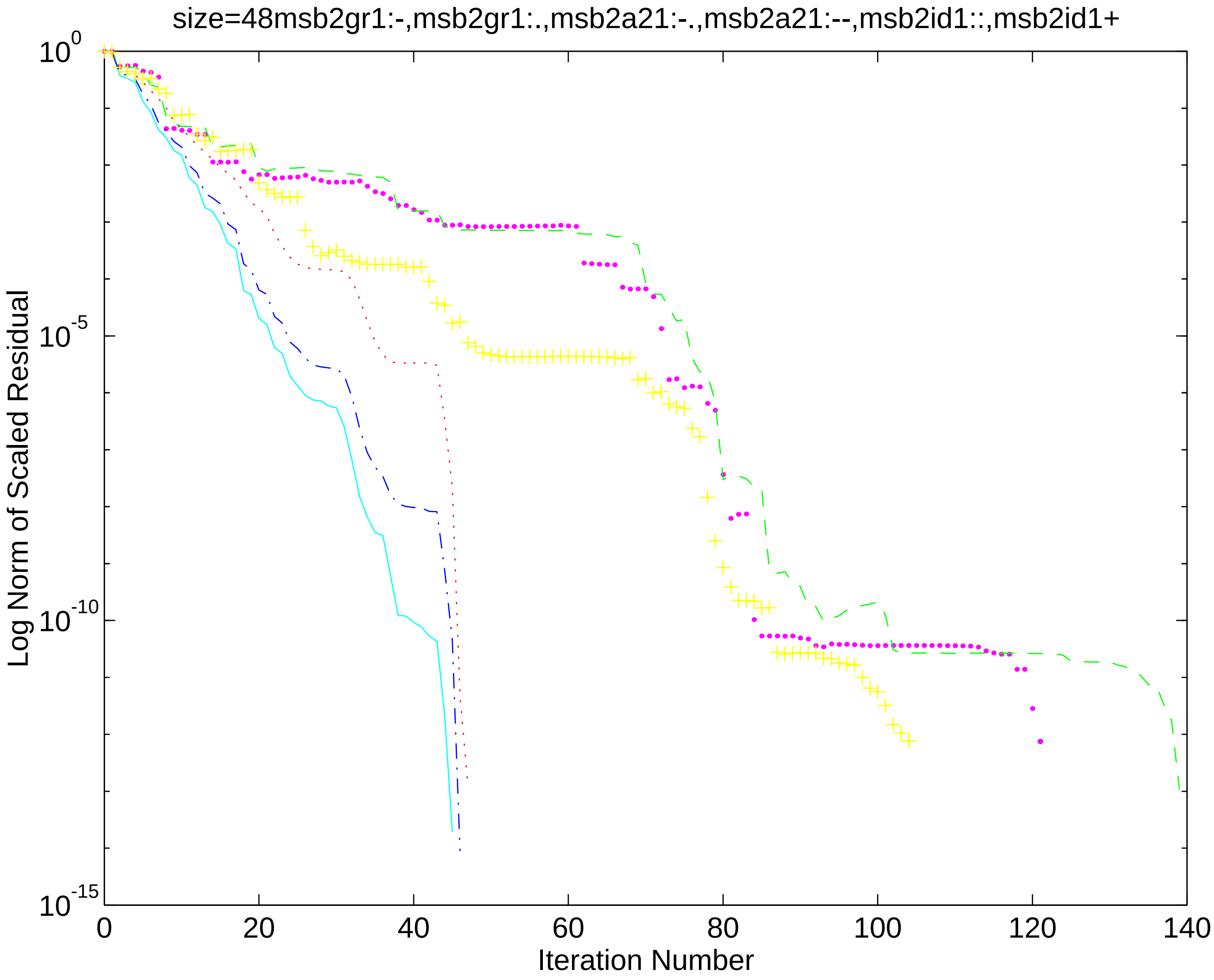

FIG. 9. *GMRES vs. QMR: Residual Norms versus Iteration Number. Eigenvalue Distribution:* Λ_{msb2}. *Singular Values Distributions:* $\Sigma_{gr48} = gr$, $\Sigma_{a2r2b0} = a2$, $\Sigma_{I_n} = id$, *Residual Norms: GMRES: solid, dot-dash, small dots. QMR: big dots, dashes, +.*

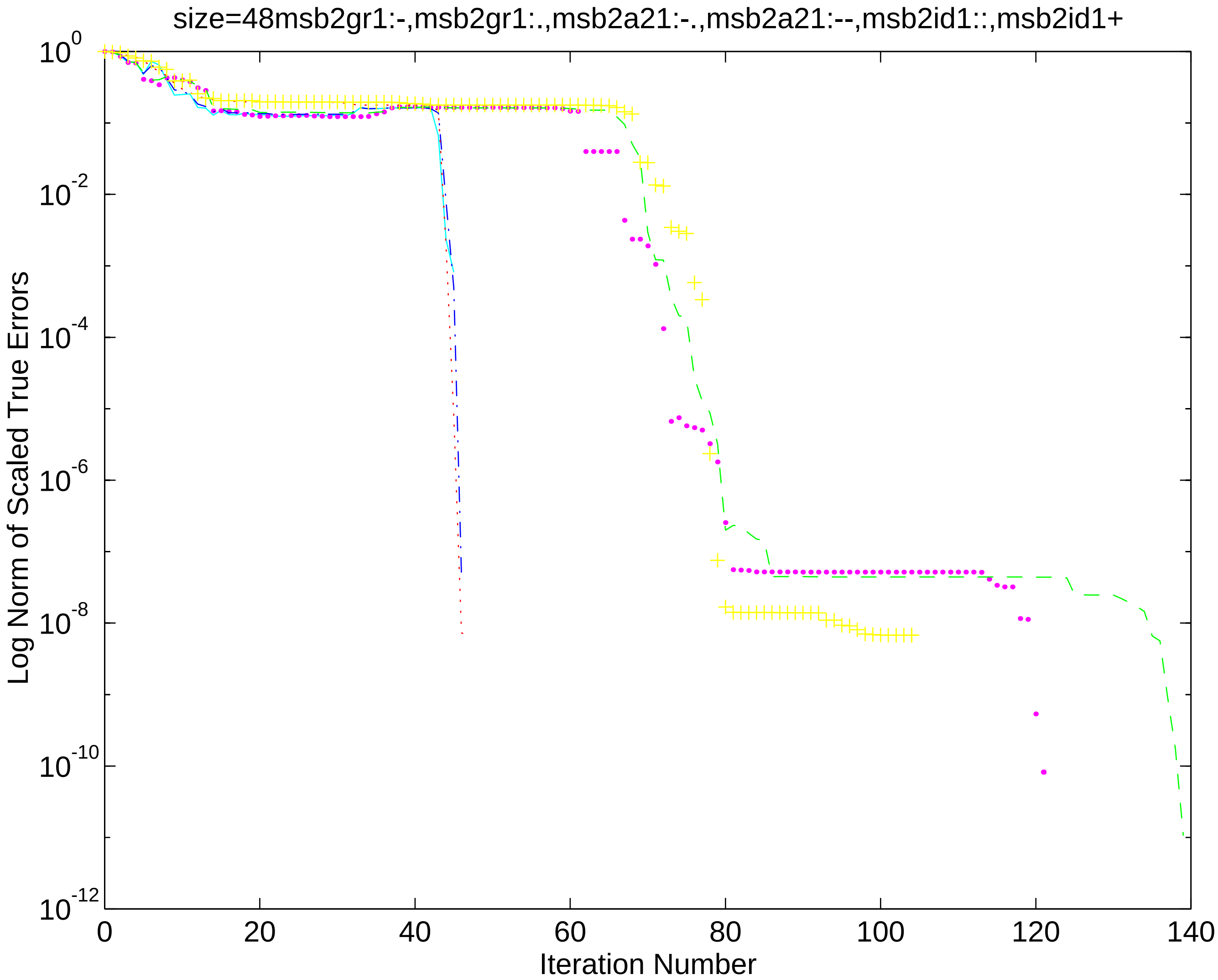

Fig. 10. *GMRES vs. QMR: Error Norms versus Iteration Number. Eigenvalue Distribution:* Λ_{msb2}. *Singular Values Distributions:* $\Sigma_{gr48} = gr$, $\Sigma_{a2r2b0} = a2$, $\Sigma_{I_n} = id$, *Error Norms: GMRES: solid, dot-dash, small dots. QMR: big dots, dashes, +.*

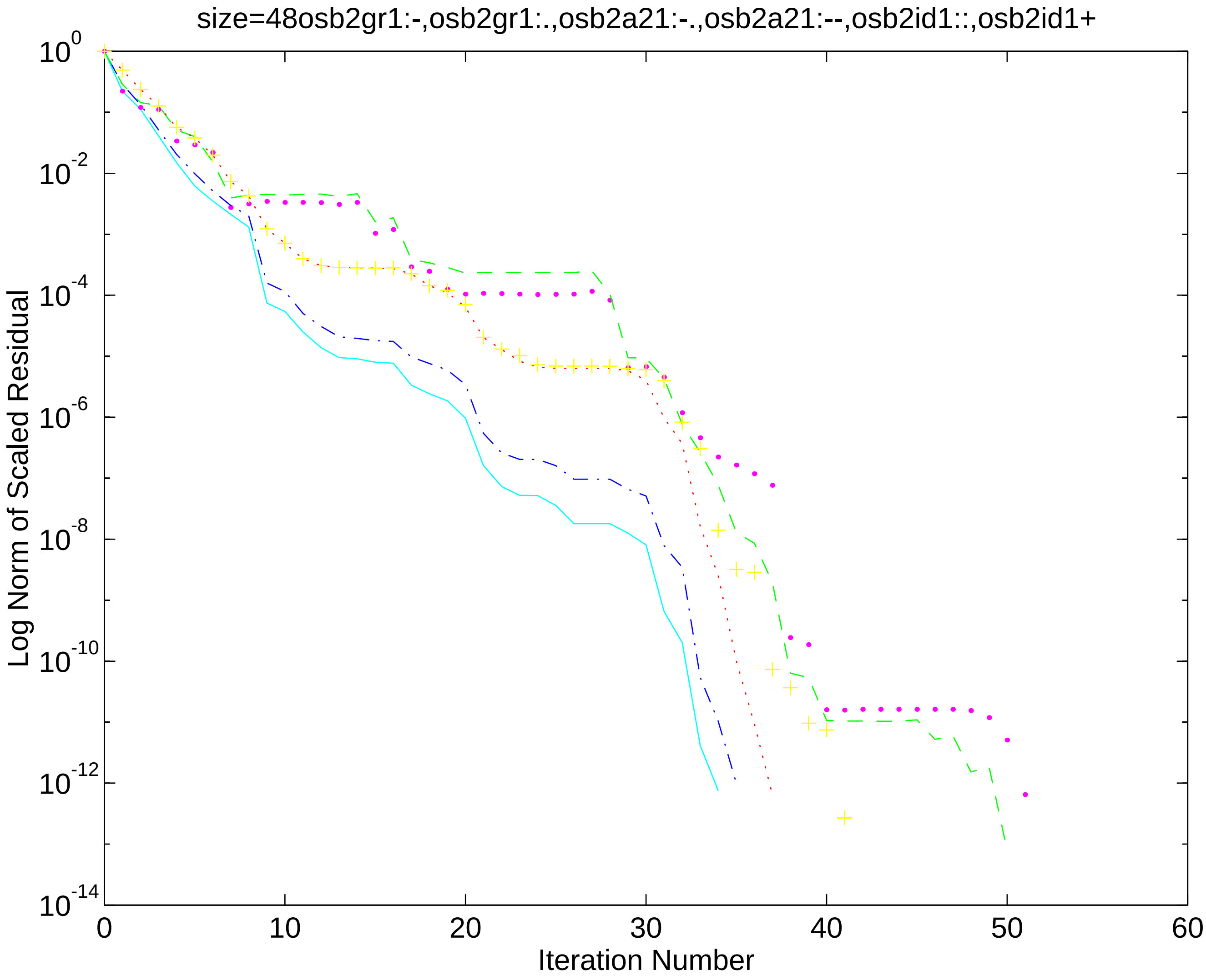

FIG. 11. *GMRES vs. QMR: Residual Norms versus Iteration Number. Eigenvalue Distribution:* Λ_{osb2}. *Singular Values Distributions:* $\Sigma_{gr48} = gr$, $\Sigma_{a2r2b0} = a2$, $\Sigma_{I_n} = id$, *Residual Norms: GMRES: solid, dot-dash, small dots. QMR: big dots, dashes, +.*

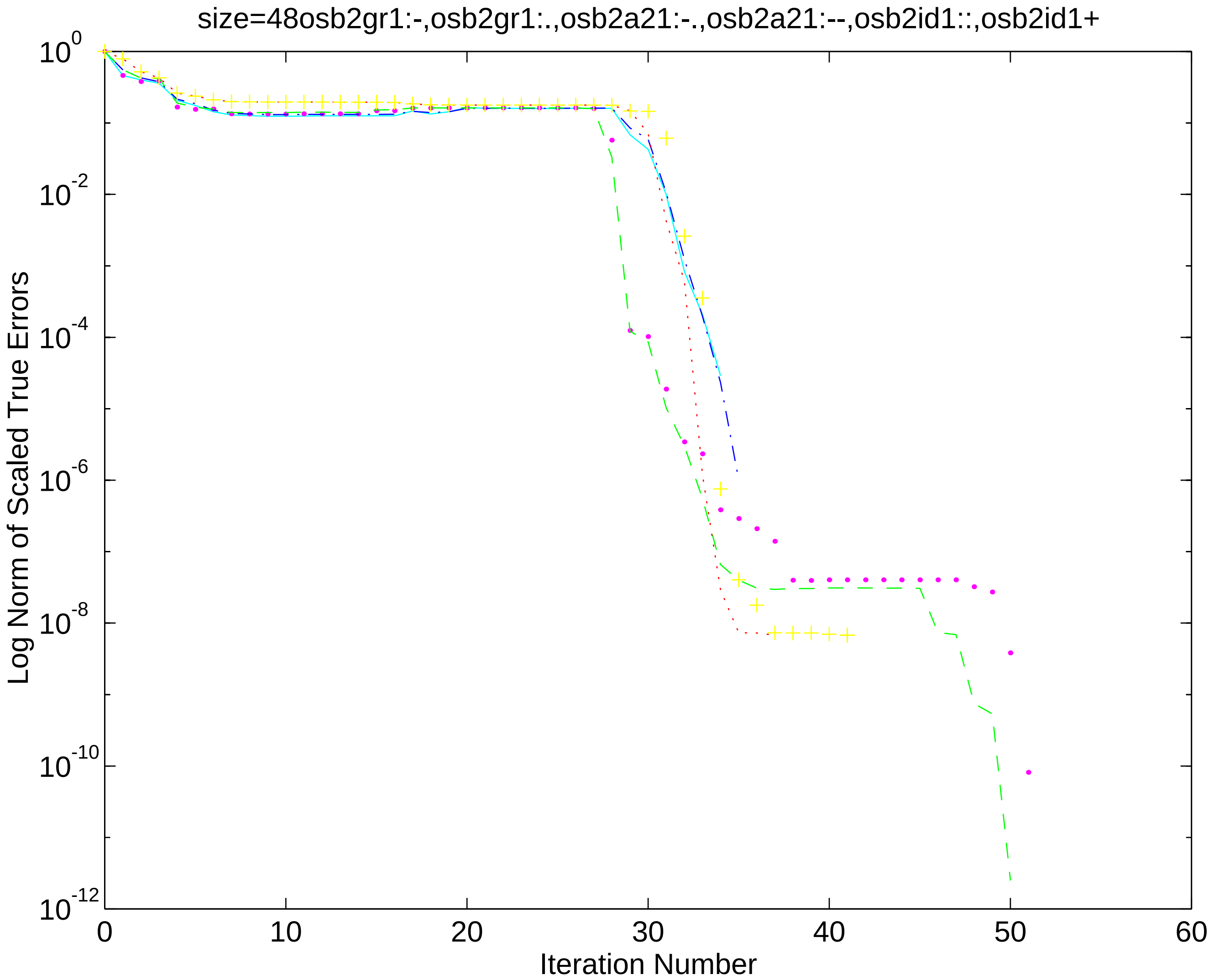

FIG. 12. *GMRES vs. QMR: Error Norms versus Iteration Number. Eigenvalue Distribution:* Λ_{osb2}. *Singular Values Distributions:* $\Sigma_{gr48} = gr$, $\Sigma_{a2r2b0} = a2$, $\Sigma_{I_n} = id$, *Error Norms: GMRES: solid, dot-dash, small dots. QMR: big dots, dashes, +.*

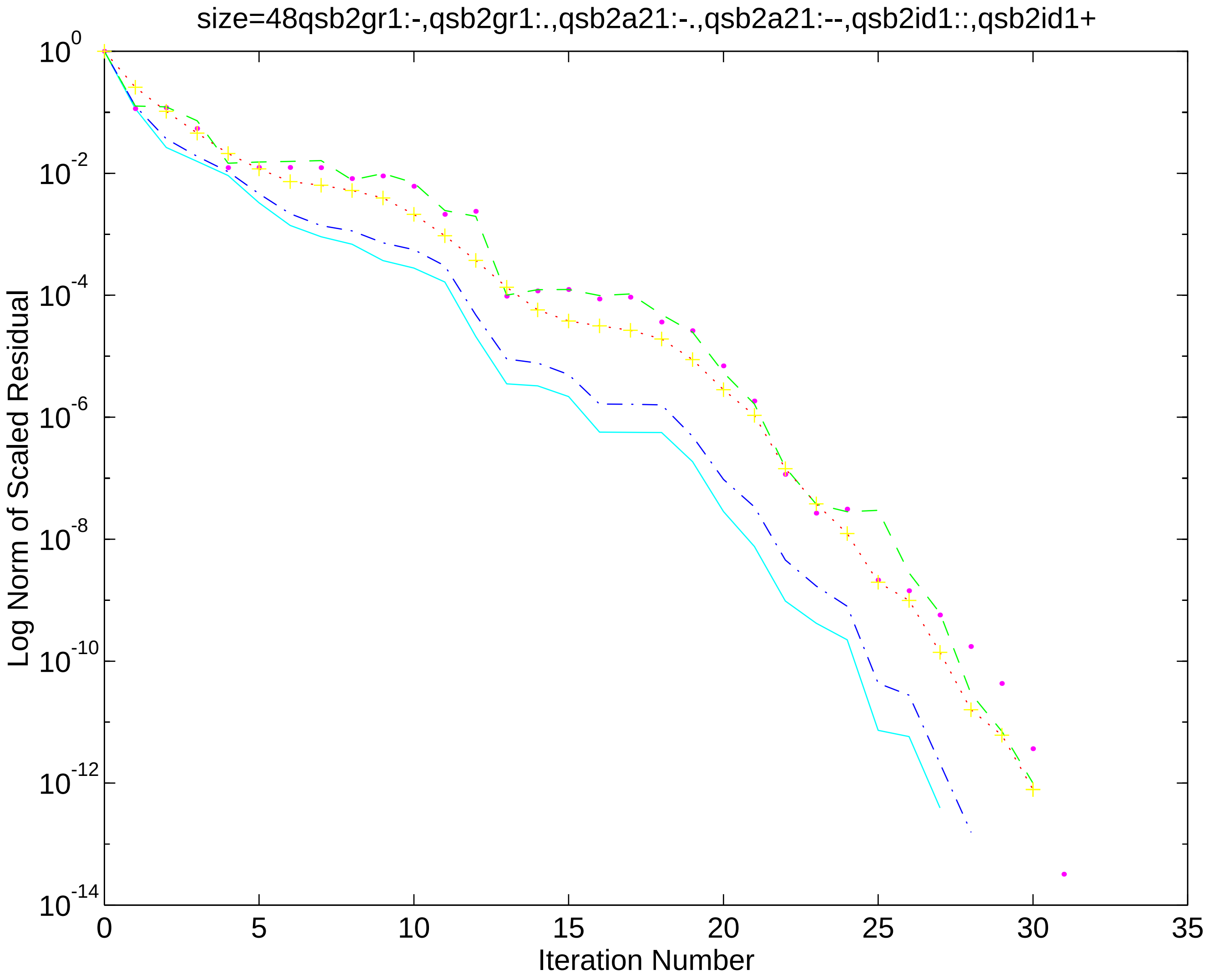

FIG. 13. *GMRES vs. QMR: Residual Norms versus Iteration Number. Eigenvalue Distribution:* Λ_{qsb2}. *Singular Values Distributions:* $\Sigma_{gr48} = gr$, $\Sigma_{a2r2b0} = a2$, $\Sigma_{I_n} = id$, *Residual Norms: GMRES: solid, dot-dash, small dots. QMR: big dots, dashes,* $+$.

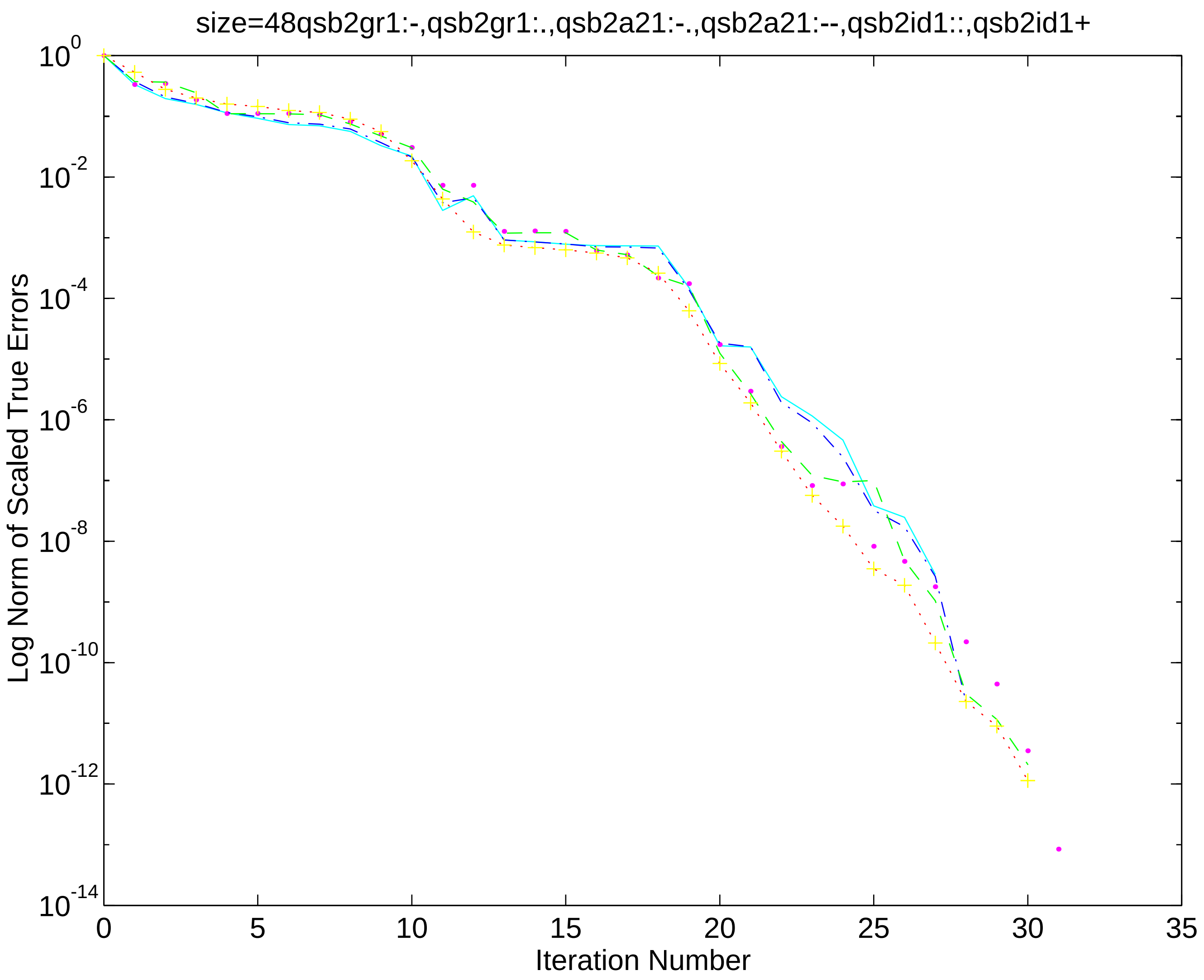

FIG. 14. *GMRES vs. QMR: Error Norms versus Iteration Number. Eigenvalue Distribution:* Λ_{qsb2}. *Singular Values Distributions:* $\Sigma_{gr48} = gr$, $\Sigma_{a2r2b0} = a2$, $\Sigma_{I_n} = id$, *Error Norms: GMRES: solid, dot-dash, small dots. QMR: big dots, dashes, +.*

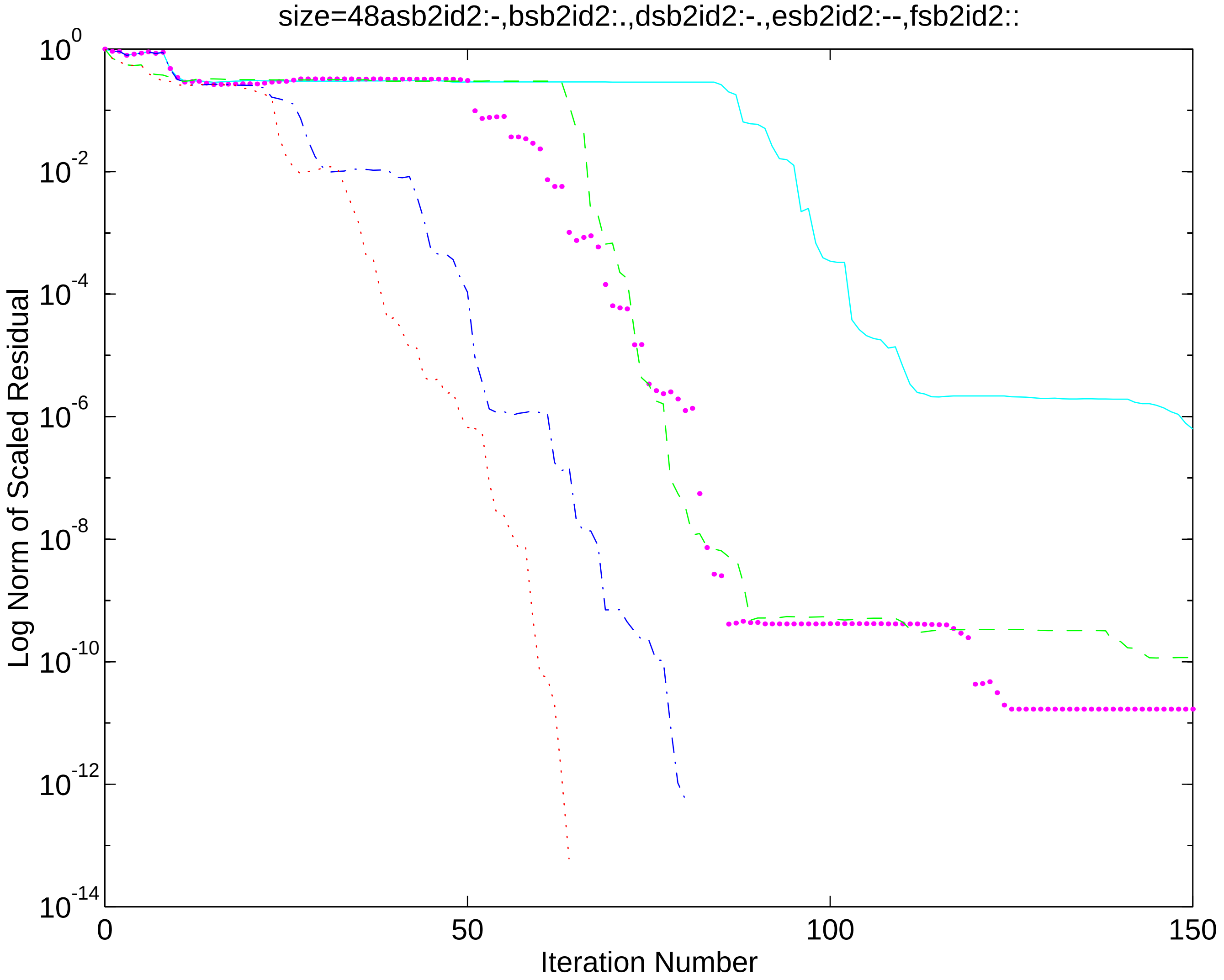

FIG. 15. *QMR: Residual Norms versus Iteration Number: Normal Case: Mixed Real and Complex Eigenvalue Distributions:* Λ_{asb2}, Λ_{bsb2}, Λ_{dsb2}, Λ_{esb2}, Λ_{fsb2}. *Singular Value Distribution:* Σ_{I_n}.

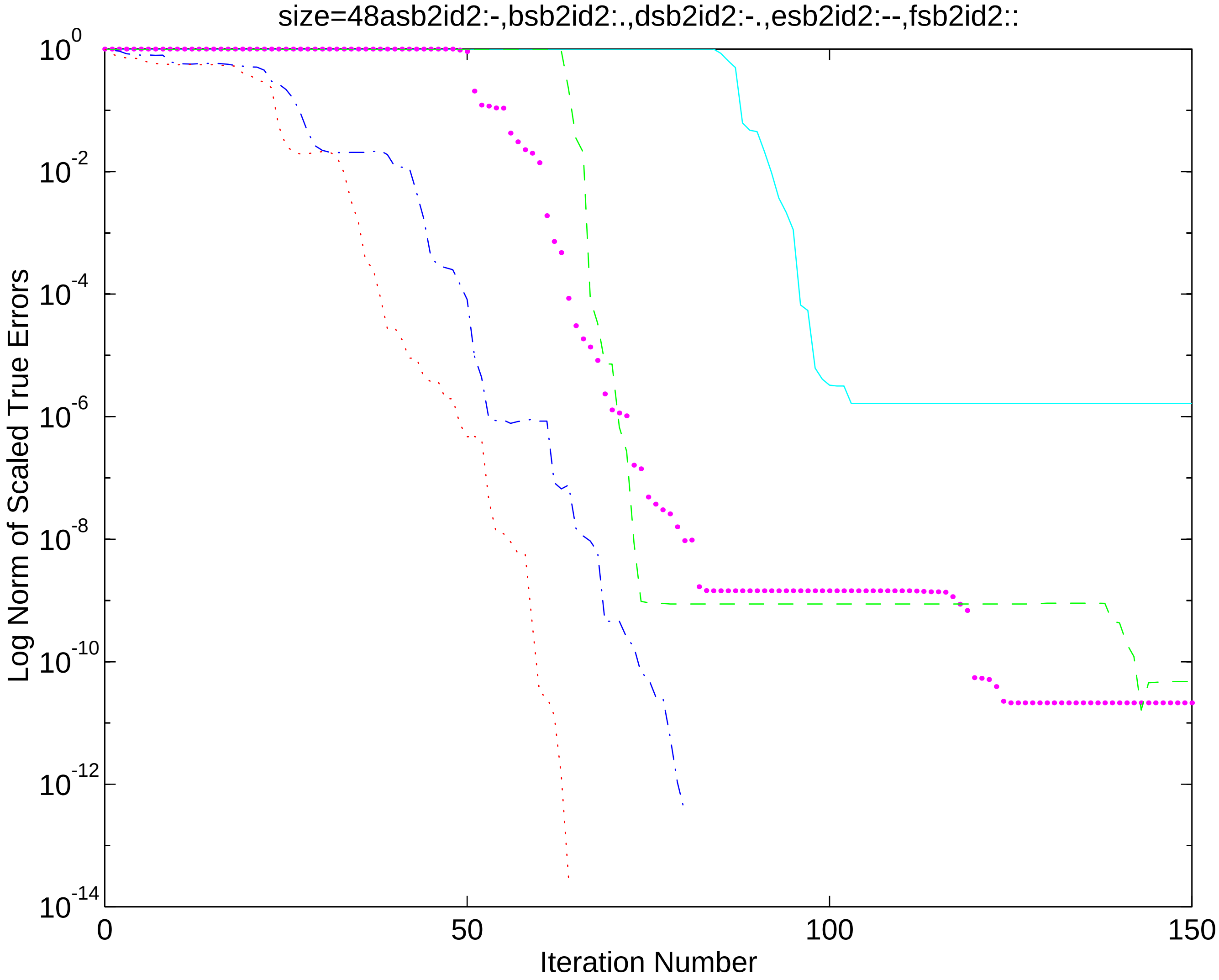

FIG. 16. *QMR: Error Norms versus Iteration Number: Normal Case: Mixed Real and Complex Eigenvalue Distributions:* Λ_{asb2}, Λ_{bsb2}, Λ_{dsb2}, Λ_{esb2}, Λ_{fsb2}. *Singular Value Distribution:* Σ_{I_n}.

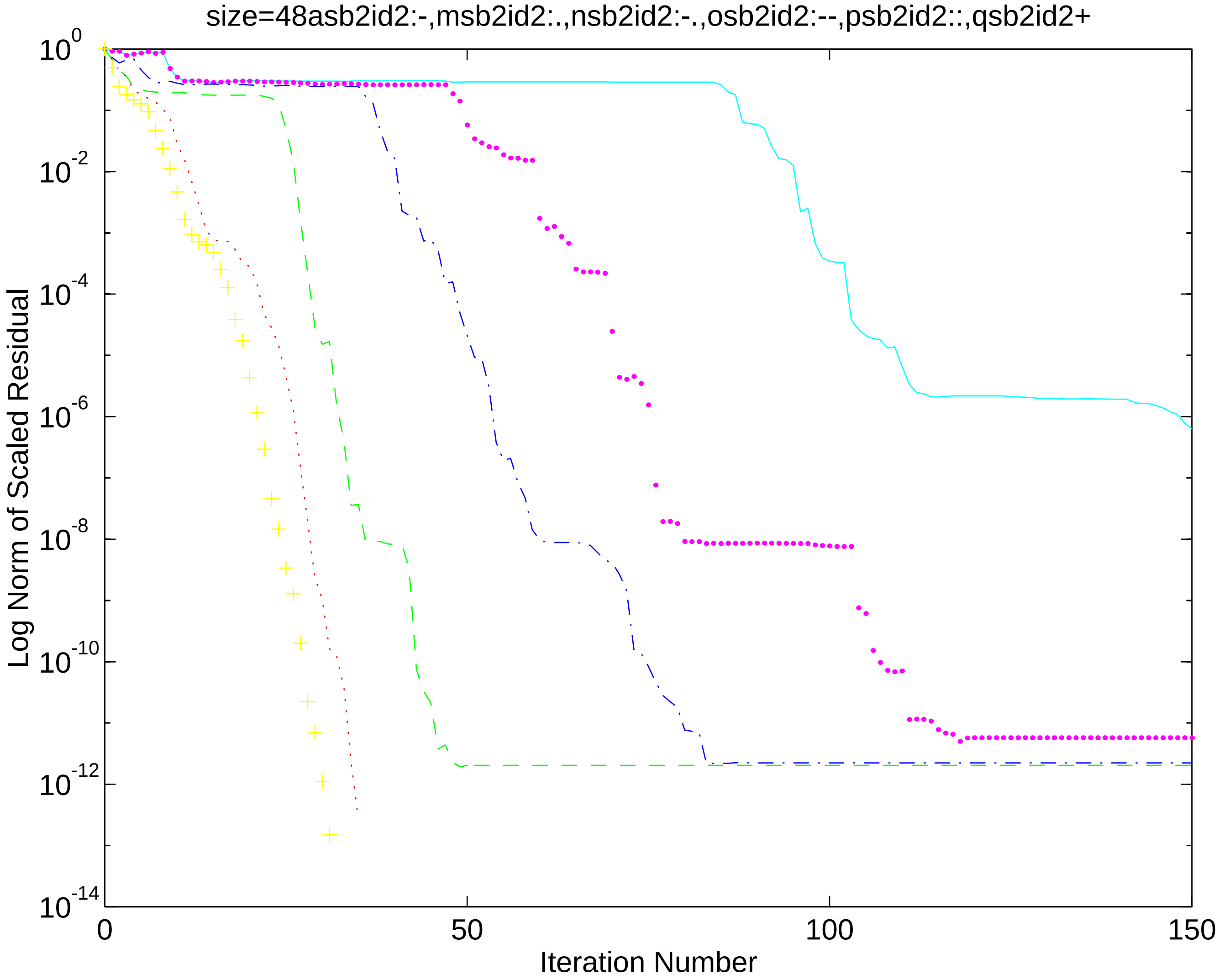

FIG. 17. *QMR: Residual Norms versus Iteration Number: Normal Case: Mixed to Nearly Real Eigenvalue Distributions:* Λ_{asb2}, Λ_{msb2}, Λ_{nsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{I_n}.

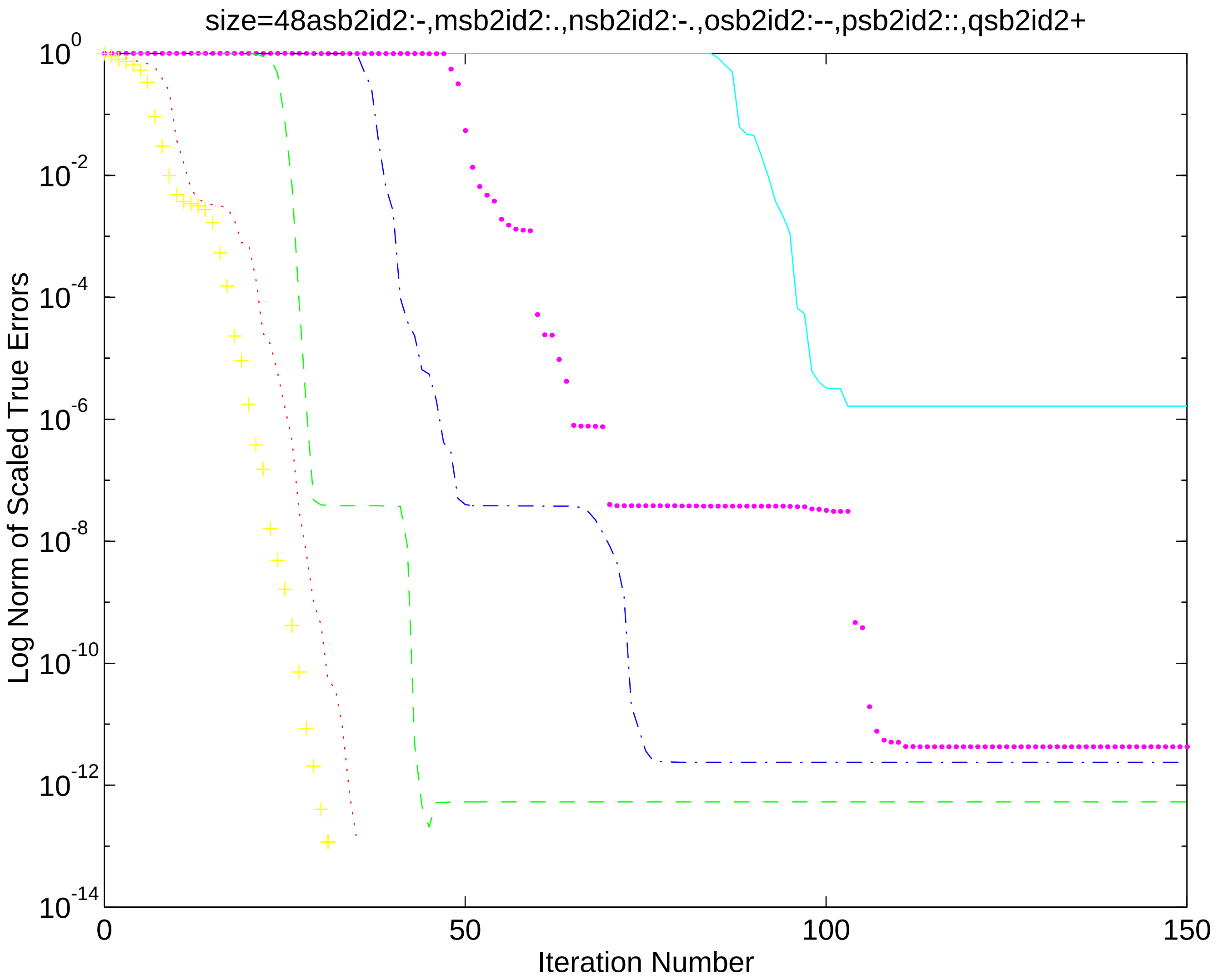

FIG. 18. *QMR: Error Norms versus Iteration Number: Normal Case: Mixed to Nearly Real Eigenvalue Distributions:* Λ_{asb2}, Λ_{msb2}, Λ_{nsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{I_n}.

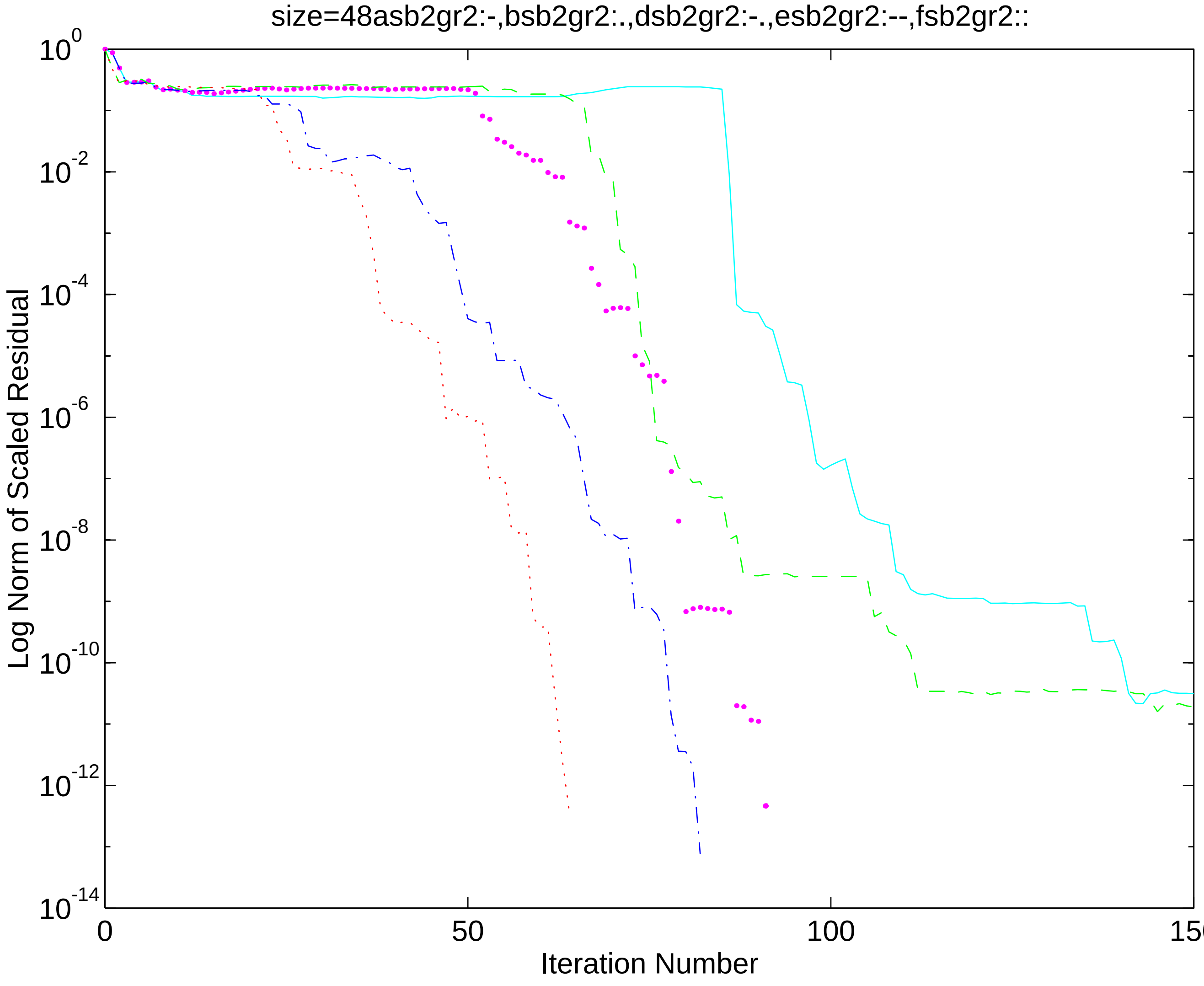

FIG. 19. *QMR: Residual Norms versus Iteration Number: Nonnormal Case: Mixed Real and Complex Eigenvalue Distributions:* Λ_{asb2}, Λ_{bsb2}, Λ_{dsb2}, Λ_{esb2}, Λ_{fsb2}. *Singular Value Distribution:* Σ_{gr48}.

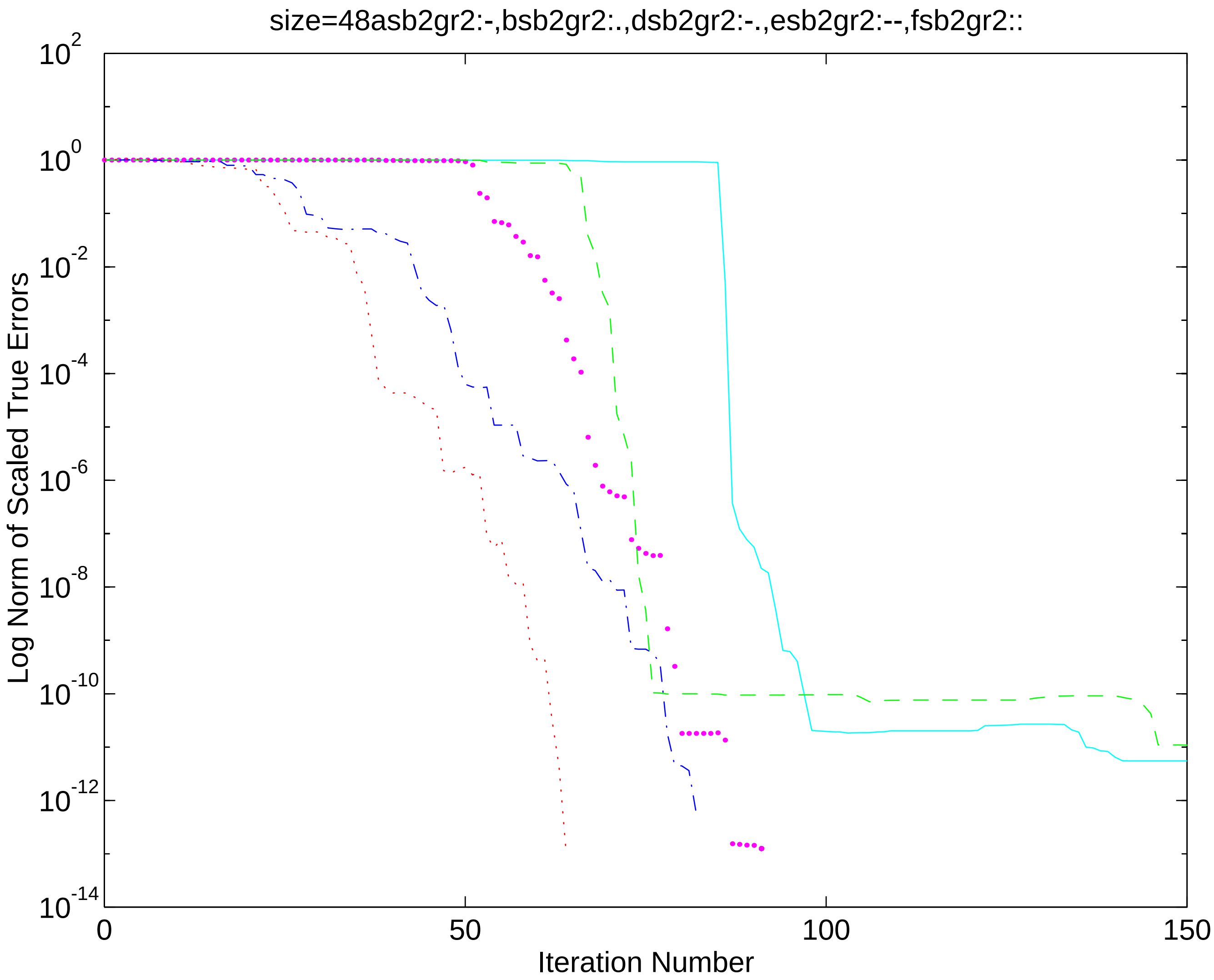

FIG. 20. *QMR: Error Norms versus Iteration Number: Nonnormal Case: Mixed Real and Complex Eigenvalue Distributions:* Λ_{asb2}, Λ_{bsb2}, Λ_{dsb2}, Λ_{esb2}, Λ_{fsb2}. *Singular Value Distribution:* Σ_{gr48}.

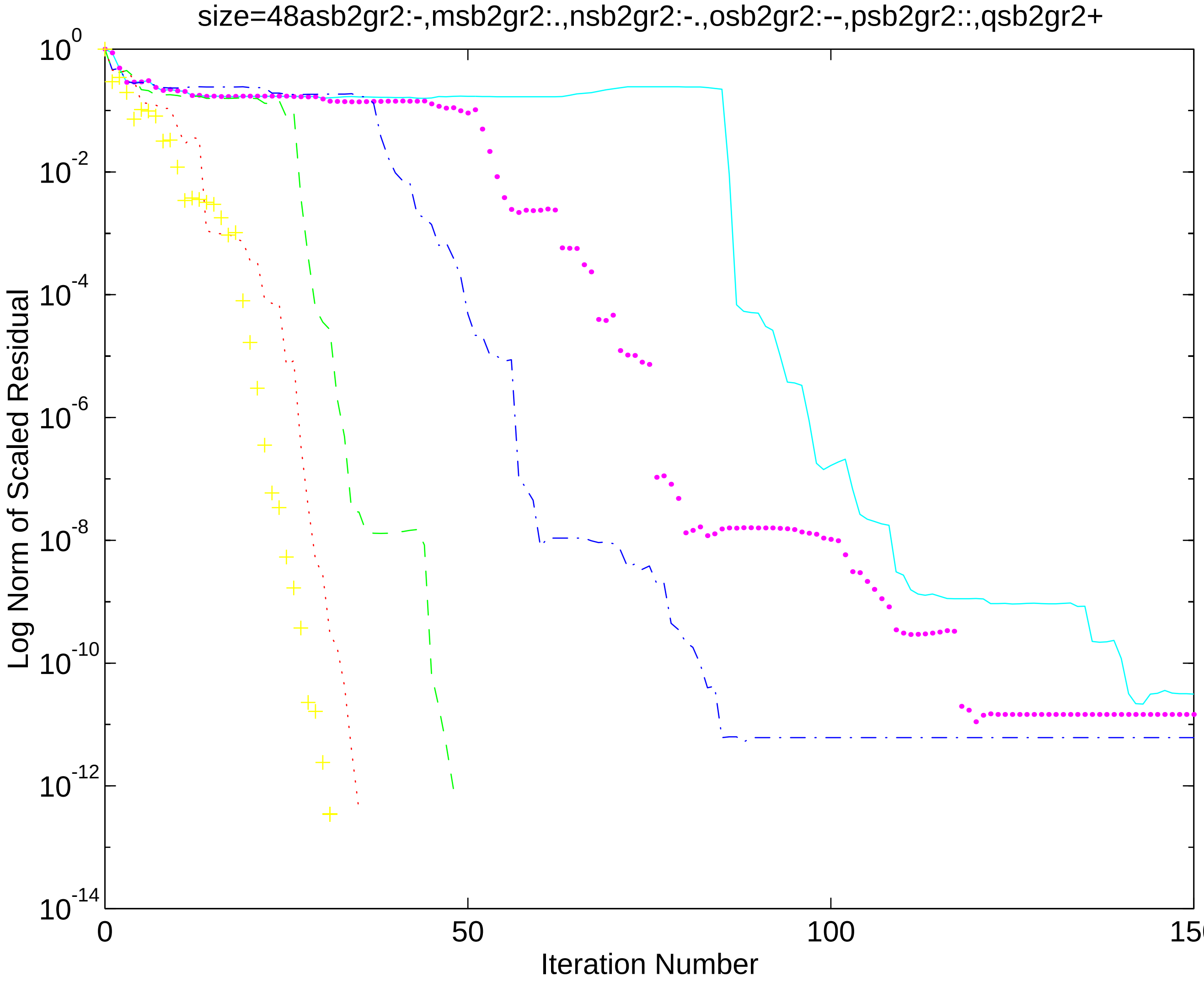

FIG. 21. *QMR: Residual Norms versus Iteration Number: Nonnormal Case: Mixed to Nearly Real Eigenvalue Distributions:* Λ_{asb2}, Λ_{msb2}, Λ_{nsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}.. *Singular Value Distribution:* Σ_{gr48}.

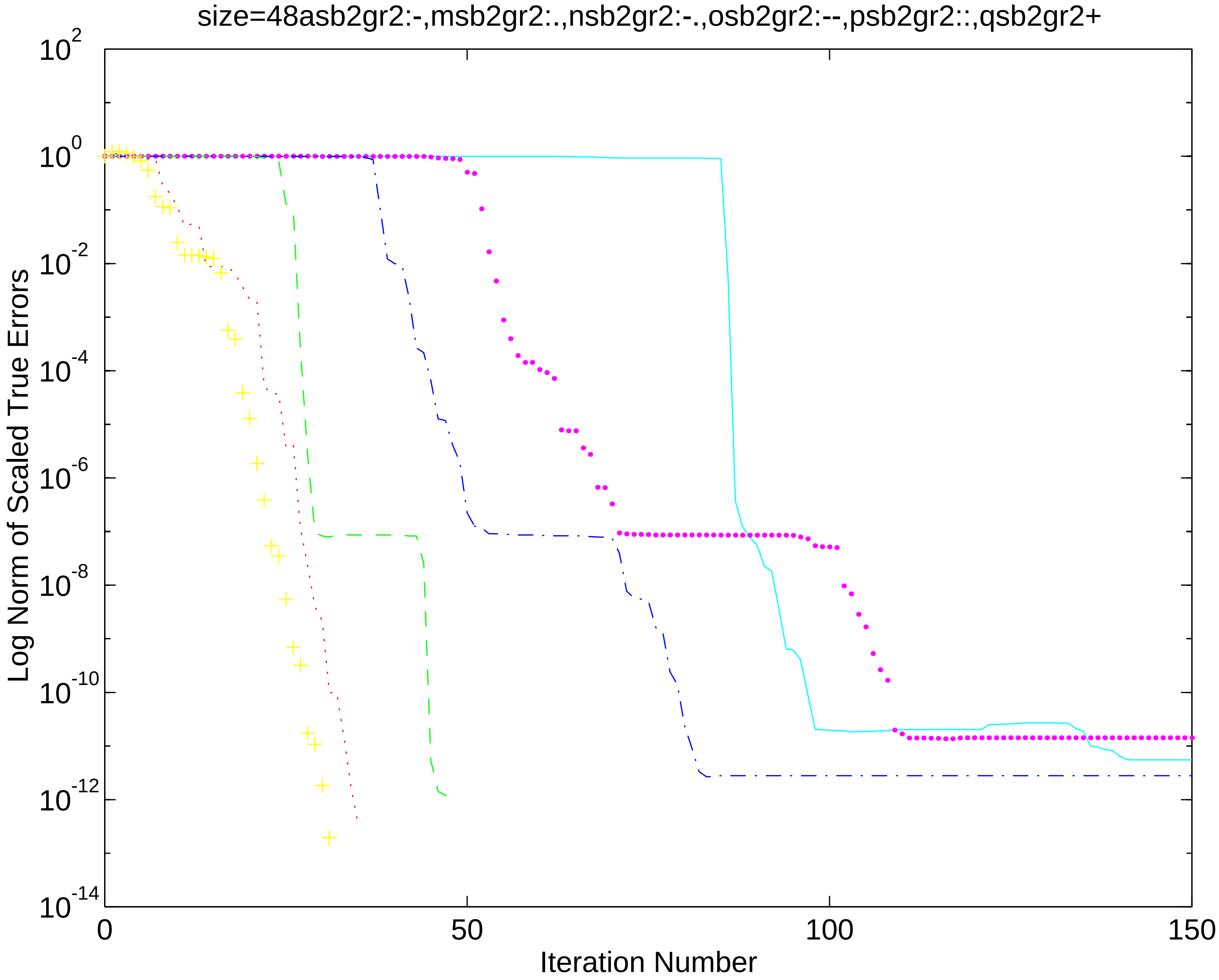

FIG. 22. *QMR: Error Norms versus Iteration Number: Nonnormal Case: Mixed to Nearly Real Eigenvalue Distributions:* Λ_{asb2}, Λ_{msb2}, Λ_{nsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}.. *Singular Value Distribution:* Σ_{gr48}.

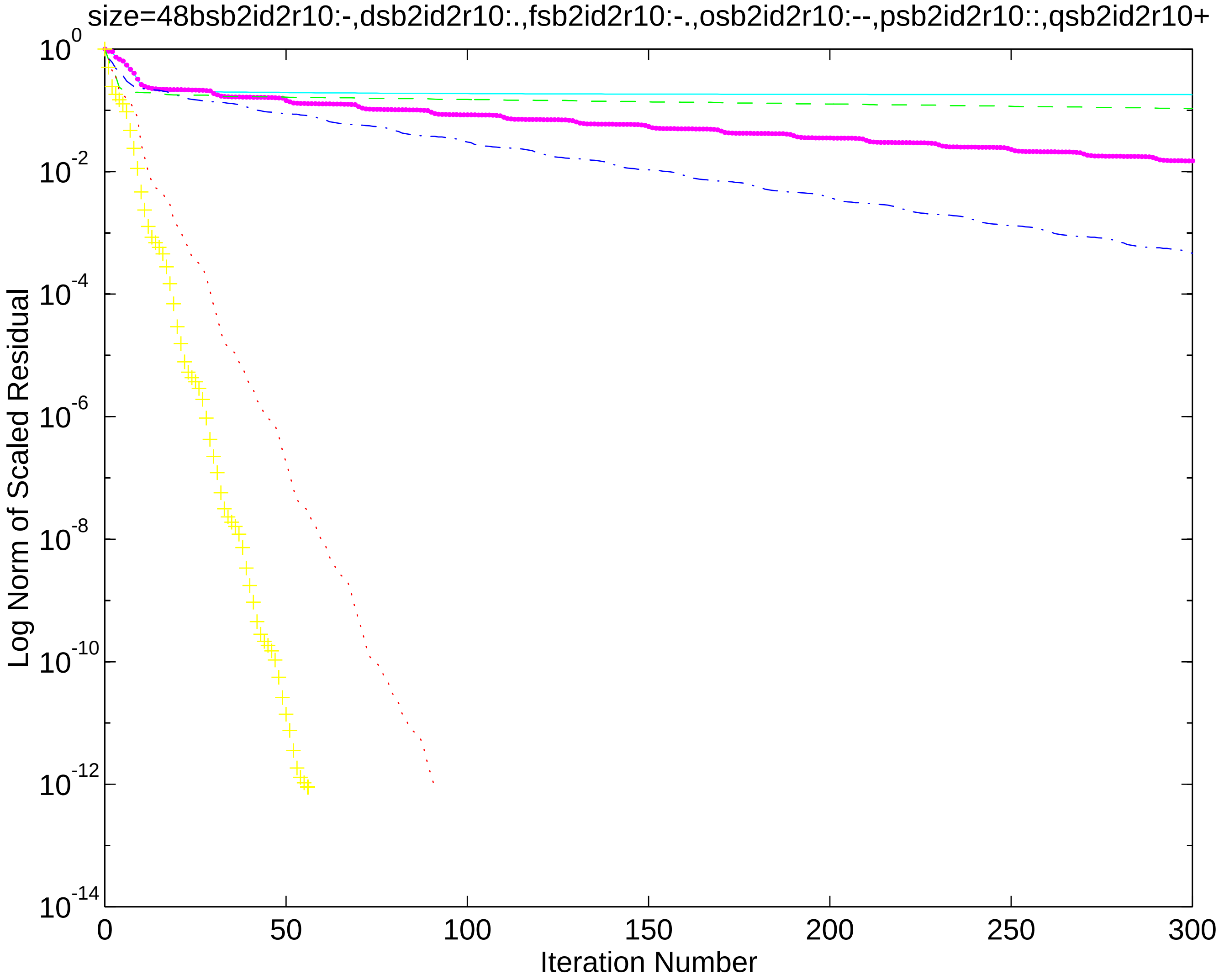

FIG. 23. *GMRES(10): Residual Norms versus Iteration Number: Normal Case: Mixed Real and Complex to Nearly Real Eigenvalue Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{I_n}.

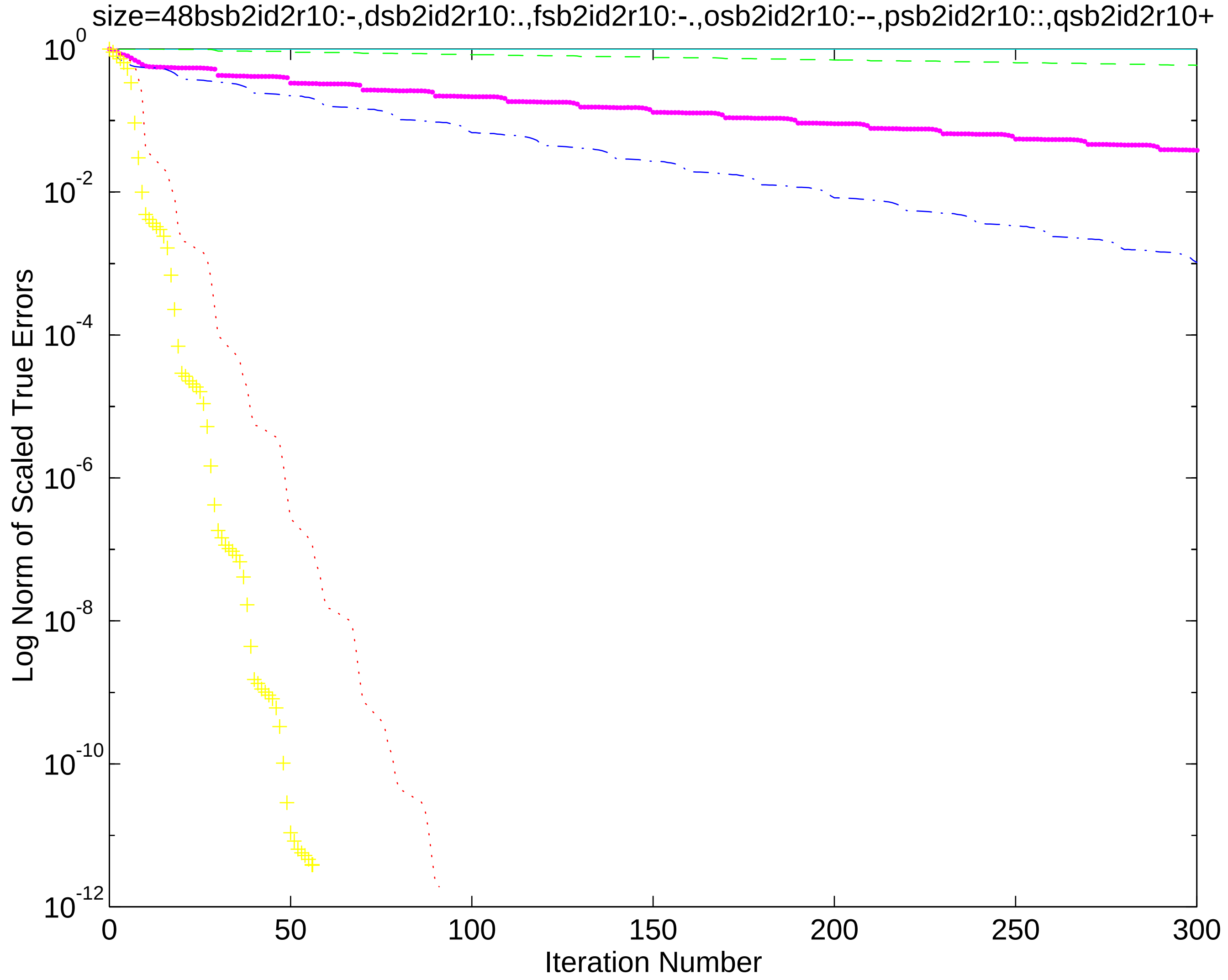

FIG. 24. *GMRES(10): Error Norms versus Iteration Number: Normal Case: Mixed Real and Complex to Nearly Real Eigenvalues Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{I_n}.

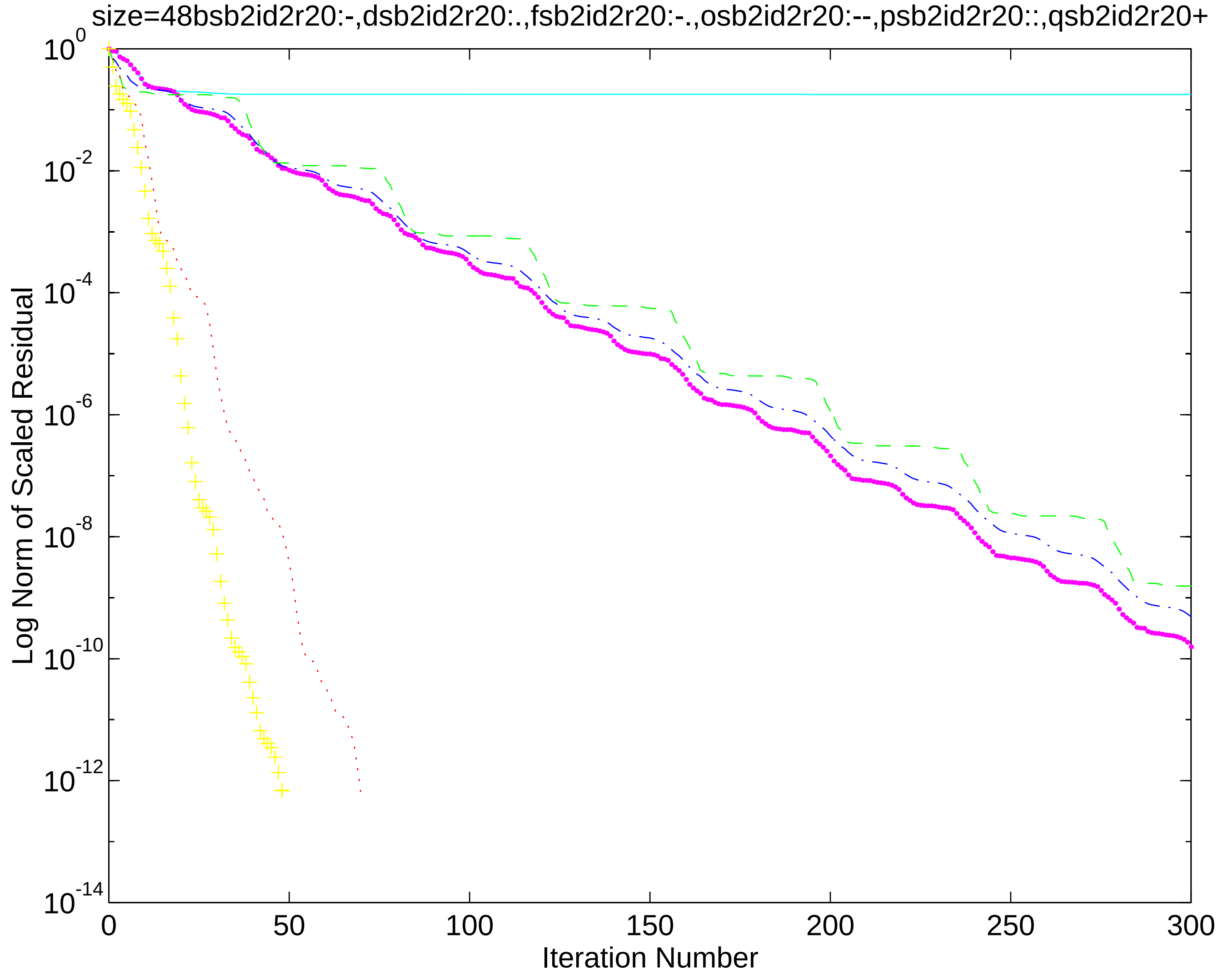

FIG. 25. *GMRES(20): Residual Norms versus Iteration Number: Normal Case: Mixed Real and Complex to Nearly Real Eigenvalue Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{I_n}.

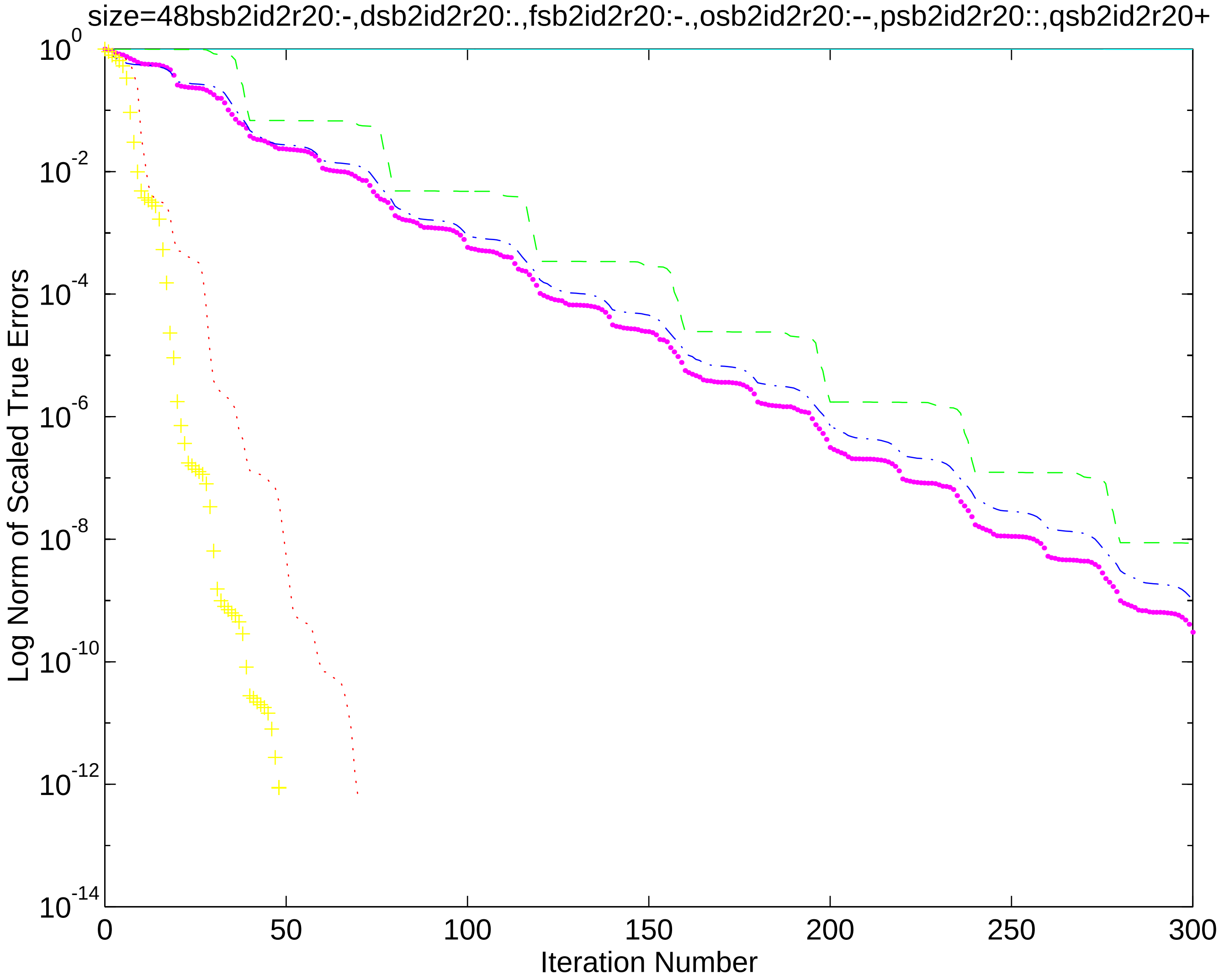

FIG. 26. *GMRES(20): Error Norms versus Iteration Number: Normal Case: Mixed Real and Complex to Nearly Real Eigenvalue Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{I_n}.

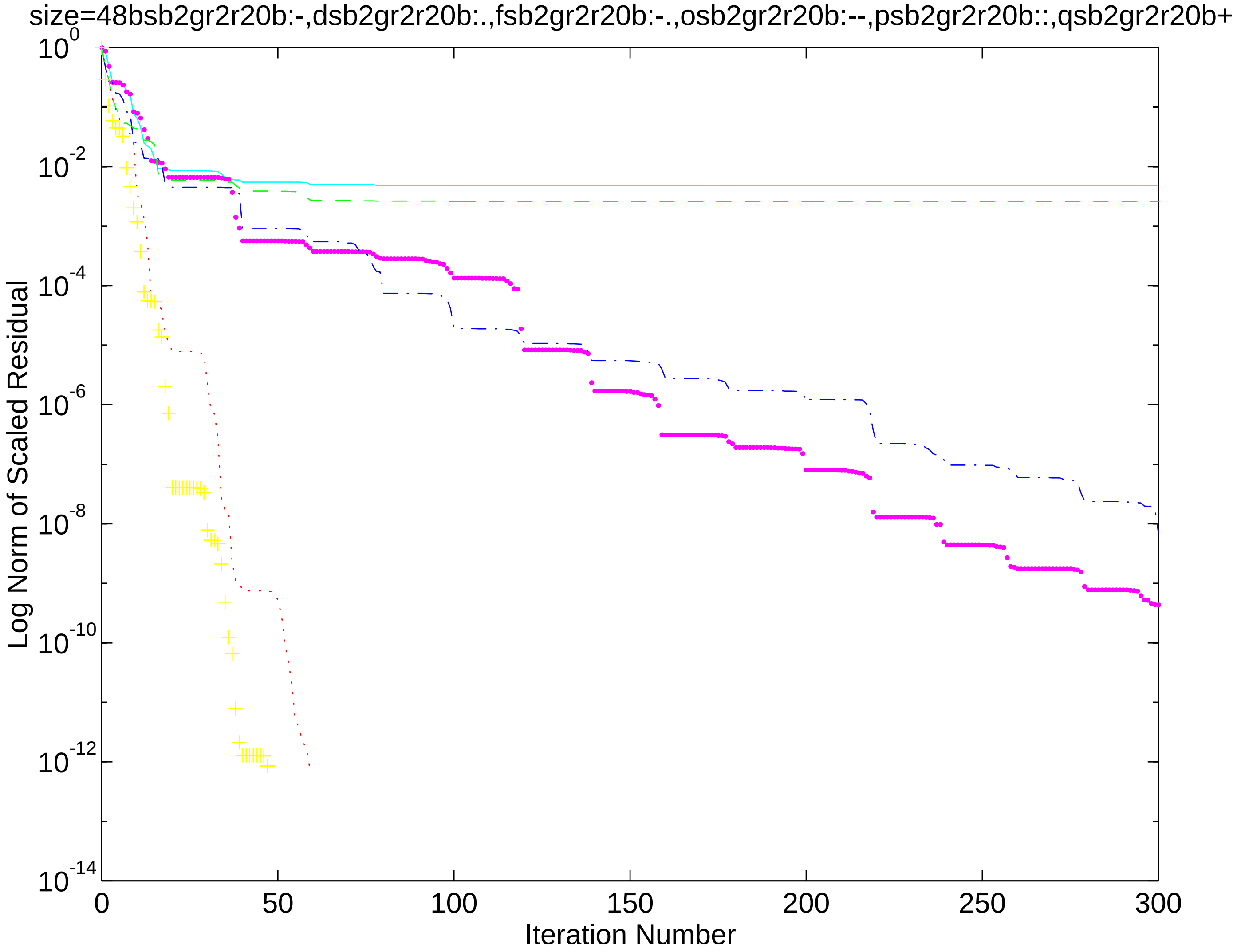

FIG. 27. *GMRES(20): Residual Norms versus Iteration Number: Nonnormal Case: Mixed Real and Complex to Nearly Real Eigenvalue Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{gr48}.

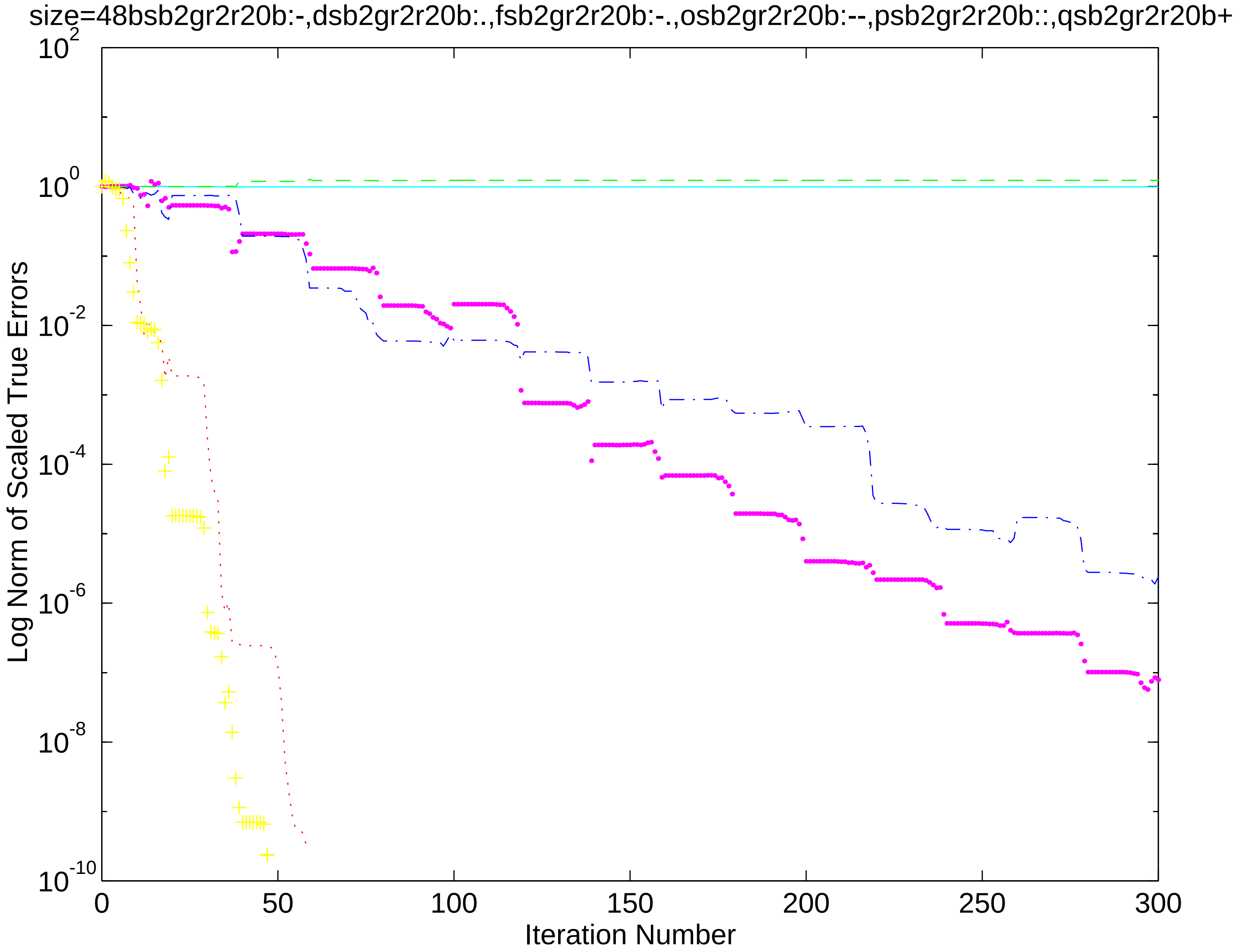

FIG. 28. *GMRES(20): Error Norms versus Iteration Number: Nonnormal Case: Mixed Real and Complex to Nearly Real Eigenvalue Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{gr48}.

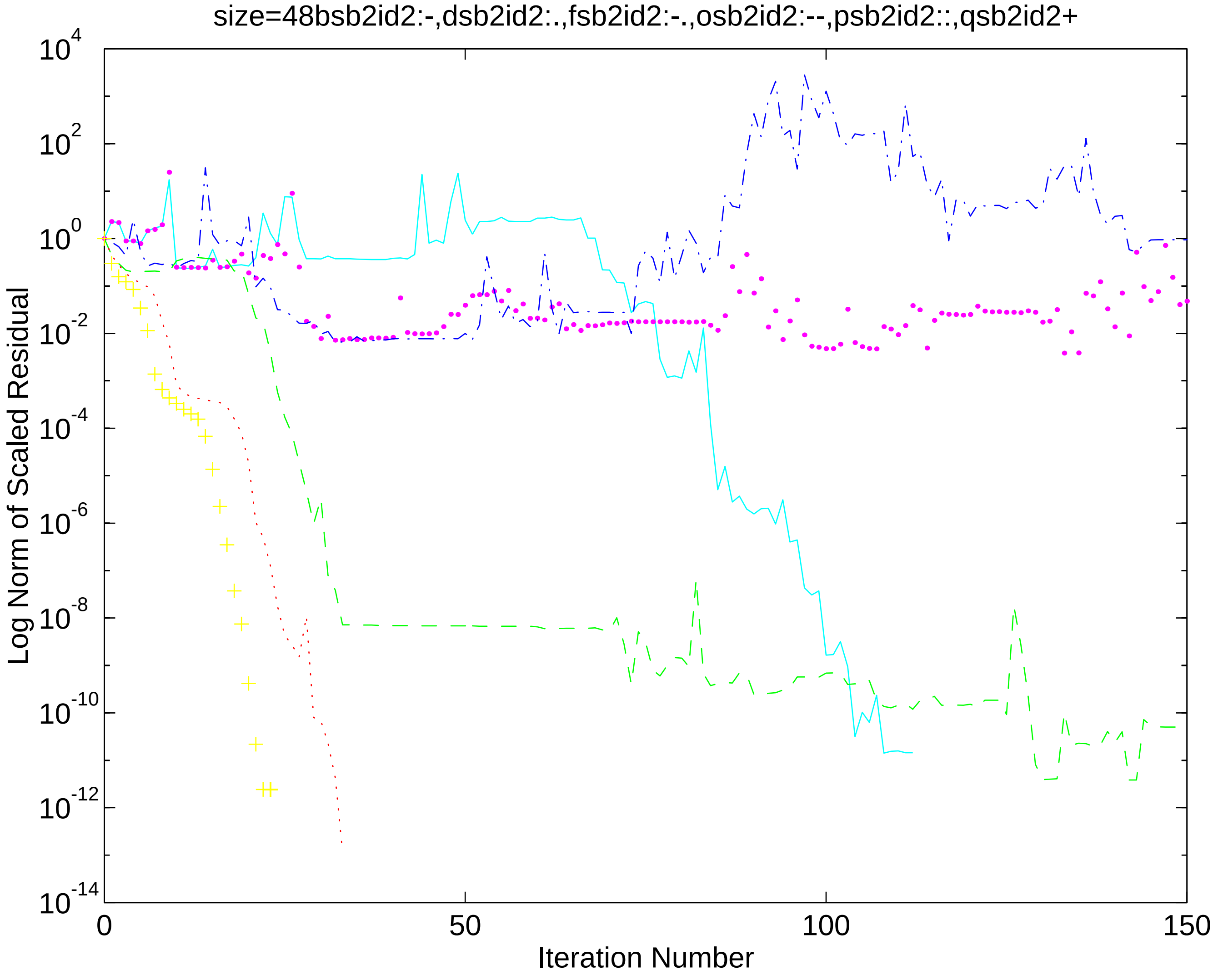

FIG. 29. *BiCGSTAB: Residual Norms versus Iteration Number: Normal Case: Mixed Real and Complex to Nearly Real Eigenvalue Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{I_n}.

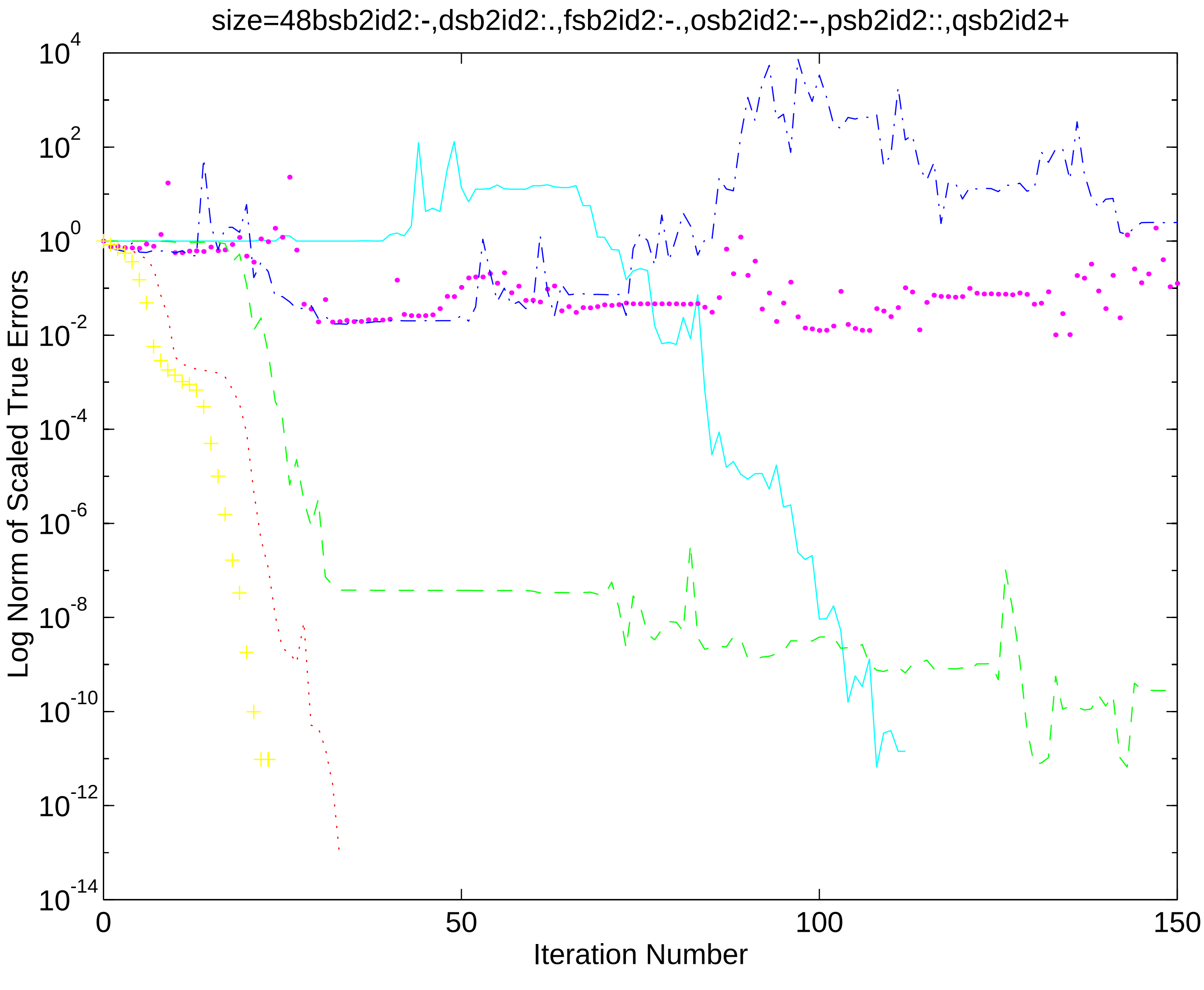

FIG. 30. *BiCGSTAB: Error Norms versus Iteration Number: Normal Case: Mixed Real and Complex to Nearly Real Eigenvalue Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{I_n}.

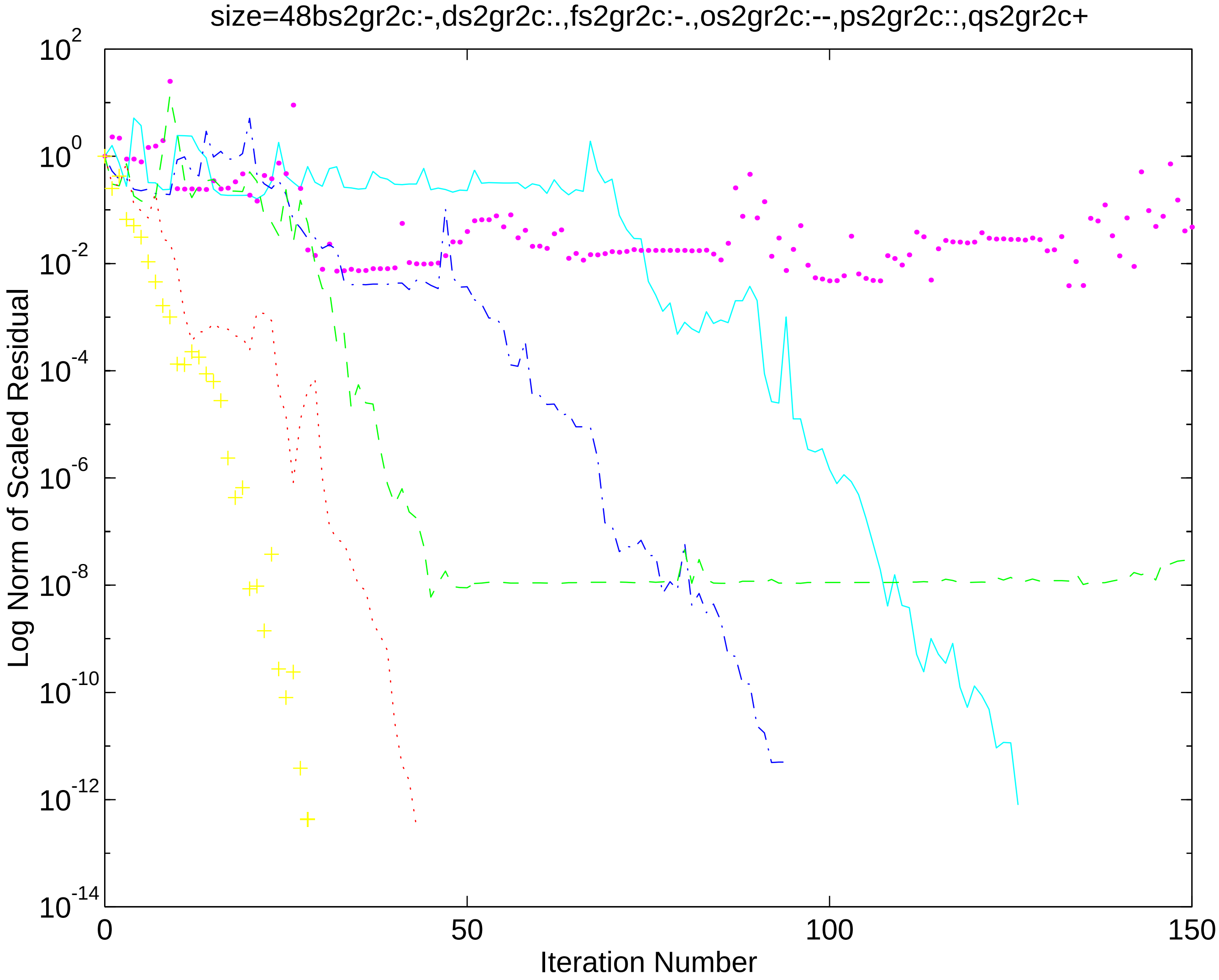

FIG. 31. *BiCGSTAB: Residual Norms versus Iteration Number: Nonnormal Case: Mixed Real and Complex to Nearly Real Eigenvalue Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{gr48}.

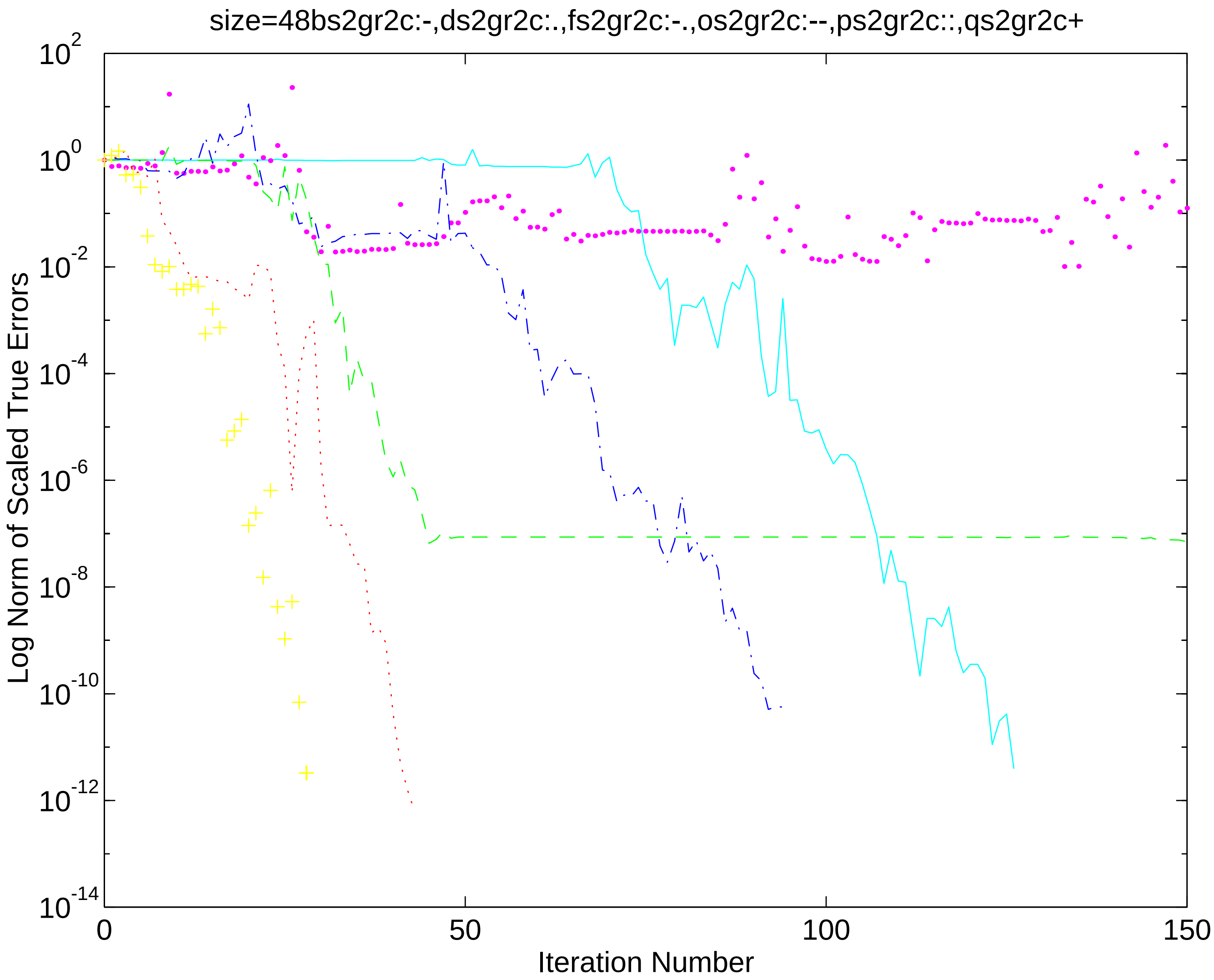

FIG. 32. *BiCGSTAB: Error Norms versus Iteration Number: Nonnormal Case: Mixed Real and Complex to Nearly Real Eigenvalue Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{gr48}.

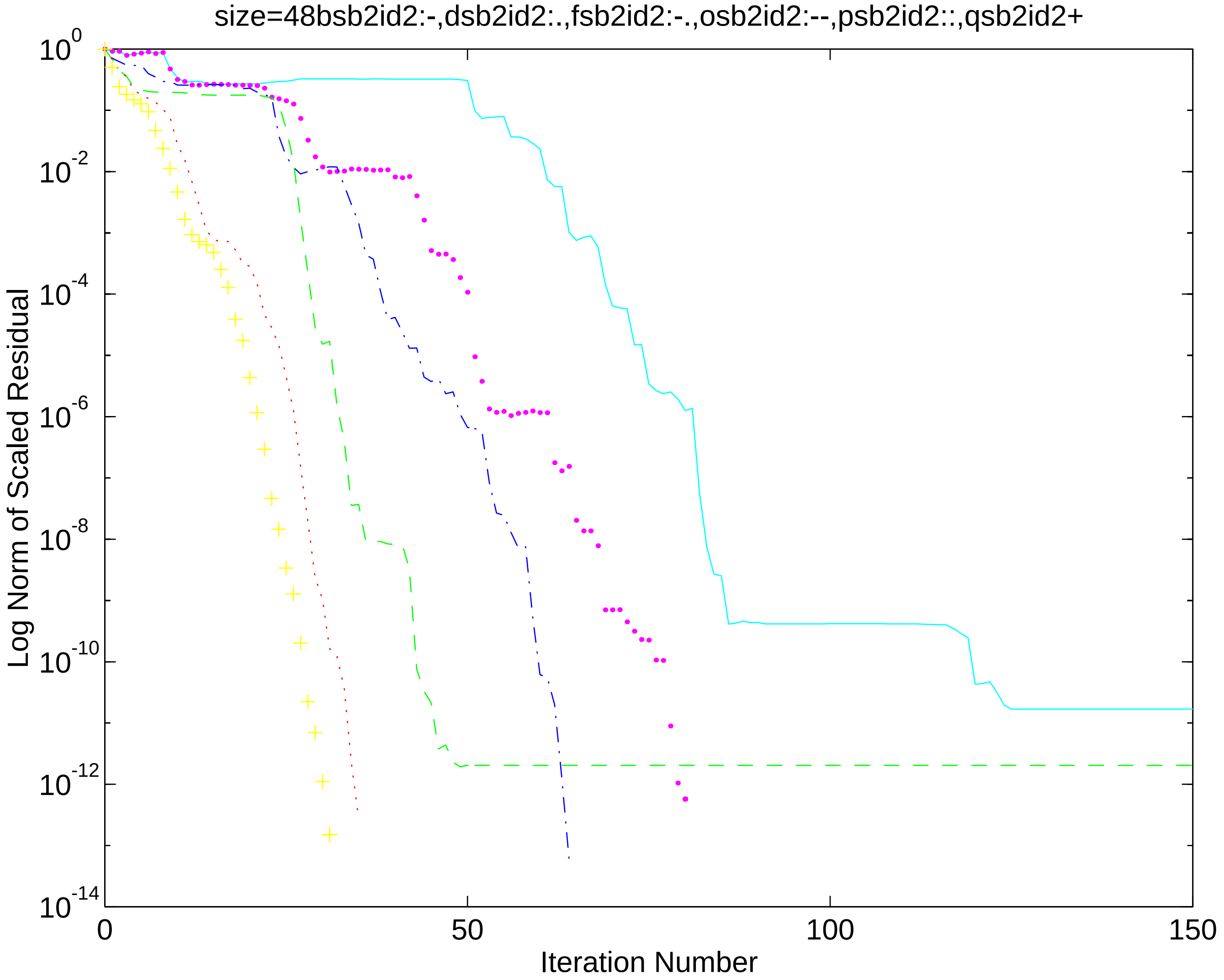

FIG. 33. *QMR: Residual Norms versus Iteration Number: Normal Case: Mixed Real and Complex to Nearly Real Eigenvalue Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{I_n}.

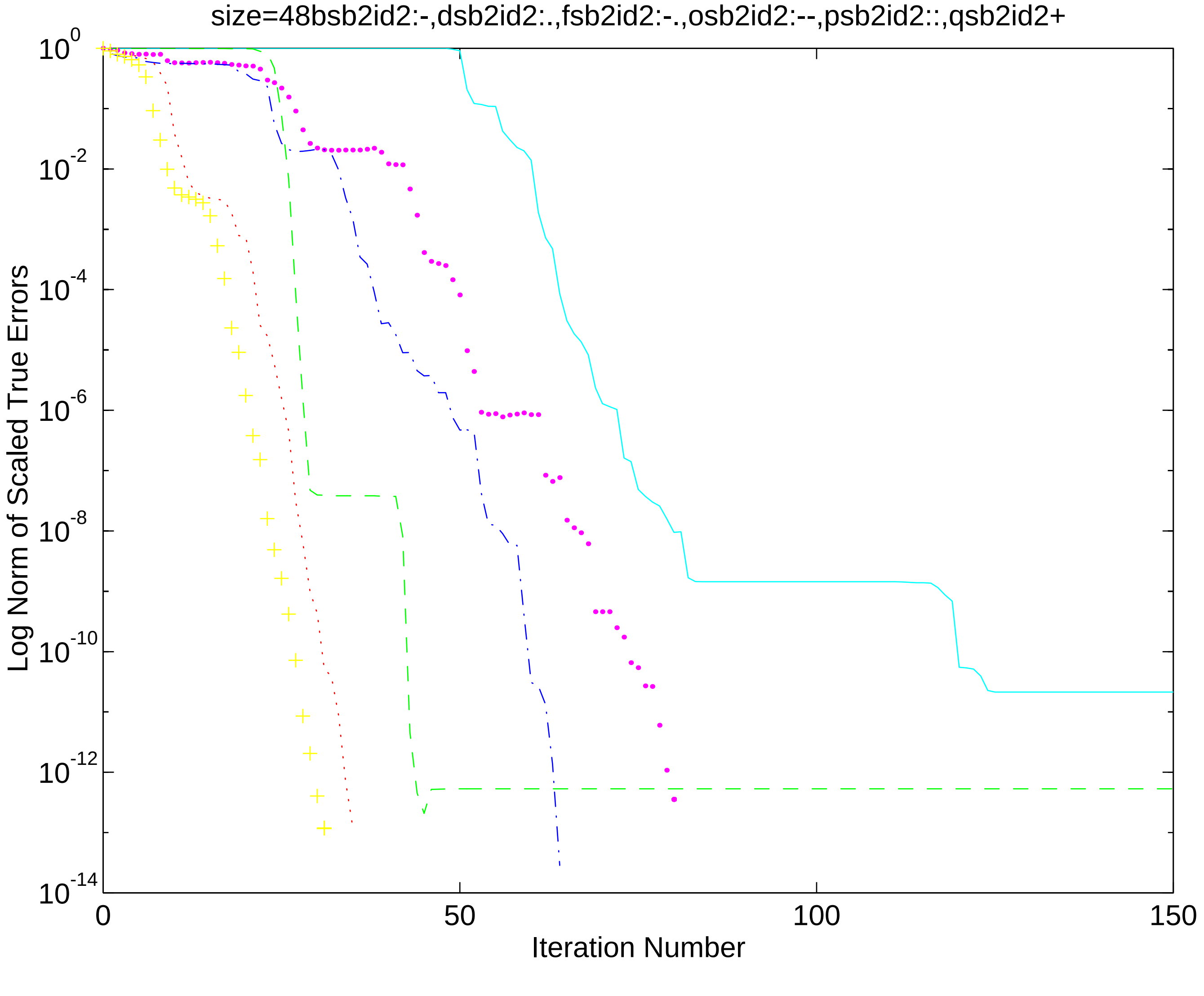

FIG. 34. *QMR: Error Norms versus Iteration Number: Normal Case: Mixed Real and Complex to Nearly Real Eigenvalue Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{I_n}.

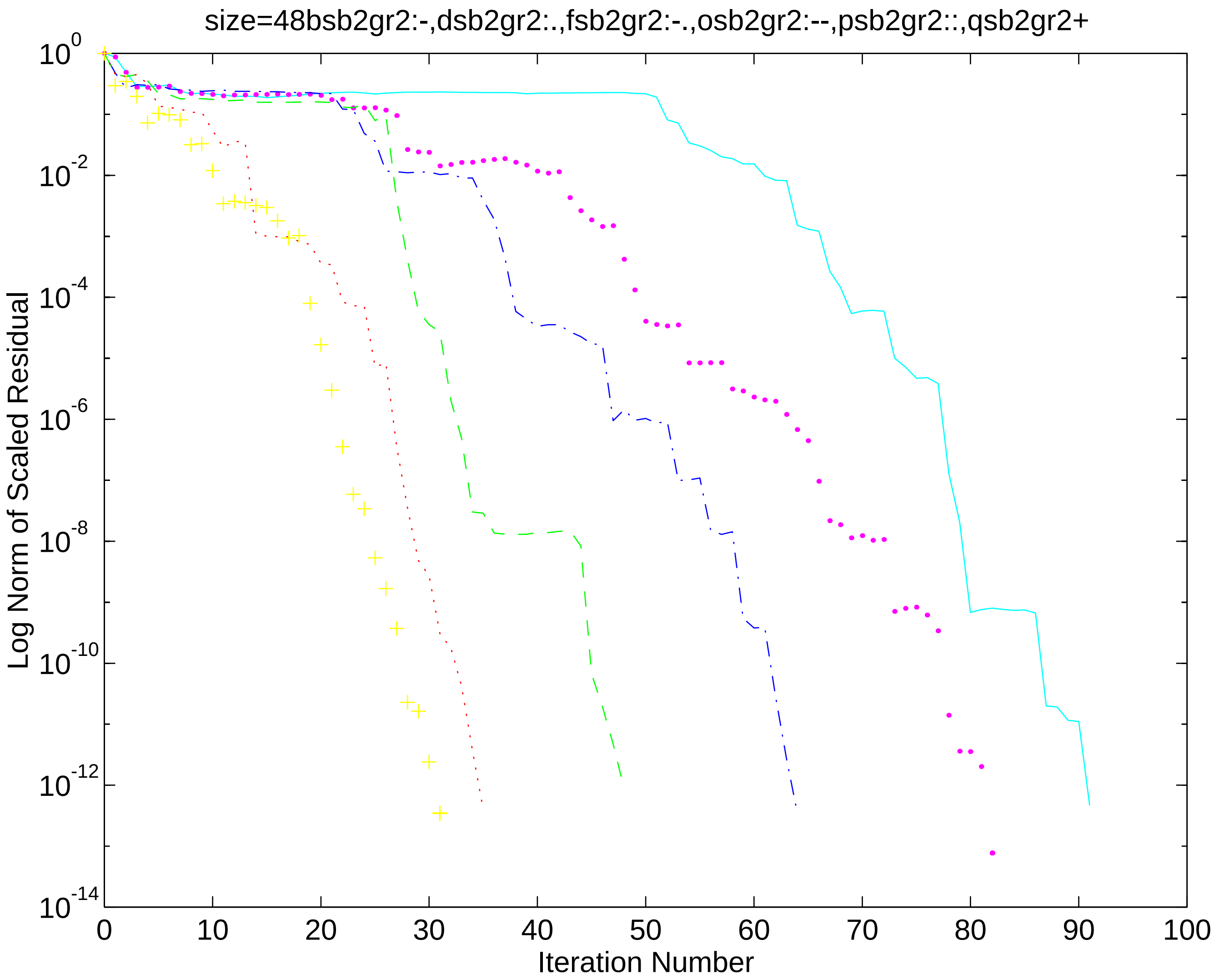

FIG. 35. *QMR: Residual Norms versus Iteration Number: Nonnormal Case: Mixed Real and Complex to Nearly Real Eigenvalue Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{gr48}.

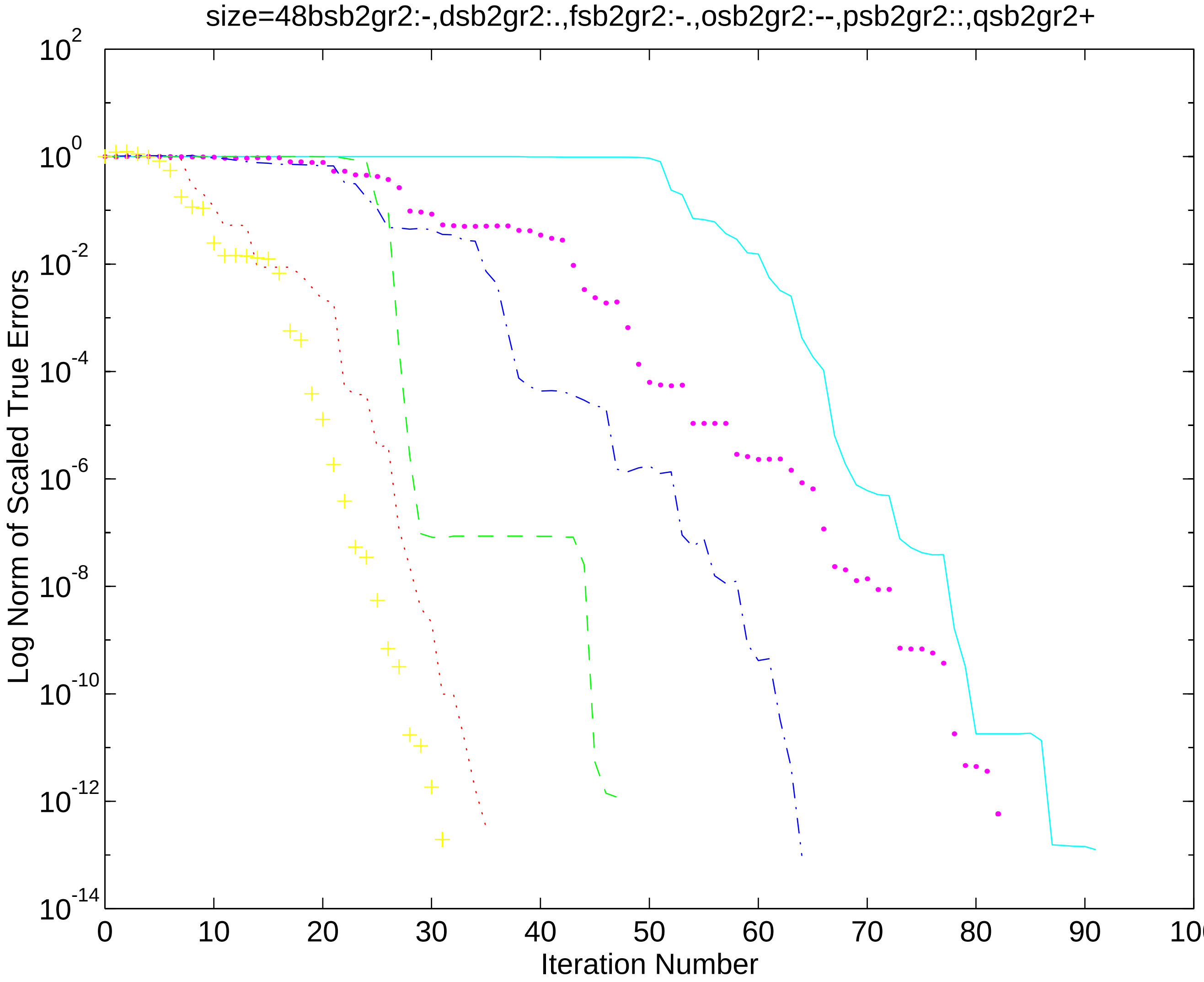

FIG. 36. *QMR: Error Norms versus Iteration Number: Nonnormal Case: Mixed Real and Complex to Nearly Real Eigenvalue Distributions:* Λ_{bsb2}, Λ_{dsb2}, Λ_{fsb2}, Λ_{osb2}, Λ_{psb2}, Λ_{qsb2}. *Singular Value Distribution:* Σ_{gr48}.