SIAM J. MATRIX ANAL. APPL.
Vol. 30, No. 4, pp. 1483–1499

HOW TO MAKE SIMPLER GMRES AND GCR MORE STABLE*

PAVEL JIRÁNEK[†], MIROSLAV ROZLOŽNÍK[‡], AND MARTIN H. GUTKNECHT[§]

Abstract. In this paper we analyze the numerical behavior of several minimum residual methods which are mathematically equivalent to the GMRES method. Two main approaches are compared: one that computes the approximate solution in terms of a Krylov space basis from an upper triangular linear system for the coordinates, and one where the approximate solutions are updated with a simple recursion formula. We show that a different choice of the basis can significantly influence the numerical behavior of the resulting implementation. While Simpler GMRES and ORTHODIR are less stable due to the ill-conditioning of the basis used, the residual basis is well-conditioned as long as we have a reasonable residual norm decrease. These results lead to a new implementation, which is conditionally backward stable, and they explain the experimentally observed fact that the GCR method delivers very accurate approximate solutions when it converges fast enough without stagnation.

Key words. large-scale nonsymmetric linear systems, Krylov subspace methods, minimum residual methods, numerical stability, rounding errors

AMS subject classifications. 65F10, 65G50, 65F35

DOI. 10.1137/070707373

1. Introduction.

In this paper we consider certain methods for solving a system of linear algebraic equations

$$Ax = b, \qquad A \in \mathbb{R}^{N\times N}, \qquad b \in \mathbb{R}^N, \tag{1.1}$$

where A is a large and sparse nonsingular matrix that is, in general, nonsymmetric. For solving such systems, Krylov subspace methods are very popular. They build a sequence of iterates x_n $(n = 0, 1, 2, \ldots)$ such that $x_n \in x_0 + \mathcal{K}_n(A, r_0)$, where $\mathcal{K}_n(A, r_0) \equiv \operatorname{span}\{r_0, Ar_0, \ldots, A^{n-1}r_0\}$ is the nth Krylov subspace generated by the matrix A from the residual $r_0 \equiv b - Ax_0$ that corresponds to the initial guess x_0. Many approaches for defining such approximations x_n have been proposed; see, e.g., the books by Greenbaum [9], Meurant [16], and Saad [22]. In particular, due to their smooth convergence behavior, minimum residual methods satisfying

$$\|r_n\| = \min_{\widetilde{x}\in x_0+\mathcal{K}_n(A,r_0)} \|b - A\widetilde{x}\|, \qquad r_n \equiv b - Ax_n, \tag{1.2}$$

are widely used; see, e.g., the GMRES algorithm of Saad and Schultz [23]. We recall that the minimum residual property (1.2) is equivalent to the orthogonality condition

$$r_n \perp A\mathcal{K}_n(A, r_0),$$

*Received by the editors November 6, 2007; accepted for publication (in revised form) by D. B. Szyld August 18, 2008; published electronically December 3, 2008.
http://www.siam.org/journals/simax/30-4/70737.html

[†]Faculty of Mechatronics and Interdisciplinary Engineering Studies, Technical University of Liberec, Hálkova 6, CZ-461 17 Liberec, Czech Republic (pavel.jiranek@tul.cz). The work of this author was supported by the MSMT CR under the project 1M0554 "Advanced Remedial Technologies."

[‡]Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 2, CZ-182 07 Prague 8, Czech Republic (miro@cs.cas.cz). The work of this author was supported by the Grant Agency of the Czech Academy of Sciences under the projects IAA100300802 and IAA1030405 and by the Institutional Research Plan AV0Z10300504 "Computer Science for the Information Society: Models, Algorithms, Applications."

[§]Seminar for Applied Mathematics, ETH Zurich, CH-8092 Zurich, Switzerland (mhg@math.ethz.ch).

where $\perp$ is the orthogonality relation induced by the Euclidean inner product $\langle \cdot, \cdot \rangle$.

The classical implementation of GMRES [23] makes use of a nested sequence of orthonormal bases of the Krylov subspaces $\mathcal{K}_n(A, r_0)$. These bases are generated by the Arnoldi process [2], and the approximate solution x_n satisfying the minimum residual property (1.2) is constructed from the transformed least squares problem with an upper Hessenberg matrix. This problem is solved via its recursive QR factorization, updated by applying Givens rotations. Once the norm of the residual is small enough, which can be seen without explicitly solving the least squares problem, the triangular system with the computed R-factor is solved, and the approximate solution x_n is computed. In [3, 11, 18] it was shown that this "classical" version of the GMRES method is backward stable provided that the Arnoldi process is implemented using the modified Gram–Schmidt algorithm or Householder reflections.

In this paper we deal with a different approach. Instead of building an orthonormal basis of $\mathcal{K}_n(A, r_0)$, we look for an orthonormal basis $V_n \equiv [v_1, \ldots, v_n]$ of $A\mathcal{K}_n(A, r_0)$. We will also consider a basis $Z_n \equiv [z_1, \ldots, z_n]$ of $\mathcal{K}_n(A, r_0)$ and assume in our analysis that the vectors Z_n have unit lengths, but they need not be orthogonal. The orthonormal basis V_n of $A\mathcal{K}_n(A, r_0)$ is obtained from the QR factorization of the image of Z_n:

$$AZ_n = V_n U_n. \tag{1.3}$$

Since $r_n \in r_0 + A\mathcal{K}_n(A, r_0) = r_0 + \mathcal{R}(V_n)$ and $r_n \perp \mathcal{R}(V_n)$, the residual $r_n = (I - V_n V_n^T) r_0$ is just the orthogonal projection of r_0 onto the orthogonal complement of $\mathcal{R}(V_n)$, which can be computed recursively as

$$r_n = r_{n-1} - \alpha_n v_n, \qquad \alpha_n \equiv \langle r_{n-1}, v_n \rangle \tag{1.4}$$

($\mathcal{R}(V_n)$ denotes the range of the matrix V_n). Let $R_{n+1} \equiv [r_0, \ldots, r_n]$, let $D_n \equiv \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$, and let $L_{n+1,n} \in \mathbb{R}^{(n+1)\times n}$ be the bidiagonal matrix with 1's on the main diagonal and -1's on the first subdiagonal; then the recursion (1.4) can be cast into a matrix relation

$$R_{n+1} L_{n+1,n} = V_n D_n. \tag{1.5}$$

Since the columns of Z_n form a basis of $\mathcal{K}_n(A, r_0)$, we can represent x_n in the form

$$x_n = x_0 + Z_n t_n, \tag{1.6}$$

so that $r_n = r_0 - AZ_n t_n = r_0 - V_n U_n t_n$. Due to $r_n \perp \mathcal{R}(V_n)$, it follows that

$$U_n t_n = V_n^T r_0 = [\alpha_1, \ldots, \alpha_n]^T. \tag{1.7}$$

Hence, once the residual norm is small enough, we can solve this upper triangular system and compute the approximate solution $x_n = x_0 + Z_n t_n$. We call this approach the *generalized simpler approach*. Its pseudocode is given in Figure 1.1. It includes, as a special case, Simpler GMRES, which was proposed by Walker and Zhou [30], where $Z_n = [\frac{r_0}{\|r_0\|}, V_{n-1}]$. We will be also interested in the case of the residual basis $Z_n = \widetilde{R}_n \equiv [\frac{r_0}{\|r_0\|}, \ldots, \frac{r_{n-1}}{\|r_{n-1}\|}]$; we will call this case RB-SGMRES (Residual-based Simpler GMRES). Recently this method was also derived and implemented by Yvan Notay [17].

```
Generalized simpler approach

choose $x_0$, compute $r_0 := b - Ax_0$
for $n = 1, 2, \ldots, m$ (until convergence)
    compute $z_n$ ($z_1 = \frac{r_0}{\|r_0\|}$) and $Az_n$;
    orthonormalize $Az_n$ with respect to $v_1, \ldots, v_{n-1}$
    to obtain $v_n$ so that $AZ_n = V_nU_n$
  $\alpha_n := r_{n-1}^T v_n$
  $r_n := r_{n-1} - \alpha_n v_n$
solve $U_m t_m = [\alpha_1, \ldots, \alpha_m]^T$
$x_m := x_0 + Z_m t_m$
```

FIG. 1.1. *Pseudocode of the generalized simpler approach.*

```
Generalized update approach

choose $x_0$, compute $r_0 := b - Ax_0$
for $n = 1, 2, \ldots, m$ (until convergence)
    compute $z_n$ ($z_1 = \frac{r_0}{\|r_0\|}$) and $Az_n$;
    orthonormalize $Az_n$ with respect to $v_1, \ldots, v_{n-1}$
    to obtain $v_n$ so that $AZ_n = V_nU_n$
    compute $p_n$ from $z_n$ and $p_1, \ldots, p_{n-1}$ so that $Z_n = P_nU_n$
  $\alpha_n := r_{n-1}^T v_n$
  $r_n := r_{n-1} - \alpha_n v_n$
  $x_n := x_{n-1} + \alpha_n p_n$
```

FIG. 1.2. *Pseudocode of the generalized update approach.*

Recursion (1.4) reveals the connection between the generalized simpler approach and yet another minimum residual approach. Let us set $p_k \equiv A^{-1}v_k$ $(k = 1, \ldots, n)$ and $P_n \equiv [p_1, \ldots, p_n]$. Then, left-multiplying (1.4) by A^{-1} yields

$$x_n = x_{n-1} + \alpha_n p_n \tag{1.8}$$

or, in matrix form, $X_{n+1}L_{n+1,n} = -P_nD_n$ with $X_{n+1} \equiv [x_0, \ldots, x_n]$. So, instead of computing the coordinates t_n of $x_n - x_0$ with respect to the basis Z_n, we can directly update x_n from x_{n-1}. However, this requires that we construct the direction vectors P_n forming an A^TA-orthogonal basis of $\mathcal{K}_n(A, r_0)$. Since U_n is known from (1.3), the recursion for p_n can be extracted from the formula

$$Z_n = P_nU_n. \tag{1.9}$$

Note that two recursions (1.3) and (1.9) can be run in the same loop, and we have to store all the direction vectors in P_n and all the orthonormal basis vectors in V_n. We will use the terminology *generalized update approach* for this case. Its pseudocode is given in Figure 1.2. The case $Z_n \equiv [\frac{r_0}{\|r_0\|}, V_{n-1}]$ of this method was proposed in [20] under the name A^TA-variant of GMRES, and up to the normalization of the vectors V_n in (1.3) it is equivalent to the ORTHODIR algorithm due to Young and Jea [33, 7]. Likewise, the case $Z_n = [r_0, \ldots, r_{n-1}]$ corresponds to the GCR (or full ORTHOMIN) method of Elman, Eisenstat, and Schultz [6, 5] (the orthogonal vectors v_n are unnormalized in the original implementation), and it is identical to

Table 1.1

Computational costs (without the cost of $m+1$ matrix-vector products) and storage requirements (without the storage of A) of the generalized simpler and update approaches after m iteration steps.

	Computational costs	Storage requirements
Generalized simpler approach	$(2N + \frac{1}{2})m^2 + (9N - \frac{1}{2})m + 4N$	$(2N + \frac{3}{2})m + \frac{1}{2}m^2 + 2N + 1$
Generalized update approach	$(3N - \frac{1}{2})m^2 + (9N - \frac{1}{2})m + 4N$	$(2N + 1)m + 2N + 2$

the GMRESR method [28] of van der Vorst and Vuik (with the choice $u_n^{(0)} = r_n$). Without normalization it was also treated in [33]. As we have already mentioned, here we will analyze the choice $Z_n = \widetilde{R}_n$. The importance of normalizing Z_n before the orthogonalization in (1.3) will be seen later.

In Table 1.1 we summarize the computational costs and storage requirements of performing m iteration steps in the generalized simpler approach and the generalized update approach, where we have excluded the storage for A and the cost of $m + 1$ matrix-vector products. In both approaches we have to store two sets of vectors—the bases V_m and Z_m (the generalized simpler approach) or V_m and P_m (the generalized update approach)—making these schemes comparable to FGMRES [21], the (flexible) preconditioned variant of the standard GMRES method [23]. This remains true also in the case of preconditioned versions of our algorithms, but we do not treat these explicitly here. In contrast to the generalized simpler approach, we do not need to store the triangular $m \times m$ matrix of orthogonalization coefficients U_m in the generalized update approach, but we have to compute the additional set of vectors P_m. Some savings are possible in special cases, as in Simpler GMRES with the particular choice of the basis $Z_m = [\frac{r_0}{\|r_0\|}, V_{m-1}]$, where the last $m - 1$ columns of Z_m need not to be stored and normalized again. Simpler GMRES is in terms of work and storage competitive to the GMRES method, which in addition was shown to be backward stable and in this context should clearly be the method of choice when preconditioning is not considered.

The paper is organized as follows. In section 2 we analyze first the maximum attainable accuracy of the generalized simpler approach based on (1.6) and (1.7). Then we turn to the generalized update approach based on (1.9) and (1.8). To keep the text readable, we assume rounding errors only in selected, most relevant parts of the computation. The bounds presented in Theorems 2.1 and 2.3 show that the conditioning of the matrix Z_n plays an important role in the numerical stability of these schemes. Both theorems give bounds on the maximum attainable accuracy measured by the normwise backward error. We also formulate analogous statements for the residual norm in terms of the condition number of the matrix U_n. While for the generalized simpler approach these bounds do not depend on the conditioning of A, the bound for the generalized update approach is proportional to $\kappa(A)$ (as we will show in our constructed numerical example, the bound is attained). However, the additional factor of $\kappa(A)$ in the generalized update approach is usually an overestimate; in practice, both approaches behave almost equally well for the same choice of basis. This is especially true for the relative errors of the computed approximate solutions, where we have essentially the same upper bound. The situation is completely analogous to results for the MINRES method [19] given by Sleijpen, van der Vorst, and Modersitzki in [25].

In section 3 we derive particular results for two choices of the basis Z_n—first for $Z_n = [\frac{r_0}{\|r_0\|}, V_{n-1}]$, leading to Simpler GMRES by Walker and Zhou [30] and to ORTHODIR, and then for $Z_n = \widetilde{R}_n$, which leads to RB-SGMRES and to a variant of GCR, respectively. It turns out that the two choices lead to a truly different behavior in the condition number of U_n, which governs the stability of the considered schemes. Since all these methods converge in a finite number of iterations, we fix the iteration index n such that $r_0 \notin A\mathcal{K}_{n-1}(A, r_0)$; that is, the exact solution has not yet been reached. Based on this we give conditions on the linear independence of the basis Z_n. It is known that the residuals are linearly dependent (or even identical) when the GMRES method stagnates (a breakdown occurs in GCR as well as in RB-SGMRES), while this does not happen for $[\frac{r_0}{\|r_0\|}, V_{n-1}]$ (Simpler GMRES and ORTHODIR are breakdown-free). On the other hand, we show that while the choice $Z_n = [\frac{r_0}{\|r_0\|}, V_{n-1}]$ leads to inherently unstable or numerically less stable schemes, the second selection $Z_n = \widetilde{R}_n$ gives rise to conditionally stable implementations provided that we have some reasonable residual decrease. In particular, we show that the RB-SGMRES implementation is conditionally backward stable. Our theoretical results are illustrated by selected numerical experiments. In section 4 we draw conclusions and give directions for future work.

Throughout the paper, we denote by $\|\cdot\|$ the Euclidean vector norm and the induced matrix norm and by $\|\cdot\|_F$ the Frobenius norm. Moreover, for $B \in \mathbb{R}^{N\times n}$ $(N \geq n)$ of rank n, $\sigma_1(B) \geq \sigma_n(B) > 0$ are the extremal singular values of B, and $\kappa(B) = \sigma_1(B)/\sigma_n(B)$ is the spectral condition number. By I we denote the unit matrix of a suitable dimension and by e_k $(k = 1, 2, \dots)$ its kth column, and we let $e \equiv [1, \dots, 1]^T$. We assume the standard model of finite precision arithmetic with the unit roundoff u (see Higham [13] for details). In our bounds, instead of distinguishing between several constants (which are in fact low-degree polynomials in N and n that can differ from place to place), we use the generic name c for constants.

2. Maximum attainable accuracy of the generalized simpler and update approaches.

In this section we analyze the final accuracy level of the generalized simpler and update approaches formulated in the previous section. In order to make our analysis readable, we assume that only the computations performed in (1.3), (1.7), and (1.9) are affected by rounding errors.

Different orthogonalization techniques for computing the columns of V_n can be applied in the QR factorization (1.3). Here we focus on such implementations where the computed R-factor U_n has been obtained in a backward stable way; i.e., there exists an orthonormal matrix $\hat{V}_n$ so that $\hat{V}_n$ and V_n satisfy

$$AZ_n = \hat{V}_n U_n + E_n, \qquad \|E_n\| \leq cu\|A\|\|Z_n\|, \tag{2.1}$$

$$AZ_n = V_n U_n + F_n, \qquad \|F_n\| \leq cu\|A\|\|Z_n\|. \tag{2.2}$$

This is certainly true for the implementation based on Householder reflections [32], the modified Gram–Schmidt process [18], or the Gram–Schmidt process with full reorthogonalization [3]. For details we refer the reader to [13, 8]. From [31, 13] we have for the computed solution $\hat{t}_n$ of (1.7) that

$$(U_n + \Delta U_n)\hat{t}_n = D_n e, \qquad |\Delta U_n| \leq cu|U_n|, \tag{2.3}$$

where the absolute value and inequalities are understood componentwise. The approximation $\hat{x}_n$ to x is then computed as

$$\hat{x}_n = x_0 + Z_n \hat{t}_n. \tag{2.4}$$

The crucial quantity for the analysis of the maximum attainable accuracy is the gap between the true residual $b - A\hat{x}_n$ of the computed approximation and the updated residual r_n obtained from the update formula (1.4) describing the projection of the previous residual; see [9, 12]. In fact, once the updated residual becomes negligible compared to the true one (and in all algorithms considered here it ultimately will), the gap will be equal to the true residual divided by $\|A\|\|\hat{x}_n\|$, which therefore can be thought of as the normwise backward error of the ultimate approximate solution $\hat{x}_n$ (after suitable normalization). Here is our basic result on this gap for the generalized simpler approach.

THEOREM 2.1. *In the generalized simpler approach, if* $cu\kappa(A)\kappa(Z_n) < 1$*, the gap between the true residual* $b - A\hat{x}_n$ *and the updated residual* r_n *satisfies*

$$\frac{\|b - A\hat{x}_n - r_n\|}{\|A\|\|\hat{x}_n\|} \leq cu\kappa(Z_n)\left(1 + \frac{\|x_0\|}{\|\hat{x}_n\|}\right).$$

Proof. From (2.4), (2.2), and (2.3) we have $b - A\hat{x}_n = r_0 - AZ_n\hat{t_n} = r_0 - (V_nU_n + F_n)(U_n + \Delta U_n)^{-1}D_ne$, and (1.4) gives $r_n = r_0 - V_nD_ne$. It is clear from (2.1) and (2.3) that the assumption $cu\kappa(A)\kappa(Z_n) < 1$ implies the invertibility of the perturbed matrix $U_n + \Delta U_n$. Using the identity $I - U_n(U_n + \Delta U_n)^{-1} = \Delta U_n(U_n + \Delta U_n)^{-1}$ and the relation $Z_n(U_n + \Delta U_n)^{-1}D_ne = Z_n\hat{t}_n = \hat{x}_n - x_0$ following from (2.4) and (2.3), we can express the gap between $b - A\hat{x}_n$ and r_n as

$$\begin{aligned}
b - A\hat{x}_n - r_n &= (V_n - (V_nU_n + F_n)(U_n + \Delta U_n)^{-1})D_ne \\
&= (V_n(I - U_n(U_n + \Delta U_n)^{-1}) - F_n(U_n + \Delta U_n)^{-1})D_ne \\
&= (V_n\Delta U_n - F_n)(U_n + \Delta U_n)^{-1}D_ne \\
&= (V_n\Delta U_n - F_n)Z_n^{\dagger}Z_n(U_n + \Delta U_n)^{-1}D_ne \\
&= (V_n\Delta U_n - F_n)Z_n^{\dagger}(\hat{x}_n - x_0).
\end{aligned}$$

Taking the norm, considering (2.1), and noting that the terms in $V_n\Delta U_n$ and F_n can be subsumed into the generic constant c, we get $\|V_n\Delta U_n - F_n\| \leq cu\|A\|\|Z_n\|$ and

$$\|b - A\hat{x}_n - r_n\| \leq cu\|A\|\kappa(Z_n)\|\hat{x}_n - x_0\|.$$

Using the triangle inequality and division by $\|A\|\|\hat{x}_n\|$ concludes the proof. □

In the previous theorem we have expressed the residual gap using the difference between the actual and initial approximations $\hat{x}_n$ and x_0, respectively. However, its norm is strongly influenced by the conditioning of the upper triangular matrix U_n. As shown in section 3, the matrix U_n can be ill-conditioned for the particular case $Z_n = [\frac{r_0}{\|r_0\|}, V_{n-1}]$, thus leading to an inherently unstable scheme, whereas (under some assumptions) the scheme with $Z_n = \widetilde{R}_n$ gives rise to a well-conditioned triangular matrix U_n. In the following corollary we give a bound for the residual gap in terms of the minimal singular values of the matrices Z_k and norms of the updated residuals r_{k-1}, $k = 1, \ldots, n$.

COROLLARY 2.2. *In the generalized simpler approach, if* $cu\kappa(A)\kappa(Z_n) < 1$*, the gap between the true residual* $b - A\hat{x}_n$ *and the updated residual* r_n *satisfies*

$$\|b - A\hat{x}_n - r_n\| \leq \frac{cu\kappa(A)}{1 - cu\kappa(A)\kappa(Z_n)} \sum_{k=1}^{n} \frac{\|r_{k-1}\|}{\sigma_k(Z_k)}.$$

Proof. The gap between the true residual $b - A\hat{x}_n$ and the updated residual r_n can be expressed as $b - A\hat{x}_n - r_n = (V_n \Delta U_n - F_n)(U_n + \Delta U_n)^{-1} D_n e$. Since $e_k^T D_n e_k = \alpha_k$ and $|\alpha_k| = \sqrt{\|r_{k-1}\|^2 - \|r_k\|^2} \leq \sqrt{2}\|r_{k-1}\|$, the norm of the term $(U_n + \Delta U_n)^{-1} D_n e$ can be estimated as follows:

$$\begin{aligned} \|(U_n + \Delta U_n)^{-1} D_n e\| &\leq \sum_{k=1}^{n} \|(U_n + \Delta U_n)^{-1} D_n e_k\| \\ &\leq \sqrt{2} \sum_{k=1}^{n} \frac{\|r_{k-1}\|}{\sigma_k([U_n + \Delta U_n]_{1:k,1:k})}, \end{aligned} \tag{2.5}$$

where $[U_n + \Delta U_n]_{1:k,1:k}$ denotes the principal $k \times k$ submatrix of $U_n + \Delta U_n$. Owing to (2.4), we can estimate the perturbation of $[U_n]_{1:k,1:k} = U_k$ as $\|[\Delta U_n]_{1:k,1:k}\| \leq cu\|U_k\|$. Perturbation theory of singular values (see, e.g., [14]) shows that

$$\sigma_k([U_n + \Delta U_n]_{1:k,1:k}) \geq \sigma_k(U_k) - cu\|U_k\| \geq \sigma_k(AZ_k) - cu\|A\|\|Z_k\| \tag{2.6a}$$

$$\geq \sigma_N(A)\sigma_k(Z_k) - cu\|A\|\|Z_k\|, \tag{2.6b}$$

which together with (2.5) concludes the proof. □

The estimates (2.5) and (2.6a) given in the previous proof that involve the minimum singular values of U_k $(k = 1, \ldots, n)$ are quite sharp. However, the estimate (2.6b) relating the minimum singular values of U_k to those of Z_k can be a large underestimate, as also observed in our numerical experiments in section 3.

Next we analyze the maximum attainable accuracy of the generalized update approach. We assume that in finite precision arithmetic the computed direction vectors satisfy

$$Z_n = P_n U_n + G_n, \qquad \|G_n\| \leq cu\|P_n\|\|U_n\|. \tag{2.7}$$

This follows from the standard rounding error analysis of the recursion for vectors P_n. Note that the norm of the matrix G_n cannot be bounded by $cu\|A\|\|Z_n\|$ as it can in the case of the QR factorization (2.2). We update then the approximate solution $\hat{x}_n$ according to (1.8):

$$\hat{x}_n = \hat{x}_{n-1} + \alpha_n p_n. \tag{2.8}$$

THEOREM 2.3. *In the generalized update approach, if* $cu\kappa(A)\kappa(Z_n) < 1$, *the gap between the true residual* $b - A\hat{x}_n$ *and the updated residual* r_n *satisfies*

$$\frac{\|b - A\hat{x}_n - r_n\|}{\|A\|\|\hat{x}_n\|} \leq \frac{cu\kappa(A)\kappa(Z_n)}{1 - cu\kappa(A)\kappa(Z_n)} \left(1 + \frac{\|x_0\|}{\|\hat{x}_n\|}\right).$$

Proof. From (2.8), (1.4), (2.2), and (2.7), $\hat{x}_n = x_0 + P_n D_n e = x_0 + (Z_n - G_n)U_n^{-1} D_n e$ and $r_n = r_0 - V_n D_n e = r_0 - (AZ_n - F_n)U_n^{-1} D_n e$, we have that

$$b - A\hat{x}_n - r_n = (AG_n - F_n)U_n^{-1} D_n e, \tag{2.9}$$

and from (2.7) and (2.1) we get $P_n = A^{-1}\hat{V}_n + A^{-1}E_n U_n^{-1} - G_n U_n^{-1}$. The norm of the matrix G_n in (2.7) can hence be bounded by

$$\|G_n\| \leq cu\kappa(A)\|Z_n\|. \tag{2.10}$$

Owing to (2.8), we have the identity $U_n^{-1}D_ne = U_n^{-1}P_n^{\dagger}P_nD_ne = (P_nU_n)^{\dagger}(\hat{x}_n - x_0)$, where $\|(P_nU_n)^{\dagger}\| \leq [1 - cu\kappa(A)\kappa(Z_n)]^{-1}\|Z_n^{\dagger}\|$, as follows from (2.7). Thus we obtain

$$\|U_n^{-1}D_ne\| \leq \frac{\|Z_n^{\dagger}\|}{1 - cu\kappa(A)\kappa(Z_n)}\|\hat{x}_n - x_0\|, \tag{2.11}$$

which together with (2.9), (2.10), and (2.2) leads to

$$\|b - A\hat{x}_n - r_n\| \leq \frac{cu\|A\|\kappa(A)\kappa(Z_n)}{1 - cu\kappa(A)\kappa(Z_n)}\|\hat{x}_n - x_0\|.$$

The proof is concluded using the triangle inequality and dividing by $\|A\|\|\hat{x}_n\|$. □

In the following we formulate an analogous corollary for the residual gap as in the case of the generalized simpler approach.

Corollary 2.4. *In the generalized update approach, if $cu\kappa(A)\kappa(Z_n) < 1$, the gap between the true residual $b - A\hat{x}_n$ and the updated residual r_n satisfies*

$$\|b - A\hat{x}_n - r_n\| \leq \frac{cu\kappa^2(A)}{1 - cu\kappa(A)\kappa(Z_n)}\sum_{k=1}^{n}\frac{\|r_{k-1}\|}{\sigma_k(Z_k)}.$$

Proof. Considering (2.2), (2.7), and (2.10) the norm of the term $AG_n - F_n$ in (2.9) can be bounded as $\|AG_n - F_n\| \leq cu\|A\|\kappa(A)$, while the term $U_n^{-1}D_ne$ can be treated as in Corollary 2.2. □

The bound on the ultimate backward error given in Theorem 2.3 is worse than the one in Theorem 2.1. We see that for the generalized simpler approach the normwise backward error is of the order of the roundoff unit, whereas for the generalized update approach we have an upper bound proportional to the condition number of A. Similarly, the bounds on the ultimate relative residual norms given in Corollaries 2.2 and 2.4 indicate that the relative residuals in the generalized simpler approach will reach the level which is approximately equal to $u\kappa(A)$, while in the generalized update approach this level becomes $u\kappa^2(A)$.

In the previous text we have given bounds in terms of the true residual $b - A\hat{x}_n$ and the updated residual r_n. It should be noted that the true residual is not available in practical computations, but for verification or for other purposes it can be estimated by the explicit evaluation of $\mathrm{fl}(b - A\hat{x}_n)$. It is clear from $\|\mathrm{fl}(b - A\hat{x}_n) - (b - A\hat{x}_n)\| \leq cu(\|b\| + \|A\|\|\hat{x}_n\|) \leq cu\|A\|(\|x\| + \|\hat{x}_n\|)$ that the error in the evaluation of the true residual (if needed) is significantly smaller than other quantities involved in our analysis.

In Theorems 2.1 and 2.3 we have estimated the attainable level of the normwise backward error of both generalized simpler and update approaches. The resulting bound is in general worse for the generalized update approach. However, as shown below, it appears that *the generalized update approach leads to an approximate solution whose forward error is essentially on the same accuracy level as the generalized simpler approach.* A similar phenomenon was also observed by Sleijpen, van der Vorst, and Modersitzki [25] in the symmetric case for two different implementations (called GMRES and MINRES in their paper).

Corollary 2.5. *If $cu\kappa(A)\kappa(Z_n) < 1$, the gap between the error $x - \hat{x}_n$ and the vector $A^{-1}r_n$ in both the generalized simpler and update approaches satisfies*

$$\frac{\|(x - \hat{x}_n) - A^{-1}r_n\|}{\|x\|} \leq \frac{cu\kappa(A)\kappa(Z_n)}{1 - cu\kappa(A)\kappa(Z_n)}\frac{\|\hat{x}_n\| + \|x_0\|}{\|x\|}.$$

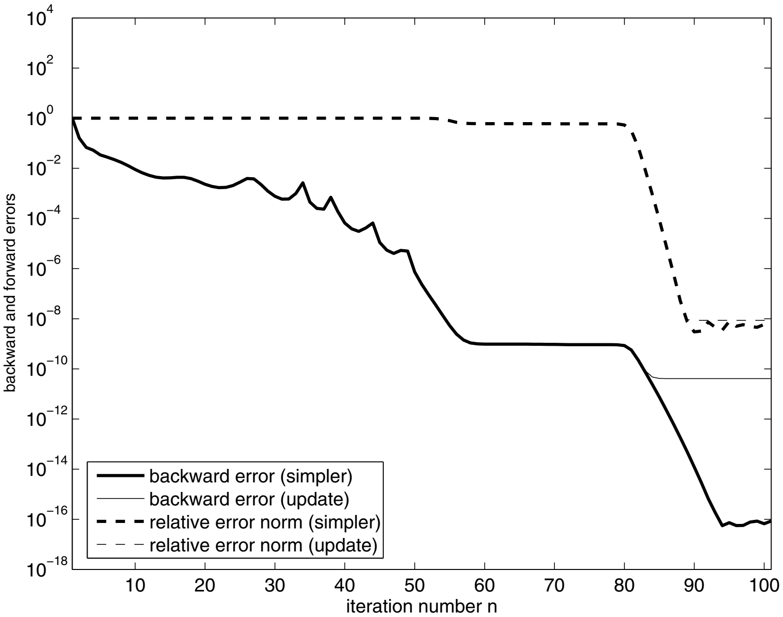

FIG. 2.1. *The test problem solved by the generalized simpler and update approaches with the almost orthonormal basis Z_n satisfying $\kappa(Z_n) \approx 1$.*

Proof. For the generalized simpler approach, the result follows directly from Theorem 2.1. For the generalized update approach, using (2.9) we have

$$(x - \hat{x}_n) - A^{-1}r_n = (-A^{-1}F_n + G_n)U_n^{-1}D_n e,$$

and the statement follows from (2.2), (2.10), and (2.11). □

Theorems 2.1 and 2.3 indicate that *as soon as the backward error of the approximate solution in the generalized simpler approach gets below $cu\kappa(A)\kappa(Z_n)$, the difference between the backward errors in the generalized simpler and update approaches may become visible and can be expected to be up to the order of $\kappa(A)$.* Based on our experience it is difficult to find an example where this difference is significant. Similarly to Sleijpen, van der Vorst, and Modersitzki [25], we use here a model example, where $A = G_1 D G_2^T \in \mathbb{R}^{100\times 100}$ with $D = \text{diag}(10^{-8}, 2\cdot 10^{-8}, 3, 4, \ldots, 100)$ and with G_1 and G_2 being Givens rotations over the angle of $\frac{\pi}{4}$ in the $(1, 10)$-plane and the $(1, 100)$-plane, respectively; finally, $b = e$. The numerical experiments were performed in MATLAB using double precision arithmetic ($u \approx 10^{-16}$) and $x_0 = 0$. In Figure 2.1 we have plotted the normwise backward errors $\|b - A\hat{x}_n\|/(\|A\|\|\hat{x}_n\|)$ (thin and thick solid lines), and the relative 2-norms of the errors $\|x - \hat{x}_n\|/\|x\|$ (thin and thick dash-dotted lines). In all our experiments the basis V_n in (1.3) is computed with the modified Gram–Schmidt orthogonalization process, where the upper triangular factor U_n is obtained in a backward stable way satisfying (2.1). In order to ensure that the difference is not affected by a possibly high condition number of Z_n, we use the implementation where the basis Z_n is computed with the modified Gram–Schmidt Arnoldi process so that $\kappa(Z_n) \approx 1$. We see that the actual backward errors are close to each other until they stagnate: for the generalized update approach this happens approximately at a level approaching $u\kappa(A)$, while for the generalized simpler approach we have stagnation on the roundoff unit level u. Similar observations could be made for the relative true residual norms (for better readability they are not shown in Figure 2.1); in the case of the generalized simpler approach the final level of the relative 2-norm of the true residual is on the level of $u\kappa(A)$, while for the generalized update approach this level is approximately one factor of $\kappa(A)$ higher. In contrast, the 2-norms of the errors stagnate on the $u\kappa(A)$ level in both approaches considered.

3. Choice of basis and stability. In this section we discuss the two main particular choices for the matrix Z_n leading to different algorithms for the generalized simpler and update approaches. For the sake of simplicity, we assume exact arithmetic here. The conditioning of Z_n plays an important role in our analysis. The effect of scaling the columns on the condition number has been analyzed by van der Sluis in [27], who showed that the normalization of columns is a nearly optimal strategy producing the condition number within the factor $\sqrt{n}$ of the minimum 2-norm condition number achievable by column scaling.

First, we choose $Z_n = [\frac{r_0}{\|r_0\|}, V_{n-1}]$, which leads to the Simpler GMRES method of Walker and Zhou [30] and to ORTHODIR by Young and Jea [33]. Hence, we choose $\{\frac{r_0}{\|r_0\|}, v_1, \ldots, v_{n-1}\}$ as a basis of $\mathcal{K}_n(A, r_0)$. To be sure that such a choice is adequate, we state the following simple lemma.

LEMMA 3.1. *Let $v_1, \ldots, v_{n-1}$ be an orthonormal basis of $A\mathcal{K}_{n-1}(A, r_0)$, $r_0 \notin A\mathcal{K}_{n-1}(A, r_0)$. Then the vectors $\frac{r_0}{\|r_0\|}, v_1, \ldots, v_{n-1}$ form a basis of $\mathcal{K}_n(A, r_0)$.*

Proof. The result follows easily from the assumption $r_0 \notin A\mathcal{K}_{n-1}(A, r_0)$. □

Note that if $r_0 \in A\mathcal{K}_n(A, r_0)$, then the condition (1.2) yields $x_n = A^{-1}b$, $r_n = 0$, and any implementation of a minimum residual method will terminate. Lemma 3.1 ensures that it makes sense to build an orthonormal basis V_n of $A\mathcal{K}_n(A, r_0)$ by the successive orthogonalization of the columns of the matrix $A[\frac{r_0}{\|r_0\|}, V_{n-1}]$ via (1.3). It reflects the fact that, for any initial residual r_0, both Simpler GMRES and ORTHODIR converge (in exact arithmetic) to the exact solution; see [33]. However, as observed by Liesen, Rozložník, and Strakoš [15], this choice of the basis is not very suitable from the stability point of view. This shortcoming is reflected by the unbounded growth of the condition number of $[\frac{r_0}{\|r_0\|}, V_{n-1}]$ discussed next. The upper bound we recall here was also derived in [30].

THEOREM 3.2. *Let $r_0 \notin A\mathcal{K}_{n-1}(A, r_0)$. Then the condition number of $[\frac{r_0}{\|r_0\|}, V_{n-1}]$ satisfies*

$$\frac{\|r_0\|}{\|r_{n-1}\|} \leq \kappa([\tfrac{r_0}{\|r_0\|}, V_{n-1}]) \leq 2\frac{\|r_0\|}{\|r_{n-1}\|}.$$

Proof. Since $r_{n-1} = (I - V_{n-1}V_{n-1}^T)r_0$, it is easy to see that r_{n-1} is the residual of the least squares problem $V_{n-1}y \approx r_0$. The statement therefore follows from [15, Theorem 3.2]. □

The conditioning of $[\frac{r_0}{\|r_0\|}, V_{n-1}]$ is thus related to the convergence of the method; in particular, it is inversely proportional to the actual relative norm of the residual. Small residuals lead to the ill-conditioning of the matrices $A[\frac{r_0}{\|r_0\|}, V_{n-1}]$ and U_n, and this negatively affects the accuracy of computed approximate solutions. This essentially means that, after some initial residual reduction, Simpler GMRES and ORTHODIR can behave unstably, which makes our analysis on maximum attainable accuracy inapplicable.

As a remedy, we now turn to the second choice, $Z_n = \widetilde{R}_n$, which leads to RB-SGMRES (proposed here as a more stable counterpart of Simpler GMRES) and to a version of GCR due to Eisenstat, Elman, and Schultz [6, 5] (see also [29]). Hence, we choose normalized residuals $r_0, \dots, r_{n-1}$ as the basis of $\mathcal{K}_n(A, r_0)$. To make sure that such a choice is adequate, we state the following result.

Lemma 3.3. *Let $v_1, \dots, v_{n-1}$ be an orthonormal basis of $A\mathcal{K}_{n-1}(A, r_0)$, $r_0 \notin A\mathcal{K}_{n-1}(A, r_0)$, and $r_k = (I - V_k V_k^T) r_0$, where $V_k \equiv [v_1, \dots, v_k]$, $k = 1, 2, \dots, n-1$. Then the following statements are equivalent:*

1. *$\|r_k\| < \|r_{k-1}\|$ for all $k = 1, \dots, n-1$,*
2. *$r_0, \dots, r_{n-1}$ are linearly independent.*

Proof. Since $r_0 \notin A\mathcal{K}_{n-1}(A, r_0) = \mathcal{R}(V_{n-1})$, we have $r_k \neq 0$ for all $k = 0, 1, \dots, n-1$. It is clear that $\|r_k\| < \|r_{k-1}\|$ if and only if $\langle r_{k-1}, v_k \rangle \neq 0$. If that holds for all $k = 1, \dots, n-1$, the diagonal matrix D_{n-1} is nonsingular. Using the relation (1.5), we find that $R_n[L_{n,n-1}, e_n] = [V_{n-1}D_{n-1}, r_{n-1}]$. Since $r_{n-1} \perp V_{n-1}$, the matrix $[V_{n-1}D_{n-1}, r_{n-1}]$ has orthogonal nonzero columns, and hence its rank equals n. Moreover, $\operatorname{rank}([L_{n,n-1}, e_n]) = n$, and thus $\operatorname{rank}(R_n) = n$; i.e., $r_0, \dots, r_{n-1}$ are linearly independent. Conversely, from the same matrix relation we find that if $r_0, \dots, r_{n-1}$ are linearly independent, then $\operatorname{rank}([V_{n-1}D_{n-1}, r_{n-1}]) = n$, and hence D_{n-1} is nonsingular, which proves that $\|r_k\| < \|r_{k-1}\|$ for all $k = 1, \dots, n-1$. □

Therefore, if the method does not stagnate, i.e., if the 2-norms of the residuals $r_0, \dots, r_{n-1}$ are strictly monotonously decreasing, then $r_0, \dots, r_{n-1}$ are linearly independent. In this case, we can build an orthonormal basis V_n of $A\mathcal{K}_n(A, r_0)$ by the successive orthogonalization of the columns of $A\widetilde{R}_n$ via (1.3). If $r_0 \in A\mathcal{K}_{n-1}(A, r_0)$, we have an exact solution of (1.1), and the method terminates with $x_{n-1} = A^{-1}b$.

Several conditions for the nonstagnation of the minimum residual method have been given in the literature. For example, Eisenstat, Elman, and Schultz [5, 6] show that GCR (and hence any minimum residual method) does not stagnate if the symmetric part of A is positive definite, i.e., if the origin is not contained in the field of values of A. See also Greenbaum and Strakoš [10] for a different proof and Eiermann and Ernst [4]. Several other conditions can be found in Simoncini and Szyld [24] and the references therein. If stagnation occurs, the residuals are no longer linearly independent, and thus the method prematurely breaks down. In particular, if $0 \in \mathcal{F}(A)$, choosing x_0 such that $\langle Ar_0, r_0 \rangle = 0$ leads to a breakdown in the first step. This was first pointed out by Young and Jea [33] with a simple 2×2 example.

However, as shown in the following theorem, when the minimum residual method does not stagnate, the columns of $\widetilde{R}_n$ are a reasonable choice for the basis of $\mathcal{K}_n(A, r_0)$.

Theorem 3.4. *If $r_0 \notin A\mathcal{K}_{n-1}(A, r_0)$ and $\|r_k\| < \|r_{k-1}\|$ for all $k = 1, \dots, n-1$, the condition number of $\widetilde{R}_n$ satisfies*

$$(3.1) \qquad 1 \leq \kappa(\widetilde{R}_n) \leq \sqrt{n}\,\gamma_n, \qquad \gamma_n \equiv \sqrt{1 + \sum_{k=1}^{n-1} \frac{\|r_{k-1}\|^2 + \|r_k\|^2}{\|r_{k-1}\|^2 - \|r_k\|^2}}.$$

Proof. From (1.5) it follows that

$$\widetilde{R}_n[\widetilde{L}_{n,n-1}, e_n] = \left[V_{n-1}, \frac{r_{n-1}}{\|r_{n-1}\|}\right], \quad \widetilde{L}_{n,n-1} \equiv \operatorname{diag}(\|r_0\|, \dots, \|r_{n-1}\|) L_{n,n-1} D_{n-1}^{-1}.$$

Since $[V_{n-1}, \frac{r_{n-1}}{\|r_{n-1}\|}]$ is an orthonormal matrix, we have from [14, Theorem 3.3.16]

$$\begin{aligned} 1 = \sigma_n\left(\left[V_{n-1}, \frac{r_{n-1}}{\|r_{n-1}\|}\right]\right) &\leq \sigma_n(\widetilde{R}_n)\|[\widetilde{L}_{n,n-1}, e_n]\| \\ &\leq \sigma_n(\widetilde{R}_n)\|[\widetilde{L}_{n,n-1}, e_n]\|_F. \end{aligned}$$

The value of $\|[\widetilde{L}_{n,n-1}, e_n]\|_F$ can be directly computed as

$$\|[\widetilde{L}_{n,n-1}, e_n]\|_F = \sqrt{1 + \sum_{k=1}^{n-1} \frac{\|r_{k-1}\|^2 + \|r_k\|^2}{\|r_{k-1}\|^2 - \|r_k\|^2}} = \gamma_n$$

since $\alpha_k^2 = \|r_{k-1}\|^2 - \|r_k\|^2$. The statement then follows using $\|\widetilde{R}_n\| \leq \|\widetilde{R}_n\|_F \leq \sqrt{n}$. $\square$

We define the quantity γ_n in (3.1) as the *stagnation factor*. The conditioning of $\widetilde{R}_n$ is thus related to the convergence of the method, but in contrast to the conditioning of $[\frac{r_0}{\|r_0\|}, V_{n-1}]$, it is related to the intermediate decrease of the residual norms and not to the residual decrease with respect to the initial residual. A different bound for the conditioning of the matrix $\widetilde{R}_n$ in terms of the residual norms of GMRES and FOM could be derived using the approach in [26].

We illustrate our theoretical results by two numerical examples using the ill-conditioned matrices FS1836 ($\|A\| \approx 1.2 \cdot 10^9$, $\kappa(A) \approx 1.5 \cdot 10^{11}$) and STEAM1 ($\|A\| \approx 2.2 \cdot 10^7$, $\kappa(A) \approx 3 \cdot 10^7$) obtained from the Matrix Market [1] with the right-hand side $b = Ae$ and with the initial guess $x_0 = 0$. In Figures 3.1, 3.2, 3.4, and 3.5 we show the normwise backward error $\|b - Ax_n\|/(\|A\|\|x_n\|)$, the relative norm of the residual $\|b - Ax_n\|/\|b\|$ and $\|r_n\|/\|b\|$, and the relative norms of the error $\|x - x_n\|/\|x\|$ for the choice $Z_n = [\frac{r_0}{\|r_0\|}, V_{n-1}]$ that corresponds to Simpler GMRES and ORTHODIR (Figures 3.1 and 3.4), and for $Z_n = \widetilde{R}_n$ corresponding to RB-SGMRES and GCR (Figures 3.2 and 3.5), respectively. In Figures 3.3 and 3.6 we report the condition numbers of the system matrix A, the basis Z_n, and the triangular matrix U_n multiplied by the unit roundoff u. We see that the backward errors, residual norms, and error norms are almost identical for corresponding implementations of the generalized simpler and update approaches. This can be observed in most cases: the differences between Simpler GMRES and ORTHODIR, and RB-SGMRES and GCR, respectively, are practically negligible. Figures 3.1 and 3.4 illustrate our theoretical considerations and show that, after some initial reduction, *the backward error of Simpler GMRES and ORTHODIR may stagnate at a significantly higher level than the backward error of RB-SGMRES or GCR, which stagnates at a level proportional to the roundoff units*, as shown in Figures 3.2 and 3.5. Due to Theorem 3.2, after some initial phase, the norms of the errors start to diverge in Simpler GMRES and ORTHODIR, while for RB-SGMRES and GCR we have a stagnation on a level approximately proportional to $u\kappa(A)$. The difference is clearly caused by the choice of the basis Z_n, which has an effect on the conditioning of the matrix U_n. We see that $\widetilde{R}_n$ remains well-conditioned up to the very end of the iteration process, while the conditioning of $[\frac{r_0}{\|r_0\|}, V_{n-1}]$ is linked to the convergence of Simpler GMRES and may lead to a very ill-conditioned triangular matrix U_n. Consequently, the approximate solution x_n computed from (1.7) becomes inaccurate, and its error starts to diverge. This problem does not occur in the RB-SGMRES method and GCR, since the matrix U_n remains well-conditioned due to the low stagnation factor. These two implementations behave almost equally to the backward stable MGS-GMRES method. For numerical experiments with MGS-GMRES on the same examples, we refer the reader to [11] and [15].

4. Conclusions. In this paper we have studied the numerical behavior of several minimum residual methods mathematically equivalent to GMRES. Two general formulations have been analyzed: the generalized simpler approach that does not require an upper Hessenberg factorization and the generalized update approach which

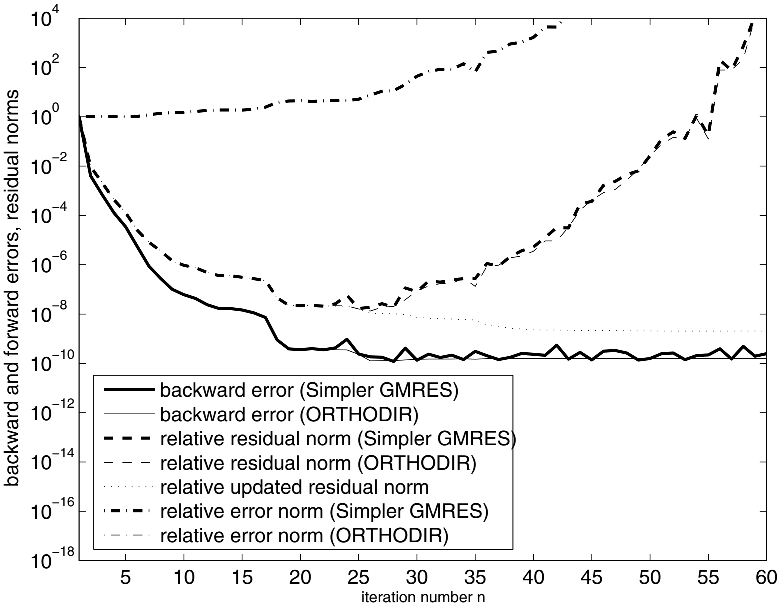

FIG. 3.1. *The test problem $FS1836$ solved by Simpler GMRES and ORTHODIR: Normwise backward error $\|b - Ax_n\|/(\|A\|\|x_n\|)$ (thick solid line: Simpler GMRES; thin solid line: ORTHODIR), relative true residual norm $\|b - Ax_n\|/\|b\|$ (thick dashed line: Simpler GMRES; thin dashed line: ORTHODIR), relative norm of the updated residual $\|r_n\|/\|b\|$ (dotted line), relative norms of the error $\|x - x_n\|/\|x\|$ (thick dash-dotted line: Simpler GMRES; thin dash-dotted line: ORTHODIR).*

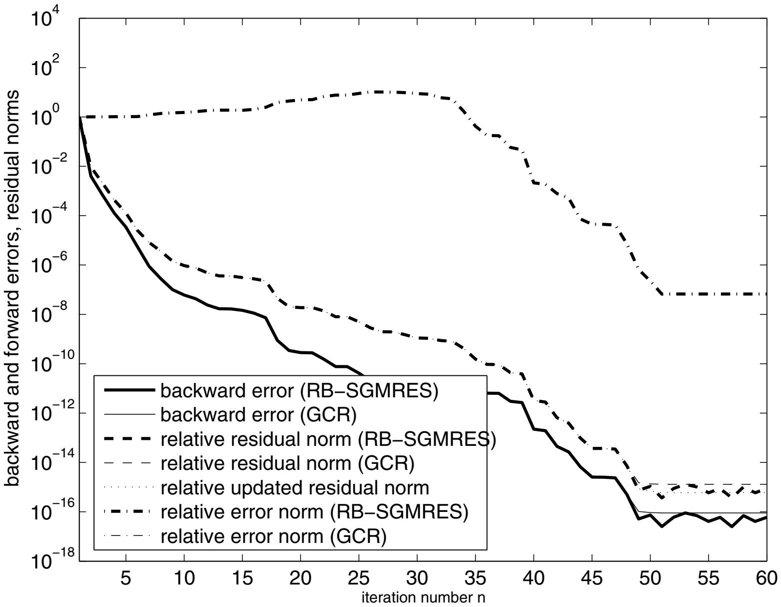

FIG. 3.2. *The test problem $FS1836$ solved by RB-SGMRES and GCR: Normwise backward error $\|b - Ax_n\|/(\|A\|\|x_n\|)$ (thick solid line: RB-SGMRES; thin solid line: GCR), relative true residual norm $\|b - Ax_n\|/\|b\|$ (thick dashed line: RB-SGMRES; thin dashed line: GCR), relative norm of the updated residual $\|r_n\|/\|b\|$ (dotted line), relative norms of the error $\|x - x_n\|/\|x\|$ (thick dash-dotted line: RB-SGMRES; thin dash-dotted line: GCR).*

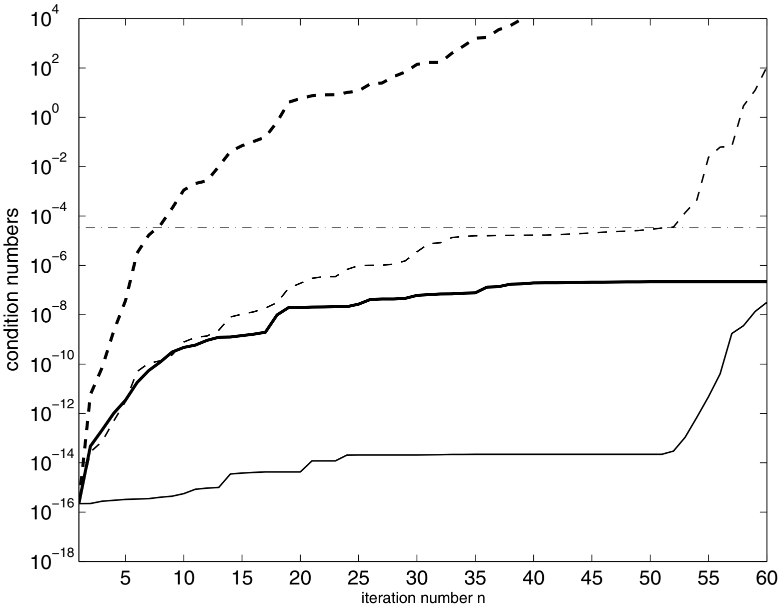

FIG. 3.3. *The test problem* $FS1836$*, condition numbers multiplied by unit roundoff* u*:* $u\kappa(A)$ *(dash-dotted line);* $u\kappa(Z_n)$ *(thick solid line) and* $u\kappa(U_n)$ *(thick dashed line) for* $Z_n = [\frac{r_0}{\|r_0\|}, V_{n-1}]$*;* $u\kappa(Z_n)$ *(thin solid line) and* $u\kappa(U_n)$ *(thin dashed line) for* $Z_n = \widetilde{R}_n$*.*

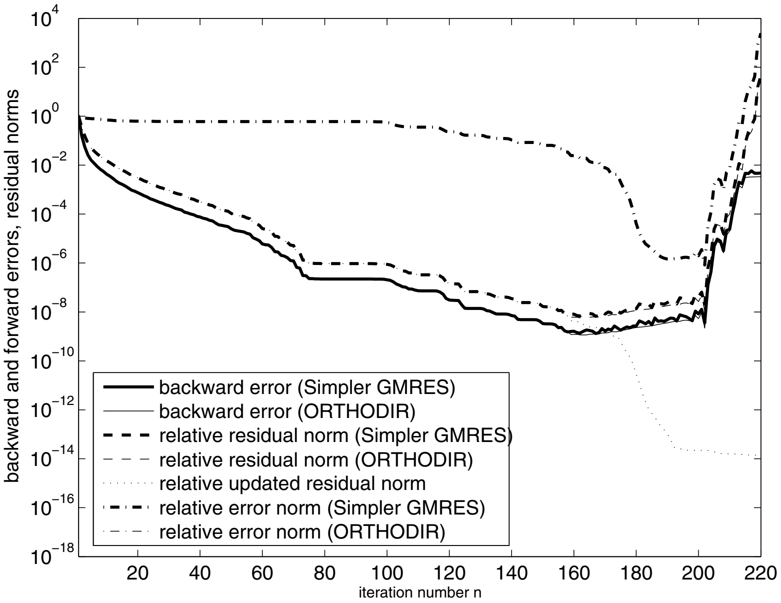

FIG. 3.4. *The test problem* $STEAM1$ *solved by Simpler GMRES and ORTHODIR: Normwise backward error* $\|b - Ax_n\|/(\|A\|\|x_n\|)$ *(thick solid line: Simpler GMRES; thin solid line: ORTHODIR), relative true residual norm* $\|b - Ax_n\|/\|b\|$ *(thick dashed line: Simpler GMRES; thin dashed line: ORTHODIR), relative norm of the updated residual* $\|r_n\|/\|b\|$ *(dotted line), relative norms of the error* $\|x - x_n\|/\|x\|$ *(thick dash-dotted line: Simpler GMRES; thin dash-dotted line: ORTHODIR).*

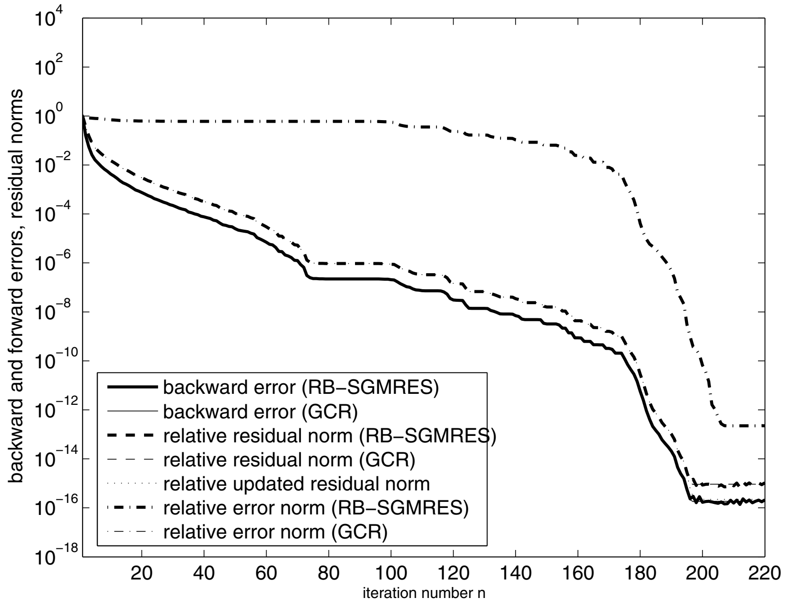

FIG. 3.5. *The test problem STEAM1 solved by RB-SGMRES and GCR: Normwise backward error* $\|b - Ax_n\|/(\|A\|\|x_n\|)$ *(thick solid line: RB-SGMRES; thin solid line: GCR), relative true residual norm* $\|b - Ax_n\|/\|b\|$ *(thick dashed line: RB-SGMRES; thin dashed line: GCR), relative norm of the updated residual* $\|r_n\|/\|b\|$ *(dotted line), relative norms of the error* $\|x - x_n\|/\|x\|$ *(thick dash-dotted line: RB-SGMRES; thin dash-dotted line: GCR).*

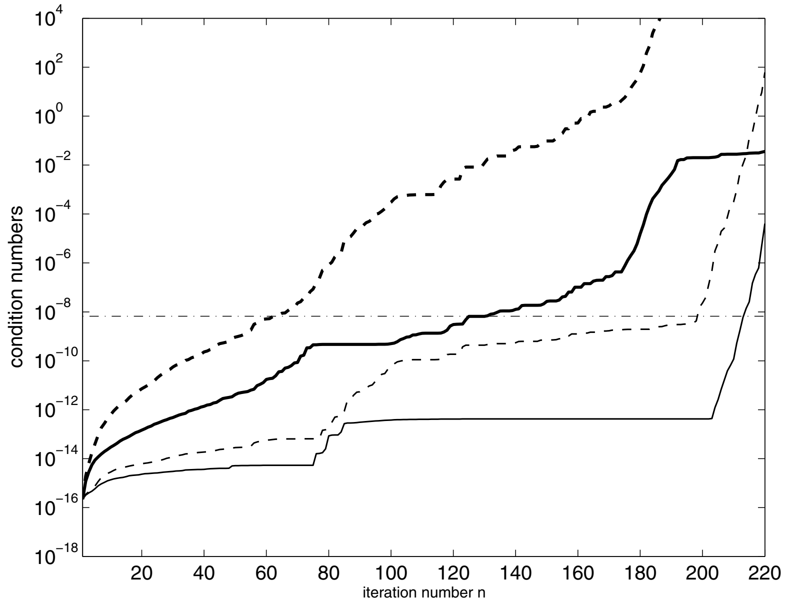

FIG. 3.6. *The test problem STEAM1, condition numbers multiplied by unit roundoff u:* $u\kappa(A)$ *(dash-dotted line);* $u\kappa(Z_n)$ *(thick solid line) and* $u\kappa(U_n)$ *(thick dashed line) for* $Z_n = [\frac{r_0}{\|r_0\|}, V_{n-1}]$*;* $u\kappa(Z_n)$ *(thin solid line) and* $u\kappa(U_n)$ *(thin dashed line) for* $Z_n = \widetilde{R}_n$*.*

is based on generating a sequence of appropriately computed direction vectors. It has been shown that for the generalized simpler approach our analysis leads to an upper bound for the backward error proportional to the roundoff unit, whereas for the generalized update approach the same quantity can be bounded by a term proportional to the condition number of A. Although our analysis suggests that the difference between both may be up to the order of $\kappa(A)$, in practice they behave very similarly, and it is very difficult to find a concrete example with a significant difference in the limiting accuracy measured by the normwise backward error of the approximate solutions x_n. Our first test problem displayed in Figure 2.1 is such a rare example. Moreover, when looking at the errors, we note that both approaches lead essentially to the same accuracy of x_n.

We have indicated that the choice of the basis Z_n is the most important issue for the stability of the considered schemes. Our analysis supports the well-known fact that, even when implemented with the best possible orthogonalization techniques, Simpler GMRES and ORTHODIR are inherently less stable due to the choice $Z_n = [\frac{r_0}{\|r_0\|}, V_{n-1}]$ for the basis. The situation becomes significantly better when we use the residual basis $Z_n = \widetilde{R}_n$. This choice leads to the popular GCR (ORTHOMIN, GMRESR) method, which is widely used in applications. Assuming some reasonable residual decrease (which happens almost always in finite precision arithmetic), we have shown that this scheme is quite efficient, and we have proposed a conditionally backward stable variant RB-SGMRES. Our theoretical results in a sense justify the use of the GCR method in practical computations. In this paper we studied only the unpreconditioned implementations. The implications for the preconditioned GCR scheme will be discussed elsewhere.

Acknowledgments. We would like to thank Julien Langou, Yvan Notay, Kees Vuik, and Gérard Meurant for valuable discussions during the preparation of the paper. We also thank the referees for their comments, which helped to improve the presentation of our results.

REFERENCES

[1] National Institute of Standards and Technology, *Matrix Market*, http://math.nist.gov/MatrixMarket.

[2] W. E. Arnoldi, *The principle of minimized iterations in the solution of the matrix eigenvalue problem*, Quart. Appl. Math., 9 (1951), pp. 17–29.

[3] J. Drkošová, A. Greenbaum, M. Rozložník, and Z. Strakoš, *Numerical stability of GMRES*, BIT, 35 (1995), pp. 309–330.

[4] M. Eiermann and O. G. Ernst, *Geometric aspects of the theory of Krylov subspace methods*, Acta Numer., 10 (2001), pp. 251–312.

[5] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (1983), pp. 345–357.

[6] H. C. Elman, *Iterative Methods for Large Sparse Nonsymmetric Systems of Linear Equations*, Ph.D. thesis, Yale University, New Haven, CT, 1982.

[7] D. K. Faddeev and V. N. Faddeeva, *Computational Methods of Linear Algebra*, Fizmatgiz, Moskow, 1960 (in Russian).

[8] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.

[9] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, Frontiers Appl. Math. 17, SIAM, Philadelphia, 1997.

[10] A. Greenbaum and Z. Strakoš, *Matrices that generate the same Krylov residual spaces*, in Recent Advances in Iterative Methods, G. H. Golub, A. Greenbaum, and M. Luskin, eds., Springer-Verlag, New York, 1994, pp. 95–119.

[11] A. Greenbaum, M. Rozložník, and Z. Strakoš, *Numerical behaviour of the modified Gram-Schmidt GMRES implementation*, BIT, 37 (1997), pp. 706–719.

[12] M. H. Gutknecht and Z. Strakoš, *Accuracy of two three-term and three two-term recurrences for Krylov space solvers*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 213–229.

[13] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, 1996.

[14] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*,, Corrected reprint of the 1991 original, Cambridge University Press, Cambridge, UK, 1994.

[15] J. Liesen, M. Rozložník, and Z. Strakoš, *Least squares residuals and minimal residual methods*, SIAM J. Sci. Comput., 23 (2002), pp. 1503–1525.

[16] G. Meurant, *Computer Solution of Large Linear Systems*, Stud. Math. Appl. 28, North–Holland, Amsterdam, 1999.

[17] Y. Notay, *Personal communication.*

[18] C. C. Paige, M. Rozložník, and Z. Strakoš, *Modified Gram–Schmidt (MGS), least squares, and backward stability of MGS-GMRES*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 264–284.

[19] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.

[20] M. Rozložník and Z. Strakoš, *Variants of residual minimizing Krylov subspace methods*, in Proceedings of the 6th Summer School Software and Algorithms of Numerical Mathematics, Ivo Marek, ed., University of West Bohemia, Pilsen, 1995, pp. 208–225.

[21] Y. Saad, *A flexible inner-outer preconditioned GMRES algorithm*, SIAM J. Sci. Comput., 14 (1993), pp. 461–469.

[22] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.

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

[24] V. Simoncini and D. B. Szyld, *New conditions for non-stagnation of minimal residual methods*, Numer. Math., 109 (2008), pp. 477–487.

[25] G. L. G. Sleijpen, H. A. van der Vorst, and J. Modersitzki, *Differences in the effects of rounding errors in Krylov solvers for symmetric indefinite linear systems*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 726–751.

[26] J. van den Eshof, G. L. G. Sleijpen, and M. B. van Gijzen, *Iterative linear system solvers with approximate matrix-vector products*, in QCD and Numerical Analysis III, Lect. Notes Comput. Sci. Eng. 47, A. Borici, A. Frommer, B. Joo, A. D. Kennedy, and B. Pendleton, eds., Springer-Verlag, Berlin, 2005, pp. 133–142.

[27] A. van der Sluis, *Condition numbers and equilibration matrices*, Numer. Math., 14 (1969), pp. 14–23.

[28] H. A. van der Vorst and C. Vuik, *GMRESR: A family of nested GMRES methods*, Numer. Linear Algebra Appl., 1 (1994), pp. 369–386.

[29] P. K. W. Vinsome, *Orthomin, an iterative method for solving sparse sets of simultaneous linear equations*, in Proceedings of the Fourth Symposium on Reservoir Simulation, SPE of AIME, Los Angeles, CA, 1976, pp. 149–159.

[30] H. F. Walker and L. Zhou, *A simpler GMRES*, Numer. Linear Algebra Appl., 1 (1994), pp. 571–581.

[31] J. H. Wilkinson, *Rounding Errors in Algebraic Processes*, Prentice–Hall, Englewood Cliffs, NJ, 1963.

[32] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, UK, 1965.

[33] D. M. Young and K. C. Jea, *Generalized conjugate gradient acceleration of nonsymmetrizable iterative methods*, Linear Algebra Appl., 34 (1980), pp. 159–194.