Numer Algor
DOI 10.1007/s11075-009-9311-2

ORIGINAL PAPER

Adaptive version of Simpler GMRES

Pavel Jiránek · Miroslav Rozložník

Received: 10 November 2008 / Accepted: 23 June 2009

Abstract In this paper we propose a stable variant of Simpler GMRES. It is based on the adaptive choice of the Krylov subspace basis at a given iteration step using the intermediate residual norm decrease criterion. The new direction vector is chosen as in the original implementation of Simpler GMRES or it is equal to the normalized residual vector as in the GCR method. We show that such an adaptive strategy leads to a well-conditioned basis of the Krylov subspace and we support our theoretical results with illustrative numerical examples.

The work of the first author was supported by the grant No. 201/09/P464 of the Grant Agency of the Czech Republic.

The work of the second author was supported by the project IAA100300802 of the Grant Agency of the Academy of Sciences of the Czech Republic and by the Institutional Research Plan AV0Z10300504 "Computer Science for the Information Society: Models, Algorithms, Applications".

P. Jiránek
Faculty of Mechatronics, Technical University of Liberec,
Studentská 2, 46117 Liberec, Czech Republic
e-mail: pavel.jiranek@tul.cz

M. Rozložník
Institute of Computer Science, Academy of Sciences of the Czech Republic,
Pod Vodárenskou věží 2, 18207 Prague 8, Czech Republic
e-mail: miro@cs.cas.cz

Present Address:
P. Jiránek (✉)
CERFACS, 42 Avenue Gaspard Coriolis, 31057 Toulouse Cedex 1, France
e-mail: jiranek@cerfacs.fr

Published online: 07 July 2009

Keywords Nonsymmetric linear systems · Krylov subspace methods · Minimum residual methods · Numerical stability · Rounding errors

Mathematics Subject Classifications (2000) 65F10 · 65G50 · 65F35

1 Introduction

We consider the solution of a large and sparse system of linear algebraic equations

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

where $A \in \mathbb{R}^{N\times N}$ is nonsingular and $b \in \mathbb{R}^N$ is a right-hand side vector. A popular method for solving such a system is the GMRES method of Saad and Schultz [13]. At the iteration step n it seeks the approximate solution x_n in the affine subspace $x_0 + \mathcal{K}_n(A, r_0)$, where

$$\mathcal{K}_n(A, r_0) := \mathrm{span}\{r_0, Ar_0, \ldots, A^{n-1}r_0\}$$

is the nth Krylov subspace generated by the matrix A and the residual vector $r_0 := b - Ax_0$ corresponding to the initial guess x_0. The GMRES method is based on the Arnoldi process [1] generating the orthonormal basis Q_n of the Krylov subspace $\mathcal{K}_n(A, r_0)$ and minimizing the Euclidean norm of the residual in $r_0 + A\mathcal{K}_n(A, r_0)$, i.e.,

$$\|b - Ax_n\| = \|b - A(x_0 + d_n)\| = \min_{d\in\mathcal{K}_n(A,r_0)} \|b - A(x_0 + d)\|. \tag{2}$$

If a stopping criterion is satisfied at some iteration step m, the coordinates y_m of d_m in the orthogonal basis Q_m are found by solving an $(m + 1) \times m$ upper Hessenberg least squares problem and the approximate solution is then computed as $x_m = x_0 + d_m = x_0 + Q_m y_m$. The GMRES method with the Householder or modified Gram-Schmidt Arnoldi implementation was proved to be backward stable in [3, 10] which means that there is an approximate solution of (1) which can be interpreted as an exact solution of a system (1) with slightly perturbed initial data A and b. See also the Higham's book [6] for details of the backward error concept.

In [16] Walker and Zhou proposed another implementation of the GMRES method. We will describe it in a slightly more general way. Let $Z_n := [z_1, \ldots, z_n]$ be a matrix such that its columns form a basis of $\mathcal{K}_n(A, r_0)$ and such that $\mathcal{R}(Z_k) = \mathcal{K}_k(A, r_0)$ for all $k = 1, \ldots, n$ and, in addition, we assume that its columns are normalized, i.e., $\|z_k\| = 1$ for $k = 1, \ldots, n$. Here $\mathcal{R}(\cdot)$ denotes the range of the matrix. The minimum residual property (2) is equivalent to the requirement of the residual vector $r_n := b - Ax_n$ being orthogonal to the subspace $A\mathcal{K}_n(A, r_0)$:

$$r_n \perp A\mathcal{K}_n(A, r_0) = \mathcal{R}(AZ_n). \tag{3}$$

The residual r_n is then easily evaluated provided we have an orthonormal basis $V_n := [v_1, \ldots, v_n]$ of $A\mathcal{K}_n(A, r_0) = \mathcal{R}(AZ_n)$, which can be computed by the QR factorization of the matrix AZ_n:

$$AZ_n = V_n U_n. \tag{4}$$

The matrix $U_n \in \mathbb{R}^{n\times n}$ is upper triangular and nonsingular if and only if the dimension of $\mathcal{K}_n(A, r_0)$ is equal to n. The residual $r_n \in r_0 + A\mathcal{K}_n(A, r_0) = r_0 + \mathcal{R}(V_n)$ satisfying the property (3) (and (2)) can then be computed as the orthogonal projection of the initial residual r_0:

$$r_n = \left(I - V_n V_n^T\right) r_0 = \left(I - v_n v_n^T\right) r_{n-1} = r_{n-1} - \alpha_n v_n,\ \alpha_n := v_n^T r_{n-1}. \tag{5}$$

The approximate solution x_n corresponding to the residual r_n has the form $x_n = x_0 + Z_n t_n$, where t_n is the solution of the upper triangular system

$$U_n t_n = V_n^T r_0 = [\alpha_1, \ldots, \alpha_n]^T. \tag{6}$$

Let $\widetilde{r}_k := r_k/\|r_k\|$ denote a normalized residual vector at the iteration step k. In [16] the matrix Z_n is chosen as $[\widetilde{r}_0, V_{n-1}]$, i.e., the normalized initial residual is extended by the first $n-1$ vectors of the orthonormal basis V_n. However, it was shown in [9, 16] that the condition number of $[\widetilde{r}_0, V_{n-1}]$ is proportional to the inverse of the relative residual norm, i.e., it grows as the residual norm decreases. Therefore the original implementation of the Simpler GMRES method [16] can suffer from numerical instability due to the ill-conditioning of the basis which moreover leads to the severe ill-conditioning of the upper triangular factor U_n in (4) possibly affected also by ill-conditioning of A; see the numerical experiments in [8, 16]. On the other hand, if the residual norm (nearly) stagnates the Simpler GMRES basis $[\widetilde{r}_0, V_{n-1}]$ remains well-conditioned. As it was shown in [8] the matrix Z_n consisting of the normalized residuals $[\widetilde{r}_0, \ldots, \widetilde{r}_{n-1}]$ remains well-conditioned provided there is a reasonable residual norm decrease at each iteration. The Simpler GMRES method with such a residual basis, called RB-SGMRES in [8], was shown to be conditionally backward stable. It is closely related to GCR by Eisenstat, Elman and Schultz [4]. See [8] and [12] for more details.

It was shown in [8] that the condition number of Z_n can affect the maximum attainable accuracy of the computed approximation. In Section 2 we propose a variant of the Simpler GMRES method (called the adaptive Simpler GMRES), which keeps the condition number of the basis Z_n at a reasonable level by adaptive selection at each iteration of a suitable direction vector based on the intermediate residual norm decrease. Whenever the residual norm (nearly) stagnates at the iteration step n we use the vector v_{n-1}. Otherwise, when we observe a sufficient residual norm decrease, we set the new direction vector equal to the normalized residual vector $\widetilde{r}_{n-1}$. A similar strategy is employed, e.g., in [11], where the Orthomin method [15] is combined with Orthodir [17] for solving saddle point problems in computational fluid dynamics. Here we show that the adaptive choice of direction vectors keeps the basis well-conditioned and that the condition number grows at most linearly with the

iteration number. Finally, we illustrate our theoretical results with numerical experiments in Section 3.

Throughout the paper, we denote by $\|\cdot\|$ the Euclidean vector norm and the induced matrix norm, and by $\|\cdot\|_F$ the Frobenius norm. 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. We denote by I_n the $n \times n$ unit matrix. If $X_i \in \mathbb{R}^{n_i\times n_i}$ $(i = 1, \ldots, m)$ are square matrices, we denote by $\mathrm{diag}(X_1, \ldots, X_m)$ the block diagonal matrix of the order $\sum_{i=1}^{m} n_i$ with diagonal blocks $X_1, \ldots, X_m$. For a vector $y \in \mathbb{R}^n$ the notation $\mathrm{diag}(y)$ or $\mathrm{diag}(y^T)$ is used in a usual manner and defines the $n \times n$ diagonal matrix with the components of y on the main diagonal and zeros elsewhere.

2 Adaptive Simpler GMRES

In this section we propose an adaptive variant of the Simpler GMRES method, which computes the basis Z_n in (4) such that its condition number is kept at a reasonably small level. This is achieved by an adaptive switching between the bases from Simpler GMRES and RB-SGMRES using an intermediate residual decrease criterion. If the residual norm at a given step sufficiently decreases the Krylov subspace basis is extended by the normalized residual vector as in RB-SGMRES or GCR; otherwise we use the last available vector of the orthonormal basis as in Simpler GMRES. In order to decide whether the residual norm is sufficiently reduced we introduce the threshold parameter $\nu \in [0, 1]$ and choose for $n > 1$ either the vector $z_n = \widetilde{r}_{n-1}$ provided that $\|r_{n-1}\| \leq \nu\|r_{n-2}\|$ or $z_n = v_{n-1}$ otherwise. We sketch the algorithm of adaptive Simpler GMRES as follows:

Algorithm 2.1 (Adaptive Simpler GMRES) *Choose x_0 and the threshold parameter $\nu \in [0, 1]$, compute $r_0 := b - Ax_0$, and for $n = 1, \ldots, m$ (until convergence) do*

1. *compute z_n:*

$$z_n = \begin{cases} \widetilde{r}_0 = r_0/\|r_0\| & \text{if } n = 1, \\ \widetilde{r}_{n-1} = r_{n-1}/\|r_{n-1}\| & \text{if } n > 1 \text{ and } \|r_{n-1}\| \leq \nu\|r_{n-2}\|, \\ v_{n-1} & \text{otherwise}, \end{cases} \tag{7}$$

2. *update the QR factorization $AZ_n = V_nU_n$,*
3. *compute $\alpha_n = v_n^T r_{n-1}$,*
4. *update $r_n := r_{n-1} - \alpha_n v_n$,*

solve $U_m t_m = [\alpha_1, \ldots, \alpha_m]^T$, compute $x_m := x_0 + Z_m t_m$.

If $\nu = 0$ then $Z_n = [\widetilde{r}_0, V_{n-1}]$ and Algorithm 2.1 is identical to Simpler GMRES [16]. The choice $\nu = 1$ results in $Z_n = [\widetilde{r}_0, \ldots, \widetilde{r}_{n-1}]$ which corresponds to RB-SGMRES [8].

Theorem 2.2 *Let A in (1) be nonsingular and n be such that the dimension of $\mathcal{K}_n(A, r_0)$ is equal to n. If $\nu \in [0, 1)$ then the columns of $Z_n = [z_1, \ldots, z_n]$ computed in Algorithm* 2.1 *form a basis of $\mathcal{K}_n(A, r_0)$ satisfying $z_k \in \mathcal{K}_k(A, r_0) \setminus \mathcal{K}_{k-1}(A, r_0)$ for all $k = 1, \ldots, n$. In particular, adaptive Simpler GMRES does not break down unless the exact solution of (1) is found.*

Proof We proceed by induction on n. For $n = 1$ the statement is clearly satisfied by setting $\mathcal{K}_0(A, r_0) := \{0\}$. Let $n > 1$ and Z_{n-1} be a basis of $\mathcal{K}_{n-1}(A, r_0)$. From (4) the columns of V_{n-1} are an orthonormal basis of $A\mathcal{K}_{n-1}(A, r_0)$. The vector v_{n-1} is computed from the vector $Az_{n-1} \in \mathcal{K}_n(A, r_0) \setminus \mathcal{K}_{n-1}(A, r_0)$ orthogonalizing it against the orthonormal basis V_{n-2} of $A\mathcal{K}_{n-2}(A, r_0)$ and thus belongs to $\mathcal{K}_n(A, r_0) \setminus \mathcal{K}_{n-1}(A, r_0)$. We consider two cases: First let $\|r_{n-1}\| > \nu\|r_{n-2}\|$ and hence by (7) $z_n = v_{n-1}$. The vector v_{n-1} extends the basis Z_{n-1} of $\mathcal{K}_{n-1}(A, r_0)$ to the basis of $\mathcal{K}_n(A, r_0)$ as follows from the discussion above. Otherwise, let $\|r_{n-1}\| \le \nu\|r_{n-2}\|$. Since $\|r_{n-1}\| < \|r_{n-2}\|$ and $r_k \in r_0 + A\mathcal{K}_k(A, r_0)$ it follows that $r_{n-1} = r_{n-2} - \alpha_{n-1}v_{n-1}$ with $\alpha_{n-1} \neq 0$ and $r_{n-1} \in \mathcal{K}_{n-1}(A, r_0) \setminus \mathcal{K}_{n-2}(A, r_0)$. Hence the columns of $Z_n = [Z_{n-1}, \widetilde{r}_{n-1}]$ are a basis of $\mathcal{K}_n(A, r_0)$. □

It is known that the residual basis can be linearly dependent if the minimum residual method does not make any progress at a given step, in particular when 0 belongs to the field of values of the matrix A it may happen that $\alpha_n = 0$ resulting in $r_n = r_{n-1}$. Therefore we have excluded the case $\nu = 1$ from Theorem 2.2 which can lead to a breakdown of the RB-SGMRES or GCR algorithms.

We recall the results on the maximum attainable accuracy of algorithms based on (4) studied in [8], which apply also for adaptive Simpler GMRES. We assume that the QR factorization at Step 2 of Algorithm 2.1 is constructed such that the upper triangular matrix U_n is computed in a backward stable way; see, e.g., [8, Equations (2.1) and (2.2)]. This is true in particular for Householder QR factorization, modified Gram-Schmidt algorithm, and (classical and modified) Gram-Schmidt algorithms with reorthogonalization. Let $\hat{x}_n$ be an approximate solution computed at iteration n of Algorithm 2.1 in finite precision arithmetic with the unit roundoff u. In addition let $cu\kappa(A)\kappa(Z_n) < 1$, where the constant c is a low-order polynomial in n and N, which guarantees that AZ_n and U_n are of full numerical rank. Then the gap between the true residual $b - A\hat{x}_n$ and the updated residual r_n can be estimated as

$$\|b - A\hat{x}_n - r_n\| \le \frac{cu\kappa(A)}{1 - cu\kappa(A)\kappa(Z_n)} \sum_{k=1}^{n} \frac{\|r_{k-1}\|}{\sigma_k(Z_k)}. \tag{8}$$

When looking at the accuracy in terms of the normwise backward error $\|b - A\hat{x}_n\|/(\|A\|\|\hat{x}_n\|)$, see, e.g., [6], it follows from [8, Theorem 2.1] that

$$\frac{\|b - A\hat{x}_n - r_n\|}{\|A\|\|\hat{x}_n\|} \le cu\kappa(Z_n)\left(1 + \frac{\|x_0\|}{\|\hat{x}_n\|}\right). \tag{9}$$

Since the norm of the updated residual r_n usually becomes orders of magnitude smaller than the norm of the true residual $b - A\hat{x}_n$, the right-hand sides of (8) and (9) represent then the bounds on the maximum attainable accuracy in terms of the residual norm and the backward error, respectively. The condition number of the basis Z_n plays therefore an important role in the numerical stability of algorithms based on (4). In the following we analyze the condition number of Z_n produced by adaptive Simpler GMRES. First we prove three auxiliary propositions, where the nonincreasing sequence $\{\rho_k\}$ represents the sequence of residual norms $\rho_k = \|r_k\|$. Note that the sequence $\{\alpha_k\}$ defined as $\alpha_k^2 = \rho_{k-1}^2 - \rho_k^2$ coincides then (up to the sign) with the orthogonalization coefficients in (5).

Lemma 2.3 *Let p and q be two integers such that $1 \le p < q$ and let $\widetilde{B}_{p,q} \in \mathbb{R}^{(q-p+1)\times(q-p+1)}$ be a lower Hessenberg matrix defined by*

$$\widetilde{B}_{p,q} := \begin{bmatrix} \boldsymbol{\alpha}_{p,q-1}/\rho_{p-1} & I_{q-p} \\ \rho_{q-1}/\rho_{p-1} & 0 \end{bmatrix},$$

where $\boldsymbol{\alpha}_{p,q-1} := [\alpha_p, \ldots, \alpha_{q-1}]^T$, $\alpha_k^2 = \rho_{k-1}^2 - \rho_k^2$ for $k = p, \ldots, q-1$, and $0 < \rho_{q-1} \le \rho_{q-2} \le \ldots \le \rho_{p-1}$. Then

$$\frac{\rho_{p-1}}{\rho_{q-1}} \le \kappa(\widetilde{B}_{p,q}) = \delta_{p,q} := \frac{\rho_{p-1} + \left(\rho_{p-1}^2 - \rho_{q-1}^2\right)^{\frac{1}{2}}}{\rho_{q-1}} \le 2\frac{\rho_{p-1}}{\rho_{q-1}}.$$

Proof The proof uses a similar technique as in [9, Theorem 2.3]. By direct computation we obtain

$$\begin{aligned} \widetilde{B}_{p,q}^T \widetilde{B}_{p,q} &= \begin{bmatrix} (\rho_{q-1}^2 + \|\boldsymbol{\alpha}_{p,q-1}\|^2)/\rho_{p-1}^2 & \boldsymbol{\alpha}_{p,q-1}^T/\rho_{p-1} \\ \boldsymbol{\alpha}_{p,q-1}/\rho_{p-1} & I_{q-p} \end{bmatrix} \\ &= \begin{bmatrix} 1 & \boldsymbol{\alpha}_{p,q-1}^T/\rho_{p-1} \\ \boldsymbol{\alpha}_{p,q-1}/\rho_{p-1} & I_{q-p} \end{bmatrix}. \end{aligned}$$

There exists an orthonormal matrix $U \in \mathbb{R}^{(q-p)\times(q-p)}$ such that

$$U\boldsymbol{\alpha}_{p,q-1} = \|\boldsymbol{\alpha}_{p,q-1}\| e_1 = \left(\rho_{p-1}^2 - \rho_{q-1}^2\right)^{\frac{1}{2}} e_1 = \rho_{p-1}\beta e_1$$

and hence

$$\begin{bmatrix} 1 & 0 \\ 0 & U \end{bmatrix} \widetilde{B}_{p,q}^T \widetilde{B}_{p,q} \begin{bmatrix} 1 & 0 \\ 0 & U^T \end{bmatrix} = \begin{bmatrix} 1 & \beta e_1^T \\ \beta e_1 & I_{q-p} \end{bmatrix} =: G,$$

where $\beta := (1 - \rho_{q-1}^2/\rho_{p-1}^2)^{\frac{1}{2}}$. The eigenvalues of G are equal to the eigenvalues of its leading principal 2×2 submatrix together with 1 of multiplicity $q - p - 1$. Since G is (orthogonally) similar to $\widetilde{B}_{p,q}^T \widetilde{B}_{p,q}$, the square roots of its eigenvalues are equal to the singular values of $\widetilde{B}_{p,q}$. The extremal singular values of $\widetilde{B}_{p,q}$ can be given as

$$\sigma_1^2(\widetilde{B}_{p,q}) = 1 + \beta, \qquad \sigma_{q-p+1}^2(\widetilde{B}_{p,q}) = 1 - \beta. \tag{10}$$

Note that $0 \leq \beta < 1$ and 1 is neither minimal nor maximal eigenvalue of G unless $\beta = 0$. A simple algebraic manipulation gives

$$\kappa(\widetilde{B}_{p,q}) = \frac{\rho_{p-1} + \left(\rho_{p-1}^2 - \rho_{q-1}^2\right)^{\frac{1}{2}}}{\rho_{q-1}} = \delta_{p,q}.$$

The upper and lower bounds for $\delta_{p,q}$ follow directly from its definition. □

Lemma 2.4 *Let q and m be two integers such that $1 \leq q < m$ and $\widetilde{C}_{q,m} \in \mathbb{R}^{(m-q+1)\times(m-q+1)}$ be a lower triangular matrix*

$$\widetilde{C}_{q,m} := \operatorname{diag}\left(\left[\boldsymbol{\alpha}_{q,m-1}^T, \rho_{m-1}\right]\right) L_{m-q+1}^{-1} \operatorname{diag}(\boldsymbol{\rho}_{q-1,m-1})^{-1},$$

where $\boldsymbol{\alpha}_{q,m-1} := [\alpha_q, \ldots, \alpha_{m-1}]^T$, $\boldsymbol{\rho}_{q-1,m-1} := [\rho_{q-1}, \ldots, \rho_{m-1}]^T$, $\alpha_k^2 = \rho_{k-1}^2 - \rho_k^2$ for $k = q, \ldots, m-1$, $0 < \rho_{m-1} < \rho_{m-2} < \ldots < \rho_{q-1}$. The matrix L_{m-q+1} of the order $m-q+1$ is lower bidiagonal with 1 on the main diagonal and -1 on the first subdiagonal. Then

$$\underline{\gamma}_{q,m} \leq \kappa(\widetilde{C}_{q,m}) \leq \overline{\gamma}_{q,m},$$

where

$$\underline{\gamma}_{q,m} := \max_{k=q,\ldots,m-1} \left(\frac{\rho_{k-1}^2 + \rho_k^2}{\rho_{k-1}^2 - \rho_k^2}\right)^{\frac{1}{2}},$$

$$\overline{\gamma}_{q,m} := (m-q+1)^{\frac{1}{2}} \left(1 + \sum_{k=q}^{m-1} \frac{\rho_{k-1}^2 + \rho_k^2}{\rho_{k-1}^2 - \rho_k^2}\right)^{\frac{1}{2}}.$$

Proof The inverse of L_{m-q+1} is the matrix with ones on the main diagonal and below and with zeros elsewhere, i.e., the lower triangular matrix with all elements in the lower triangular equal to one. By direct computation it can be verified that the columns of the matrix $\widetilde{C}_{q,m}$ have unit norms and thus

$$\|\widetilde{C}_{q,m}\| \leq \|\widetilde{C}_{q,m}\|_F = (m-q+1)^{\frac{1}{2}}, \tag{11}$$

The lower bound follows by considering the definition of the matrix norm

$$\|\widetilde{C}_{q,m}\| = \max_{\|z\|=1} \|\widetilde{C}_{q,m} z\| \geq \max_{k=1,\ldots,m-q+1} \|\widetilde{C}_{q,m} e_k\| = 1, \tag{12}$$

where e_k denotes the k-th column of the unit matrix. The inverse of $\widetilde{C}_{q,m}$ exists, since $\alpha_k \neq 0$ for $k = q, \ldots, m-1$ and $\rho_{m-1} \neq 0$, and it is a lower bidiagonal matrix

$$\widetilde{C}_{q,m}^{-1} = \operatorname{diag}(\boldsymbol{\rho}_{q-1,m-1}) L_{m+q-1} \operatorname{diag}\left(\left[\boldsymbol{\alpha}_{q,m-1}^T, \rho_{m-1}\right]\right)^{-1}.$$

The minimal singular value of $\widetilde{C}_{q,m}$ can be estimated from below and above in the similar way by considering the inverse of $\widetilde{C}_{q,m}$ leading to

$$\max_{k=q,\ldots,m-1}\left(\frac{\rho_{k-1}^2+\rho_k^2}{\rho_{k-1}^2-\rho_k^2}\right)^{\frac{1}{2}} \le \|\widetilde{C}_{q,m}^{-1}\| \le \|\widetilde{C}_{q,m}^{-1}\|_F = \left(1+\sum_{k=q}^{m-1}\frac{\rho_{k-1}^2+\rho_k^2}{\rho_{k-1}^2-\rho_k^2}\right)^{\frac{1}{2}},$$

which together with (11) and (12) concludes the proof. □

The proof of the previous lemma was already given (for $q = 1$) in [8, Theorem 3.4] in the context of RB-SGMRES, which was shown to be conditionally backward stable provided that the upper bound $\overline{\gamma}_{1,m}$ is reasonably small. Another estimate on the condition number of the residual basis can be established using the Gershgorin theorem [5] (see also, e.g., [14, Theorem 1.11]).

Lemma 2.5 *Let the assumptions of Lemma* 2.4 *be satisfied and let*

$$\underline{\lambda}_k := \begin{cases} \frac{\rho_k}{\rho_{k+1}+\rho_k} & \text{for} \quad k = q-1, \\ \frac{\rho_k}{\rho_{k+1}+\rho_k} - \frac{\rho_k}{\rho_k+\rho_{k-1}} & \text{for} \quad k = q,\ldots,m-2, \\ \frac{\rho_{k-1}}{\rho_k+\rho_{k-1}} & \text{for} \quad k = m-1, \end{cases}$$

$$\overline{\lambda}_k := \begin{cases} \frac{\rho_k}{\rho_k-\rho_{k+1}} & \text{for} \quad k = q-1, \\ \frac{\rho_k}{\rho_{k-1}-\rho_k} + \frac{\rho_k}{\rho_k-\rho_{k+1}} & \text{for} \quad k = q,\ldots,m-2, \\ \frac{\rho_{k-1}}{\rho_{k-1}-\rho_k} & \text{for} \quad k = m-1. \end{cases}$$

Then

$$\kappa(\widetilde{C}_{q,m}) \le \left(\frac{\max\limits_{k=q-1,\ldots,m-1}\overline{\lambda}_k}{\min\limits_{k=q-1,\ldots,m-1}\underline{\lambda}_k}\right)^{\frac{1}{2}}.$$

Proof The matrix $\widetilde{C}_{q,m}^{-1}\widetilde{C}_{q,m}^{-T}$ can be written in the form

$$\widetilde{C}_{q,m}^{-1}\widetilde{C}_{q,m}^{-T} = \operatorname{diag}(\boldsymbol{\rho}_{q-1,m-1})\, T \operatorname{diag}(\boldsymbol{\rho}_{q-1,m-1}),$$

where

$$T = \begin{bmatrix} \frac{1}{\alpha_q^2} & -\frac{1}{\alpha_q^2} & & & \\ -\frac{1}{\alpha_q^2} & \frac{1}{\alpha_q^2}+\frac{1}{\alpha_{q+1}^2} & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \frac{1}{\alpha_{m-2}^2}+\frac{1}{\alpha_{m-1}^2} & -\frac{1}{\alpha_{m-1}^2} & \\ & & -\frac{1}{\alpha_{m-1}^2} & \frac{1}{\alpha_{m-1}^2}+\frac{1}{\rho_{m-1}^2} & \end{bmatrix}.$$

It is straightforward to show that the matrix $\widetilde{C}_{q,m}^{-1}\widetilde{C}_{q,m}^{-T}$ is diagonally dominant. Let $\eta_{q-1},\ldots,\eta_{m-1}$ and $\delta_{q-1},\ldots,\delta_{m-1}$ denote the diagonal entries and the sum of absolute values of the off-diagonal entries in rows $1,\ldots,m-q+1$. Note that the diagonal entries are positive, while the off-diagonal ones are negative. Since $\widetilde{C}_{q,m}^{-1}\widetilde{C}_{q,m}^{-T}$ is symmetric, its eigenvalues are real and belong to $\cup_{k=q-1}^{m-1}[\eta_k-\delta_k,\eta_k+\delta_k]$ due to the Gershgorin theorem [5]. We find that $\eta_k-\delta_k=\underline{\lambda}_k$ and $\eta_k+\delta_k=\overline{\lambda}_k$ and the proof of the statement is finished. □

The bound using the Gershgorin theorem will be employed below in order to establish the a priori estimate on the condition number of the basis provided we have a prescribed value of the threshold parameter ν. Whenever we want to explain the local contributions of intermediate residual norm decreases to the condition number of the Krylov subspace basis, the estimate in Lemma 2.4 is however more useful. We exploit it in the following theorem where we consider the case where at the steps $n=2,\ldots,q$ of Algorithm 2.1 the vector z_n is chosen as in Simpler GMRES, i.e., $z_n=v_{n-1}$, and $z_n=\widetilde{r}_{n-1}$ as in RB-SGMRES for $n=q+1,\ldots,m$. It corresponds to adaptive Simpler GMRES applied to a problem with some initial stagnation of the minimum residual norm and a fast convergence afterwards.

Theorem 2.6 *Let* $Z_m=[\widetilde{r}_0,v_1,\ldots,v_{q-1},\widetilde{r}_q,\ldots,\widetilde{r}_{m-1}]$ *for some integer* q *such that* $1<q<m$ *and let* $0<\|r_{m-1}\|<\cdots<\|r_{q-1}\|$. *Then*

$$Z_m=[V_{m-1},\widetilde{r}_{m-1}]H_m, \tag{13}$$

with $H_m=C_mB_m$, $B_m:=\mathrm{diag}(\widetilde{B}_{1,q},I_{m-q})$, $C_m:=\mathrm{diag}(I_{q-1},\widetilde{C}_{q,m})$. *The vectors* $\boldsymbol{\alpha}_{1,q-1}$, $\boldsymbol{\alpha}_{q,m-1}$, *and* $\boldsymbol{\rho}_{q-1,m-1}$ *and the matrices* $\widetilde{B}_{1,q}$ *and* $\widetilde{C}_{q,m}$ *are defined as in Lemmas* 2.3 *and* 2.4 *(with* $p=1$ *and with* $\rho_k:=\|r_k\|$*). The condition number of* Z_m *can then be bounded as follows:*

$$\max\left\{1,\frac{\delta_{1,q}}{\overline{\gamma}_{q,m}},\frac{\underline{\gamma}_{q,m}}{\delta_{1,q}}\right\}\le\kappa(Z_m)\le\delta_{1,q}\overline{\gamma}_{q,m} \tag{14}$$

with $\delta_{1,q}$, $\overline{\gamma}_{q,m}$, *and* $\underline{\gamma}_{q,m}$ *defined in Lemmas* 2.3 *and* 2.4.

Proof From (5) we have

$$\widetilde{r}_0=[V_{q-1},\widetilde{r}_{q-1}]\begin{bmatrix}\boldsymbol{\alpha}_{1,q-1}/\rho_0\\ \rho_{q-1}/\rho_0\end{bmatrix}.$$

Hence $[\widetilde{r}_0,V_{q-1}]=[V_{q-1},\widetilde{r}_{q-1}]\widetilde{B}_{1,q}$ and

$$Z_m=[V_{q-1},\widetilde{r}_{q-1},\ldots,\widetilde{r}_{m-1}]B_m. \tag{15}$$

Again using (5) we find that $[\widetilde{r}_{q-1},\ldots,\widetilde{r}_{m-1}]=[v_q,\ldots,v_{m-1},\widetilde{r}_{m-1}]\widetilde{C}_{q,m}$ and

$$[V_{q-1},\widetilde{r}_{q-1},\ldots,\widetilde{r}_{m-1}]=[V_{m-1},\widetilde{r}_{m-1}]C_m. \tag{16}$$

Combining (15) and (16) proves (13) and since $[V_{m-1}, \widetilde{r}_{m-1}]$ has orthogonal columns, the singular values of Z_m are equal to the singular values of $H_m = C_m B_m$. Assumptions of the theorem imply that C_m and B_m are nonsingular. Using the definition of the condition number $\kappa(H_m) = \sigma_1(C_m B_m)/\sigma_m(C_m B_m)$ and the inequalities for the singular values (see, e.g., [7, Theorem 3.3.16]), we obtain

$$\max\left\{\frac{\kappa(B_m)}{\kappa(C_m)}, \frac{\kappa(C_m)}{\kappa(B_m)}\right\} \leq \kappa(H_m) \leq \kappa(B_m)\kappa(C_m).$$

Applying Lemmas 2.3 and 2.4 concludes the proof. □

Corollary 2.7 *Let the assumptions of Theorem* 2.6 *be satisfied. In addition, let* $\|r_k\| > \nu\|r_{k-1}\|$ *for* $k = 1, \ldots, q-1$ *and* $\|r_k\| \leq \nu\|r_{k-1}\|$ *for* $k = q, \ldots, m-1$ *for some* $\nu < 1$. *Then*

$$\kappa(Z_m) \leq \frac{2\sqrt{2}}{\nu^{q-1}} \frac{1+\nu}{1-\nu}. \tag{17}$$

Proof From $\|r_k\| \leq \nu\|r_{k-1}\|$ it follows:

$$\frac{1}{(1+\nu)\|r_{k-1}\|} \leq \frac{1}{\|r_{k-1}\| + \|r_k\|} \leq \frac{1}{2\|r_k\|}$$

and

$$\frac{1}{\|r_{k-1}\|} \leq \frac{1}{\|r_{k-1}\| - \|r_k\|} \leq \frac{1}{(1-\nu)\|r_{k-1}\|},$$

and therefore

$$\frac{1}{2}\frac{1-\nu}{1+\nu} \leq \underline{\lambda}_k \leq \overline{\lambda}_k \leq \frac{1+\nu}{1-\nu}$$

for $k = q-1, \ldots, m-1$. Then the result follows from Theorem 2.6 and Lemma 2.5 (with $\rho_k = \|r_k\|$) and from the assumption $\|r_k\| > \nu\|r_{k-1}\|$ ($k = 1, \ldots, q-1$), which implies that $\|r_{q-1}\| > \nu^{q-1}\|r_0\|$ and $1 \leq \delta_{1,q} \leq 2/\nu^{q-1}$. □

Theorem 2.6 and Corollary 2.7 indicate that at the iteration steps where the residual norm (nearly) stagnates, the contribution of vectors from Simpler GMRES to the condition number of Z_m is given approximately by the inverse of the relative residual norm decrease during the (near) stagnation. At steps where the residual norm is sufficiently reduced, the condition number of corresponding vectors in Z_m is bounded by the stagnation factors $\underline{\gamma}_{q,m}$ and $\overline{\gamma}_{q,m}$ from below and above. Considering (9) and (17) we can estimate the backward error of adaptive Simpler GMRES from

$$\frac{\|b - A\hat{x}_m - r_m\|}{\|A\|\|\hat{x}_m\|} \leq cu \frac{1}{\nu^{q-1}} \frac{1+\nu}{1-\nu}\left(1 + \frac{\|x_0\|}{\|\hat{x}_m\|}\right).$$

Provided that the factor dependent on ν in the right-hand side of the inequality is not large, the adaptive variant of Simpler GMRES is backward stable. Ultimately it means that the approximate solution $\hat{x}_m$ is an exact solution of $(A + \Delta A)x_m = b$ with slightly perturbed data $A + \Delta A$, where $\|\Delta A\| = O(u)\|A\|$.

In the inequality (17) of Corollary 2.7 we can find a quasi-optimal value of $\nu = \nu_{\text{opt}}$ minimizing the right-hand side of the bound (i.e., not the actual value of $\kappa(Z_m)$). It is clear that $q - 1 \le m$, so

$$\kappa(Z_m) \le \frac{2\sqrt{2}}{\nu^m} \frac{1+\nu}{1-\nu}. \tag{18}$$

The value of ν minimizing the right-hand side of (18) is given by

$$\nu_{\text{opt}}(m) = \frac{\sqrt{1+m^2} - 1}{m}.$$

It can be shown that the first term $[\nu_{\text{opt}}(m)]^{-m}$ in (18) grows with m and approaches $e \approx 2.7183$ as $m \to \infty$. For the second term we have $\frac{1+\nu_{\text{opt}}(m)}{1-\nu_{\text{opt}}(m)} \sim 2m$ with $m \to \infty$. Hence the quasi-optimal bound in (18) behaves like $\kappa(Z_m) \lesssim 4\sqrt{2}\, e\, m$ for $m \to \infty$. The threshold parameter $\nu_{\text{opt}}(m)$ is asymptotically reaching the value 1 for growing m, where m can be associated with the maximum number of iterations or the restart parameter. We observed in numerical experiments that $\nu_{\text{opt}}(m)$ minimizing the right-hand side of (18) does not always lead to optimal condition number of the basis and the smaller value, say $\nu = 0.9$, can do better. On the other hand, we have shown that the quasi-optimal value $\nu_{\text{opt}}(m)$ leads to at worst linearly growing $\kappa(Z_m)$.

Theorem 2.6 can be generalized to the case with multiple switching between the bases from Simpler GMRES basis and RB-SGMRES. Such situation is more realistic since it can happen that there are intermediate stagnations of the residual norm followed by a fast convergence phase. Therefore we introduce the sequences of indices $\{q_j\}_{j=1}^{\ell}$ and $\{m_j\}_{j=1}^{\ell}$ corresponding to ℓ stagnation and convergence phases; see Fig. 1 for the illustration and the explanation of the notation in the theorem below. The quantities δ, $\underline{\gamma}$, and $\overline{\gamma}$ play a similar role as the factors $\delta_{1,q}$, $\underline{\gamma}_{q,m}$, and $\overline{\gamma}_{q,m}$ in Theorem 2.6.

Theorem 2.8 *Let Z_m has the block form $Z_m = [\widetilde{Z}_1, \ldots, \widetilde{Z}_\ell, \widetilde{r}_{m-1}]$, where*

$$\widetilde{Z}_j := [\widetilde{r}_{m_{j-1}-1}, v_{m_{j-1}}, \ldots, v_{q_j-1}, \widetilde{r}_{q_j}, \ldots, \widetilde{r}_{m_j-2}] \in \mathbb{R}^{N\times(m_j-m_{j-1})},$$

$m_0 = 1$, $m_\ell = m$, $m_{j-1} < q_j < m_j$ *and* $0 < \|r_{m_j-1}\| < \cdots < \|r_{q_j-1}\|$ *for* $j = 1, \ldots, \ell$. *Then*

$$\max\left\{1, \frac{\delta}{\overline{\gamma}}, \frac{\underline{\gamma}}{\delta}\right\} \le \kappa(Z_m) \le \overline{\gamma}\delta, \tag{19}$$

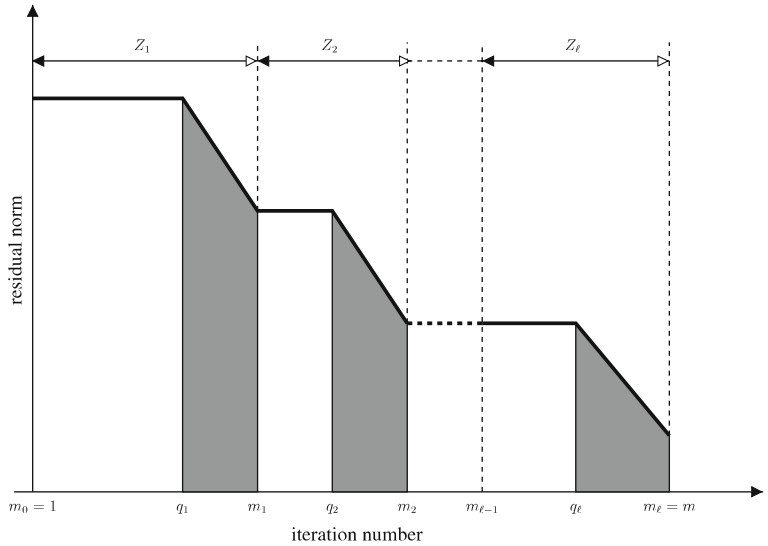

Fig. 1 Multiple switching between the Simpler GMRES basis and the residual basis in the case of the occurrence of local stagnations in the residual norm. In the run of adaptive Simpler GMRES, the *white areas* correspond to the Simpler GMRES basis, while the *gray areas* correspond to the normalized residual basis of RB-SGMRES

where

$$\delta := \max_{j=1,\ldots,\ell} \frac{\|r_{m_{j-1}}\| + \left(\|r_{m_{j-1}}\|^2 - \|r_{q_j}\|^2\right)^{\frac{1}{2}}}{\|r_{q_j}\|},$$

$$\underline{\gamma} := \max_{j=1,\ldots,\ell} \max_{i=q_j,\ldots,m_j-1} \left(\frac{\|r_{i-1}\|^2 + \|r_i\|^2}{\|r_{i-1}\|^2 - \|r_i\|^2}\right)^{\frac{1}{2}},$$

$$\overline{\gamma} := \left(1 + \sum_{j=1}^{\ell}(m_j - q_j)\right)^{\frac{1}{2}} \left(1 + \sum_{j=1}^{\ell}\sum_{i=q_j}^{m_j-1} \frac{\|r_{i-1}\|^2 + \|r_i\|^2}{\|r_{i-1}\|^2 - \|r_i\|^2}\right)^{\frac{1}{2}}.$$

Proof As in Theorem 2.6 we use (5) repeatedly in order to relate Z_m (which forms the basis of $\mathcal{K}_m(A, r_0)$ due to the assumptions of the theorem) to the orthonormal basis $[V_{m-1}, \widetilde{r}_{m-1}]$. In each $\widetilde{Z}_j$ we relate the first residual $\widetilde{r}_{m_{j-1}-1}$ to $\widetilde{r}_{q_j-1}$ using the vectors $v_{m_{j-1}}, \ldots, v_{q_j-1}$. Thus we obtain

$$\begin{aligned}\widetilde{Z}_j &= [\widetilde{r}_{m_{j-1}-1}, v_{m_{j-1}}, \ldots, v_{q_j-1}, \widetilde{r}_{q_j}, \ldots, \widetilde{r}_{m_j-2}] \\ &= [v_{m_{j-1}}, \ldots, v_{q_j-1}, \widetilde{r}_{q_j-1}, \widetilde{r}_{q_j}, \ldots, \widetilde{r}_{m_j-2}]\mathrm{diag}(\widetilde{B}_{m_{j-1},q_j}, I_{m_j-q_j-1}) \\ &= \widetilde{Y}_j B_{m_{j-1},q_j},\end{aligned}$$

where $\widetilde{Y}_j := [v_{m_{j-1}}, \ldots, v_{q_j-1}, \widetilde{r}_{q_j-1}, \widetilde{r}_{q_j}, \ldots, \widetilde{r}_{m_j-2}]$, $\widetilde{B}_{m_{j-1},q_j}$ is defined in Lemma 2.3 (with $\rho_k = \|r_k\|$), and $B_{m_{j-1},q_j} := \mathrm{diag}(\widetilde{B}_{m_{j-1},q_j}, I_{m_j-q_j-1})$. Hence it follows that

$$Z_m = [\widetilde{Y}_1, \ldots, \widetilde{Y}_\ell, \widetilde{r}_{m-1}]B_m, \tag{20}$$

with the matrix B_m defined by $B_m := \mathrm{diag}(B_{m_0,q_1}, \ldots, B_{m_{\ell-1},q_\ell}, 1)$. We now relate $[\widetilde{Y}_1, \ldots, \widetilde{Y}_\ell, \widetilde{r}_{m-1}]$ to $[V_{m-1}, \widetilde{r}_{m-1}]$. More precisely, we express the columns of $[V_{m-1}, \widetilde{r}_{m-1}] = [\widetilde{V}_1, \ldots, \widetilde{V}_\ell, \widetilde{r}_{m-1}]$ in terms of $[\widetilde{Y}_1, \ldots, \widetilde{Y}_\ell, \widetilde{r}_{m-1}]$, where $\widetilde{V}_j := [v_{m_{j-1}}, \ldots, v_{m_j-1}]$. From (5) we have

$$[r_{q_j-1}, \ldots, r_{m_j-1}]L_{m_j-q_j+1,m_j-q_j} = [v_{q_j}, \ldots, v_{m_j-1}]\mathrm{diag}(\boldsymbol{\alpha}_{q_j,m_j-1}). \tag{21}$$

Here $L_{m_j-q_j+1,m_j-q_j}$ is defined as $L_{m_j-q_j+1,m_j-q_1} := [L^T_{m_j-q_j}, -e_{m_j-q_j}]^T$, where $e_{m_j-q_j}$ stands for the last column of $I_{m_j-q_j}$. From (21) it follows that

$$[v_{q_j}, \ldots, v_{m_j-1}] = [\widetilde{r}_{q_j-1}, \ldots, \widetilde{r}_{m_j-2}]\widetilde{G}_{q_j,m_j} - \frac{1}{\alpha_{m_j-1}} r_{m_j-1} e^T_{m_j-q_j}, \tag{22}$$

with $\widetilde{G}_{q_j,m_j}$ defined by $\widetilde{G}_{q_j,m_j} := \mathrm{diag}(\boldsymbol{\rho}_{q_j-1,m_j-2})L_{m_j-q_j}[\mathrm{diag}(\boldsymbol{\alpha}_{q_j,m_j-1})]^{-1}$. Since r_{m_j-1} (or $\widetilde{r}_{m_j-1}$) is not in $[Y_1, \ldots, Y_\ell, \widetilde{r}_{m-1}]$ (for $j = 1, \ldots, \ell-1$), we express it in terms of the residual $r_{q_{j+1}-1}$ and the vectors $v_{m_j}, \ldots, v_{q_{j+1}-1}$ as

$$r_{m_j-1} = r_{q_{j+1}-1} + [v_{m_j}, \ldots, v_{q_{j+1}-1}]\boldsymbol{\alpha}_{m_j,q_{j+1}-1} = \widetilde{Y}_{j+1}\begin{bmatrix} \boldsymbol{\alpha}_{m_j,q_{j+1}-1} \\ \rho_{q_{j+1}-1} \\ \mathbf{0}_{m_{j+1}-q_{j+1}-1} \end{bmatrix}$$

for $j = 1, \ldots, \ell-1$. Here $\mathbf{0}_{m_{j+1}-q_{j+1}-1}$ denotes the column zero vector of the given dimension. From (22) we hence obtain

$$\begin{aligned} \widetilde{V}_j &= \widetilde{Y}_j\,\mathrm{diag}(I_{q_j-m_{j-1}}, \widetilde{G}_{q_j,m_j}) - \frac{1}{\alpha_{m_j-1}} r_{m_j-1} e^T_{m_j-m_{j-1}} \\ &= \widetilde{Y}_j\,\mathrm{diag}(I_{q_j-m_{j-1}}, \widetilde{G}_{q_j,m_j}) - \frac{1}{\alpha_{m_j-1}} \widetilde{Y}_{j+1}\begin{bmatrix} \boldsymbol{\alpha}_{m_j,q_{j+1}-1} \\ \rho_{q_{j+1}-1} \\ \mathbf{0}_{m_{j+1}-q_{j+1}-1} \end{bmatrix} e^T_{m_j-m_{j-1}} \\ &= \widetilde{Y}_j\widetilde{D}_{j,j} + \widetilde{Y}_{j+1}\widetilde{D}_{j+1,j}, \qquad j = 1, \ldots, \ell-1, \end{aligned} \tag{23}$$

and

$$\begin{aligned} \widetilde{V}_\ell &= \widetilde{Y}_\ell\,\mathrm{diag}(I_{q_\ell-m_{\ell-1}}, \widetilde{G}_{q_\ell,m_\ell}) - \frac{\rho_{m_\ell-1}}{\alpha_{m_\ell-1}} \widetilde{r}_{m_\ell-1} e^T_{m_\ell-m_{\ell-1}} \\ &= \widetilde{Y}_\ell\widetilde{D}_{\ell,\ell} + \widetilde{r}_{m-1}\widetilde{D}_{\ell+1,\ell}. \end{aligned} \tag{24}$$

Combining (23) and (24), we get

$$[V_{m-1}, \widetilde{r}_{m-1}] = [\widetilde{Y}_1, \ldots, \widetilde{Y}_\ell, \widetilde{r}_{m-1}]D_m, \tag{25}$$

where D_m is a lower block bidiagonal matrix

$$D_m := \begin{bmatrix} \widetilde{D}_{1,1} & & & & \\ \widetilde{D}_{2,1} & \widetilde{D}_{2,2} & & & \\ & \ddots & \ddots & & \\ & & & \widetilde{D}_{\ell,\ell} & \\ & & & \widetilde{D}_{\ell+1,\ell} & 1 \end{bmatrix}.$$

Using (20) and (25) we get the desired relation $Z_m = [V_{m-1}, \widetilde{r}_{m-1}] D_m^{-1} B_m$. Since $[V_{m-1}, \widetilde{r}_{m-1}]$ has orthonormal columns, it follows that

$$\max\left\{\frac{\kappa(B_m)}{\kappa(D_m)}, \frac{\kappa(D_m)}{\kappa(B_m)}\right\} \le \kappa(Z_m) \le \kappa(D_m)\kappa(B_m). \tag{26}$$

Due to Lemma 2.3 and (10) we have

$$\kappa(B_m) = \max_{j=1,\ldots,\ell} \kappa(\widetilde{B}_{m_{j-1},q_j}) = \max_{j=1,\ldots,\ell} \delta_{m_{j-1},q_j} = \delta. \tag{27}$$

To estimate the norm of D_m we find a permutation matrix Π such that

$$\Pi D_m \Pi^T = \begin{bmatrix} I & 0 \\ 0 & \widetilde{D}_m \end{bmatrix}, \tag{28}$$

where we moved the identities from the matrices $\widetilde{D}_{j,j}$ into the leading principal identity matrix of $\Pi D_m \Pi^T$. It follows that $\|D_m\| = \max\{1, \|\widetilde{D}_m\|\} \le \max\{1, \|\widetilde{D}_m\|_F\}$. Since

$$\begin{aligned} \|\widetilde{D}_m\|_F^2 &= 1 + \sum_{j=1}^{\ell} \left(\|\widetilde{G}_{q_j,m_j}\|_F^2 + \|\widetilde{D}_{j+1,j}\|_F^2\right) \\ &= 1 + \sum_{j=1}^{\ell} \sum_{i=q_j}^{m_j-1} \frac{\rho_{i-1}^2 + \rho_i^2}{\rho_{i-1}^2 - \rho_i^2}, \end{aligned}$$

and $\|\widetilde{D}_m\|_F \ge 1$, we can bound the norm of the matrix D_m as

$$\|D_m\|^2 \le 1 + \sum_{j=1}^{\ell} \sum_{i=q_j}^{m_j-1} \frac{\rho_{i-1}^2 + \rho_i^2}{\rho_{i-1}^2 - \rho_i^2}. \tag{29}$$

The inverse of D_m can be computed either directly from D_m or by making the relation between $[\widetilde{Y}_1, \ldots, \widetilde{Y}_\ell, \widetilde{r}_{m-1}]$ and $[V_{m-1}, \widetilde{r}_{m-1}]$ in the opposite direction, which is simpler. Taking into account (5) we can express the residuals $\widetilde{r}_k$ in $\widetilde{Y}_j$ $(j = 1, \ldots, \ell)$ using $\widetilde{r}_{m-1}$ and the corresponding direction vectors $v_{k+1}, \ldots, v_{m-1}$ from V_{m-1}. But since $[V_{m-1}, \widetilde{r}_{m-1}]$ has orthonormal columns and the residuals in $\widetilde{Y}_j$ are normalized, it follows that their coordinates in the basis $[V_{m-1}, \widetilde{r}_{m-1}]$ have unit norms. Considering the same permutation matrix

Π as in (28) we can show that the columns of the lower triangular matrix $\widetilde{D}_m^{-1}$ contain the permuted coordinates of the residuals in $\widetilde{Y}_j$ $(j = 1, \ldots, \ell)$ in the basis $[V_{m-1}, \widetilde{r}_{m-1}]$, and thus they have unit norms. Hence we obtain the bound

$$\|D_m^{-1}\|^2 \le 1 + \sum_{i=1}^{\ell} (m_j - q_j). \tag{30}$$

Similarly as in Lemma 2.4 we get the lower bounds

$$\|D_m\| \ge \max_{j=1,\ldots,\ell} \max_{i=q_j,\ldots,m_j-1} \left(\frac{\rho_{i-1}^2 + \rho_i^2}{\rho_{i-1}^2 - \rho_i^2} \right)^{\frac{1}{2}}, \quad \|D_m^{-1}\| \ge 1. \tag{31}$$

Combining (26), (27), (29), (30), and (31) concludes the proof. □

3 Numerical experiments

We illustrate our theoretical results on numerical examples selected from the Matrix Market [2] and performed in MATLAB using double precision

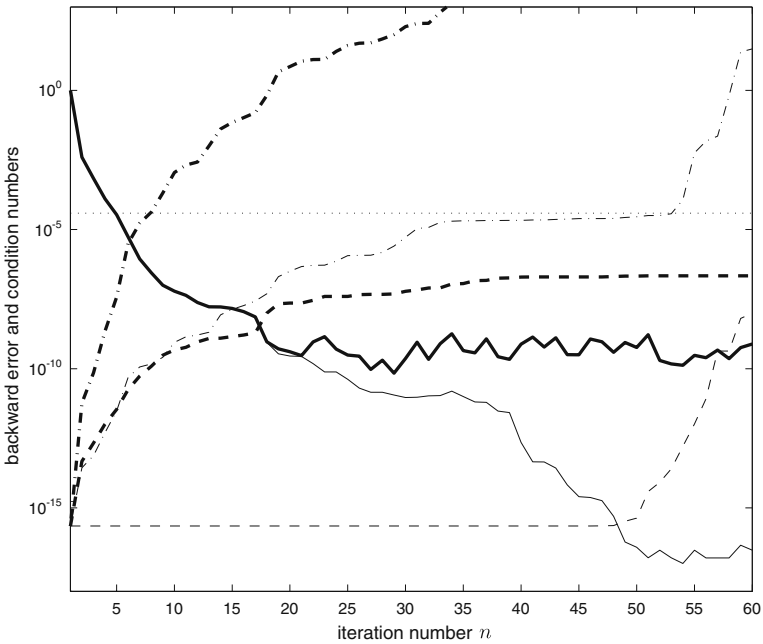

Fig. 2 Test problem with `FS1836` and $b = A[1, \ldots, 1]^T$ solved by Simpler GMRES ($\nu = 0$): backward errors (*bold solid line*) and condition numbers $u\kappa(Z_n)$ (*bold dashed line*) and $u\kappa(U_n)$ (*bold dash-dotted line*); GMRES: normwise backward errors (*solid line*) and condition numbers $u\kappa(Q_n)$ (*dashed line*) and $u\kappa(AQ_n)$ (*dash-dotted line*); $u\kappa(A)$ (*dotted line*)

arithmetic with $u \approx 10^{-16}$. Results for the adaptive Simpler GMRES and modified Gram-Schmidt implementation of GMRES applied to the system with the matrix `FS1836` ($N = 183$, $\|A\| \approx 1.2 \cdot 10^9$, $\kappa(A) \approx 1.7 \cdot 10^{11}$) are illustrated on Figs. 2, 3, 4 and 5. The right-hand side vector b is equal either to $A[1, \ldots, 1]^T$ (Figs. 2 and 3) or to the left singular vector corresponding to the smallest singular value of A (Figs. 4 and 5). On each plot we show the backward error $\|b - Ax_n\|/(\|A\|\|x_n\| + \|b\|)$ (bold solid lines) associated to the approximate solutions x_n computed by adaptive Simpler GMRES with three considered values of the threshold parameter: $\nu = 0$ (Fig. 2) where adaptive Simpler GMRES is equivalent to Simpler GMRES of Walker and Zhou [16], $\nu = 1$ (Fig. 4) leading to RB-SGMRES [8], and $\nu = 0.9$ (Figs. 3 and 5). We also plot the backward errors of approximate solutions computed by GMRES of Saad and Schultz [13] (solid lines). The actual values of condition numbers of Z_n and U_n (multiplied by the unit roundoff u) are plotted by bold dashed and bold dash-dotted lines, respectively. For comparison we report also the condition numbers of Q_n and AQ_n from GMRES (dashed and dash-dotted lines). The condition number of the system matrix A multiplied by the unit roundoff u is depicted by dotted lines.

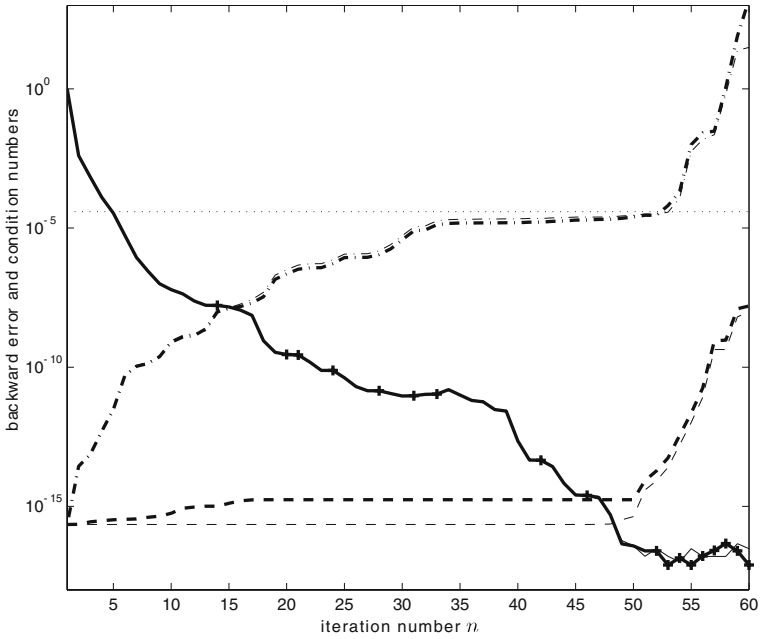

Fig. 3 Test problem with `FS1836` and $b = A[1, \ldots, 1]^T$ solved by adaptive Simpler GMRES with $\nu = 0.9$: normwise backward errors (*bold solid line*) and condition numbers $u\kappa(Z_n)$ (*bold dashed line*) and $u\kappa(U_n)$ (*bold dash-dotted line*); GMRES: backward errors (*solid line*) and condition numbers $u\kappa(Q_n)$ (*dashed line*) and $u\kappa(AQ_n)$ (*dash-dotted line*); $u\kappa(A)$ (*dotted line*). The steps where the Simpler GMRES basis is used are denoted by plus signs

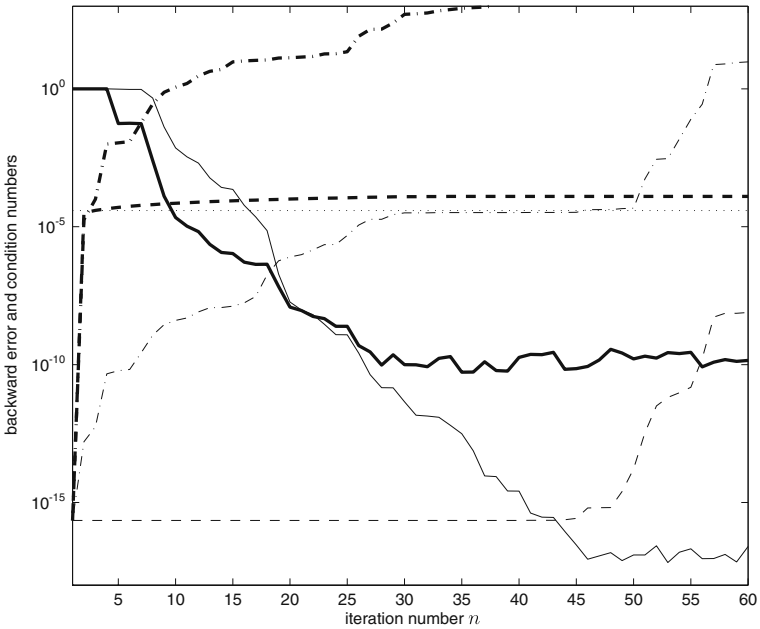

Fig. 4 Test problem with `FS1836` and b equal to the left singular vector corresponding to the smallest singular value of A solved by RB-SGMRES ($\nu = 1$): backward errors (*bold solid line*) and condition numbers $u\kappa(Z_n)$ (*bold dashed line*) and $u\kappa(U_n)$ (*bold dash-dotted line*); GMRES: normwise backward errors (*solid line*) and condition numbers $u\kappa(Q_n)$ (*dashed line*) and $u\kappa(AQ_n)$ (*dash-dotted line*); $u\kappa(A)$ (*dotted line*)

It is clear from our experiments that both Simpler GMRES and RB-SGMRES may lead to low accuracy of the computed approximate solution due to the ill-conditioning of Z_n in the case of a rapid initial convergence with $b = A[1, \ldots, 1]^T$ or in the case of a long initial stagnation in the residual norm with b equal to the left singular vector corresponding to the smallest singular value, respectively; see Figs. 2 and 4. However, as it can be observed from Figs. 3 and 5, the adaptive version of Simpler GMRES with the threshold value $\nu = 0.9$ leads to reasonably conditioned bases for both right-hand sides. In Figs. 3 and 5 we denote by plus signs the steps where the Simpler GMRES basis is used instead of the normalized residual.

When we keep the condition number of the basis Z_n small, one can observe that $\kappa(Z_n) \approx \kappa(Q_n)$ and $\kappa(U_n) \approx \kappa(AQ_n) \lesssim \kappa(A)$. Indeed, by reducing the condition number of Z_n, the proposed adaptive strategy tries to imitate the ideal basis generated by GMRES. This ensures that the condition number of the matrix U_n (or AZ_n) in (4) is less than or equal to $\kappa(A)\kappa(Z_n)$ (see Figs. 3 and 5) and therefore guarantees the validity of the bounds (8) and (9), which rely on the assumption of the numerical nonsingularity of the basis AZ_n.

Figure 6 shows the dependence of $\kappa(Z_m)$ with respect to the threshold parameter ν for several real problems with various condition numbers and

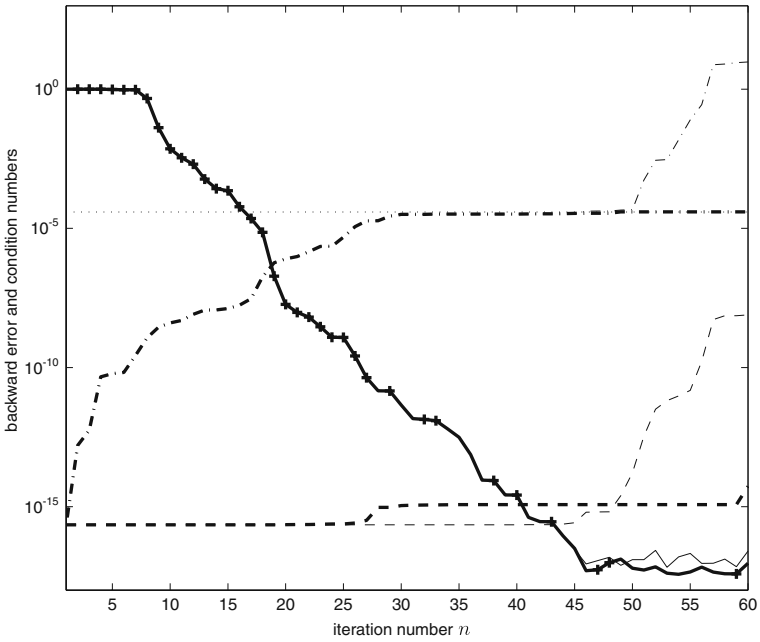

Fig. 5 Test problem with `FS1836` and b equal to the left singular vector corresponding to the smallest singular value of A solved by adaptive Simpler GMRES with $\nu = 0.9$: backward errors (*bold solid line*) and condition numbers $u\kappa(Z_n)$ (*bold dashed line*) and $u\kappa(U_n)$ (*bold dash-dotted line*); GMRES: normwise backward errors (*solid line*) and condition numbers $u\kappa(Q_n)$ (*dashed line*) and $u\kappa(AQ_n)$ (*dash-dotted line*); $u\kappa(A)$ (*dotted line*), The steps where the Simpler GMRES basis is used are denoted by plus signs

dimensions from 225 up to 1080. For each problem we stop the method at the iteration step m, where the normwise backward error dropped below the level 10^{-14}. Note that for each problem and each value of ν varying between 0 and 1, the adaptive version of Simpler GMRES was able to reach such a high accuracy and thus the ill-conditioning of the basis does not necessarily lead to a poor accuracy in terms of the backward error. This phenomenon can be explained using (8), which shows that large $\kappa(Z_k)$ can be damped with the small residual norm $\|r_{k-1}\|$. However, we were not able to prove this for the normwise backward error in [8]. Nevertheless, as we have shown there are examples where ill-conditioning of Z_m leads to low attainable accuracy of the computed approximate solution; cf. Figs. 2 and 4. It is therefore reasonable to keep the condition number of the basis Z_m at a reasonably small level and consequently to keep the columns of AZ_m linearly independent as well as the matrix U_m numerically nonsingular. It is clear from Fig. 6 that, for our examples, the value of ν close (but not equal) to 1 leads to a nearly optimal condition number of Z_m in adaptive Simpler GMRES and, therefore, the residual vectors should be preferred in Z_m even for a moderate intermediate residual norm decrease.

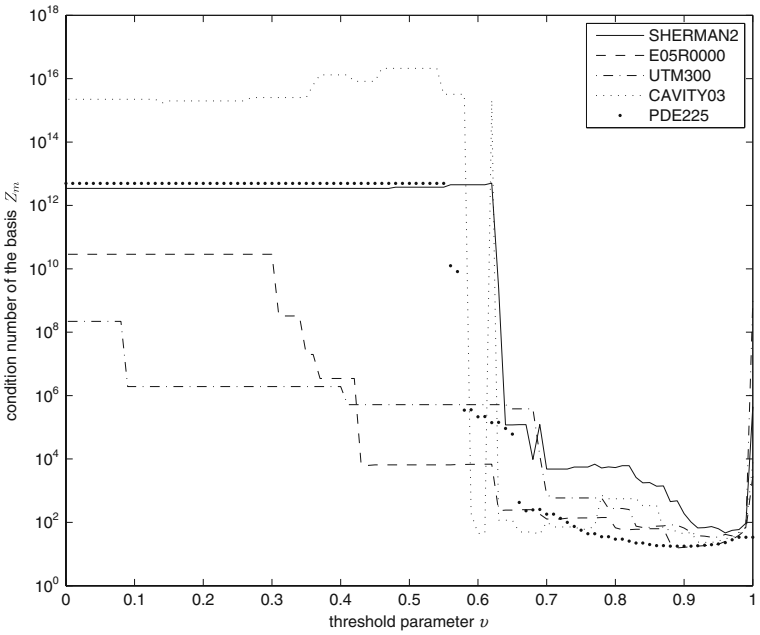

Fig. 6 The dependence of the condition number of Z_m on the choice of the threshold parameter ν for various matrices and right-hand sides taken from Matrix Market

4 Concluding remarks

The classical GMRES method [13] with Householder QR or modified Gram-Schmidt implementation of the Arnoldi process was shown to be backward stable in [3, 10] and in this sense should be the method of choice. The numerical stability of minimum residual methods based on (4) is strongly influenced by the condition number of the Krylov subspace basis Z_n. The Simpler GMRES method (with $Z_n = [\widetilde{r}_0, V_{n-1}]$ by Walker and Zhou [16] is not backward stable due to the relation of $\kappa(Z_n)$ to the inverse of the relative residual norm [9, 16]. Indeed, the rapid residual norm decrease leads to an ill-conditioned Krylov subspace basis and vice versa. On the other hand, the normalized residuals $Z_n = [\widetilde{r}_0, \ldots, \widetilde{r}_{n-1}]$ form a well-conditioned basis of the Krylov subspace provided there is a sufficient residual norm decrease at each iteration step. This choice leads to a backward stable variant of Simpler GMRES–RB-SGMRES [8], which in almost all cases computes very accurate approximate solutions (actually it is very difficult to find a problem where RB-SGMRES or GCR behaves poorly). To overcome the potential weakness caused by the initial or intermediate stagnation of the residual, in this paper we proposed the adaptive variant of Simpler GMRES, which keeps the basis well-conditioned leading to a maximum attainable accuracy similar to classical GMRES method of Saad and Schultz [13].

Acknowledgements We would like to thank the anonymous referees and G. Meurant for their comments and suggestions, which helped to improve the presentation of our results.

References

1. Arnoldi, W.E.: The principle of minimized iteration in the solution of the matrix eigenproblem. Q. Appl. Math. **9**, 17–29 (1951)
2. Boisvert, R.F., Pozo, R., Remington, K., Barret, R., Dongarra, J.J.: The Matrix Market: a web resource for test matrix collections web resource for test matrix collections. In: Boisvert, R.F. (ed.) Quality of Numerical Software, Assessment and Enhancement. Chapman & Hall, London (1997)
3. Drkošová, J., Greenbaum, A., Rozložník, M., Strakoš, Z.: Numerical stability of GMRES. BIT **35**(3), 309–330 (1995)
4. Eisenstat, S.C., Elman, H.C., Schultz, M.H.: Variational iterative methods for nonsymmetric systems of linear equations. SIAM J. Numer. Anal. **20**(2), 345–357 (1983)
5. Gershgorin, S.A.: Über die Abgrenzung der Eigenwerte einer matrix. Izv. Akad. Nauk. SSSR Ser. Mat. **7**, 749–754 (1931)
6. Higham, N.J.: Accuracy and Stability of Numerical Algorithms. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (1996)
7. Horn, R.A., Johnson, C.R.: Topics in Matrix Analysis. Cambridge University Press, Cambridge (1991)
8. Jiránek, P., Rozložník, M., Gutknecht, M.H.: How to make Simpler GMRES and GCR more stable. SIAM J. Matrix Anal. Appl. **30**(4), 1483–1499 (2008)
9. Liesen, J., Rozložník, M., Strakoš, Z.: Least squares residuals and minimal residual methods. SIAM J. Sci. Comput. **23**(5), 1503–1525 (2002)
10. Paige, C.C., Rozložník, M., Strakoš, Z.: Modified Gram-Schmidt (MGS), least squares, and backward stability of MGS-GMRES. SIAM J. Matrix Anal. Appl. **28**(1), 264–284 (2006)
11. Ramage, A., Wathen, A.J.: Iterative solution techniques for the Stokes and Navier-Stokes equations. Int. J. Numer. Methods Fluids **19**, 67–83 (1994)
12. Rozložník, M., Strakoš, Z.: Variants of the residual minimizing Krylov subspace methods. In: Marek, I. (ed.) Proceedings of the 11th Summer School Software and Analysis of Numerical Mathematics, pp. 208–225. Pilsen, University of West Bohemia (1995)
13. Saad, Y., Schultz, M.H.: GMRES: a generalized minimal residual algorithm for solving non-symmetric linear systems. SIAM J. Sci. Stat. Comput. **7**(3), 856–869 (1986)
14. Varga, R.S.: Matrix Iterative Analysis, 2nd edn. Springer, Berlin (2000)
15. Vinsome, P.K.W.: Orthomin, an iterative method for solving sparse sets of simultaneous linear equations. In: Proceedings of the Fourth Symposium on Reservoir Simulation, pp. 149–159. Society of Petroleum Engineers of the American Institute of Mining, Metallurgical, and Petroleum Engineers (1976)
16. Walker, H.F., Zhou, L.: A simpler GMRES. Numer. Linear Algebra Appl. **1**(6), 571–581 (1994)
17. Young, D.M., Kang, J.C.: Generalized conjugate-gradient acceleration of nonsymmetrizable iterative methods. Linear Algebra Appl. **34**, 159–194 (1980)