ELSEVIER

Applied Numerical Mathematics 19 (1995) 279-286

APPLIED NUMERICAL MATHEMATICS

Residual smoothing and peak/plateau behavior in Krylov subspace methods

Homer F. Walker *

Department of Mathematics and Statistics, Utah State University, Logan, UT 84322-3900, USA

Abstract

Recent results on residual smoothing are reviewed, and it is observed that certain of these are equivalent to results obtained by different means that relate "peaks" and "plateaus" in residual norm sequences produced by certain pairs of Krylov subspace methods.

Keywords: Residual smoothing techniques; Iterative linear algebra methods; Krylov subspace methods; Orthogonal and minimal residual methods; Conjugate gradient and cqnjugate residual methods; SYMMLQ and MINRES methods; Arnoldi and GMRES methods; ORTHORES and ORTHODIR methods; Biconjugate gradient (BCG) and conjugate gradient squared (CGS) methods; Bi-CGSTAB methods; Quasi-minimal residual (QMR) methods

1. Introduction

Let $\{x_k\}$ be a sequence of approximate solutions of a linear system $Ax = b$, $A \in \mathbb{R}^{n\times n}$, and let $\{r_k \equiv b - Ax_k\}$ be the associated sequence of residuals. We consider the following general *residual smoothing* technique:

$$y_0 = x_0, \quad s_0 = r_0,$$
$$\left.\begin{array}{l} y_k = y_{k-1} + \eta_k(x_k - y_{k-1}) \\ s_k = s_{k-1} + \eta_k(r_k - s_{k-1}) \end{array}\right\} \quad k = 1, 2, \ldots. \tag{1.1}$$

The name derives from the fact that if the original residual sequence $\{r_k\}$ is irregularly behaved, then the parameters η_k can be chosen to produce $\{y_k\}$ with a more "smoothly" behaved residual sequence $\{s_k = b - Ay_k\}$. More general forms of residual smoothing are considered by Brezinski and Redivo-Zaglia [1], but (1.1) will suffice here.

* This work was supported in part by United States Air Force Office of Scientific Research Grant AFOSR-91-0294, United States Department of Energy Grants DE-FG02-92ER25136 and DE-FG03-94ER25221, and National Science Foundation Grant DMS-9400217, all with Utah State University.

SSDI 0168-9274(95)00087-9

A *Krylov subspace method* for solving $Ax = b$ begins with some x_0 and, at the kth step, determines an iterate $x_k = x_0 + z_k$ through a correction z_k in the kth *Krylov subspace*

$$\mathcal{K}_k \equiv \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}.$$

The purpose of this brief exposition is to point out the equivalence of certain results for residual smoothing techniques (1.1) to observations that relate "peaks" and "plateaus" in residual norm sequences produced by certain pairs of Krylov subspace methods. By a "peak" we mean a marked increase in the residual norms followed by a decrease; by a "plateau" we mean a period of very slow decrease or stagnation. Some results quoted here can be phrased more generally; we have tried to stick to the most basic and important cases for the sake of brevity and clarity. Also, our interest is in theoretical properties, and we do not consider issues having to do with finite-precision arithmetic.

2. Minimal residual smoothing and pairs of orthogonal and minimal residual methods

A natural choice of η_k in (1.1) is that for which $\|s_k\|_2$ is minimal. This η_k is easily characterized by

$$\eta_k \equiv \arg\min_{\eta} \|s_{k-1} + \eta(r_k - s_{k-1})\|_2 = -\frac{s_{k-1}^{\mathrm{T}}(r_k - s_{k-1})}{\|r_k - s_{k-1}\|_2^2}. \tag{2.1}$$

The resulting smoothing technique, [1] which we refer to as *minimal residual smoothing*, was introduced by Schönauer [14] and studied extensively by Weiss [17]. Note that it clearly gives $\|s_k\|_2 \leqslant \|s_{k-1}\|_2$ and $\|s_k\|_2 \leqslant \|r_k\|_2$ for each k. We summarize as follows:

Algorithm MRS. Minimal residual smoothing [14,17].
INITIALIZE: SET $s_0 = r_0$ AND $y_0 = x_0$.
ITERATE: FOR $k = 1, 2, \ldots$, DO:
COMPUTE x_k AND r_k.
COMPUTE $\eta_k = -s_{k-1}^{\mathrm{T}}(r_k - s_{k-1})/\|r_k - s_{k-1}\|_2^2$.
SET $s_k = s_{k-1} + \eta_k(r_k - s_{k-1})$ AND $y_k = y_{k-1} + \eta_k(x_k - y_{k-1})$.

When the residuals r_k are mutually orthogonal, i.e., when $r_i^{\mathrm{T}} r_j = 0$ for $i \neq j$, Weiss [17] has shown a number of important results for $\{s_k\}$ generated by Algorithm MRS (see also Weiss [18]). Slightly specialized or adapted versions of those most useful here are the following:

$$s_k = \frac{1}{\sum_{j=0}^{k} \frac{1}{\|r_j\|_2^2}} \sum_{i=0}^{k} \frac{1}{\|r_i\|_2^2} r_i, \tag{2.2}$$

$$\frac{1}{\|s_k\|_2^2} = \sum_{j=0}^{k} \frac{1}{\|r_j\|_2^2}, \tag{2.3}$$

[1] It has been noted in [20] that the straightforward implementation of (1.1) using η_k defined by (2.1) may suffer numerical inaccuracy; however, this is not important here. See [20] for an alternative formulation that may have better numerical performance.

$$\|s_k\|_2 = \min_{\eta_1,\dots,\eta_k} \left\| s_0 + \sum_{i=1}^{k} \eta_i (r_i - s_{i-1}) \right\|_2 . \tag{2.4}$$

Using the orthogonality of the residuals r_k, one can establish (2.2)-(2.4) by induction on k. Weiss [18] has termed (2.3) "Kirchoff's rule" because of its similarity to the rule of that name for determining resistance in parallel electrical circuits. We note for later use that (2.3) is equivalent to

$$\|r_k\|_2 = \frac{\|s_k\|_2}{\sqrt{1 - (\|s_k\|_2/\|s_{k-1}\|_2)^2}}. \tag{2.5}$$

These results have particularly important application to pairs of *orthogonal residual* and *minimal residual* Krylov subspace methods. Recall that, at the kth step, an orthogonal residual method produces a residual r_k^{OR} that is orthogonal to $\mathcal{K}_k$ (provided it is possible to do so) while a minimal residual method produces a residual r_k^{MR} that has minimal norm over all corrections in $\mathcal{K}_k$. It would be inappropriate to review all of the many implementations of these methods for special and general A, but we note several particularly well-known orthogonal/minimal residual method pairs: the conjugate gradient (Hestenes and Stiefel [9]) and conjugate residual (Stiefel [15]) methods for symmetric positive-definite A; SYMMLQ and MINRES (Paige and Saunders [11]) for symmetric indefinite A; ORTHORES and ORTHODIR (Young and Jea [19]) for general A; and, also for general A, the Arnoldi method (or full orthogonalization method) (Saad [12]) and the generalized minimal residual (GMRES) method (Saad and Schultz [13]).

Weiss [17] has shown that *applying minimal residual smoothing to an orthogonal residual method results in a minimal residual method.* Indeed, the residuals r_k^{OR} are clearly mutually orthogonal; furthermore, the vectors $r_j^{\mathrm{OR}} - r_{j-1}^{\mathrm{MR}}$, $j = 1, \dots, k$, can be seen to form a basis of $A(\mathcal{K}_k)$ for each k. Then the result follows inductively from (2.4) above. Specializing (2.2) and (2.3) to this case yields the following theorem.

Theorem 2.1 ([17]). *The residuals produced by orthogonal and minimal residual methods satisfy*

$$r_k^{\mathrm{MR}} = \frac{1}{\sum_{j=0}^{k} \frac{1}{\|r_j^{\mathrm{OR}}\|_2^2}} \sum_{i=0}^{k} \frac{1}{\|r_i^{\mathrm{OR}}\|_2^2} r_i^{\mathrm{OR}}, \tag{2.6}$$

$$\frac{1}{\|r_k^{\mathrm{MR}}\|_2^2} = \sum_{j=0}^{k} \frac{1}{\|r_j^{\mathrm{OR}}\|_2^2}. \tag{2.7}$$

Note that (2.6) expresses r_k^{MR} as a convex combination of $r_0^{\mathrm{OR}}, \dots, r_k^{\mathrm{OR}}$, in which each r_i^{OR} is weighted in proportion to the reciprocal of the square of its norm. Furthermore, rearranging (2.7) in the manner of (2.5) gives the following corollary.

Corollary 2.2. *The norms of the residuals produced by orthogonal and minimal residual methods satisfy*

$$\|r_k^{\mathrm{OR}}\|_2 = \frac{\|r_k^{\mathrm{MR}}\|_2}{\sqrt{1 - \left(\|r_k^{\mathrm{MR}}\|_2/\|r_{k-1}^{\mathrm{MR}}\|_2\right)^2}}. \tag{2.8}$$

This corollary is an extension to general orthogonal/minimal residual methods of a result of Brown [2, Theorem 5.1] relating the performance of the Arnoldi and GMRES methods. The result of [2, Theorem 5.1] is given in a somewhat different form but has been rephrased in essentially the form (2.8) by Cullum and Greenbaum [4, Theorem 3]. This result provides the basis of observations by Brown [2] and Cullum and Greenbaum [4,5] that, taken together, in effect correlate peaks in $\{\|r_k^{OR}\|_2\}$ with plateaus in $\{\|r_k^{MR}\|_2\}$.

Equations (2.7) and (2.8) are equivalent but give somewhat different perspectives on how $\{\|r_k^{MR}\|_2\}$ and $\{\|r_k^{OR}\|_2\}$ are related. Equation (2.7) makes plain the "global" dependence of $\|r_k^{MR}\|_2$ on $\|r_0^{OR}\|_2, \ldots, \|r_k^{OR}\|_2$; note that $\{\|r_k^{MR}\|_2\}$ is small if and only if at least one of $\|r_0^{OR}\|_2, \ldots, \|r_k^{OR}\|_2$ is small. Equation (2.8) brings out the "local" dependence of $\|r_k^{OR}\|_2$ on $\|r_k^{MR}\|_2$ and $\|r_{k-1}^{MR}\|_2$. We always have $\|r_k^{OR}\|_2 \geqslant \|r_k^{MR}\|_2$, of course, and (2.8) further shows that equality holds if and only if $\|r_k^{OR}\|_2 = \|r_k^{MR}\|_2 = 0$, i.e., both methods have reached the solution. Before the solution is reached, we have from (2.8) that the factor by which $\|r_k^{OR}\|_2$ exceeds $\|r_k^{MR}\|_2$ is determined by $\|r_k^{MR}\|_2/\|r_{k-1}^{MR}\|_2$, which measures the progress made by the minimal residual method at the current step. In particular, (2.8) shows clearly that, as observed by Brown [2] and Cullum and Greenbaum [4] for the Arnoldi and GMRES methods, if one method performs either very well or very poorly, then so does the other.

3. Quasi-minimal residual smoothing and certain pairs of Lanczos-based methods

Another kind of smoothing of the form (1.1), called *quasi-minimal residual smoothing*, has been considered by Zhou and Walker [20]. The name derives from the fact that when this smoothing is applied to the residuals and iterates produced by the biconjugate gradient (BCG) method (Lanczos [10], Fletcher [6]), then the resulting smoothed residuals and iterates are just those of the quasi-minimal residual (QMR) method [2] of Freund and Nachtigal [8]; see [20, §3]. We will discuss this relationship between BCG and QMR in more depth after formulating quasi-minimal residual smoothing and reviewing some of its general properties.

The general formulation is as follows: [3]

Algorithm QMRS. Quasi-minimal residual smoothing [20].

INITIALIZE: SET $s_0 = r_0$, $y_0 = x_0$, AND $\tau_0 = \|r_0\|_2$.
ITERATE: FOR $k = 1, 2, \ldots,$ DO:
COMPUTE x_k AND r_k.
DEFINE τ_k BY $1/\tau_k^2 = 1/\tau_{k-1}^2 + 1/\|r_k\|_2^2$.
SET $s_k = s_{k-1} + (\tau_k^2/\|r_k\|_2^2)(r_k - s_{k-1})$ AND $y_k = y_{k-1} + (\tau_k^2/\|r_k\|_2^2)(x_k - y_{k-1})$.

The following results for $\{s_k\}$ and $\{\tau_k\}$ produced by Algorithm QMRS are easily shown by induction [20, §3]:

$$s_k = \frac{1}{\sum_{j=0}^{k} \frac{1}{\|r_j\|_2^2}} \sum_{i=0}^{k} \frac{1}{\|r_i\|_2^2} r_i, \tag{3.1}$$

[2] Here, we mean the basic QMR method *without* look-ahead.

[3] It is noted in [20] that this formulation may have numerical weaknesses similar to those of Algorithm MRS. See [20] for an alternative formulation that avoids these weaknesses.

$$\frac{1}{\tau_k^2} = \sum_{j=0}^{k} \frac{1}{\|r_j\|_2^2}. \tag{3.2}$$

Note that (3.1) is the same as (2.2), although now the residuals r_k are not mutually orthogonal, and that (3.2) is an analogue of (2.3), in which τ_k plays the role of a *quasi-residual norm*. Also, (3.2) is equivalent to

$$\|r_k\|_2 = \frac{\tau_k}{\sqrt{1 - (\tau_k/\tau_{k-1})^2}}, \tag{3.3}$$

which is an analogue of (2.5).

From (3.1) and (3.2), one obtains ([20, §3])

$$\|s_k\|_2 \leqslant \sqrt{k+1}\,\tau_k = \sqrt{\frac{1}{\frac{1}{k+1}\sum_{j=0}^{k} \frac{1}{\|r_j\|_2^2}}}. \tag{3.4}$$

Thus $\|s_k\|_2$ is bounded by the square root of the harmonic mean of $\|r_0\|_2^2, \ldots, \|r_k\|_2^2$.

To specialize these results to BCG and QMR, let $\{r_k^{\mathrm{BCG}}\}$ and $\{r_k^{\mathrm{QMR}}\}$ denote the residual sequences generated by those methods and suppose that Algorithm QMRS is applied with $\{r_k\} = \{r_k^{\mathrm{BCG}}\}$. By the results of Zhou and Walker [20, §3] cited earlier, we have $\{s_k\} = \{r_k^{\mathrm{QMR}}\}$. It is further shown in [20, §3] that

$$\tau_k = \sqrt{\frac{1}{\sum_{j=0}^{k} \frac{1}{\|r_j^{\mathrm{BCG}}\|_2^2}}} = \min_{y \in \mathbb{R}^k} \big\| \|r_0\|_2 e_1 - H_k y \big\|_2, \tag{3.5}$$

in which $e_1 = (1, 0, \ldots, 0)^{\mathrm{T}} \in \mathbb{R}^{k+1}$ and $H_k \in \mathbb{R}^{(k+1)\times k}$ is the tridiagonal matrix satisfying $AV_k = V_{k+1}H_k$, where, for each k, $V_k \in \mathbb{R}^{n\times k}$ is the matrix the columns of which are the first k normalized Lanczos vectors. Then (3.4) becomes

$$\|r_k^{\mathrm{QMR}}\|_2 \leqslant \sqrt{k+1}\,\tau_k = \sqrt{\frac{1}{\frac{1}{k+1}\sum_{j=0}^{k} \frac{1}{\|r_j^{\mathrm{BCG}}\|_2^2}}} \tag{3.6}$$

for τ_k given by (3.5). The bound $\|r_k^{\mathrm{QMR}}\|_2 \leqslant \sqrt{k+1}\,\tau_k$ has been given by Freund and Nachtigal [8, Proposition 4.1] and is used as a preliminary termination criterion in QMR. From (3.1), (3.2), and (3.5), we have the following analogue of Theorem 2.1.

Theorem 3.1 ([20]). *The residuals produced by BCG and QMR satisfy*

$$r_k^{\mathrm{QMR}} = \frac{1}{\sum_{j=0}^{k} \frac{1}{\|r_j^{\mathrm{BCG}}\|_2^2}} \sum_{i=0}^{k} \frac{1}{\|r_i^{\mathrm{BCG}}\|_2^2} r_i^{\mathrm{BCG}}, \tag{3.7}$$

$$\frac{1}{\tau_k^2} = \sum_{j=0}^{k} \frac{1}{\|r_j^{\mathrm{BCG}}\|_2^2}, \tag{3.8}$$

where τ_k satisfies (3.5).

In analogy with (2.6), (3.7) shows that $r_k^{\rm QMR}$ is a convex combination of $r_0^{\rm BCG}, \ldots, r_k^{\rm BCG}$, in which each $r_i^{\rm BCG}$ is weighted in proportion to the reciprocal of the square of its norm. This gives considerable insight into why the QMR residual norms tend to decrease fairly smoothly, if not monotonically. Indeed, if $\|r_k^{\rm BCG}\|_2$ is small for some k, then $r_k^{\rm BCG}$ is given large weight in (3.7) and $\|r_k^{\rm QMR}\|_2$ is small. If subsequent BCG residual norms increase, then the effect of this increase is mollified by relatively small weights in (3.7) and any increase in $\|r_k^{\rm QMR}\|_2$ is correspondingly small.

Rearranging (3.8) in the manner of (3.3) gives the following corollary:

Corollary 3.2. *The norms of the BCG residuals satisfy*

$$\|r_k^{\rm BCG}\|_2 = \frac{\tau_k}{\sqrt{1-(\tau_k/\tau_{k-1})^2}}, \tag{3.9}$$

where τ_k satisfies (3.5).

This result has been given by Cullum and Greenbaum [4, Theorem 6] using closely related results of Freund and Nachtigal [8]. It is used in [4] to correlate peaks in $\{\|r_k^{\rm BCG}\|_2\}$ with plateaus in the quasi-residual norm sequence $\{\tau_k\}$.

Although the quasi-residual norms are not the true QMR residual norms (we have only the bound (3.6) in general), it has been experimentally observed by Cullum and Greenbaum [4] that the quasi-residual norms tend to be good approximations of the true QMR residual norms. Thus (3.8) and (3.9) indicate approximate relationships

$$\frac{1}{\|r_k^{\rm QMR}\|_2^2} \approx \sum_{j=0}^{k} \frac{1}{\|r_j^{\rm BCG}\|_2^2}, \tag{3.10}$$

$$\|r_k^{\rm BCG}\|_2 \approx \frac{\|r_k^{\rm QMR}\|_2}{\sqrt{1-\left(\|r_k^{\rm QMR}\|_2/\|r_{k-1}^{\rm QMR}\|_2\right)^2}} \tag{3.11}$$

that suggest correlations of peaks in $\{\|r_k^{\rm BCG}\|_2\}$ with plateaus in $\{\|r_k^{\rm QMR}\|_2\}$. Such correlations are clearly seen in experiments by Cullum and Greenbaum [4] and Zhou and Walker [20].

Equations (3.8)–(3.9) and approximations (3.10)–(3.11) give perspectives on the relationship of $\{\|r_k^{\rm BCG}\|_2\}$ to $\{\tau_k\}$ and $\{\|r_k^{\rm QMR}\|_2\}$ similar to those given by (2.7)–(2.8) on the relationship of $\{\|r_k^{\rm OR}\|_2\}$ to $\{\|r_k^{\rm MR}\|_2\}$; see the last paragraph of Section 2.

Algorithm QMRS can be used to derive a QMR-type method from any given method. Subsequently, (3.1)–(3.3) and the bound (3.4) can be used to draw relationships between the residuals of the given method and the residuals and quasi-residual norms of the resulting QMR-type method. In particular, these relationships can be used to correlate peaks in the residual norms of the given method with plateaus in the residual and quasi-residual norms of the QMR-type method as in the case of BCG and QMR above.

We mention two examples from Zhou and Walker [20, §3] to which these observations are relevant. In these examples, application of Algorithm QMRS to certain methods results in QMR-type methods that have been introduced elsewhere, and the above observations apply. First, the TFQMR method of Freund [7] can be obtained by applying Algorithm QMRS to the residuals, iterates, and certain intermediate quantitites generated by CGS. Second, the QMRCGSTAB method of Chan et al. [3] can be similarly obtained by applying Algorithm QMRS to the residuals, iterates, and certain other quantitites generated by the Bi-CGSTAB method of Van der Vorst [16].

4. Minimal residual smoothing for nonorthogonal residuals

We conclude by briefly revisiting minimal residual smoothing. Weiss [18, Theorem 2] has observed that for arbitrary (not necessarily mutually orthogonal) $\{r_k\}$, the bound (3.4), with the quasi-residual norm τ_k given by (3.2), holds for $\{s_k\}$ generated by Algorithm MRS as well as by Algorithm QMRS. (This springs directly from the fact that s_k generated by Algorithm MRS satisfies $\|s_k\|_2 \leqslant \min_{i=0,\dots,k} \|r_i\|_2$; see [18, Theorem 2].) Furthermore, since the quasi-residual norms satisfy (3.3), we have the same correlation of peaks in $\{\|r_k\|_2\}$ with plateaus in $\{\tau_k\}$ as in the case of Algorithm QMRS, which suggests a correlation of peaks in $\{\|r_k\|_2\}$ with plateaus in $\{\|s_k\|_2\}$ as in that case. In fact, it is seen in a number of experiments reported by Zhou and Walker [20] that Algorithms MRS and QMRS give fairly similar results when applied to the residuals and iterates produced by BCG and CGS. In particular, peaks in $\{\|r_k\|_2\}$ are strongly correlated with plateaus in $\{\|s_k\|_2\}$ when either Algorithm MRS or Algorithm QMRS is applied to BCG or CGS.

References

[1] C. Brezinski and M. Redivo-Zaglia, Hybrid procedures for solving systems of linear equations, *Numer. Math.* 67 (1994) 1-19.

[2] P.N. Brown, A theoretical comparison of the Arnoldi and GMRES algorithms, *SIAM J. Sci. Statist. Comput.* 20 (1992) 58-78.

[3] T.F. Chan, E. Gallopoulos, V. Simoncini, T. Szeto and C.H. Tong, A quasi-minimal residual variant of the Bi-CGSTAB algorithm for nonsymmetric systems, *SIAM J. Sci. Comput.* 15 (1994) 338-347.

[4] J.K. Cullum and A. Greenbaum, Residual relationships within three pairs of iterative algorithms for solving $Ax = b$, Tech. Rept. RC 18672, IBM T. J. Watson Research Center, Yorktown Heights, NY (1993); also: *SIAM J. Matrix Anal. Appl.* (to appear).

[5] J.K. Cullum and A. Greenbaum, Peaks, plateaus, numerical instabilities in Galerkin and minimal residual pairs of methods for solving $Ax = b$, Preprint (1994).

[6] R. Fletcher, Conjugate gradient methods for indefinite systems, in: *Numerical Analysis Dundee 1975*, Lecture Notes in Mathematics 506 (Springer, Berlin, 1976).

[7] R.W. Freund, A transpose-free quasi-minimal residual algorithm for non-Hermitian linear systems, *SIAM J. Sci. Comput.* 14 (1993) 470-482.

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

[9] M. Hestenes and E. Stiefel, Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bur. Stand.* 49 (1952) 409-436.

[10] C. Lanczos, Solution of systems of linear equations by minimized iterations, *J. Res. Nat. Bur. Stand.* 49 (1952) 33-53.

[11] C.C. Paige and M.A. Saunders, Solution of sparse indefinite systems of linear equations, *SIAM J. Numer. Anal.* 12 (1975) 617-624.

[12] Y. Saad, Krylov subspace methods for solving large unsymmetric linear systems, *Math. Comput.* 37 (1981) 105-126.

[13] Y. Saad and M.H. Schultz, GMRES: a generalized minimal residual method for solving nonsymmetric linear systems, *SIAM J. Sci. Statist. Comput.* 7 (1986) 856-869.

[14] W. Schönauer, *Scientific Computing on Vector Computers* (North-Holland, Amsterdam, 1987).

[15] E. Stiefel, Relaxationsmethoden bester Strategie zur Lösung linearer Gleichungssysteme, *Comment. Math. Helv.* 29 (1955) 157–179.
[16] H.A. van der Vorst, Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems, *SIAM J. Sci. Statist. Comput.* 13 (1992) 631–644.
[17] R. Weiss, Convergence behavior of generalized conjugate gradient methods, Ph.D. Thesis, University of Karlsruhe, Karlsruhe (1990).
[18] R. Weiss, Relations between smoothing and QMR, Tech. Rept., Universität Karlsruhe, Numerikforschung für Supercomputer, Karlsruhe (1993); also: *Numer Math.* (to appear).
[19] D.M. Young and K.C. Jea, Generalized conjugate gradient acceleration of nonsymmetrizable iterative methods, *Linear Algebra Appl.* 34 (1980) 159–194.
[20] L. Zhou and H.F. Walker, Residual smoothing techniques for iterative methods, *SIAM J. Sci. Comput.* 15 (1994) 297–312.