STOPPING CRITERIA FOR THE ITERATIVE SOLUTION OF LINEAR LEAST SQUARES PROBLEMS

X.-W. CHANG[†], C. C. PAIGE[†], AND D. TITLEY-PELOQUIN[†]

Abstract. We explain an interesting property of minimum residual iterative methods for the solution of the linear least squares (LS) problem. Our analysis demonstrates that the stopping criteria commonly used with these methods can in some situations be too conservative, causing any chosen method to perform too many iterations or even fail to detect that an acceptable iterate has been obtained. We propose a less conservative criterion to determine if a given iterate is an acceptable LS solution. This is merely a sufficient condition, but it approaches a necessary condition in the limit as the given iterate approaches the exact LS solution. We also propose a necessary and sufficient condition to determine if a given approximate LS solution is an acceptable LS solution, based on recent results on backward perturbation analysis of the LS problem. Although both of the above new conditions use quantities that are too expensive to compute in practical situations, we suggest potential approaches for estimating some of these quantities efficiently. We illustrate our results with several numerical examples.

Key words. Linear least squares, iterative methods, large sparse matrix problems, stopping criteria, backward perturbation analysis.

AMS subject classifications. 65F10, 65F20, 65F50, 65G50.

1. Introduction.

Given $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, the linear least squares (LS) problem is

$$\min_x \|b - Ax\|_2. \tag{1.1}$$

See for example [2] and [6] for useful background. We assume throughout that A has full column rank. Under this assumption, $\hat{x}$ is the unique solution of (1.1) if and only if $A^T(b - A\hat{x}) = 0$.

In this paper we discuss stopping criteria for the iterative solution of large sparse LS problems. To make the exposition concrete we concentrate on the widely used algorithm LSQR of Paige and Saunders [12, 13]. Note, however, that much of our discussion, and our stopping criteria, are applicable to other iterative methods for the solution of large sparse LS problems. (The practical implementations of these stopping criteria will of course vary from method to method.)

In Section 2 we define what we mean by an *acceptable LS solution* and a *backward stable LS solution.* In Section 3 we briefly summarize algorithm LSQR and state the stopping criteria originally proposed for LSQR in [12, §5]. These are based on sufficient (but not necessary) conditions for a given iterate to be an acceptable LS solution. Note that LSQR's stopping criteria are also frequently used with other iterative methods for the solution of large sparse LS problems; see for example [4, §2.4 & §3.3].

In Section 4 we analyze these stopping criteria. We explain an interesting property of minimum residual iterative methods and use this to show that LSQR's stopping criteria can in some situations be too conservative. The use of these criteria can cause any chosen iterative method to perform too many iterations, or in the worst case, to fail to detect that an acceptable iterate has been obtained.

[†] School of Computer Science, McGill University, 3480 University Street, McConnell Engineering Building, Room 318, Montreal, Quebec, Canada, H3A 2A7, {chang,paige,dtitle}@cs.mcgill.ca. The research of the first author was supported by NSERC of Canada Grant RGPIN217191-07. The research of the second author was supported by NSERC of Canada Grant RGPIN9236. The research of the third author was supported by NSERC of Canada PGS-D Fellowship.

In Section 5 and Section 6 we propose two conditions to determine if a given iterate $x_k \in \mathbb{R}^n$ is an acceptable LS solution. The condition in Section 5 is merely sufficient, but it approaches a necessary condition in the limit as x_k approaches the exact LS solution $\hat{x}$. The condition in Section 6 is both necessary and sufficient. To our knowledge, the latter new result is the only known such condition.

We give some numerical examples in Section 7 in which we compare these conditions with LSQR's stopping criteria. Section 8 contains our discussion and conclusions.

1.1. Notation.

We generally use upper-case letters for matrices, lower-case Roman letters for vectors and indices, and lower-case Greek letters for scalars. e_j denotes the j-th column of the unit matrix I. The *true* LS solution of (1.1) is denoted $\hat{x}$ with $\hat{r} \equiv b - A\hat{x}$, whereas x_k is used for the k-th iterate of an algorithm (often LSQR here) with $r_k \equiv b - Ax_k$. For vectors, $\|\cdot\|$ denotes the 2-norm. For matrices we use $\|\cdot\|_2$ and $\|\cdot\|_F$ for the 2- and Frobenius norm, respectively, while $\|\cdot\|_{2,F}$ denotes the use of either (consistently within an expression). We use $\mathcal{R}(A)$ to denote the range of A, and P_A and $P_A^\perp$ for the orthogonal projectors onto $\mathcal{R}(A)$ and the orthogonal complement of $\mathcal{R}(A)$, respectively. Assuming that A has full column rank, its Moore-Penrose generalized inverse is given by $A^\dagger = (A^TA)^{-1}A^T$, and thus for a nonzero vector v, $v^\dagger = v^T/\|v\|^2$. Finally $\kappa_{2,F}(A) \equiv \|A\|_{2,F}\|A^\dagger\|_{2,F}$.

For the reader's convenience we give a reference table of the important quantities used in the stopping criteria discussed in this paper, together with the first equation number where each appears, and an indication of its use, or what it is:

$\xi_{2,F}(x_k,\alpha,\beta)$	(2.5)	(used in testing) for acceptable LS solutions
$\eta_{2,F}(x_k,\alpha,\beta)$	(3.3)	for acceptable nearly compatible system solutions
$\psi_{2,F}(x_k,\alpha,\beta)$	(5.1)	an asymptotically tight upper bound on $\xi_{2,F}(x_k,\alpha,\beta)$
$\omega(x_k,\theta)$	(6.1)	the minimal backward error for a compatible system
$\mu(x_k,\theta)$	(6.3)	the minimal backward error for a LS problem.

2. Acceptable and backward stable least squares solutions.

Most practical LS problems contain uncertainties in the data, and instead of solving (1.1) with ideal data A and b we can at best solve a nearby problem

$$\min_x \|(b+f)-(A+E)x\|, \tag{2.1}$$

where E and f are small in some sense. Commonly E and f have small norms relative to the norms of A and b, and we thus only consider the case

$$\|E\|_{2,F} \le \alpha\|A\|_{2,F} \quad \text{and} \quad \|f\| \le \beta\|b\| \tag{2.2}$$

for some α and β satisfying $0 \le \alpha, \beta \ll 1$ (where we hope that estimates of α and β are known). In practical applications it is often the case that $\alpha \ll \beta$, because b is often a vector of measurements that is subject to much larger measurement errors than the matrix A.

We say that an iterate $x_k \in \mathbb{R}^n$ is an *acceptable LS solution* when it is the exact LS solution of a problem within the accepted range of relative errors in the data. In other words for any given α and β, an iterate x_k is an acceptable LS solution if and only if there exist perturbations E and f satisfying

$$(A+E)^T[b+f-(A+E)x_k] = 0, \quad \|E\|_{2,F} \le \alpha\|A\|_{2,F}, \quad \|f\| \le \beta\|b\|. \tag{2.3}$$

Obviously this is the case if and only if

$$\xi_{2,F}(x_k, \alpha, \beta) \leq 1, \tag{2.4}$$

where

$$\begin{aligned}\xi_{2,F}(x_k, \alpha, \beta) \equiv \min_{E,f,\xi} \Big\{ \xi : (A+E)^T[b+f-(A+E)x_k] = 0, \\ \|E\|_{2,F} \leq \xi\alpha\|A\|_{2,F},\ \|f\| \leq \xi\beta\|b\| \Big\}.\end{aligned} \tag{2.5}$$

To summarize, for any chosen α and β, an iterate x_k is an acceptable LS solution if and only if it satisfies (2.4).

Even if we have perfect data A and b, we cannot expect to solve (1.1) exactly in floating-point arithmetic. The best we can generally hope to do is to solve a problem of the form (2.1) with $\alpha = \mathcal{O}(u)$ and $\beta = \mathcal{O}(u)$ in (2.2), u being the machine's unit roundoff. We say that an iterate x_k is a *backward stable LS solution* when it satisfies (2.4) with $\alpha = \mathcal{O}(u)$ and $\beta = \mathcal{O}(u)$. A backward stable LS solution is thus simply an acceptable LS solution with a specific choice of α and β.

Note that for any scalar $\tau > 0$, we can verify that $\xi_{2,F}(x_k, \tau, \tau) \cdot \tau = \xi_{2,F}(x_k, 1, 1)$ by using (2.5) to define $\xi_{2,F}(x_k, \tau, \tau) \cdot \tau$, and then replacing each quantity $\xi\tau$ in the resulting right-hand side by the new variable $\tilde{\xi}$, giving $\xi_{2,F}(x_k, 1, 1)$. Thus

$$\xi_{2,F}(x_k, 1, 1) \leq \tau \quad \Leftrightarrow \quad \xi_{2,F}(x_k, \tau, \tau) \leq 1. \tag{2.6}$$

A backward stable iterate x_k therefore satisfies $\xi_{2,F}(x_k, 1, 1) = \mathcal{O}(u)$. Following the nomenclature in [9, §7.1 & §20.7], we call the quantity $\xi_{2,F}(x_k, 1, 1)$ the optimal *normwise relative backward error for LS problems.*

If $\xi_{2,F}(x_k, \alpha, \beta) \leq 1$, many known upper bounds on the relative error $\|\hat{x}-x_k\|/\|\hat{x}\|$ exist as a function of α and β; see for example [9, §20.1] and [6, §5.3.7]. Therefore, in most practical applications (but not necessarily ill-posed problems) we can be satisfied with a given $x_k \in \mathbb{R}^n$ as an approximate solution to (1.1) when it satisfies $\xi_{2,F}(x_k, \alpha, \beta) \leq 1$ with an appropriate choice of α and β.

Unfortunately, finding an analytical expression for $\xi_{2,F}(x_k, \alpha, \beta)$ in (2.5) remains an open question. Some easily computable upper bounds on $\xi_{2,F}(x_k, \alpha, \beta)$ are known, and these can be used to give sufficient conditions for $\xi_{2,F}(x_k, \alpha, \beta) \leq 1$. Such conditions are commonly used as stopping criteria for the iterative solution of large sparse LS problems; see for example [12, §5], [4, §2.4 & §3.3] and [2, p.309]. We outline some of these in the next section.

3. Algorithm LSQR and its stopping criteria.

In this section we give a brief overview of algorithm LSQR [12, 13] and its stopping criteria. The bidiagonalization "Bidiag 2" in [12, §3] is that given by Golub and Kahan in [5, (2.4)], but with the initial vector $q_1 = A^Tb/\|A^Tb\|$. For LS solutions it is preferable to use the variant "Bidiag 1" in [12, §3], which in theory after k steps produces matrices $U_{k+1} \in \mathbb{R}^{m\times(k+1)}$ and $V_k \in \mathbb{R}^{n\times k}$ such that $U_{k+1}(e_1\beta_1) = b$ where $\beta_1 \equiv \|b\|$, and

$$AV_k = U_{k+1}B_{k+1,k}, \qquad A^TU_{k+1} = V_kB_{k+1,k}^T + v_{k+1}\alpha_{k+1}e_{k+1}^T = V_{k+1}B_{k+1}^T,$$

$$B_{k+1,k} \equiv \begin{bmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \beta_3 & \ddots & \\ & & \ddots & \alpha_k \\ & & & \beta_{k+1} \end{bmatrix} \in \mathbb{R}^{(k+1)\times k}, \quad B_{k+1} \equiv [B_{k+1,k}|e_{k+1}\alpha_{k+1}].$$

In theory U_{k+1} and V_k both have orthonormal columns, but in practice rounding errors cause a loss of orthogonality.

In the k-th step we find the minimum residual approximate solution of the form $x_k = V_k y_k$, where $y_k \in \mathbb{R}^k$ solves

$$\min_y \|b - AV_k y\| = \min_y \|U_{k+1}(e_1\beta_1 - B_{k+1,k}y)\| = \min_y \|e_1\beta_1 - B_{k+1,k}y\|. \tag{3.1}$$

Thus $x_k = V_k y_k$ and $r_k \equiv b - Ax_k$ (for $k = 1, 2, \dots$) are successive approximations to the true solution $\hat{x}$ and residual $\hat{r}$. The bidiagonal LS problem in (3.1) can easily be solved for y_k using the QR factorization of $B_{k+1,k}$. A careful implementation [13] of LSQR requires only two matrix-vector products per iteration and stores only the latest columns of U_{k+1} and V_k.

Listed below are the stopping criteria used in LSQR. Given an iterate x_k with corresponding residual $r_k \equiv b - Ax_k$, the algorithm stops if one of the following three conditions is satisfied:

$$\left\{\begin{array}{ll} 1.\ \|r_k\| \le \alpha\|A\|_{2,F}\|x_k\| + \beta\|b\| & \text{(a test for compatible systems)} \\ 2.\ \|A^T r_k\|/\|r_k\| \le \alpha\|A\|_{2,F} & \text{(a criterion for LS problems)} \\ 3.\ \kappa_{2,F}(A) \ge \gamma & \text{(a regularizing criterion).} \end{array}\right. \tag{3.2}$$

The parameters α and β (distinct from the elements α_k and β_k of $B_{k+1,k}$) are set according to the accuracy of the data; see (2.2). From now on we assume the sensible case that $0 < \alpha, \beta \ll 1$. If rough estimates of these relative errors are not known, α and β could be set to a small multiple of the unit roundoff u. The parameter γ is the maximum condition number we are willing to tolerate. (In LSQR, $\kappa_F(B_{k+1,k})$, which is no greater than $\kappa_F(A)$, is checked against γ.) Note that criterion 2 in (3.2) assumes $r_k \neq 0$. If the residual is zero then x_k is clearly the exact solution $\hat{x}$ of (1.1).

LSQR provides the user with cheap estimates of $\|r_k\|$ and $\|A^T r_k\|$ at each iteration; see for example [12, §5]. Cheap lower bounds on $\|A\|_F$ and $\kappa_F(A)$ are also available, and lower bounds on $\|A\|_2$ and $\kappa_2(A)$ can be computed reasonably cheaply at each iteration. These lower bounds are monotonically increasing with k and are thus successively better approximations to $\|A\|_{2,F}$ and $\kappa_{2,F}(A)$. Similar estimates are also usually available in other iterative methods for the solution of LS problems, such as CGLS (see for example [2, §7]). When using such estimates, one should keep in mind that they are not always accurate and can give misleading results. We demonstrate this with an example in Section 4.3.

We now show that LSQR's stopping criteria 1 and 2 correspond to particular upper bounds on $\xi_{2,F}(x_k, \alpha, \beta)$ in (2.5). LSQR's stopping criteria 1 and 2 thus give sufficient but not necessary conditions for x_k to be an acceptable LS solution.

Criterion 1 in (3.2) can be obtained by tightening the normal equations constraint for LS in (2.5) to an equality constraint for compatible systems. Since this cannot decrease the resulting minimum value of ξ, it follows that

$$\begin{aligned} &\xi_{2,F}(x_k, \alpha, \beta) \le \eta_{2,F}(x_k, \alpha, \beta) \\ &\qquad \equiv \min_{E,f,\eta} \Big\{\eta :\ (A+E)x_k = b + f,\ \|E\|_{2,F} \le \eta\alpha\|A\|_{2,F},\ \|f\| \le \eta\beta\|b\|\Big\}. \end{aligned} \tag{3.3}$$

Rigal and Gaches [14] showed that

$$\eta_{2,F}(x_k, \alpha, \beta) = \frac{\|r_k\|}{\alpha\|A\|_{2,F}\|x_k\| + \beta\|b\|}. \tag{3.4}$$

They proposed the condition $\eta_{2,F}(x_k,\alpha,\beta) \le 1$, in other words

$$\|r_k\| \le \alpha \|A\|_{2,F}\|x_k\| + \beta\|b\|, \tag{3.5}$$

as a stopping criterion for the iterative solution of *compatible* linear systems. Note that for any scalar $\tau > 0$,

$$\eta_{2,F}(x_k,\tau,\tau) \le 1 \quad \Leftrightarrow \quad \eta_{2,F}(x_k,1,1) \le \tau. \tag{3.6}$$

The quantity

$$\eta_{2,F}(x_k,1,1) \equiv \frac{\|r_k\|}{\|A\|_{2,F}\|x_k\| + \|b\|}$$

is known in the literature as the optimal *normwise relative backward error* for compatible systems; see [9, Theorem 7.1].

Since the LS method is often used for solving nearly compatible overdetermined linear systems, the condition (3.5) can be used as a stopping criterion for the iterative solution of LS problems, hence criterion 1 in (3.2). From (3.3) and (3.4), if (3.5) holds then $\xi_{2,F}(x_k,\alpha,\beta) \le 1$ and the iterate x_k is an acceptable LS solution. Similarly if $\eta_{2,F}(x_k,1,1) = \mathcal{O}(u)$ then $\xi_{2,F}(x_k,1,1) = \mathcal{O}(u)$ and x_k is a backward stable LS solution; see (2.6).

Criterion 2 in (3.2) can be obtained from the fact that any feasible perturbations E and f in (2.5) must lead to an upper bound on the minimum $\xi_{2,F}(x_k,\alpha,\beta)$. Stewart [16, §3] observed that the perturbations $E_0 = -r_k r_k^{\dagger} A$ and $f_0 = 0$ satisfy the normal equations constraint in (2.5). With $E = E_0$ and $f = f_0$, (2.5) gives

$$\xi_{2,F}(x_k,\alpha,\beta) \le \min_{\xi} \left\{ \xi : \|E_0\|_{2,F} \le \xi\alpha\|A\|_{2,F} \right\} = \frac{\|E_0\|_{2,F}}{\alpha\|A\|_{2,F}} = \frac{\|A^T r_k\|}{\alpha\|A\|_{2,F}\|r_k\|}. \tag{3.7}$$

Therefore if criterion 2 in (3.2) is satisfied, then $\xi_{2,F}(x_k,\alpha,\beta) \le 1$, and from (2.4) x_k is an acceptable LS solution. Furthermore if $\|A^T r_k\|/\|r_k\| = \mathcal{O}(u)\|A\|_{2,F}$ then $\xi_{2,F}(x_k,1,1) = \mathcal{O}(u)$, and x_k is a backward stable LS solution—see the line following (2.6). Note that $A^T r_k$ is the residual vector of the normal equations at x_k, and the quantity $\|A^T r_k\|/\|r_k\|$ is the norm of a backward perturbation matrix in A only.

Criterion 3 in (3.2) tells us to stop if our "reduced representation" (3.1) of the problem becomes too ill-conditioned; it is an attempt to regularize ill-conditioned problems. As the focus of this paper is not on regularization of ill-conditioned problems, we will not discuss criterion 3 further.

In the next section we give new insights into the behavior of the quantities $\|r_k\|$ and $\|A^T r_k\|/\|r_k\|$ that are used in LSQR's stopping criteria 1 and 2 in (3.2).

4. Analysis of LSQR's stopping criteria 1 and 2.

4.1. An interesting observation. On all problems we have tested we have made the following observation:

LSQR first reduces the residual norm $\|r_k\|$ while $\|A^T r_k\|/\|r_k\|$ remains roughly constant or tends to oscillate in ill-conditioned problems. The residual norm $\|r_k\|$ then reaches a plateau (after which it remains almost constant) at which point the quantity $\|A^T r_k\|/\|r_k\|$ starts to decrease, and decreases until it too reaches a plateau.

This surprising behavior, for which we propose an explanation below, is clearly illustrated in Figure 4.1 (repeated in more detail as "Test Problem 3" in Figure 7.1).

The oscillation of $\|A^T r_k\|$ has often been observed in various iterative methods for the solution of large sparse LS problems. For example it is stated in [12, §6.2] that in practice in LSQR "$\|A^T r_k\|/\|r_k\|$ can vary rather dramatically with k, but it does tend to stabilize for large k". Björck [2, p.289] states that in CGLS "$\|A^T r_k\|$ will often exhibit large oscillations when $\kappa(A)$ is large". In [4] Choi uses MINRES (see [11, §6]) and a new variant thereof to solve singular symmetric linear systems. Since these have either no solution or infinitely many solutions, they can be solved as (possibly compatible, rank-deficient) LS problems. Like LSQR but unlike MINRES, algorithm MINRES-QLP given in [4] converges to the minimum 2-norm LS solution. It is remarked [4, p.27] that $\|A^T r_k\|$ "is often observed to be oscillating".

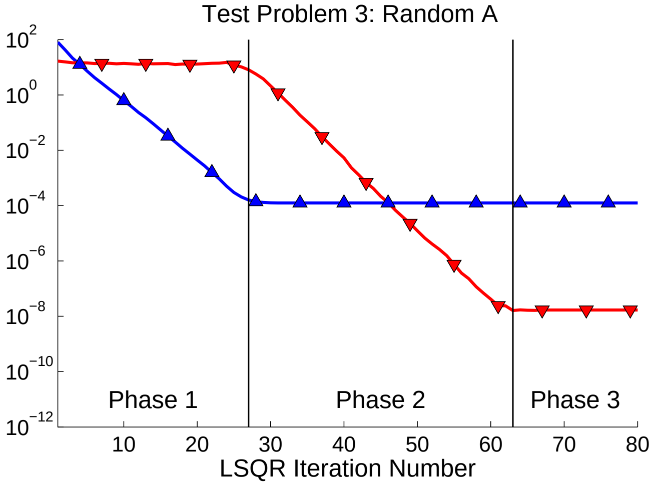

FIG. 4.1. *Behavior of* $\|r_k\|$ (▲) *and* $\|A^T r_k\|/\|r_k\|$ (▼).

4.2. A possible explanation. Assume that $m \times n$ A has full column rank and let the singular value decomposition (SVD) of A be

$$A = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^T = U_1 \Sigma V^T = \sum_{i=1}^{n} u_i \sigma_i v_i^T$$

where $U \equiv \begin{bmatrix} U_1 & U_2 \end{bmatrix} \equiv [u_1, \ldots, u_m] \in \mathbb{R}^{m\times m}$ and $V \equiv [v_1, \ldots, v_n] \in \mathbb{R}^{n\times n}$ (distinct from U_k and V_k in Section 3) are orthogonal matrices and $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \cdots \geq \sigma_n > 0$. With this notation, the orthogonal projectors onto the range of A and onto the orthogonal complement of the range of A are, respectively,

$$P_A = U_1 U_1^T \quad \text{and} \quad P_A^\perp = U_2 U_2^T.$$

If $\hat{x}$ is the true LS solution with corresponding residual $\hat{r} \equiv b - A\hat{x}$, then

$$P_A^\perp r_k = U_2 U_2^T (b - Ax_k) = U_2 U_2^T b = P_A^\perp b = \hat{r} \tag{4.1}$$

for all k, so that

$$\|r_k\|^2 = \|P_A r_k\|^2 + \|P_A^\perp r_k\|^2 = \|P_A r_k\|^2 + \|\hat{r}\|^2. \tag{4.2}$$

Note that in theory LSQR decreases $\|r_k\|$ every step (see [12, (5.2) & p.50]) so it also decreases $\|P_A r_k\|$ every step. Furthermore, because V_n in Section 3 is theoretically orthogonal, in theory $\|P_A r_k\| = 0$ when $k = n$ (and possibly even for some $k < n$). Thus in LSQR $\|P_A r_k\|$ converges strictly monotonically to 0. We need the following lemma.

Lemma 4.1. *Given $A \in \mathbb{R}^{m\times n}$ with the above-defined SVD, $b \in \mathbb{R}^m$ and $x_k \in \mathbb{R}^n$, define $r_k \equiv b - Ax_k$. Then*

$$\|A^T r_k\| = \bar{\sigma}_k \|P_A r_k\| \tag{4.3}$$

for some $\bar{\sigma}_k$ in the closed interval $[\sigma_n, \sigma_1]$.

Proof. Using the SVD of A,

$$\|A^T r_k\|^2 = r_k^T U_1 \Sigma^2 U_1^T r_k = \sum_{i=1}^{n} (u_i^T r_k)^2 \sigma_i^2 = \bar{\sigma}_k^2 \sum_{i=1}^{n} (u_i^T r_k)^2$$

for some $\bar{\sigma}_k \in [\sigma_n, \sigma_1]$. Now because

$$\|P_A r_k\|^2 = r_k^T U_1 U_1^T r_k = \sum_{i=1}^{n} (u_i^T r_k)^2,$$

it immediately follows that $\|A^T r_k\| = \bar{\sigma}_k \|P_A r_k\|$. □

In well-conditioned problems the singular values of A are all very roughly of similar orders of magnitude. Therefore in well-conditioned problems the order of magnitude of $\bar{\sigma}_k$ is very roughly constant as a function of k. In ill-conditioned problems, $\bar{\sigma}_k$ can oscillate wildly but always lies between the extreme singular values of A. Note that the behavior of $\bar{\sigma}_k$ as a function of k depends on the size of the residual norms and on how the residuals are aligned with respect to the left singular vectors u_1 to u_n.

We can now describe what appears to be the main basis for the interesting observation. We do so by dividing the LSQR iteration process into three phases, illustrated in Figure 4.1, any of which need not exist.

Phase 1. This phase is defined by those iterates for which in (4.2)

$$\|P_A r_k\| > \|P_A^{\perp} r_k\| = \|\hat{r}\|, \tag{4.4}$$

and so from (4.2)

$$\|r_k\|^2 = \|P_A r_k\|^2 + \|\hat{r}\|^2 \approx \|P_A r_k\|^2. \tag{4.5}$$

Thus $\|r_k\| \approx \|P_A r_k\|$ and LSQR decreases $\|r_k\|$, while from (4.5) and Lemma 4.1

$$\frac{\|A^T r_k\|}{\|r_k\|} = \frac{\bar{\sigma}_k \|P_A r_k\|}{\|r_k\|} \approx \bar{\sigma}_k, \tag{4.6}$$

which must lie between the extreme singular values of A. Thus in this phase Stewart's $\|A^T r_k\|/\|r_k\|$ is roughly constant in well-conditioned problems and can oscillate in ill-conditioned problems.

The sum of squares in (4.5) makes (4.6) a particularly good approximation even when $\|P_A r_k\|$ is not very much larger than $\|P_A^{\perp} r_k\| = \|\hat{r}\|$. For example if $\|P_A r_k\| = 2\|\hat{r}\|$ we get $\|r_k\| = (\sqrt{5}/2)\|P_A r_k\|$, leading to a relative error of only $(\|r_k\| - \|P_A r_k\|)/\|P_A r_k\| \approx 12\%$. If $\|P_A r_k\| = 10\|\hat{r}\|$, the relative error becomes approximately 0.5%.

Usually the iteration starts in phase 1, because usually $\|r_0\| \gg \|\hat{r}\|$ and thus $\|P_A r_0\| \gg \|\hat{r}\|$; see (4.5) with $k = 0$. However it may happen that $\|P_A r_0\| \le \|\hat{r}\|$, for example if x_0 is a very good approximation to $\hat{x}$. In this case there is no phase 1 and the iteration starts in either phase 2 or phase 3 below.

Phase 2. First suppose that the linear system is not compatible to machine precision. We consider the compatible case afterwards.

As LSQR decreases $\|P_A r_k\|$, there is a first k such that

$$\|P_A r_k\| \leq \|P_A^{\perp} r_k\| = \|\hat{r}\|; \tag{4.7}$$

see (4.1) and (4.2). This is the start of phase 2, in which the residuals are now dominated by their projection onto the orthogonal complement of $\mathcal{R}(A)$. In this phase $\|P_A r_k\|$ continues to decrease but $\|r_k\|$ hardly decreases because it is dominated by $\|P_A^{\perp} r_k\| = \|\hat{r}\|$, which is constant.

Because $\|P_A r_k\|$ still decreases while $\|r_k\|$ remains roughly constant, from (4.3) the quantity $\|A^T r_k\|/\|r_k\| = \bar{\sigma}_k \|P_A r_k\|/\|r_k\|$ tends to decrease. Thus in phase 2 it is $\|r_k\|$ that remains nearly constant, while $\|A^T r_k\|/\|r_k\|$ tends to decrease.

As LSQR continues to decrease $\|P_A r_k\|$ (recall that $\|P_A r_k\| \to 0$ in theory) there is a first k such that

$$\|P_A r_k\| = \mathcal{O}(u)(\|A\|_{2,F}\|x_k\| + \|b\|). \tag{4.8}$$

This implies that x_k is a backward stable LS solution; see Section 5 for a detailed explanation.

Phase 3. This phase begins when LSQR has decreased $\|P_A r_k\|$ to the level in (4.8). In all our numerical experiments we have found that $\|P_A r_k\|$ does not decrease below this level, even though in theory $\|P_A r_k\| \to 0$. This is true even for compatible systems where in theory $\|r_k\| \to 0$. The numerical examples in Section 7 illustrate this behavior.

Now suppose that the linear system is compatible to the level of machine precision, meaning that the true LS solution $\hat{x}$ satisfies $\eta_{2,F}(\hat{x}, 1, 1) = \mathcal{O}(u)$; see (3.3) and (3.6). In this case, in apparently all but the most extreme circumstances, LSQR converges to an iterate x_k that also solves $Ax = b$ to the level of machine precision. In other words, for these problems there is usually a k such that $\eta_{2,F}(x_k, 1, 1) = \mathcal{O}(u)$ and thus

$$\|r_k\| = \mathcal{O}(u)(\|A\|_{2,F}\|x_k\| + \|b\|); \tag{4.9}$$

see (3.4) and (3.6). In this case there is effectively no phase 2 (and phase 3 starts immediately after phase 1) because if (4.9) is satisfied then clearly so is (4.8).

In phase 3 neither $\|r_k\|$ nor $\|A^T r_k\|/\|r_k\|$ decreases further. If the system is not compatible to machine precision, then as in phase 2 $\|r_k\|$ is dominated by $\|P_A^{\perp} r_k\| = \|\hat{r}\|$, which is constant. If the linear system is compatible to machine precision then $\|r_k\|$ satisfies (4.9) and is therefore roughly constant. Thus we have

$$\|r_k\| \approx \max\big\{\|\hat{r}\|, \mathcal{O}(u)(\|A\|_{2,F}\|x_k\| + \|b\|)\big\}, \tag{4.10}$$

and from (4.3) and (4.8) we obtain

$$\frac{\|A^T r_k\|}{\|r_k\|} = \frac{\bar{\sigma}_k \|P_A r_k\|}{\|r_k\|} = \bar{\sigma}_k \frac{\mathcal{O}(u)(\|A\|_{2,F}\|x_k\| + \|b\|)}{\|r_k\|}. \tag{4.11}$$

In phase 3 the ratio $\bar{\sigma}_k$ usually remains almost constant (even for ill-conditioned problems; see the convergence plots in Figure 7.2). This is due to r_k being nearly constant in this phase, so from Lemma 4.1 $\bar{\sigma}_k$ is also almost constant. Therefore $\|A^T r_k\|/\|r_k\|$ also remains nearly constant.

Summary. Initially $\|A^T r_k\|/\|r_k\| \approx \bar{\sigma}_k \in [\sigma_n, \sigma_1]$ is roughly constant in well-conditioned problems and oscillates between the extreme singular values of A in ill-conditioned problems. This quantity only starts to decrease once $\|r_k\|$ is no longer dominated by $\|P_A r_k\|$, which happens when $\|r_k\|$ reaches the plateau in (4.10). Eventually $\|A^T r_k\|/\|r_k\|$ also reaches the plateau in (4.11) where it too remains almost constant.

4.3. Relation to the stopping criteria 1 and 2. We can relate the above observations to the stopping criteria 1 and 2 in (3.2), which give sufficient but not necessary conditions for x_k to be an acceptable LS solution. Since the theoretically strictly monotonically decreasing $\|r_k\|$ no longer decreases significantly after phase 1 and plateaus at the level given in (4.10), LSQR's criterion 1 may never be triggered in significant-residual problems (for which the maximum in (4.10) is given by $\|\hat{r}\|$). On the other hand, because $\|A^T r_k\|/\|r_k\|$ reaches a plateau (4.11) at the start of phase 3, LSQR's criterion 2 may never be triggered in nearly compatible systems (for which (4.9) holds in phase 3 and (4.11) is thus roughly of the order of $\bar{\sigma}_k$). So both stopping criteria are needed—but even then there can be difficulties.

We illustrate this with an example in Figure 4.2 (repeated with different detail as "Test Problem 3" in Figure 7.1). Suppose we would like to obtain a backward stable LS solution, so we set $\alpha = \beta = u$ in (2.5) and (3.2). It is easy to see from the plotted tolerances in Figure 4.2 that neither stopping criterion in (3.2) is ever triggered, regardless of the number of iterations performed. This demonstrates that LSQR's stopping criteria can be much too conservative, and can lead a user to incorrect conclusions about whether or not LSQR (or any other algorithm) has converged to a required tolerance.

In the next section we give a new tighter estimate of $\xi_{2,F}(x_k, \alpha, \beta)$. This result indicates that in the above example in fact $\xi_2(x_k, 1, 1) = \mathcal{O}(u)$ (see (2.6)) when $k = 63$. In other words a backward stable iterate (in the 2-norm) is obtained at the end of phase 2.

We note that although the quantity $\|A^T r_k\|/\|r_k\|$ plateaus in what we have called phase 3, in practice LSQR's *approximation* to $\|A^T r_k\|/\|r_k\|$ generally does not, as shown in Figure 4.2. Therefore stopping criterion 2 in (3.2) may be triggered if we use LSQR's approximation in our computation, even though the actual $\|A^T r_k\|/\|r_k\| \gg \alpha\|A\|_{2,F}$. In the above example this leads to stopping criterion 2 (in the 2-norm with $\alpha = u$) being triggered at the iteration $k = 83$, late in phase 3 and well after a backward stable iterate has actually been obtained.

5. A new upper bound on $\xi_{2,F}(x_k, \alpha, \beta)$.

We now give a new upper bound on $\xi_{2,F}(x_k, \alpha, \beta)$ defined in (2.5), and show that it becomes asymptotically tight in the limit as x_k approaches the true LS solution $\hat{x}$ of (1.1). This bound could be used to improve LSQR's stopping criteria significantly.

THEOREM 5.1. *Given $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and $x_k \in \mathbb{R}^n$, define $r_k \equiv b - Ax_k$ and use the definition of $\xi_{2,F}(x_k, \alpha, \beta)$ in (2.5). Then*

$$\xi_{2,F}(x_k, \alpha, \beta) \le \psi_{2,F}(x_k, \alpha, \beta) \equiv \frac{\|P_A r_k\|}{\alpha\|A\|_{2,F}\|x_k\| + \beta\|b\|}. \tag{5.1}$$

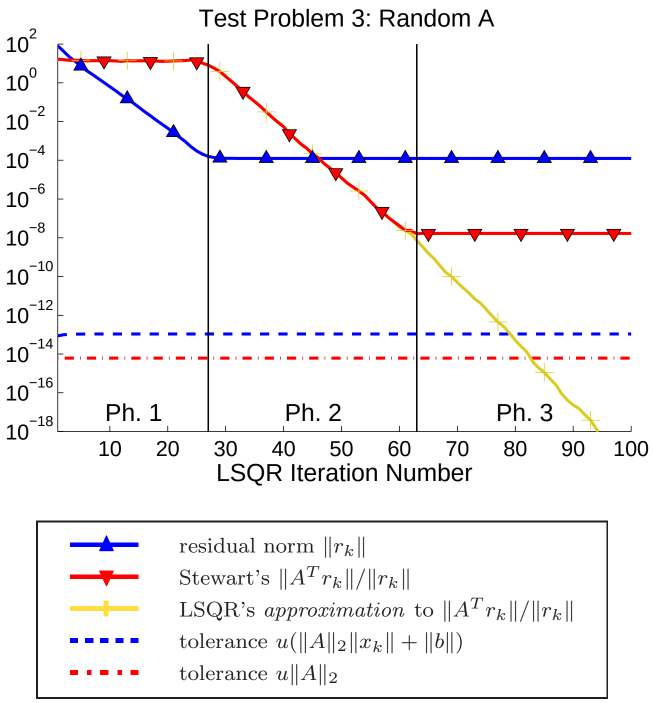

FIG. 4.2. *Behavior of* $\|r_k\|$ *and* $\|A^T r_k\|/\|r_k\|$.

Proof. Consider the perturbations

$$E^* \equiv \frac{\alpha\|A\|_{2,F}\|x_k\|}{\alpha\|A\|_{2,F}\|x_k\| + \beta\|b\|} P_A r_k x_k^{\dagger},$$

$$f^* \equiv -\frac{\beta\|b\|}{\alpha\|A\|_{2,F}\|x_k\| + \beta\|b\|} P_A r_k,$$

so that

$$\frac{\|E^*\|_{2,F}}{\alpha\|A\|_{2,F}} = \frac{\|f^*\|}{\beta\|b\|} = \frac{\|P_A r_k\|}{\alpha\|A\|_{2,F}\|x_k\| + \beta\|b\|} = \psi_{2,F}(x_k, \alpha, \beta). \tag{5.2}$$

Also notice that $b + f^* - (A + E^*)x_k = r_k - P_A r_k = P_A^{\perp} r_k$. From this we see that E^* and f^* satisfy the normal equations constraint in (2.5). It then follows from (2.5) and (5.2) that

$$\xi_{2,F}(x_k, \alpha, \beta) \le \min_{\xi} \left\{ \xi : \ \|E^*\|_{2,F} \le \xi\alpha\|A\|_{2,F}, \ \|f^*\| \le \xi\beta\|b\| \right\} = \psi_{2,F}(x_k, \alpha, \beta),$$

so that (5.1) holds. $\square$

Recall that an iterate x_k is an acceptable LS solution if and only if it satisfies $\xi_{2,F}(x_k, \alpha, \beta) \le 1$ (see (2.5)). As a result of Theorem 5.1, if

$$\boxed{\text{LStest1:} \qquad \psi_{2,F}(x_k, \alpha, \beta) \le 1} \tag{5.3}$$

or equivalently

$$\|P_A r_k\| \le \alpha\|A\|_{2,F}\|x_k\| + \beta\|b\|, \tag{5.4}$$

then $\xi_{2,F}(x_k, \alpha, \beta) \leq 1$ and x_k is an acceptable LS solution. Comparing the above to LSQR's stopping criterion 1 in (3.2), we immediately see that (5.4) gives a less pessimistic criterion. In fact the upper bound $\psi_{2,F}$ in (5.1) can be much tighter than $\eta_{2,F}$ in (3.3) (see also (3.4)) especially during what we have called phases 2 and 3 when $\|P_A r_k\|$ is no longer the main component of $\|r_k\|$.

We can use Theorem 5.1 to explain why $\|P_A r_k\|$ levels off at the end of phase 2, as illustrated in Figure 7.1 and noted after (4.8). If (4.8) holds, then from (5.1) we have $\xi_{2,F}(x_k, 1, 1) = \mathcal{O}(u)$ and thus from (2.6) x_k is a backward stable LS solution. As discussed in Section 2, this is the best we can generally hope to achieve in floating-point arithmetic. Therefore, we cannot generally expect that $\|P_A r_k\|$ will decrease below the level given in (4.8).

We now show that our new upper bound $\psi_{2,F}(x_k, \alpha, \beta)$ in (5.1) becomes asymptotically tight with $\xi_{2,F}(x_k, \alpha, \beta)$ in (2.5) in the limit as x_k approaches the true LS solution $\hat{x}$. Note that in theory as $x_k \to \hat{x}$, $\psi_{2,F}(x_k, \alpha, \beta) \to 0$ and so $\xi_{2,F}(x_k, \alpha, \beta) \to 0$; see (5.1). The following theorem shows that both converge at the same rate.

Theorem 5.2. *Using the notation of Theorem 5.1 and letting $\hat{x}$ denote the true LS solution of (1.1),*

$$\lim_{x_k \to \hat{x}} \frac{\xi_{2,F}(x_k, \alpha, \beta)}{\psi_{2,F}(x_k, \alpha, \beta)} = 1. \tag{5.5}$$

Proof. We have shown in Theorem 5.1 that $\xi_{2,F}(x_k, \alpha, \beta) \leq \psi_{2,F}(x_k, \alpha, \beta)$ for all $x_k \in \mathbb{R}^n$. Therefore

$$\lim_{x_k \to \hat{x}} \frac{\xi_{2,F}(x_k, \alpha, \beta)}{\psi_{2,F}(x_k, \alpha, \beta)} \leq 1.$$

On the other hand notice that the optimal perturbations $\hat{E}_k$ and $\hat{f}_k$ in (2.5) must satisfy

$$(A + \hat{E}_k)^T (r_k + \hat{f}_k - \hat{E}_k x_k) = 0,$$

so that $P_{A+\hat{E}_k}(r_k + \hat{f}_k - \hat{E}_k x_k) = 0$ and thus $P_{A+\hat{E}_k} r_k = P_{A+\hat{E}_k}(\hat{E}_k x_k - \hat{f}_k)$. Using the fact that $\|P_{A+\hat{E}_k}\| \leq 1$ along with the other constraints in (2.5), it follows that the optimal $\xi_{2,F}(x_k, \alpha, \beta)$ in (2.5) must satisfy

$$\|P_{A+\hat{E}_k} r_k\| \leq \|\hat{E}_k\|_{2,F} \|x_k\| + \|\hat{f}_k\| \leq \xi_{2,F}(x_k, \alpha, \beta) \cdot (\alpha \|A\|_{2,F} \|x_k\| + \beta \|b\|),$$

so that with (5.1)

$$\xi_{2,F}(x_k, \alpha, \beta) \geq \frac{\|P_{A+\hat{E}_k} r_k\|}{\alpha \|A\|_{2,F} \|x_k\| + \beta \|b\|} = \psi_{2,F}(x_k, \alpha, \beta) \frac{\|P_{A+\hat{E}_k} r_k\|}{\|P_A r_k\|}.$$

In the limit as $x_k \to \hat{x}$ we have $\|\hat{E}_k\|_{2,F} \to 0$ and thus $P_{A+\hat{E}_k} r_k \to P_A r_k$; see for example [15, §3 & 4]. Therefore

$$\lim_{x_k \to \hat{x}} \frac{\xi_{2,F}(x_k, \alpha, \beta)}{\psi_{2,F}(x_k, \alpha, \beta)} \geq 1.$$

□

REMARK 5.1. *From the above proof we can easily observe that if we impose the constraint $\mathcal{R}(A+E) = \mathcal{R}(A)$ on E in (2.5), then $\xi_{2,F}(x_k, \alpha, \beta) = \psi_{2,F}(x_k, \alpha, \beta)$. This constraint may make sense in some situations.*

Stewart [15, §5] proved a result similar to but weaker than Theorem 5.1. Using $\hat{r} \equiv b - A\hat{x}$ to denote the true LS residual, he observed that the perturbations $E_1 = (r_k - \hat{r})x_k^\dagger$ and $f_1 = 0$ satisfy the normal equations constraint in (2.5), so that

$$\xi_{2,F}(x_k, \alpha, \beta) \leq \min_{\xi} \left\{ \xi : \ \|E_1\|_{2,F} \leq \xi\alpha\|A\|_{2,F} \right\} = \frac{\|E_1\|_{2,F}}{\alpha\|A\|_{2,F}} = \frac{\|r_k - \hat{r}\|}{\alpha\|A\|_{2,F}\|x_k\|}. \quad (5.6)$$

Since $r_k = P_A r_k + P_A^\perp r_k = P_A r_k + \hat{r}$, the bound in (5.6) is equivalent to

$$\xi_{2,F}(x_k, \alpha, \beta) \leq \frac{\|P_A r_k\|}{\alpha\|A\|_{2,F}\|x_k\|}.$$

By also considering perturbations in b, we obtain the new and tighter upper bound $\psi_{2,F}(x_k, \alpha, \beta)$ in (5.1), which is asymptotically tight with $\xi_{2,F}(x_k, \alpha, \beta)$ for all values of α and β.

Stewart noted [16, §3] that his bound in (5.6) could only be computed if the LS problem were contrived so that $\hat{r}$ was known *a priori*. Since this is generally not the case in practical applications, the bound in (5.6) cannot generally be used in practice. Of course in practical applications $\|P_A r_k\|$ is also not available *a priori* and is too expensive to compute directly. However by thinking of our new bound $\psi_{2,F}(x_k, \alpha, \beta)$ in (5.1) in terms of the projection $\|P_A r_k\|$ instead of the quantity $\|r_k - \hat{r}\|$, we can try to find new ways to estimate $\xi_{2,F}(x_k, \alpha, \beta)$ by estimating $\|P_A r_k\|$. The following bounds, for example, follow immediately from Lemma 4.1:

$$\frac{\|A^T r_k\|}{\sigma_1} \leq \|P_A r_k\| \leq \frac{\|A^T r_k\|}{\sigma_n}, \quad (5.7)$$

where σ_1 and σ_n are the largest and smallest singular value of A, respectively. Note that $\|A^T r_k\|$ is generally easily computable and estimates of the extreme singular values of A are available in LSQR.

For very well-conditioned problems, corresponding to $\sigma_1 \approx \sigma_n$, the bounds in (5.7) are fairly tight. In our numerical tests we found that even for ill-conditioned problems the lower bound in (5.7) is usually much tighter than the upper bound. In other words $\bar{\sigma}_k$ in Lemma 4.1 usually lies close to the largest singular value of A, especially late in phase 2 and in phase 3 of the iteration process. A better understanding of the behavior of $\bar{\sigma}_k$ with k could lead to an efficiently computable estimate of $\|P_A r_k\|$ in (5.4). We leave this for a future investigation.

Here we suggest another potential approach to estimating the projection $\|P_A r_k\|$ efficiently in LSQR. Using the fact that $P_A = AA^\dagger$ and $P_A A = A$, and letting $\hat{x} = A^\dagger b$ denote the true LS solution of (1.1), we obtain

$$\|P_A r_k\| = \|AA^\dagger(b - Ax_k)\| = \|A(\hat{x} - x_k)\| \equiv \|\hat{x} - x_k\|_{A^T A}.$$

The quantity $\|P_A r_k\|$ is therefore the so-called energy norm of the error at step k. Several estimates of the energy norm of the error have been proposed for the method of conjugate gradients; see for example the discussion in [17] and an extension to CGLS in [1]. We expect that it will be possible to use these ideas to estimate $\|P_A r_k\|$ efficiently in LSQR. We leave the details for future work.

6. A necessary and sufficient condition for $\xi_{2,F}(x_k,\alpha,\beta)\le 1$.

We now show how the groundbreaking theoretical results of Waldén, Karlson and Sun, namely [18, Theorem 2.2] and [18, Corollary 2.1], can be used to give a necessary and sufficient condition to determine if an iterate x_k is an acceptable LS solution.

The minimization problems in Lemma 6.1 are commonly called *normwise backward error* or *minimal backward error* problems; see [9, §7] and [18].

Lemma 6.1. *Given $A\in\mathbb{R}^{m\times n}$, $b\in\mathbb{R}^m$, $0\neq x_k\in\mathbb{R}^n$ and $\theta>0$, define $r_k\equiv b-Ax_k$. Then for compatible systems we have* ([9, p. 134, Exercise 7.8])

$$\omega(x_k,\theta)\equiv\min_{\Delta A,\Delta b}\{\|[\Delta A,\theta\Delta b]\|_F:(A+\Delta A)x_k=b+\Delta b\}=\frac{\theta\|r_k\|}{\sqrt{1+\theta^2\|x_k\|^2}},\tag{6.1}$$

$$\omega(x_k,\infty)\equiv\lim_{\theta\to\infty}\omega(x_k,\theta)=\min_{\Delta A}\{\|\Delta A\|_F:\ (A+\Delta A)x_k=b\}=\frac{\|r_k\|}{\|x_k\|}.\tag{6.2}$$

If we replace the above equality constraints by the LS normal equations, then with the above-defined $\omega(x_k,\theta)$ and $\omega(x_k,\infty)$ we have ([18]; see also [9, §20.7])

$$\begin{aligned}&\mu(x_k,\theta)\\&\quad\equiv\min_{\Delta A,\Delta b}\Big\{\|[\Delta A,\theta\Delta b]\|_F:\ (A+\Delta A)^T[(b+\Delta b)-(A+\Delta A)x_k]=0\Big\}\qquad(6.3)\\&\quad=\min\Big\{\omega(x_k,\theta),\ \sigma_{\min}\Big(\big[A,\omega(x_k,\theta)\cdot(I-r_kr_k^\dagger)\big]\Big)\Big\},\\&\mu(x_k,\infty)\equiv\lim_{\theta\to\infty}\mu(x_k,\theta)\\&\quad=\min_{\Delta A}\Big\{\|\Delta A\|_F:\ (A+\Delta A)^T[b-(A+\Delta A)x_k]=0\Big\}\qquad(6.4)\\&\quad=\min\Big\{\omega(x_k,\infty),\ \sigma_{\min}\Big(\big[A,\omega(x_k,\infty)\cdot(I-r_kr_k^\dagger)\big]\Big)\Big\},\end{aligned}$$

where $\sigma_{\min}(\cdot)$ denotes the smallest singular value.

The question is how to use $\mu(x_k,\theta)$ or $\mu(x_k,\infty)$ to determine if x_k is an acceptable LS solution as defined in (2.4) and (2.5). Let $\widehat{\Delta A}$ be the optimal perturbation in (6.4). If $\mu(x_k,\infty)\equiv\|\widehat{\Delta A}\|_F\le\alpha\|A\|_F$, then clearly $E=\widehat{\Delta A}$ and $f=0$ satisfy the constraints in (2.3) in the Frobenius norm and give $\xi_F(x_k,\alpha,\beta)\le 1$ in (2.4). On the other hand Gu [8, Theorem 3.1] showed that for any E, f, α and β such that

$$(A+E)^T[b+f-(A+E)x_k]=0,\quad \|E\|_{2,F}\le\alpha\|A\|_{2,F},\quad \|f\|\le\beta\|b\|,$$

so that $\xi_F(x_k,\alpha,\beta)\le 1$ (see (2.5)), there exists a ΔA satisfying the constraint in (6.4) with

$$\frac{\|\Delta A\|_{2,F}}{\|A\|_{2,F}}\le\alpha+2\beta\frac{1+\alpha}{1-2\beta},\tag{6.5}$$

ensuring that $\mu(x_k,\infty)\lesssim(\alpha+2\beta)\|A\|_F$ in (6.4) when $\alpha,\beta\ll 1$. So for any α and β satisfying $0\le\alpha,\beta\ll 1$ we have

$$\begin{aligned}&\mu(x_k,\infty)\le\alpha\|A\|_F\ \Rightarrow\ \xi_F(x_k,\alpha,\beta)\le 1;\\&\xi_F(x_k,\alpha,\beta)\le 1\ \Rightarrow\ \mu(x_k,\infty)\lesssim(\alpha+2\beta)\|A\|_F.\end{aligned}\tag{6.6}$$

From this we see that if $\beta\lesssim\alpha$ then checking the condition $\mu(x_k,\infty)\le\alpha\|A\|_F$ is a reliable way to determine if x_k is an acceptable solution in the Frobenius norm, and in particular for checking if it is a backward stable LS solution.

Unfortunately when $\alpha \ll \beta$ the condition $\mu(x_k, \infty) \leq \alpha\|A\|_F$ is only sufficient for x_k to be an acceptable solution. In other words when $\alpha \ll \beta$

$$\xi_F(x_k, \alpha, \beta) \leq 1 \;\; \not\Rightarrow \;\; \mu(x_k, \infty) \lesssim \alpha\|A\|_F$$

and in this case the criterion $\mu(x_k, \infty) \leq \alpha\|A\|_F$ might not detect that an acceptable iterate has been obtained until many unnecessary iterations have been performed, if it detects it at all. Recall from Section 2 that the case $\alpha \ll \beta$ often occurs in practical LS applications, because b is often a measurement vector that is subject to much larger measurement errors than the matrix A.

We now consider how to use $\mu(x_k, \theta)$ with an appropriate finite θ in (6.3) to determine if x_k is an acceptable LS solution for any choice of α and β in (2.5). The following lemma for compatible systems was proven in [3, §2].

Lemma 6.2. *Using the notation of Lemma 6.1, if $\beta\|b\| > 0$ then the choice*

$$\theta = \hat{\theta}_{2,F} \equiv \left(\frac{\alpha\|A\|_{2,F}}{\beta\|b\| \cdot \|x_k\|}\right)^{1/2}$$

makes the optimal ΔA and Δb in (6.1) equal to the optimal E and f in (3.3).

Note that the relationship between $\omega(x_k, \theta)$ in (6.1) and $\eta_{2,F}(x_k, \alpha, \beta)$ in (3.3) for compatible systems parallels that between $\mu(x_k, \theta)$ in (6.3) and $\xi_{2,F}(x_k, \alpha, \beta)$ in (2.5) for LS problems. We have not found an exact equivalence to Lemma 6.2 for the LS case, but we do have a very strong result, which we give in the following theorem.

Theorem 6.3. *Given full column rank $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and $0 \neq x_k \in \mathbb{R}^n$, define $r_k \equiv b - Ax_k$, $\xi_F(x_k, \alpha, \beta)$ as in (2.5) and $\mu(x_k, \theta)$ as in (6.3). Let*

$$\theta = \hat{\theta} \equiv \frac{\alpha\|A\|_F}{\beta\|b\|}.$$

Then

$$\xi_F(x_k, \alpha, \beta) \leq \frac{\mu(x_k, \hat{\theta})}{\alpha\|A\|_F} \leq \sqrt{2}\xi_F(x_k, \alpha, \beta). \tag{6.7}$$

Proof. Let $\widehat{\Delta A}$ and $\widehat{\Delta b}$ represent the optimal perturbations in (6.3) with $\theta = \hat{\theta}$. Clearly $E = \widehat{\Delta A}$ and $f = \widehat{\Delta b}$ cannot improve on the optimal $\xi_F(x_k, \alpha, \beta)$ in (2.5), and so either $\|\widehat{\Delta A}\|_F \geq \xi_F(x_k, \alpha, \beta)\alpha\|A\|_F$ or $\|\widehat{\Delta b}\| \geq \xi_F(x_k, \alpha, \beta)\beta\|b\|$, or both, giving

$$\max\left\{\frac{\|\widehat{\Delta A}\|_F}{\alpha\|A\|_F}, \frac{\|\widehat{\Delta b}\|}{\beta\|b\|}\right\} \geq \xi_F(x_k, \alpha, \beta).$$

Thus we have

$$\begin{aligned}\mu(x_k, \hat{\theta}) = \|[\widehat{\Delta A}, \hat{\theta}\widehat{\Delta b}]\|_F &= \sqrt{\|\widehat{\Delta A}\|_F^2 + \frac{\alpha^2\|A\|_F^2}{\beta^2\|b\|^2}\|\widehat{\Delta b}\|^2} \\ &\geq \max\left\{\|\widehat{\Delta A}\|_F, \frac{\alpha\|A\|_F}{\beta\|b\|}\|\widehat{\Delta b}\|\right\} \\ &\geq \xi_F(x_k, \alpha, \beta)\alpha\|A\|_F,\end{aligned}$$

proving the first inequality in (6.7).

On the other hand for the optimal $\hat{E}$ and $\hat{f}$ in (2.5) we have

$$\|\hat{E}\|_F \le \xi_F(x_k,\alpha,\beta)\alpha\|A\|_F, \quad \|\hat{f}\| \le \xi_F(x_k,\alpha,\beta)\beta\|b\|,$$

where $\hat{E}$ and $\hat{f}$ cannot improve on the optimal $\widehat{\Delta A}$ and $\widehat{\Delta b}$ in (6.3) with $\theta = \hat{\theta}$. Therefore

$$\begin{aligned}\mu(x_k,\hat{\theta}) &= \|[\widehat{\Delta A}, \hat{\theta}\widehat{\Delta b}]\|_F \le \|[\hat{E}, \hat{\theta}\hat{f}]\|_F \\ &\le \sqrt{\xi_F^2(x_k,\alpha,\beta)\alpha^2\|A\|_F^2 + \frac{\alpha^2\|A\|_F^2}{\beta^2\|b\|^2}\xi_F^2(x_k,\alpha,\beta)\beta^2\|b\|^2} \\ &= \sqrt{2}\xi_F(x_k,\alpha,\beta)\alpha\|A\|_F,\end{aligned}$$

leading to the second inequality in (6.7). □

Theorem 6.3 dealt with $\xi_F(x_k,\alpha,\beta)$ and $\mu(x_k,\hat{\theta})$ defined for the perturbed LS normal equations $(A{+}E)^T[b{+}f{-}(A{+}E)x_k] = 0$. But the same analysis can be applied to $\eta_F(x_k,\alpha,\beta)$ in (3.3) and $\omega(x_k,\hat{\theta})$ in (6.1), which were defined for the perturbed compatible system $(A+E)x_k = b+f$. Carrying out the argument in Theorem 6.3 with $\xi_F(x_k,\alpha,\beta)$ replaced by $\eta_F(x_k,\alpha,\beta)$, and $\mu(x_k,\hat{\theta})$ replaced by $\omega(x_k,\hat{\theta})$, gives

$$\eta_F(x_k,\alpha,\beta) \le \frac{\omega(x_k,\hat{\theta})}{\alpha\|A\|_F} \le \sqrt{2}\eta_F(x_k,\alpha,\beta).$$

Recall from Section 2 that x_k is an acceptable LS solution if and only if it satisfies $\xi_{2,F}(x_k,\alpha,\beta) \le 1$, but that at present no explicit formula is known for computing $\xi_{2,F}(x_k,\alpha,\beta)$. Theorem 6.3 implies that

$$\begin{aligned}&\mu(x_k,\hat{\theta}) \le \alpha\|A\|_F \;\Rightarrow\; \xi_F(x_k,\alpha,\beta) \le 1; \\ &\xi_F(x_k,\alpha,\beta) \le 1 \;\Rightarrow\; \mu(x_k,\hat{\theta}) \le \sqrt{2}\alpha\|A\|_F.\end{aligned} \tag{6.8}$$

(Compare this to (6.6) in which β is present in the second expression—a subtle but very important difference when $\alpha \ll \beta$.) The following is therefore a nearly optimal test for an acceptable Frobenius-norm LS solution for any choice of α and β:

$$\boxed{\text{LStest2:} \qquad \mu(x_k,\hat{\theta}) \le \alpha\|A\|_F, \quad \text{where } \hat{\theta} = \frac{\alpha\|A\|_F}{\beta\|b\|}.} \tag{6.9}$$

As a consequence of Theorem 6.3 we can now determine almost exactly when an iterate x_k is an acceptable LS solution by using the result of Waldén, Karlson and Sun in Lemma 6.1. Unfortunately Lemma 6.1 gives an expression for $\mu(x_k,\theta)$ that costs $\mathcal{O}(m^3)$ flops to compute, and is thus too expensive to be useful in large sparse applications. Nevertheless we can still use (6.9) to test the effectiveness of LSQR's stopping criteria 1 and 2 as well as our new condition LStest1 in (5.3). We can also try to develop more effective stopping criteria for the iterative solution of large sparse LS problems by finding bounds on or estimates for $\mu(x_k,\hat{\theta})$ instead of $\xi_{2,F}(x_k,\alpha,\beta)$.

A few estimates of $\mu(x_k,\infty)$ exist in the literature; see for example [7, §4] and the references therein. We point out that some of these can be generalized to estimate $\mu(x_k,\hat{\theta})$, but to our knowledge none of these estimates is *both* provably reliable *and* computable in $\mathcal{O}(mn)$ flops, so at present no estimate of $\mu(x_k,\hat{\theta})$ is truly suitable for use in large sparse applications. We leave the efficient and reliable estimation of $\mu(x_k,\hat{\theta})$ for a future investigation.

7. Illustrations.

To illustrate our results we run LSQR on various test problems and plot at each iteration:

(i) the residual norm $\|r_k\|$;
(ii) Stewart's $\|A^T r_k\|/\|r_k\|$;
(iii) the norm of the projections $\|P_A r_k\|$ and $\|P_A^\perp r_k\|$;
(iv) the ratio $\bar{\sigma}_k = \|A^T r_k\|/\|P_A r_k\|$ from (4.3).

The projections are computed by obtaining the QR factorization of A using Matlab's built-in function `qr`, and the quantities $\|r_k\|$, $\|A^T r_k\|$, $\|A\|_2$ and $\|A\|_F$ are computed explicitly. (In other words we do not use LSQR's approximation to the above quantities, because our goal here is to illustrate the actual convergence behavior, described in Section 4.1, of the true $\|r_k\|$ and $\|A^T r_k\|/\|r_k\|$.) We also show at which iteration phases 1 and 2 end with vertical lines indicating the first k for which (4.7) and (4.8) hold, respectively. Admittedly the $\mathcal{O}(u)$ term in (4.8) is somewhat vague; for our illustrations we plot the vertical line at the first k for which $\|P_A r_k\|$ settles near $u(\|A\|_2\|x_k\| + \|b\|)$. Also note that only one vertical line is plotted in test problems 1 and 4 because there is effectively no phase 2 or 1, respectively, in these problems.

We use test problems 1 to 4 to illustrate how the effectiveness of LSQR's stopping criteria 1 and 2 in (3.2) depends on the size of the true residual norm $\|\hat{r}\|$. For these simple test problems each element of the matrix $A \in \mathbb{R}^{300\times 120}$ is randomly chosen from a normal distribution with mean 0 and standard deviation 1, so that A is almost certainly well-conditioned. This exhibits the interesting observation discussed in Section 4.1, and supports our explanation beautifully. Let s_n represent an n-vector containing all ones and let each element of an m-vector t_m be randomly chosen from a normal distribution with mean 0 and standard deviation 1.

(i) In test problem 1, $b = As_{120} + 10^{-15}t_{300}$ and $\|\hat{r}\| \approx 10^{-14}$.
(ii) In test problem 2, $b = As_{120} + 10^{-10}t_{300}$ and $\|\hat{r}\| \approx 10^{-9}$.
(iii) In test problem 3, $b = As_{120} + 10^{-5}t_{300}$ and $\|\hat{r}\| \approx 10^{-4}$.
(iv) In test problem 4, $b = As_{120} + 10^{0}t_{300}$ and $\|\hat{r}\| \approx 10^{1}$.

Results for these test problems are illustrated in Figure 7.1.

We use test problems 5 to 8 to illustrate how increasing the condition number $\kappa_2(A)$ affects the behavior of $\|r_k\|$ and $\|A^T r_k\|/\|r_k\|$ in (3.2) and $\bar{\sigma}_k$ in (4.3). These problems are set up so that they all have roughly the same true LS residual $\|\hat{r}\| \approx 10^{-7}$, and are presented by increasing condition number, with $\kappa_2(A) \approx 7.1$, 4.5×10^3, 1.5×10^8 and 3.8×10^{12}, respectively. Although the focus of this paper is not on regularization of ill-conditioned problems, we include these examples to illustrate the numerical behavior of LSQR, and our observation from Section 4.1, on these increasingly ill-conditioned problems.

For these test problems 5 to 8 we create A and b as in [12, §8]. These problems are called $P(m, n, d, p)$. The matrix $A \in \mathbb{R}^{m\times n}$ has singular values

$$\sigma_i \equiv \Big(\lfloor (i-1+d)/d \rfloor \cdot d/n\Big)^p,$$

where integer division by d is used to obtain repeated singular values. The true solution $\hat{x}$ and residual $\hat{r}$ are fixed, after which b is set to $b = A\hat{x} + \hat{r}$. For all the details see [12, p.63]. Test problems 5 to 8 are $P(500, 200, 4, 1/2)$, $P(800, 200, 3, 2)$, $P(750, 300, 7, 5)$ and $P(400, 150, 6, 9)$, respectively. Results for these problems are illustrated in Figure 7.2.

For test problems 9 to 12 we use the large sparse sample problems Well1033, Well1850, Illc1033 and Illc1850 from the Matrix Market [10], respectively. The problems starting with "Well" denote well-conditioned problems, whereas those starting

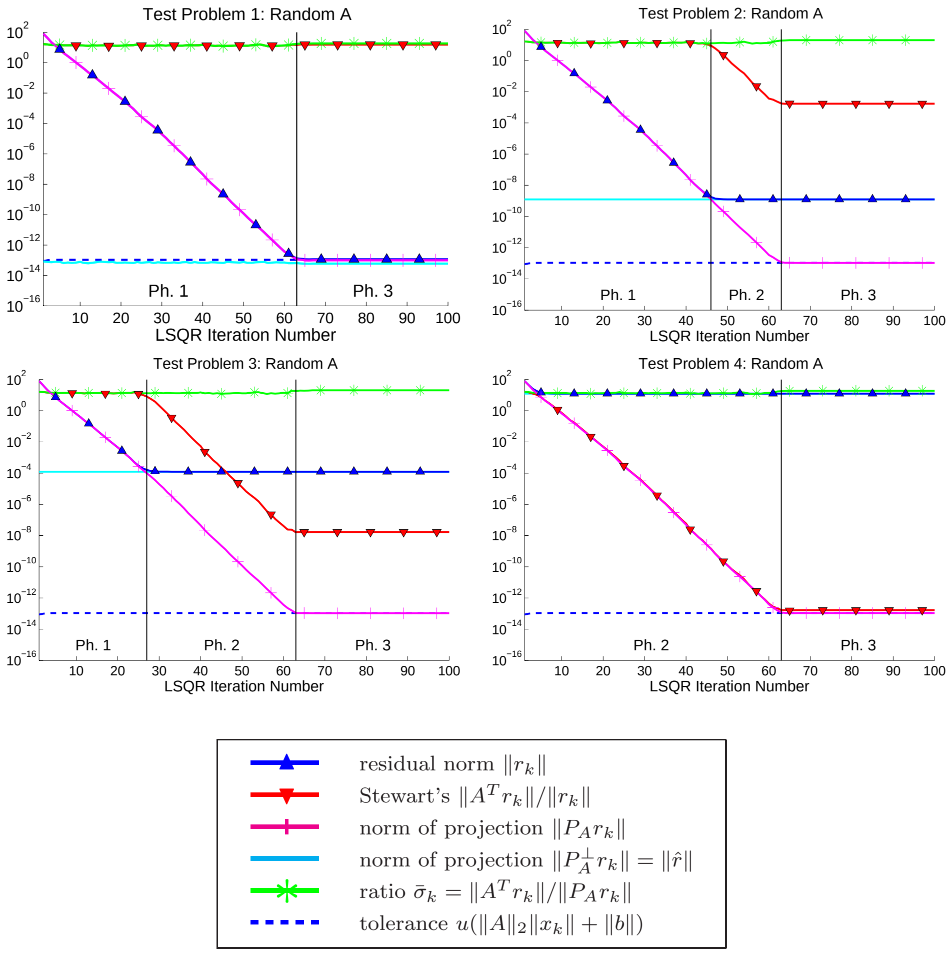

Fig. 7.1. *Test problems 1 to 4: well-conditioned examples with increasing true* $\|\hat{r}\|$.

with "Illc" denote ill-conditioned problems. The problems ending with the number 1033 involve a matrix $A \in \mathbb{R}^{1033\times 320}$, whereas those ending with 1850 involve a matrix $A \in \mathbb{R}^{1850\times 712}$. In all these test problems we create the vector b as follows: $b = A[1, 1, \ldots, 1]^T + 10^{-8}[m, m-1, \ldots, 1]^T$. Results for these problems are illustrated in Figure 7.3.

Each convergence plot in Figures 7.1 to 7.3 corresponds to one instance of a LS problem. Tables 7.1 and 7.2 below give the number of iterations required to trigger

(i) LSQR's stopping criteria 1 and 2 from (3.2), in the Frobenius norm;
(ii) the new condition LStest1 (5.3);
(iii) the new condition LStest2 (6.9);

for various values of α and β. The iteration counts in (i) above are given when the relevant norms in (3.2) are computed explicitly (True) and using LSQR's approximation (App.). Each row corresponds to an average number of iterations required, rounded to the nearest integer, for the same matrix A with 100 different noisy vectors $b = As_n + 10^{-p}t_m$ (where s_n and t_m are defined above). The symbol ∞ is used when a particular condition is *never* satisfied in *any* of the 100 tests, regardless of the number

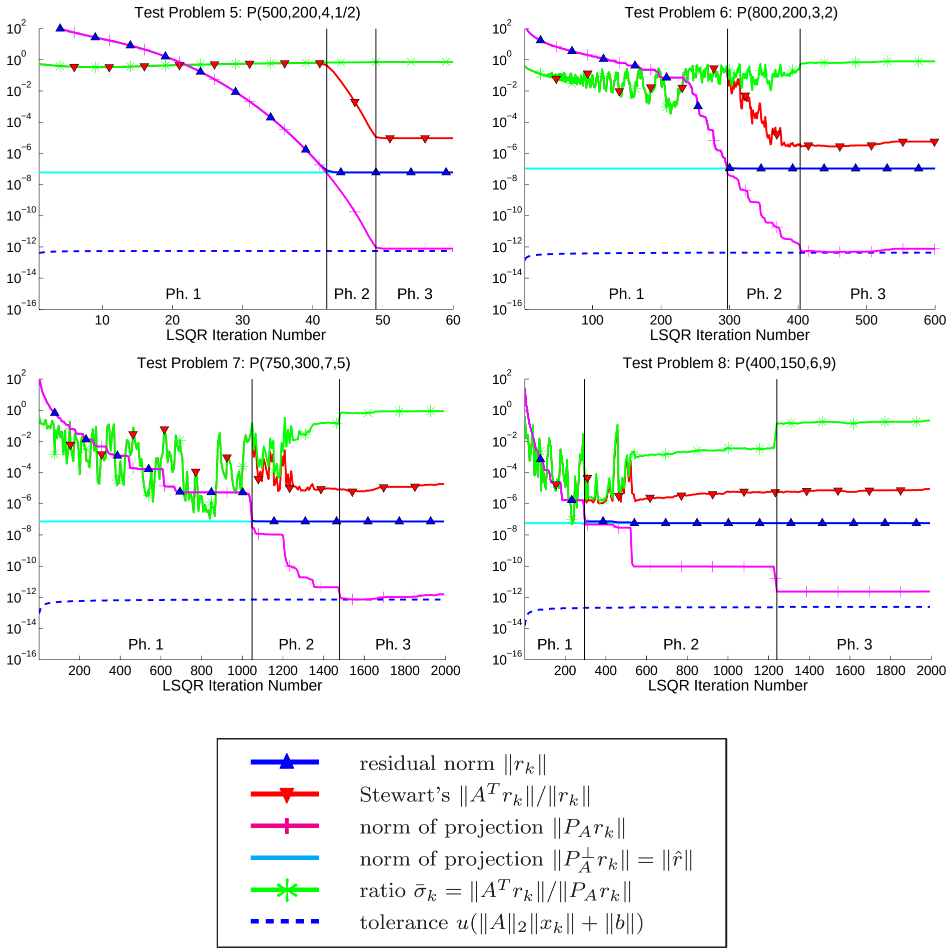

FIG. 7.2. *Test problems 5 to 8: increasingly ill-conditioned examples with* $\|\hat{r}\| \approx 10^{-7}$.

of LSQR iterations performed. We compute the quantity $\mu(x_k, \hat{\theta})$ in (6.3), with $\hat{\theta}$ in Theorem 6.3, using Matlab's built-in command `svd`.

We use test problems 1 to 4 in Table 7.1 (with $p = 15$, 10, 5 and 0, respectively) to demonstrate the impact of the size of the true residual norm on the effectiveness of the stopping criteria. We use test problems 9 and 11 in Table 7.2 (with $p = 7$) to test LSQR's stopping criteria on matrices from the Matrix Market. Ideally methods such as LSQR are applied to systems which (possibly after preconditioning) are well-conditioned, but this is not always possible, and an understanding of the behavior of these methods on ill-conditioned problems can be helpful in some practical cases, as well as giving us greater insight into the general numerical behavior of these algorithms. For this reason, we give numerical results for both a well-conditioned and an ill-conditioned test problem in Table 7.2.

8. Discussion and conclusions.

As a general trend, in the well-conditioned test problems 1 to 5 the ratio $\bar{\sigma}_k = \|A^T r_k\|/\|P_A r_k\|$ remains relatively constant. As the problems become more and more ill-conditioned, here in test problems 6 to 8, $\bar{\sigma}_k$ has a much more oscillatory behavior, as expected.

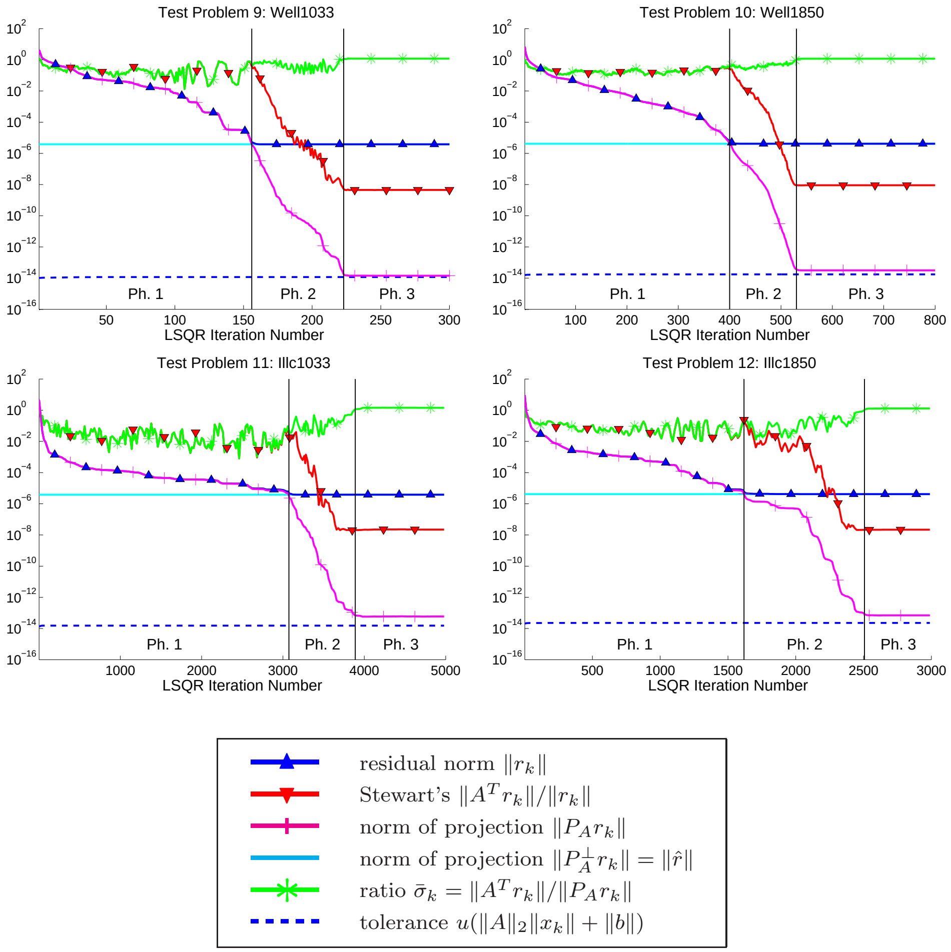

FIG. 7.3. *Test problems 9 to 12: sparse examples from the Matrix Market with* $\|\hat{r}\| \approx 10^{-6}$.

In most of the test problems there is a very clear visual distinction in the convergence plots between phases 1 and 2, and also between phases 2 and 3. In phase 1 we clearly observe that $\|r_k\| \approx \|P_A r_k\|$ decreases while Stewart's $\|A^T r_k\|/\|r_k\| \approx \bar{\sigma}_k$ oscillates between the extreme singular values of A (which means it remains roughly constant for the well-conditioned problems and can vary wildly for the ill-conditioned problems). In phase 2 on the other hand, we see that $\|r_k\| \approx \|P_A^{\perp} r_k\|$ remains nearly constant while $\|A^T r_k\|/\|r_k\|$ starts to decrease. Note that in phase 2 $\|A^T r_k\|/\|r_k\|$ decreases at almost exactly the same rate as $\|P_A r_k\|$, as suggested by Lemma 4.1 and the fact that $\|r_k\| \approx \|\hat{r}\|$ is nearly constant in phase 2. Also note that there is no phase 2 in test problem 1 because $\|r_k\|$ satisfies (4.9) when $k \approx 63$, and no phase 1 in test problem 4 because $\|r_0\| \approx \|\hat{r}\|$; see the comments after (4.6). Finally in phase 3, both $\|r_k\|$ and $\|A^T r_k\|/\|r_k\|$ remain nearly constant.

The above patterns are most obvious in the well-conditioned problems and less so in the very ill-conditioned problems. For example in test problems 7 and 8, $\bar{\sigma}_k$ oscillates a great deal and $\|P_A r_k\|$ decreases very slowly and in a staircase pattern, making the boundaries between successive phases less clear.

Table 7.1

The effectiveness of LSQR's stopping criteria depends on the size of $\|\hat{r}\|$.

Test	Parameters		LSQR 1		LSQR 2		LStest1	LStest2
Problem	α	β	True	App.	True	App.	see (5.3)	see (6.9)
1 $\|\hat{r}\| \approx 10^{-14}$	10^{-4}	10^{-4}	13	13	∞	73	13	13
	10^{-8}	10^{-4}	18	18	∞	85	18	18
	10^{-8}	10^{-8}	30	30	∞	85	30	30
	10^{-12}	10^{-8}	34	34	∞	98	34	34
	10^{-14}	10^{-14}	53	53	∞	103	53	53
2 $\|\hat{r}\| \approx 10^{-9}$	10^{-4}	10^{-4}	13	13	58	58	13	13
	10^{-8}	10^{-4}	18	18	∞	72	18	18
	10^{-8}	10^{-8}	30	30	∞	72	30	30
	10^{-12}	10^{-8}	34	34	∞	84	34	34
	10^{-14}	10^{-14}	∞	∞	∞	91	53	53
3 $\|\hat{r}\| \approx 10^{-4}$	10^{-4}	10^{-4}	13	13	38	38	13	13
	10^{-8}	10^{-4}	18	18	54	54	18	18
	10^{-8}	10^{-8}	∞	∞	54	54	30	30
	10^{-12}	10^{-8}	∞	∞	∞	69	34	34
	10^{-14}	10^{-14}	∞	∞	∞	75	53	53
4 $\|\hat{r}\| \approx 10^{1}$	10^{-4}	10^{-4}	∞	∞	13	13	14	13
	10^{-8}	10^{-4}	∞	∞	30	30	18	18
	10^{-8}	10^{-8}	∞	∞	30	30	31	30
	10^{-12}	10^{-8}	∞	∞	46	46	35	35
	10^{-14}	10^{-14}	∞	∞	54	54	54	53

Table 7.2

Testing LSQR's stopping criteria on problems from the Matrix Market.

Test	Parameters		LSQR 1		LSQR 2		LStest1	LStest2
Problem	α	β	True	App.	True	App.	see (5.3)	see (6.9)
9 Well1033 $\|\hat{r}\| \approx 10^{-6}$	10^{-4}	10^{-4}	69	69	172	172	69	69
	10^{-8}	10^{-4}	111	111	207	207	111	111
	10^{-8}	10^{-8}	159	159	207	207	158	158
	10^{-12}	10^{-8}	∞	∞	∞	242	164	164
	10^{-14}	10^{-14}	∞	∞	∞	267	206	206
11 Illc1033 $\|\hat{r}\| \approx 10^{-6}$	10^{-4}	10^{-4}	43	43	1346	1346	43	43
	10^{-8}	10^{-4}	110	110	3563	3562	110	110
	10^{-8}	10^{-8}	3066	3066	3563	3562	3045	3049
	10^{-12}	10^{-8}	∞	∞	∞	4049	3154	3154
	10^{-14}	10^{-14}	∞	∞	∞	4610	3610	3614

We notice that the amplitude of the phase 1 oscillations in $\bar{\sigma}_k$ (and from (4.6) also in $\|A^T r_k\|/\|r_k\|$) are sometimes very large when LSQR is near stalling; see for example the plots for test problems 6 to 9. It would be interesting (but probably very difficult) to try to understand this phenomenon more clearly.

It is important to notice that in every plot $\|P_A r_k\|$ decreases monotonically to a level determined by the machine precision (see the comments following (4.2), (4.8) and Theorem 5.1) after which it remains nearly constant. This is a property of the minimum residual method, here LSQR. One contribution of this paper is to show how this decrease is first exhibited in the decrease of $\|r_k\|$ until it reaches its computational plateau, and then exhibited in the decrease of the previously fairly level on average, but sometimes quite oscillatory, $\|A^T r_k\|/\|r_k\|$. Since $\|P_A r_k\|$ itself is not in practice directly available (see the comments in Section 5) this might in itself be useful information. Note from (4.2) and Lemma 4.1 that this is essentially a property of $\|\hat{r}\|$ and the decreasing $\|r_k\|$, not a property reserved to LSQR.

We note that these examples and this analysis shed light on the convergence of all well-behaved minimum residual iterative methods for the LS problem. In particular they give a good understanding of why Stewart's $\|A^T r_k\|/\|r_k\|$ usually shows no significant improvement at all during what we have described as phase 1. In fact only when $\|r_k\|$ has effectively reached its plateau does Stewart's $\|A^T r_k\|/\|r_k\|$ start to decrease—which it tends to do throughout phase 2—until it too finally reaches a plateau in phase 3. We have thus provided an explanation for this previously puzzling result for minimum residual LS iterations. Thinking of the convergence of LSQR in terms of projections of the residuals has also led to the potentially very useful Theorems 5.1 and 5.2.

Theorem 6.3 allows us to determine almost exactly at which iteration x_k is an acceptable solution—although this comes at a very high computational cost. As mentioned in section 7, Tables 7.1 and 7.2 give the number of iterations required to trigger the various stopping criteria. By examining these two tables to find at which iteration the condition LStest2 is first satisfied, we see that LSQR's stopping criteria 1 and 2 in (3.2) can be much too pessimistic. In the nearly compatible test problem 1, LSQR's stopping criterion 1 is triggered exactly when an acceptable solution is obtained. This is not surprising considering that this criterion is ideal for compatible systems; see (3.3) and (3.5). On the other hand in the very large-residual test problem 4, it is criterion 2 that is more reliable because there is effectively no phase 1 and $\|A^T r_k\|/\|r_k\|$ starts decreasing from the first few iterations. It is still triggered a little late when we set $\alpha \ll \beta$ (which is usually reasonable in practical applications, as discussed in Section 6). This is to be expected because the quantity $\|A^T r_k\|/\|r_k\|$ is a backward error in A only (see the comments after (3.7)) and thus LSQR's stopping criterion 2 does not use β in any way.

Away from these two extremes, however, *both* of LSQR's stopping criteria (with norms computed explicitly) can be much too pessimistic. In fact in all problems we have tested in which

$$u(\|A\|_F\|\hat{x}\| + \|b\|) \ll \|\hat{r}\| \ll \|r_0\|,$$

neither the residual norm $\|r_k\|$ nor Stewart's $\|A^T r_k\|/\|r_k\|$ reach their respective tolerances in (3.2) when α and β are chosen sufficiently small. (For example in all test problems except problems 1 and 4, in Tables 7.1 and 7.2 there are cases where both "True" LSQR stopping criteria fail to detect a backward stable iterate.) In contrast, our two new conditions LStest1 (5.3) and LStest2 (6.9) detect acceptable iterates for all choices of α and β satisfying $\alpha, \beta \geq \mathcal{O}(u)$.

We note that stopping criterion 2 in (3.2) is usually triggered if LSQR's *approximation* to $\|A^T r_k\|/\|r_k\|$ is used to test the criterion, because this approximation does not usually plateau at the end of phase 2 (while $\|A^T r_k\|/\|r_k\|$ actually does); see Figure 4.2. In this case, convergence is reported even though in fact the actual $\|A^T r_k\|/\|r_k\| \gg \alpha\|A\|_{2,F}$. Criterion 2 using LSQR's approximation to $\|A^T r_k\|/\|r_k\|$ is usually triggered in practice in what we have called phase 3, some iterations after LSQR has actually converged to a backward stable LS iterate.

Each iteration count in Tables 7.1 and 7.2 corresponds to an average using 100 different noisy vectors b. The iteration count for each individual test was almost always within $\approx 5\%$ of the average. In almost all our numerical tests, our new conditions LStest1 from and LStest2 are triggered at almost exactly the same iteration. This seems to indicate that the asymptotically optimal criterion LStest1 ($\psi_F(x_k, \alpha, \beta) \leq 1$ from (5.3)) is just as reliable as the criterion LStest2 ($\mu(x_k, \hat{\theta}) \leq \alpha\|A\|_F$ from (6.9)), which we proved necessary and sufficient to within a factor of $\sqrt{2}$ in Section 6.

Finally we note that although the above new conditions are very reliable for detecting when an acceptable LS solution has been obtained, neither can at present be estimated both reliably and efficiently enough to be used in practical large sparse applications. In the future we intend to examine whether reliable estimates of $\|P_A r_k\|$ and $\mu(x_k, \hat{\theta})$ can be computed efficiently. We are optimistic that this is the case for $\|P_A r_k\|$, as noted in the last three paragraphs of Section 5. Such estimates could be used in LStest1 and LStest2 and would make ideal stopping criteria for the iterative solution of large sparse LS problems.

Acknowledgments. The authors are very grateful to Sou-Cheng Choi, Michael Saunders, and two anonymous referees. All their detailed and very helpful comments and suggestions have greatly improved the paper.

REFERENCES

[1] M. Arioli and S. Gratton, *Least-squares problems, normal equations, and stopping criteria for the conjugate gradient method*, Technical Report RAL-TR-2008-008, Rutherford Appleton Laboratory, Oxfordshire, UK, 2008.

[2] A. Björck, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, PA, 1996.

[3] X.-W. Chang, C. C. Paige and D. Titley-Peloquin, *Characterizing matrices that are consistent with given solutions*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1406-1420.

[4] S.-C. Choi, *Iterative Methods for Singular Linear Equations and Least-Squares Problems*, Ph.D. thesis, Stanford University, 2006.

[5] G. H. Golub and W. Kahan, *Calculating the singular values and pseudo-inverse of a matrix*, SIAM J. Numer. Anal., 2 (1965), pp. 205–224.

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

[7] J. F. Grcar, *Optimal sensitivity analysis of linear least squares*, Technical Report LBNL-52434, Lawrence Berkeley National Laboratory, Berkeley, CA, 2003.

[8] M. Gu, *Backward perturbation bounds for linear least squares problems*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 363–372.

[9] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed, SIAM, Philadelphia, PA, 2002.

[10] The Matrix Market, *http://math.nist.gov/MatrixMarket*.

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

[12] C. C. Paige and M. A. Saunders, LSQR: *an algorithm for sparse linear equations and sparse least squares*, ACM Trans. Math. Software, 8 (1982), pp. 43–71.

[13] C. C. Paige and M. A. Saunders, *Algorithm 583,* LSQR: *sparse linear equations and sparse least squares problems*, ACM Trans. Math. Software, 8 (1982), pp. 195–209.

[14] J. L. Rigal and J. Gaches, *On the compatibility of a given solution with the data of a linear system*, J. ACM, 14 (1967), pp. 543–548.

[15] G. W. Stewart, *On the perturbation of pseudo-inverses, projections, and linear least squares problems*, SIAM Rev., 19 (1977), pp. 634–662.

[16] G. W. Stewart, *Research, development, and* LINPACK, in Mathematical Software III, J. R. Rice ed., Academic Press, New York, NY, pp. 1–14, 1977.

[17] Z. Strakoš and P. Tichy, *On error estimation in the conjugate gradient method and why it works in finite precision computations*, ETNA, 13 (2002), pp. 56–80.

[18] B. Waldén, R. Karlson and J.-G. Sun, *Optimal backward perturbation bounds for the linear least squares problem*, Numer. Linear Algebra Appl., 2 (1995), pp. 271–286.