BIT
1999, Vol. 39, No. 4, pp. 740–756

0006-3835/99/3904-0740 $15.00

BOUNDS ON SINGULAR VALUES REVEALED BY QR FACTORIZATIONS *†

CHING-TSUAN PAN[1] and PING TAK PETER TANG[2] ‡

[1] *Department of Mathematical Sciences, Northern Illinois University De Kalb, IL 60115-2888. email: pan@math.niu.edu*

[2] *Computational Software Laboratory, SC12-301, Intel Corporation 2200 Mission College Blvd., Santa Clara, CA 94502. email: Peter.Tang@intel.com*

Abstract.

We introduce a pair of dual concepts: pivoted blocks and reverse pivoted blocks. These blocks are the outcome of a special column pivoting strategy in QR factorization. Our main result is that under such a column pivoting strategy, the QR factorization of a given matrix can give tight estimates of any two a priori-chosen consecutive singular values of that matrix. In particular, a rank-revealing QR factorization is guaranteed when the two chosen consecutive singular values straddle a gap in the singular value spectrum that gives rise to the rank degeneracy of the given matrix. The pivoting strategy, called cyclic pivoting, can be viewed as a generalization of Golub's column pivoting and Stewart's reverse column pivoting. Numerical experiments confirm the tight estimates that our theory asserts.

AMS subject classification: 65F30, 15A23, 15A42, 15A15.

Key words: Rank-revealing QR factorization, singular values, cyclic column pivoting.

1 Introduction.

Recently, Hong and Pan [5] showed that for any matrix $A \in \mathbb{R}^{m\times n}$, $m \geq n$ and given k, $1 \leq k < n$, there exists a column permutation Π such that

$$(1.1) \quad A\Pi = QR = Q \begin{array}{c} \begin{array}{cc} k & n-k \end{array} \\ \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix} \end{array} = Q \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & 0 & \ddots & \vdots \\ & & & r_{nn} \end{pmatrix},$$

and

$$\sigma_{\min}(R_{11}) \approx \sigma_k(A), \qquad \sigma_{\max}(R_{22}) \approx \sigma_{k+1}(A),$$

*Received February 1999. Revised May 1999. Communicated by Åke Björck.

†This paper is a revised version of [7].

‡The second author was supported by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U.S. Department of Energy, under Contract W-31-109-Eng-38.

where $Q \in \mathbb{R}^{m\times n}$, $Q^TQ = I$, $R \in \mathbb{R}^{n\times n}$ is upper triangular, and

$$\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_n(A)$$

are A's (or equivalently, R's) singular values. In particular, if k is the numerical rank of R (cf. [4, 5]), the factorization is rank revealing. Hong and Pan showed that the permutation Π in (1.1) can be chosen by examining the $n-k$ right singular vectors of A that correspond to

$$\sigma_{k+1}(A), \sigma_{k+2}(A), \ldots, \sigma_n(A).$$

In this paper, we investigate bounds on $\sigma_k(A)$ and $\sigma_{k+1}(A)$ by quantities computable from R using only QR factorizations, without the need for any singular vector.

We introduce a pair of dual concepts: pivoted blocks and reverse pivoted blocks. They are the natural generalization of what we called pivoted magnitudes and reverse pivoted magnitudes, which are the results of the well-known Golub's column pivoting [3] and a less well-known column pivoting strategy proposed by Stewart [9]. The main result is that a pivoted block R_{11} (or equivalently a reverse pivoted block R_{22}) ensures

$$\sigma_{\min}(R_{11}) \approx \sigma_k(A) \quad \text{and} \quad \sigma_{\max}(R_{22}) \approx \sigma_{k+1}(A).$$

Moreover, the reverse pivoted magnitude of R_{11} and the pivoted magnitude of R_{22} also approximate $\sigma_k(A)$ and $\sigma_{k+1}(A)$, respectively. This result allows us to devise a column pivoting strategy, which we call cyclic pivoting, that produces the pivoted and reverse pivoted blocks. And, as by-products of this proposed pivoting strategy, we arrive at an upper triangular factor R such that its corresponding kth and $(k+1)$th diagonal elements satisfy

$$|r_{kk}| \approx \sigma_k(A) \quad \text{and} \quad |r_{k+1,k+1}| \approx \sigma_{k+1}(A).$$

In particular, if k is the numerical rank of A, the cyclic pivoting strategy guarantees a rank-revealing QR factorization.

As Stewart pointed out in [10], all the algorithms proposed in this paper as well as all the Hybrid algorithms proposed in [6] depend on a pre-chosen integer k. Although both papers indicate that any existing heuristic (no rigorous bound proved or provable) RRQR algorithm [1, 3, 8, 11] could provide us with a good guess of the numerical rank k.

The rest of the paper is organized as follows. Section 2 establishes bounds on singular values derived from the properties of pivoted and reverse pivoted magnitudes. Section 3 presents our main theorems on pivoted blocks and reverse pivoted blocks. Section 4 presents the cyclic column pivoting algorithms used to find the pivoted and reverse pivoted blocks. Section 5 presents some numerical experiments used to confirm the theoretical results. Section 6 concludes the paper.

2 Pivoted magnitudes.

Given a matrix A, there are two well-known pivoting strategies for the QR factorization that produce tight bounds on $\sigma_{\max}(A)$ and $\sigma_{\min}(A)$ from particular entries in R. Those particular entries of R are important to our subsequent discussion. Let e_j denote the standard basis vector in $\mathbb{R}^n$. We begin with the following definition:

Definition 2.1. *Given an $m \times n$ matrix A, $m \geq n$, let $\Pi_{i,j}$ be the permutation such that $A\Pi_{i,j}$ interchanges columns i and j of A. The pivoted magnitude of A, $\eta(A)$, is defined to be the maximum magnitude of the $(1,1)$ entry of the R factors of $A\Pi_{1,l}$, $l = 1, 2, \ldots, n$, thus:*

$$\eta(A) \stackrel{\text{def}}{=} \max \left\{ |e_1^T R e_1| : A\Pi_{1,l} = QR,\ l = 1, 2, \ldots, n \right\}.$$

The reverse pivoted magnitude of A, $\tau(A)$, is defined to be the minimum magnitude of the (n,n) entry of the R factors of $A\Pi_{l,n}$, $l = 1, 2, \ldots, n$, thus:

$$\tau(A) \stackrel{\text{def}}{=} \min \left\{ |e_n^T R e_n| : A\Pi_{l,n} = QR,\ l = 1, 2, \ldots, n \right\}.$$

Algorithmically, one can think of applying QR factorization with column pivoting [3] to A. Then, the magnitude of r_{11} of the resulting R factor is $\eta(A)$. Clearly,

$$\eta(A) = \max_j \|Ae_j\|_2, \quad j = 1, 2, \ldots, n.$$

If A is nonsingular, we also have

$$\tau(A) = 1/\max_j \|e_j^T A^{-1}\|_2,$$

as shown in [9]. In [9], Stewart calls a related column pivoting strategy the *reverse pivoting* strategy.

The following lemma is not new. The result for $\tau(A)$ is proved in [9] and [2]. The result for $\eta(A)$ is rather straightforward. We therefore only state the results:

Lemma 2.1. *Let A be an $m \times n$ matrix, $m \geq n$. Then,*

$$\eta(A) \leq \sigma_{\max}(A) \leq \sqrt{n}\eta(A), \tag{2.1}$$

and

$$(1/\sqrt{n})\tau(A) \leq \sigma_{\min}(A) \leq \tau(A). \tag{2.2}$$

Now we consider the QR factorization

$$A = QR = Q \begin{array}{c} \begin{array}{cc} k & n-k \end{array} \\ \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix} \end{array}.$$

Two related submatrices are important in subsequent discussions: $\bar{R}_{11}$, the $(k+1) \times (k+1)$ leading principal submatrix of R, and $\bar{R}_{22}$, the $(n-k+1) \times (n-k+1)$ trailing principal submatrix of R. The next two lemmas facilitate the discussions to follow.

LEMMA 2.2. *If* $\tau(R_{11}) \geq f\,\eta(\bar{R}_{22})$ *for some* $0 < f \leq 1$, *then*

$$\sigma_{\min}(R_{11}) \leq \sigma_k(A) \leq \left(\sqrt{k(n-k+1)}/f\right) \cdot \sigma_{\min}(R_{11}) \tag{2.3}$$

and

$$(1/\sqrt{k}) \cdot \tau(R_{11}) \leq \sigma_k(A) \leq \left(\sqrt{k(n-k+1)}/f\right) \cdot \tau(R_{11}). \tag{2.4}$$

PROOF. From the interlacing properties of singular values, we have $\sigma_{\min}(R_{11}) \leq \sigma_k(A)$ and $\sigma_{\max}(\bar{R}_{22}) \geq \sigma_k(A)$. Furthermore,

$$\begin{aligned}\sigma_{\min}(R_{11}) &\geq \tau(R_{11})/\sqrt{k} \\ &\geq f\,\eta(\bar{R}_{22})/\sqrt{k} \\ &\geq f\,\sigma_{\max}(\bar{R}_{22})/\sqrt{k(n-k+1)} \\ &\geq f\,\sigma_k(A)/\sqrt{k(n-k+1)}.\end{aligned}$$

Thus (2.3) is proved. From (2.3) and Lemma 2.1, (2.4) follows easily. $\square$

LEMMA 2.3. *If* $\tau(\bar{R}_{11}) \geq f\,\eta(R_{22})$ *for some* $0 < f \leq 1$, *then*

$$\left(f/\sqrt{(n-k)(k+1)}\right) \cdot \sigma_{\max}(R_{22}) \leq \sigma_{k+1}(A) \leq \sigma_{\max}(R_{22}), \tag{2.5}$$

and

$$\left(f/\sqrt{(k+1)(n-k)}\right) \cdot \eta(R_{22}) \leq \sigma_{k+1}(A) \leq \sqrt{n-k} \cdot \eta(R_{22}). \tag{2.6}$$

PROOF. From the interlacing properties of singular values, we have $\sigma_{k+1}(A) \leq \sigma_{\max}(R_{22})$ and $\sigma_{\min}(\bar{R}_{11}) \leq \sigma_{k+1}(A)$. Furthermore,

$$\begin{aligned}\sigma_{\max}(R_{22}) &\leq \sqrt{n-k}\,\eta(R_{22}) \\ &\leq (1/f)\,\sqrt{(n-k)}\,\tau(\bar{R}_{11}) \\ &\leq (1/f)\,\sqrt{(n-k)(k+1)}\,\sigma_{\min}(\bar{R}_{11}) \\ &\leq (1/f)\,\sqrt{(n-k)(k+1)}\,\sigma_{k+1}(A).\end{aligned}$$

Thus (2.5) is proved. From (2.5) and Lemma 2.1, (2.6) follows easily. $\square$

Roughly speaking, Lemma 2.2 says that if $\tau(R_{11}) \geq f\eta(\bar{R}_{22})$ with f not too small, then

$$\tau(R_{11}) \approx \sigma_{\min}(R_{11}) \approx \sigma_k(A); \tag{2.7}$$

and Lemma 2.3 says that if $\tau(\bar{R}_{11}) \geq f\eta(R_{22})$ with f not too small, then

$$\eta(R_{22}) \approx \sigma_{\max}(R_{22}) \approx \sigma_{k+1}(A). \tag{2.8}$$

Similar results can be found in [6], where two main algorithms based on Lemmas 2.2 and 2.3 are presented. There, the algorithm "Hybrid-I" guarantees (2.7) and the algorithm "Hybrid-II" guarantees (2.8). In order to achieve (2.7) and (2.8) simultaneously, the algorithm "Hybrid-III" is proposed, which is an iteration where the inner loop invokes Hybrid-I followed by Hybrid-II. Clearly, Hybrid-III yields a rank-revealing QR factorization whenever there is a numerical-rank gap between $\sigma_k(A)$ and $\sigma_{k+1}(A)$. We achieve (2.7) and (2.8) quite differently than does Hybrid-III. This different approach is the main subject of the next two sections.

3 Pivoted blocks.

Let A have the QR factorization

$$A = QR = Q \begin{array}{c} \begin{array}{cc} \;\;k & \;\; n-k \end{array} \\ \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix} \end{array}.$$

Our main result is that if R_{11} is a *pivoted block*, or equivalently, if R_{22} is a *reverse pivoted block* (both to be defined momentarily), then (2.7) and (2.8) hold simultaneously.

Definition 3.1. *Consider the column permutations $\Pi_{l,k}$, $l = 1, 2, \ldots, k$, and*

$$A\Pi_{l,k} = Q^{(l)} R^{(l)} = Q^{(l)} \begin{array}{c} \begin{array}{cc} \;\;k & \;\; n-k \end{array} \\ \begin{pmatrix} R_{11}^{(l)} & R_{12}^{(l)} \\ 0 & R_{22}^{(l)} \end{pmatrix} \end{array} = Q^{(l)} \begin{pmatrix} r_{11}^{(l)} & r_{12}^{(l)} & \cdots & r_{1n}^{(l)} \\ & r_{22}^{(l)} & \cdots & r_{2n}^{(l)} \\ & 0 & \ddots & \vdots \\ & & & r_{nn}^{(l)} \end{pmatrix}.$$

If

$$\left|r_{kk}^{(l)}\right| = \eta\left(\bar{R}_{22}^{(l)}\right), \quad l = 1, 2, \ldots, k,$$

then we call R_{11} a pivoted block.

Note that, although $R_{22}^{(1)} = R_{22}^{(2)} = \cdots = R_{22}^{(k)}$, the $\bar{R}_{22}^{(l)}$ are different in general. Therefore, by itself, the condition $\tau\left(R_{11}^{(j)}\right) = \eta\left(\bar{R}_{22}^{(j)}\right)$ for some j, $1 \le j \le k$, cannot guarantee a pivoted block R_{11}.

When $k = 1$ the pivoted block is $r_{1,1}^{(l)}$ and $\eta\left(\bar{R}_{22}^{(l)}\right) = \eta(A)$. Thus one can view pivoted block as a generalization of the concept of pivoted magnitude defined in Section 2.

We also consider pivoted blocks relaxed by a factor f (we call them f-pivoted blocks), $0 < f \le 1$, where the equalities above are replaced by the inequalities

$$|r_{kk}^{(l)}| \ge f\,\eta\left(\bar{R}_{22}^{(l)}\right), \quad l = 1, 2, \ldots, k.$$

Among other benefits, the f-factor provides us with flexibility in algorithm implementation without sacrificing theoretic rigor. This is particularly applicable in the case of reverse pivoting (see Algorithms 2 and 3 in Section 4 for details). We will elaborate more on the use of f-factors in Section 5.

THEOREM 3.1. *Let*

$$\begin{array}{cc} & \begin{array}{cc} k & n-k \end{array} \\ A = QR = Q & \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix}. \end{array}$$

If R_{11} *is a pivoted block, then*

$$\sigma_{\min}(R_{11}) \le \sigma_k(A) \le \sqrt{k(n-k+1)}\,\sigma_{\min}(R_{11}),$$

and

$$\sqrt{(n-k)(k+1)}\,\sigma_{\max}(R_{22}) \le \sigma_{k+1}(A) \le \sigma_{\max}(R_{22}).$$

Moreover, $\sigma_k(A)$ *and* $\sigma_{k+1}(A)$ *can be estimated by the quantities* $\tau(R_{11})$ *and* $\eta(R_{22})$*:*

$$(1/\sqrt{k})\,\tau(R_{11}) \le \sigma_k(A) \le \sqrt{k(n-k+1)}\,\tau(R_{11}),$$

and

$$\big(1/\sqrt{(k+1)(n-k)}\,\big)\,\eta(R_{22}) \le \sigma_{k+1}(A) \le \sqrt{n-k}\,\eta(R_{22}).$$

In other words, the conclusions (2.3), (2.4), (2.5) *and* (2.6) *of Lemmas* 2.2 *and* 2.3 *are all satisfied with* $f = 1$. *More generally, if* R_{11} *is an* f*-pivoted block,* $0 < f \le 1$, *then* (2.3), (2.4), (2.5) *and* (2.6) *are all satisfied with the relaxation factor* f.

In particular, if A *has numerical rank* k, *we have a rank-revealing QR factorization with*

$$\sigma_{\min}(R_{11}) \ge \big(f/\sqrt{k(n-k+1)}\,\big)\,\sigma_k(A)$$

and

$$\sigma_{\max}(R_{22}) \le \big(\sqrt{(n-k)(k+1)}/f\big)\,\sigma_{k+1}(A).$$

PROOF. It suffices to prove the more general case when R_{11} is an f-pivoted block. This condition clearly implies that for some j, $1 \le j \le k$,

$$\tau(R_{11}) = \big|r_{kk}^{(j)}\big| = \tau\big(R_{11}^{(j)}\big) \ge f\,\eta\big(\bar{R}_{22}^{(j)}\big).$$

Thus, the condition of Lemma 2.2 is satisfied for $R^{(j)}$. Notice that $\sigma_{\min}(R_{11}^{(j)}) = \sigma_{\min}(R_{11})$ and $\tau(R_{11}^{(j)}) = \tau(R_{11})$, for $1 \le j \le k$, and thus the conclusions in Lemma 2.2, namely, (2.3) and (2.4), are both true. To prove the rest of the theorem, we need only to show that for some j, $1 \le j \le k$,

$$\tau(\bar{R}_{11}^{(j)}) \ge f\,\eta(R_{22}^{(j)}). \tag{3.1}$$

Then the conclusions in Lemma 2.3, (2.5) and (2.6), are both true since $R_{22} = R_{22}^{(l)}$ for each l. To prove (3.1), without loss of generality we assume that

$$|r_{kk}| = \tau(R_{11}) \quad \text{and} \quad |r_{k+1\,k+1}| = \eta(R_{22}).$$

To ease the notation further, let

$$\bar{R}_{11} \equiv \begin{pmatrix} R_{11} & b \\ 0 & \beta \end{pmatrix}.$$

Now, (3.1) is equivalent to

$$\tau(\bar{R}_{11}) \geq f\,|\beta|. \tag{3.2}$$

Let

$$\beta' = \tau(\bar{R}_{11}),$$

and assume that it is obtained by permuting the lth column of $\bar{R}_{11}$ with the $(k+1)$th column, followed by a triangularization. We can assume $l \neq k+1$ since otherwise $\beta' = \beta$ and (3.2) is trivial. We permute the lth column with the $(k+1)$th column in two steps. We first permute the lth with the kth column and triangularize. We then permute the kth with the $(k+1)$th column and triangularize. In the first step, the lower-right 2×2 block of $\bar{R}_{11}$

$$\begin{pmatrix} r_{kk} & b_k \\ 0 & \beta \end{pmatrix}$$

is transformed to

$$\begin{pmatrix} \alpha & \gamma \\ 0 & \beta \end{pmatrix}. \tag{3.3}$$

Note that β is unchanged in this process. We now must have

$$|\alpha| \geq f\sqrt{\gamma^2 + \beta^2}, \tag{3.4}$$

because R_{11} is an f-pivoted block. In the second step, the block (3.3) is transformed to

$$\begin{pmatrix} \alpha' & \gamma' \\ 0 & \beta' \end{pmatrix}.$$

Note that this orthogonal transformation affects only rows k and $k+1$ of the matrix. Consequently, the magnitude of the determinant of the lower-right 2×2 block of $\bar{R}_{11}$ is unchanged:

$$|\alpha'\beta'| = |\alpha\beta| \quad \text{and} \quad |\alpha'| = \sqrt{\gamma^2 + \beta^2}.$$

From (3.4), $|\beta'| \geq f\,|\beta|$, and thus $\tau(\bar{R}_{11}) \geq f\,|\beta|$. This completes the proof of the theorem. $\square$

Now we define what we call a *reverse pivoted block.*

DEFINITION 3.2. *Consider the column permutations* $\Pi_{l,k+1}$, $l = k+1, k+2, \ldots, n$, *and*

$$A\Pi_{l,k+1} = Q^{(l)}R^{(l)} = Q^{(l)} \begin{array}{c} \begin{array}{cc} k & n-k \end{array} \\ \begin{pmatrix} R_{11}^{(l)} & R_{12}^{(l)} \\ 0 & R_{22}^{(l)} \end{pmatrix} \end{array} = Q^{(l)} \begin{pmatrix} r_{11}^{(l)} & r_{12}^{(l)} & \cdots & r_{1n}^{(l)} \\ & r_{22}^{(l)} & \cdots & r_{2n}^{(l)} \\ & 0 & \ddots & \vdots \\ & & & r_{nn}^{(l)} \end{pmatrix}.$$

If

$$\tau\big(\bar{R}_{11}^{(l)}\big) = |r_{k+1,k+1}^{(l)}|, \quad l = k+1, k+2, \ldots, n,$$

we call R_{22} *a reverse pivoted block.*

Note that when $k = n-1$ the reverse pivoted block is $r_{n,n}^{(l)}$ and $\tau(\bar{R}_{11}^{(l)}) = \tau(A)$. Thus one can also view reverse pivoted blocks as a generalization of the concept of reverse pivoted magnitude defined in last section.

We also consider reverse pivoted blocks relaxed by a factor f (we call them reverse f-pivoted blocks), $0 < f \leq 1$, where the equalities above are replaced by the inequalities

$$\tau\big(\bar{R}_{11}^{(l)}\big) \geq f\,\big|r_{k+1,k+1}^{(l)}\big|, \quad l = k+1, k+2, \ldots, n.$$

It turns out that pivoted blocks and reverse pivoted blocks are dual concepts.

THEOREM 3.2. *Let* A *have the QR factorization*

$$A = QR = Q \begin{array}{c} \begin{array}{cc} k & n-k \end{array} \\ \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix} \end{array}.$$

Then R_{11} *is a pivoted block if and only if* R_{22} *is a reverse pivoted block. More generally,* R_{11} *is an* f*-pivoted block if and only if* R_{22} *is a reverse* f*-pivoted block, where* $0 < f \leq 1$.

PROOF. It suffices to prove the general case. Suppose R_{11} is an f-pivoted block. We wish to show that

$$\tau\big(\bar{R}_{11}^{(l)}\big) \geq f\,\big|r_{k+1,k+1}^{(l)}\big|, \quad l = k+1, k+2, \ldots, n,$$

where

$$A\Pi_{l,k+1} = Q^{(l)}R^{(l)} = Q^{(l)} \begin{array}{c} \begin{array}{cc} k & n-k \end{array} \\ \begin{pmatrix} R_{11}^{(l)} & R_{12}^{(l)} \\ 0 & R_{22}^{(l)} \end{pmatrix} \end{array}.$$

Fix an arbitrary l, $k+1 \leq l \leq n$, and denote $\bar{R}_{11}^{(l)}$ by

$$\bar{R}_{11}^{(l)} \equiv \begin{pmatrix} R_{11} & a \\ 0 & \beta \end{pmatrix}.$$

We will show that $\tau(\bar{R}_{11}^{(l)}) \geq f\,|\beta|$. This amounts to showing that when any column of R_{11} is permuted to the last column of $\bar{R}_{11}^{(l)}$, creating a new trailing element β' after retriangularization, we must have $|\beta'| \geq f\,|\beta|$. To this end, we first permute that column of R_{11} to the kth column and retriangularize. We denote the 2×2 submatrix consisting of rows and columns k and $k+1$ by

$$\begin{pmatrix} r_{kk} & b_k \\ 0 & \beta \end{pmatrix} \equiv \begin{pmatrix} \alpha & \gamma \\ 0 & \beta \end{pmatrix}.$$

Note that the transformation does not alter β. From the assumption that R_{11} is a f-pivoted block, we have

$$|\alpha| \geq f\sqrt{\gamma^2+\beta^2}.$$

Now we interchange the kth and $(k+1)$th columns. After retriangularization the block becomes

$$\begin{pmatrix} \alpha & \gamma \\ 0 & \beta \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} \alpha' & \gamma' \\ 0 & \beta' \end{pmatrix}$$

Thus, $|\alpha| \geq f\,|\alpha'|$. Using again the determinant argument employed in the previous proof, we obtain $|\alpha\beta| = |\alpha'\beta'|$, and we have $|\beta'| \geq f\,|\beta|$ as desired. The first half of the theorem is thus established.

Next we assume that R_{22} is a reverse f-pivoted block. We wish to show that for any $l = 1, 2, \ldots, k$,

$$\left|r_{kk}^{(l)}\right| \geq f\,\eta\big(\bar{R}_{22}^{(l)}\big), \tag{3.5}$$

where $A\Pi_{l,k} = Q^{(l)}R^{(l)}$. To this end, we permute any jth column of $R^{(l)}$, $k+1 \leq j \leq n$, with the $(k+1)$th column and retriangularize. Consider the resulting $(k+1)\times(k+1)$ principal submatrix

$$\bar{R}_{11}^{(l)} \equiv \begin{pmatrix} R_{11}^{(l)} & b \\ 0 & \beta \end{pmatrix},$$

and denote the 2×2 lower-right block by

$$\begin{pmatrix} r_{kk}^{(l)} & b_k \\ 0 & \beta \end{pmatrix} \equiv \begin{pmatrix} \alpha & \gamma \\ 0 & \beta \end{pmatrix}.$$

Now we wish to show that

$$|\alpha| \geq f\sqrt{\gamma^2+\beta^2}. \tag{3.6}$$

Suppose we interchange the kth and $(k+1)$th columns. After retriangularization. the corresponding block becomes

$$\begin{pmatrix} \alpha & \gamma \\ 0 & \beta \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} \alpha' & \gamma' \\ 0 & \beta' \end{pmatrix}.$$

We note that $|\alpha'| = \sqrt{\gamma^2+\beta^2}$. Moreover, since R_{22} is a reverse f-pivoted block, $|\beta'| \geq f\,|\beta|$. Finally, using the argument about the determinant $|\alpha\beta| = |\alpha'\beta'|$, we have

$$|\alpha| \geq f\,|\alpha'| = f\sqrt{\gamma^2+\beta^2}.$$

Thus (3.6) and (3.5) are proved. $\square$

4 Algorithms.

In the previous section, the main properties of pivoted and reverse pivoted blocks were discussed. Do these pivoted blocks always exist? The following algorithms answer this question positively. The three algorithms apply a kind of pivoting strategy that we call cyclic pivoting to produce f-pivoted blocks (Algorithm 1) and reverse f-pivoted blocks (Algorithms 2 and 3). We use the notation $\mathcal{R}(M)$ to denote the R factor of a matrix M.

ALGORITHM 1. **Cyclic Pivoting.** Given k, $1 \le k < n$, and a threshold f, $0 < f \le 1$, this algorithm produces a column permutation Π such that $A\Pi = QR$ and R_{11} is a f-pivoted block. As a result,

$$\sigma_{\min}(R_{11}) \le \sigma_k(A) \le \left(\sqrt{k(n-k+1)}/f\right)\sigma_{\min}(R_{11}),$$

and

$$\left(f/\sqrt{(n-k)(k+1)}\,\right)\sigma_{\max}(R_{22}) \le \sigma_{k+1}(A) \le \sigma_{\max}(R_{22}).$$

Step 0. Initialization: $R := \mathcal{R}(A\Pi)$ with (Golub) column pivoting, where Π is the column permutation. Set $i := k-1$.

Step 1. Iteration; cyclic pivoting

Step 1.1. If $i = 0$, exit algorithm.
Step 1.2. Set $R := \mathcal{R}(R \cdot \Pi_{i,k})$; $\Pi := \Pi \cdot \Pi_{i,k}$.
Step 1.3. If $|r_{kk}| \ge f\,\eta(\bar{R}_{22})$, then set $i := i-1$. Otherwise, perform exchange as follows: Find an l, $k+1 \le l \le n$, such that

$$\eta(\bar{R}_{22}) = \|[r_{k,l}, r_{k+1,l}, \ldots, r_{n,l}]^T\|_2$$

Set $R := \mathcal{R}(R\Pi_{k,l})$, $\Pi := \Pi \cdot \Pi_{k,l}$, and $i := k-1$.
Step 1.4. Go back to Step 1.1.

The iteration will terminate, because whenever an exchange takes place in Step 1.3, the value $|\det(R_{11})|$ strictly increases by a factor of at least $1/f$. Therefore this exchange can happen only a finite number of times. Then, at most $k-1$ iterations can take place after the final exchange. Clearly, at termination, R_{11} is an f-pivoted block.

Algorithm 1 produces an f-pivoted block R_{11}. Likewise we can also propose an algorithm to make R_{22} a reverse f-pivoted block. However, since computing $\tau(\bar{R}_{11})$ can cost $O(k^3)$, we propose the following two more practical algorithms. Algorithm 2 uses an estimate of $\tau(\bar{R}_{11})$, denoted by $\tilde{\tau}(\bar{R}_{11})$; Algorithm 3 uses an estimate of $\sigma_{\min}(\bar{R}_{11})$, denoted by $\tilde{\sigma}_{\min}(\bar{R}_{11})$.

ALGORITHM 2. **Reverse Cyclic Pivoting** using $\tilde{\tau}(\bar{R}_{11})$. Given $1 \le k < n$ and a threshold $0 < f \le 1$, this algorithm produces a column permutation Π such that $A\Pi = QR$ and R_{22} is a reverse $(f/\sqrt{k+1})$-pivoted block. As a result,

$$\sigma_{\min}(R_{11}) \ge \left(f/\sqrt{k(k+1)(n-k+1)}\,\right)\sigma_k(A),$$

and

$$\sigma_{\max}(R_{22}) \leq \left((k+1)\sqrt{n-k}/f\right)\sigma_{k+1}(A).$$

Step 0. Initialization: $R := \mathcal{R}(A\Pi)$ with (Golub) column pivoting, where Π is the column permutation. Set $i := k+1$.

Step 1. Iteration; reverse cyclic pivoting

Step 1.1. If $i = n+1$, exit algorithm.

Step 1.2. Set $R := \mathcal{R}(R \cdot \Pi_{k+1,i})$; $\Pi := \Pi \cdot \Pi_{k+1,i}$. Find the right singular vector $v = [v_1 v_2 \dots v_{k+1}]^T$ corresponding to $\sigma_{\min}(\bar{R}_{11})$. Choose l such that $|v_l| = \max_j |v_j|$. Then set

$$\gamma := \left(\mathcal{R}(\bar{R}_{11} \cdot \Pi_{l,k+1})\right)_{k+1,k+1} \quad \text{and} \quad \tilde{\tau}(\bar{R}_{11}) := |\gamma|.$$

Step 1.3. If $\tilde{\tau}(\bar{R}_{11}) \geq f\,|r_{k+1,k+1}|$, then set $i := i+1$. Otherwise, l must satisfy $1 \leq l \leq k$. Perform $R := \mathcal{R}(R \cdot \Pi_{l,k+1})$, giving

$$|r_{k+1,k+1}| = \tilde{\tau}(\bar{R}_{11}).$$

Set $\Pi := \Pi \cdot \Pi_{l,k+1}$ and $i := k+2$.

Step 1.4. Go back to Step 1.1.

The iteration will terminate because whenever an exchange takes place in Step 1.3, the value $|\det(R_{22})|$ strictly decreases by a factor of f or less. Therefore this exchange can only happen a finite number of times. Then, at most $n-k-1$ iterations can take place after the final exchange.

To show that the resulting block R_{22} is a reverse $(f/\sqrt{k+1})$-pivoted block, one only needs to observe that at the termination we have for $A\Pi_{k+1,l} = Q^{(l)}R^{(l)}$, $l = k+1, k+2, \dots, n$,

$$\begin{aligned}\tau\big(\bar{R}_{11}^{(l)}\big) &\geq \sigma_{\min}\big(\bar{R}_{11}^{(l)}\big)\\ &\geq (1/\sqrt{k+1})\tilde{\tau}\big(\bar{R}_{11}^{(l)}\big)\\ &\geq (1/\sqrt{k+1})f\,\big|r_{k+1,k+1}^{(l)}\big|.\end{aligned}$$

The next algorithm uses $\tilde{\sigma}_{\min}(\bar{R}_{11})$. In this case, in order to guarantee the algorithm's termination, we have to impose the restriction $f \leq 1/\sqrt{k+1}$.

ALGORITHM 3. **Reverse Cyclic Pivoting** using $\tilde{\sigma}_{\min}(\bar{R}_{11})$. Given $1 \leq k < n$ and a threshold $0 < f \leq 1/\sqrt{k+1}$, this algorithm produces a column permutation Π such that $A\Pi = QR$ and R_{22} is an reverse f/M-pivoted block where $M \geq 1$. In general, $M \leq 4$. As a result,

$$\sigma_{\min}(R_{11}) \geq \left(f/M\sqrt{k(n-k+1)}\right)\sigma_k(A),$$

and

$$\sigma_{\max}(R_{22}) \leq \left(M\sqrt{(n-k)(k+1)}/f\right)\sigma_{k+1}(A).$$

Step 0. Initialization: $R := \mathcal{R}(A\Pi)$ with (Golub) column pivoting, where Π is the column permutation. Set $i := k+1$.

Step 1. Iteration; reverse cyclic pivoting

Step 1.1. If $i = n+1$, exit algorithm.
Step 1.2. Set $R := \mathcal{R}(R \cdot \Pi_{k+1,i})$; $\Pi := \Pi \cdot \Pi_{k+1,i}$; and $\sigma := \tilde{\sigma}_{\min}(\bar{R}_{11})$.
Step 1.3. If $\sigma \geq f\,|r_{k+1,k+1}|$, then set $i := i+1$. Otherwise, Find an l, $1 \leq l \leq k$, such that $R := \mathcal{R}(R \cdot \Pi_{l,k+1})$ gives

$$|r_{k+1,k+1}| \leq \sqrt{k+1}\sigma_{\min}(\bar{R}_{11}).$$

Set $\Pi := \Pi \cdot \Pi_{l,k+1}$ and $i := k+2$. Note that one way to obtain l is to solve for the right singular vector $v = [v_1 v_2 \ldots v_{k+1}]^T$ corresponding to σ. Then, l is chosen so that $|v_l| = \max_j |v_j|$.
Step 1.4. Go back to Step 1.1.

To see that the algorithm does terminate, we first point out that the singular value estimation we use guarantees $\tilde{\sigma}_{\min} \geq \sigma_{\min}$. Now, it suffices to observe that after an exchange in Step 1.3, we have

$$\begin{aligned} |r_{k+1,k+1}| &\leq \sqrt{k+1}\sigma_{\min}(\bar{R}_{11}) \\ &\leq \sqrt{k+1}\tilde{\sigma}_{\min}(\bar{R}_{11}) \\ &< f\sqrt{k+1}(\,|r_{k+1,k+1}|_{\text{previous}}). \end{aligned}$$

Thus, the value $|\det(R_{22})|$ strictly decreases by a factor of $f\sqrt{k+1}$ or less whenever an exchange is performed.

At termination, we arrive at a situation where for $A\Pi_{k+1,l} = Q^{(l)}R^{(l)}$, $l = k+1, k+2, \ldots, n$,

$$\tau\big(\bar{R}_{11}^{(l)}\big) \geq \sigma_{\min}\big(\bar{R}_{11}^{(l)}\big) \geq \frac{1}{M}\tilde{\sigma}_{\min}\big(\bar{R}_{11}^{(l)}\big) \geq (f/M)\big|r_{k+1,k+1}^{(l)}\big|.$$

Thus, R_{22} is a reverse f/M-pivoted block.

We make some comments on the singular value estimation. Note that each time through Step 1.2 in Algorithms 2 and 3 we need to compute $\tilde{\sigma}_{\min}(\bar{R}_{11})$ and the associated right singular vector v, where the $\bar{R}_{11}$ in question differs from the one in the previous iteration only in the last column. At present, there are various dynamic condition estimators (see [11] for a detailed discussion) that can update an estimate $\tilde{\sigma}$ in $O(k)$ work. Such an estimate satisfies

$$\tilde{\sigma} \approx \sigma \quad \text{in general,} \quad \text{and} \quad \tilde{\sigma} \geq \sigma \quad \text{always.}$$

Moreover, numerous experiments suggest that

$$\tilde{\sigma}_{\min} \leq 4\sigma_{\min}$$

in general, that is to say $1 \leq M$ always and $M \leq 4$ in general.

To conclude this section, we indicate that the algorithms proposed here appear to be similar to the ones proposed in [6], but they are different. For example, Algorithm 1 never uses any kind of reverse pivoting, while the reverse pivoting (or Stewart II) is the fundamental component of every Hybrid algorithm in [6]. The advantage in avoiding reverse pivoting is obvious: it is expensive to form the inverse explicitly. There seems to be no cheaper and yet exact way, competitive with Golub pivoting, for reverse pivoting. Algorithms 2 and 3 are examples of implementing reverse pivoting approximately.

The "cyclic pivoting" strategy is directly motivated by the proof of Lemma 2.1 in [5]. We transplanted it to deal with the R-factor itself, and call it cyclic pivoting since in each iteration we have to permute each of the first k columns to the kth column of R.

5 Numerical examples.

In this section we present our preliminary numerical results illustrating the theory and algorithms proposed. All computations were done on a SUN SPARC IPC, using MATLAB 3.5i.

We only present results of Algorithm 1 here as the results from the other algorithms are similar. Our first experiment tests the ability of the algorithms proposed in detecting the gaps between two consecutive singular values. We applied Algorithm 1 with $f = 0.99$ to the following example:

EXAMPLE 5.1. Let $A \in \mathbb{R}^{12\times 10}$ be

$$A = H_{12} \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} H_{10}, \qquad H_n = I - \frac{2}{n} ee^T, \tag{5.1}$$

with $e^T = (1, 1, \ldots, 1)$ and

$$\Sigma = \mathrm{diag}(100, 10, 8, 4, 1, 0.2, 0.1, 0.05, 0.01, 0.0001).$$

The numerical results are summarized in Table 5.1.

In Table 5.1, the superscript of $R_{ii}^{(l)}$ indicates the corresponding $R_{11} \in \mathbb{R}^{l\times l}$. From the results we know that the gaps (100, 10), (1, .2), (0.05, 0.01), and (0.01, 0.0001) are well detected.

Our second group of experiments deals with random examples defined by Example 5.2. The example has an obvious numerical rank five. The purpose of this experiment is to test how accurate and efficient these algorithms in finding RRQR factorizations.

EXAMPLE 5.2. Let $A \in \mathbb{R}^{12\times 10}$ be

$$A = Q_{12} \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} Q_{10} E, \tag{5.2}$$

where

$$\Sigma = \mathrm{diag}(2, 4, 6, 8, 9, \mu, \mu, \mu, \mu, \mu), \qquad \mu = 0.0001,$$

$Q_n \in \mathbb{R}^{n\times n}$ ($n = 10, 12$), are randomly generated orthogonal matrices, and E is a pre-specified permutation matrix. In the following, when we mention

Table 5.1: Gap detection using Algorithm 1.

k	$\sigma_{\min}(R_{11}^{(k)})$	$\sigma_k(A)$	$\sigma_{\max}(R_{22}^{(k-1)})$
1	80.0453	100.0000	*
2	7.5904	10.0000	10.3392
3	4.3518	8.0000	8.5638
4	1.3744	4.0000	5.0602
5	0.4588	1.0000	2.1754
6	0.1016	0.2000	0.2887
7	0.0763	0.1000	0.2517
8	0.0450	0.0500	0.0804
9	0.0097	0.0100	0.0137
10	*	0.0001	0.0001

the algorithms proposed in last section, we mean the algorithms without Golub pivoting initialization unless otherwise indicated.

First we run each algorithm on 500 different random matrices of Example 5.2 with a fixed value $f = 0.65$ and a fixed permutation matrix E. For each matrix A, we use the column pivoting strategies specified by corresponding algorithms ($k = 5$) to find the column permutation Π such that $A\Pi = QR$ is an RRQR factorization. We record four values for each factorization: $\sigma_{\min}(R_{11})$, $\tau(R_{11})$, $\sigma_{\max}(R_{22})$, and $\eta(R_{22})$. The first two values are used to approximate $\sigma_5(A) = 2$, and the other two values are used to approximate $\sigma_6(A) = 0.0001$.

Figures 5.1 and 5.2 present the result obtained by Algorithm 1. In Figure 5.1, $\sigma_{\min}(R_{11})$ and $\tau(R_{11})$ are presented for all 500 test matrices. We sort the test matrices so that $\sigma_{\min}(R_{11})$ is increasing.

In Figure 5.2, $\sigma_{\max}(R_{22})$ and $\eta(R_{22})$ are presented for all 500 test matrices. We sort the test matrices so that $\sigma_{\max}(R_{22})$ is increasing. Note that the pivoted magnitudes $\tau(R_{11})$ and $\eta(R_{22})$ approximate the singular values in question better than $\sigma_{\min}(R_{11})$ and $\sigma_{\max}(R_{22})$ do. We do not know if this phenomenon holds in general.

However, these preliminary experiments suggest that $\tau(R_{11})$ and $\eta(R_{22})$ are good approximations of the two consecutive singular values if R_{11} is an f-pivoted block. Note in particular that in Algorithm 1, these two pivoted magnitudes are by-products of the pivoting strategy.

The f-factor allows more flexibility in algorithm implementation without sacrificing the rigor of the corresponding theoretical bounds derived. For example, without the f-factor, we do not see how one could avoid explicit inverse in implementing reverse pivoting such that Lemma 2.2 and 2.3 could still be used to prove a guaranteed bound.

However, choosing a right relaxation factor f is subtle. First, because of round-off errors effect, we never choose f to be 1. Second, a large f generally

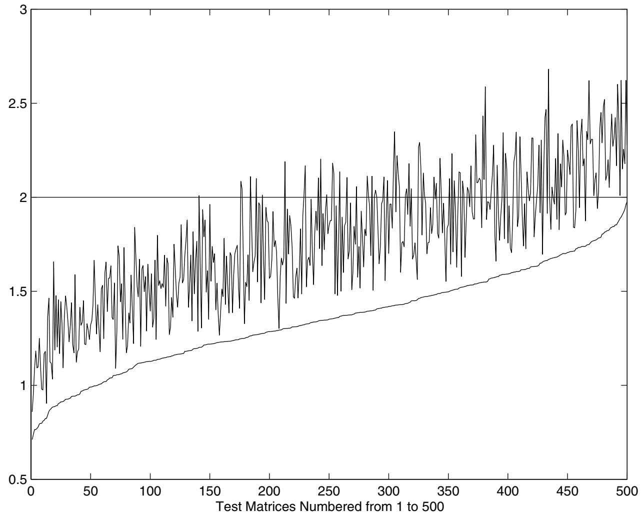

Figure 5.1: Estimating $\sigma_5 = 2$ by $\tau(R_{11})$ (upper jagged curve) and $\sigma_{\min}(R_{11})$ (lower smooth curve).

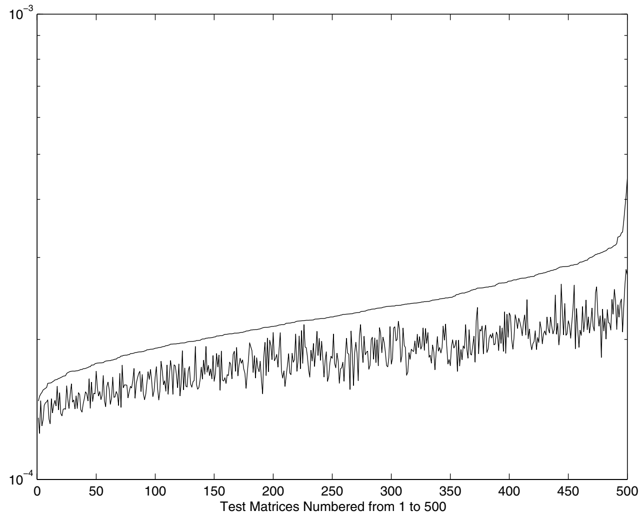

Figure 5.2: Estimating $\sigma_6 = 10^{-4}$ by $\eta(R_{22})$ (lower jagged curve) and $\sigma_{\max}(R_{22})$ (upper smooth curve).

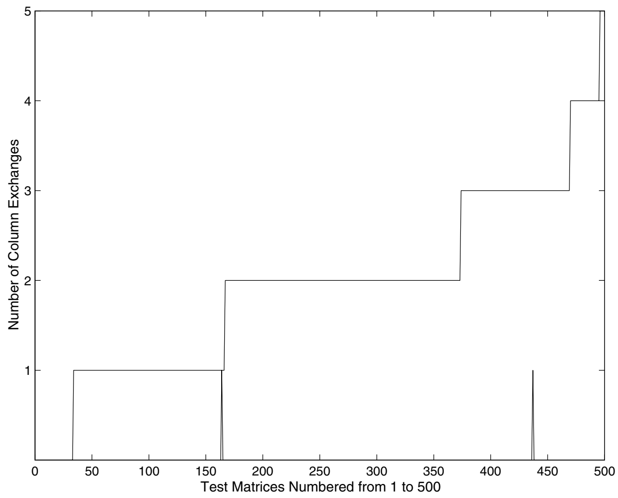

Figure 5.3: Effect of Golub pivoting (Step 0) in Algorithm 1: upper curve without Step 0; lower curve with Step 0;

means a small increase in $|\det(R_{11})|$ when there is a column exchange between R_{11} and R_{22}, and consequently, this generally means more iterations. On the other hand, too small an f is undesirable because in general this leads to poor approximation to singular values. Each user should weigh these factors carefully before choosing an appropriate f-factor.

The last experiment points out the importance of Step 0, that is, an initialization using Golub pivoting. A similar conclusion is also observed in [6]. Figure 5.3 shows the difference in work required in terms of number of exchange steps with and without Golub pivoting as an initialization. We perform Algorithm 1 on 500 test matrices of Example 5.2 with and without Step 0. We choose $f = 0.65$. The y-axis represents the number of column exchanges made between R_{11} and R_{22}. From this result, once again we confirm that for practical purposes Golub pivoting is still the most cost effective scheme with almost 100 percent reliability. But then again, to get provable bounds, Algorithm 1 is still required, even if column exchange is unnecessary most of the time.

When the singular value spectrum has several gaps and k is not known, Step 0 is also a good starting point. After running a Golub's pivoting, one can almost always locate possible gaps. Then, cyclic pivoting strategies can be employed to confirm the conjectures.

Finally, for low rank deficient matrices, using Chan's algorithm in Step 0 combined with Algorithm 2 or 3 is recommended. For low rank matrices, Algorithm 1 is recommended.

6 Concluding remarks.

We wish to point out that the key strategy used in this paper has much similarity with the one used in [5, 6]. We call it the "determinant strategy". In [5], a block whose smallest singular value is large is identified by the block whose determinant has the maximum magnitude. For Hybrid III in [6] and for Algorithms 1, 2, and 3 in this paper, a common thread is to increase the determinant's magnitude of the leading $k \times k$ block R_{11} to the point that it cannot be further increased. The result is an R_{11} block that we called a pivoted block. Actually, in Hong and Pan's paper [5], a block having absolute maximum determinant is stronger than necessary from an algorithmic point of view. We plan to address this determinant issue more systematically in our future work.

REFERENCES

1. C. H. Bischof, *Incremental condition estimation*, SIAM J. Matrix Anal. Appl., 11:2 (1990), pp. 312–322.
2. T. F. Chan, *Rank revealing QR factorizations*, Linear Algebra Appl, 88/89 (1987), pp. 67–82.
3. G. H. Golub, *Numerical methods for solving least squares problems*, Numer. Math. 7 (1965), pp. 206–216.
4. G. H. Golub, V. Klema, and G. W. Stewart, *Rank degeneracy and least squares problems*, Tech. Report TR-456, Dept. of Computer Science, University of Maryland, College Park, MD, 1976.
5. Y. P. Hong and C.-T. Pan, *Rank-revealing QR factorizations and the singular value decomposition*, Math. Comp. 58 (1992), pp. 213–232.
6. S. Chandrasekaran and I. Ipsen, *On rank-revealing QR factorizations*, SIAM J. Matrix Anal. Appl., 15:2 (1994), pp. 592–622.
7. C.-T. Pan and P. T. P. Tang, *Bounds on singular values revealed by QR factorizations*, Preprint MCS-P332-1092, Oct. 1992, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL, 1992.
8. D. J. Pierce and R. J. Plemmons, *Fast adaptive condition estimation*, SIAM J. Matrix Anal. Appl., 13:1 (1992), pp. 274–291.
9. G. W. Stewart, *Rank degeneracy*, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 403–413.
10. G. W. Stewart, *Determining rank in the presence of error*, in Linear Algebra for Large Scale and Real-Time Applications, M. S. Moonen, G. H. Golub, and B. L. R. De Moor eds., Kluwer Academic Publishers, Dordrecht, 1992, pp. 275–292.
11. P. T. P. Tang, *Dynamic condition estimation and Rayleigh–Ritz approximation*, SIAM J. Matrix Anal. Appl., 15:1 (1994), pp 331–346.