

# VALIDATED SOLUTIONS OF SADDLE POINT LINEAR SYSTEMS\*

TAKUMA KIMURA<sup>†</sup> AND XIAOJUN CHEN<sup>‡</sup>

**Abstract.** We propose a fast verification method for saddle point linear systems where the (1,1) block is singular. The proposed verification method is based on an algebraic analysis of a block diagonal preconditioner and rounding mode controlled computations. Numerical comparison of several verification methods with various block diagonal preconditioners is given.

**Key words.** saddle point matrix, numerical verification, block preconditioning

**AMS subject classifications.** 65F05, 65G05, 65G20

**1. Introduction.** We consider the system of saddle point linear systems

$$(1.1) \quad \mathcal{H}u = b \equiv \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix},$$

where  $A$  is an  $n \times n$  symmetric positive semidefinite matrix,  $B$  is an  $n \times m$  matrix, with  $m \leq n$ . We denote  $l = n + m$ . We assume that the coefficient matrix  $\mathcal{H}$  is nonsingular, which implies that  $B$  has full-column rank.

In recent years, saddle point problems have received considerable attention. A large amount of work has been devoted to developing efficient algorithms for solving saddle point problems. In a recent comprehensive survey [1], Benzi, Golub and Liesen discussed a large selection of numerical methods for saddle point problems. We are aware that it is very important to verify the accuracy of approximate solutions obtained by the numerical methods. However, there is little discussion on validated solution of saddle point problems by taking all possible effects of rounding errors into account.

Standard validation methods for the solution of a system of linear equations use an approximation of the inverse of the coefficient matrix. These methods are not efficient for the saddle point problem (1.1) when the dimension  $l$  is large or the condition number of  $\mathcal{H}$  is large, due to the indefiniteness of  $\mathcal{H}$  and the singularity of  $A$ .

In [2], a numerical validation method is proposed for verifying the accuracy of approximation solutions of the saddle point problem (1.1) without using an approximation of the inverse  $\mathcal{H}^{-1}$ , under the assumption that  $A$  is symmetric positive definite. The method uses the special structure of the saddle point problem to represent the variable  $x$  by the inverse of  $A$  and the variable  $y$ . For the case that  $A$  is singular and the size of the problem is large, it is a significant challenge to compute a rigorous upper bound for the norm  $\|\mathcal{H}^{-1}\|$  without using an approximation of the inverse  $\mathcal{H}^{-1}$ . In this paper, we present a fast method to compute rigorous error bounds for the saddle point problem (1.1) where  $A$  is symmetric positive semidefinite. In particular, we present a fast method to compute a constant  $\alpha$  such that

$$(1.2) \quad \|u^* - u\|_2 \leq \alpha \|b - \mathcal{H}u\|_2 \quad \text{for } u \in R^l,$$

---

\* The first version of this paper was completed while the second author was working in Hirosaki University. This work was supported partly by the Scientific Research Grant-in-Aid from Japan Society for Promotion of Science.

<sup>†</sup> Department of Mathematical Sciences, Hirosaki University, Hirosaki, 036-8561, Japan. (h07gs801@stu.hirosaki-u.ac.jp).

<sup>‡</sup> Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong, P.R. China. (maxjchen@polyu.edu.hk). This author's work is supported partly by the Research Grants Council of Hong Kong.

where  $u^*$  is the exact solution of (1.1). This method is based on an algebraic analysis of a block diagonal preconditioner for saddle point systems studied in a recent paper [6] by Golub, Greif and Varah. Instead of approximating the inverse of the  $l \times l$  indefinite matrix  $\mathcal{H}$ , we use approximations of inverses of two symmetric positive definite matrices in  $R^{n \times n}$  and  $R^{m \times m}$  to define the constant  $\alpha$  in the error bound (1.2). Moreover, we present fast methods to estimate upper bounds for the norms of inverses of the two symmetric positive definite matrices based on fast validated matrix computation by Oishi and Rump [8].

In section 2, we define the error constant  $\alpha$ . In section 3, we discuss how to compute an upper bound of  $\alpha$  efficiently and accurately by taking all possible effects of rounding errors into account. In section 4, we compare our verification method with the Krawczyk method [13], the LU decomposition method [8], `verifylss` function of INTLAB, and a block component verification method proposed in [2] using examples from CUTER [6], optimal surface fitting [3, 10], mixed finite element discretization of Stokes equations [1] and image restoration [5, 15].

**2. A new error bound.** Let  $W$  be an  $m \times m$  symmetric positive semidefinite matrix such that

$$M(W) = A + BWB^T$$

is a symmetric positive definite matrix. Note that  $BWB^T$  is singular for any symmetric positive definite matrix  $W$  when  $m < n$ . However, if  $W$  is symmetric positive definite, we can show that  $M(W)$  is symmetric positive definite under the condition that  $A$  is positive semidefinite and  $\mathcal{H}$  is nonsingular. To see it, let  $\bar{x} \neq 0$  be a solution of  $M(W)\bar{x} = 0$ . Then we have

$$\bar{x}^T A \bar{x} + \bar{x}^T B W B^T \bar{x} = 0.$$

Since  $A$  and  $BWB^T$  are positive semidefinite, we obtain

$$\bar{x}^T A \bar{x} = 0 \quad \text{and} \quad \bar{x}^T B W B^T \bar{x} = 0,$$

which implies

$$A \bar{x} = 0 \quad \text{and} \quad B^T \bar{x} = 0.$$

Therefore, we find that  $\mathcal{H}z = 0$  with  $z = (\bar{x}, 0)$ . This contradicts that  $\mathcal{H}$  is nonsingular.

Recently, Golub, Greif and Varah [6] performed an algebraic study of the block diagonal positive definite preconditioner

$$(2.1) \quad \mathcal{M}(W) = \begin{pmatrix} M(W) & 0 \\ 0 & B^T M(W)^{-1} B \end{pmatrix}.$$

They showed that  $\mathcal{M}(W)$  has the attractive property that the eigenvalues of the associated preconditioned matrix  $\mathcal{M}(W)^{-1}\mathcal{H}$  are bounded in a small range.

LEMMA 2.1. [6] The eigenvalues of the preconditioned matrix  $\mathcal{M}(W)^{-1}\mathcal{H}$  are bounded within the two intervals

$$\left[ -1, \frac{1 - \sqrt{5}}{2} \right] \cup \left[ 1, \frac{1 + \sqrt{5}}{2} \right].$$

Lemma 2.1 makes possible for us to present a rigorous error bounds for the saddle point problem (1.1).

THEOREM 2.2. Let  $u^*$  be the exact solution of (1.1). For any  $u \in R^l$ , we have

$$(2.2) \|u^* - u\|_2 \leq \frac{2}{\sqrt{5} - 1} \max(\|M(W)^{-1}\|_2, \|M(W)\|_2 \| (B^T B)^{-1} \|_2) \|b - \mathcal{H}u\|_2.$$

*Proof.* Obviously, we have,

$$\|u^* - u\|_2 \leq \|\mathcal{H}^{-1}\|_2 \|b - \mathcal{H}u\|_2.$$

Now, we use Lemma 2.1 to give an upper bound of  $\|\mathcal{H}^{-1}\|_2$ . Let  $\mathcal{L}$  be a nonsingular matrix such that  $\mathcal{L}\mathcal{L}^T = \mathcal{M}(W)$  and let

$$(2.3) \quad \mathcal{S} = \mathcal{L}^{-1}\mathcal{H}\mathcal{L}^{-T}.$$

Then the inverse  $\mathcal{H}^{-1}$  can be given as

$$\mathcal{H}^{-1} = \mathcal{L}^{-T}\mathcal{S}^{-1}\mathcal{L}^{-1}.$$

Since  $\mathcal{H}$  and  $\mathcal{S}$  are symmetric, we have

$$(2.4) \quad \begin{aligned} \|\mathcal{H}^{-1}\|_2 &= \max_{\substack{v \in R^l \\ v \neq 0}} \left| \frac{v^T \mathcal{L}^{-T} \mathcal{S}^{-1} \mathcal{L}^{-1} v}{v^T v} \right| \\ &= \max_{\substack{v \in R^l \\ v \neq 0}} \left| \frac{v^T \mathcal{L}^{-T} \mathcal{S}^{-1} \mathcal{L}^{-1} v}{v^T \mathcal{L}^{-T} \mathcal{L}^{-1} v} \frac{v^T \mathcal{L}^{-T} \mathcal{L}^{-1} v}{v^T v} \right| \\ &\leq \max_{\substack{w \in R^l \\ w \neq 0}} \left| \frac{w^T \mathcal{S}^{-1} w}{w^T w} \right| \max_{\substack{v \in R^l \\ v \neq 0}} \left| \frac{v^T \mathcal{M}(W)^{-1} v}{v^T v} \right| \\ &= \|\mathcal{S}^{-1}\|_2 \|\mathcal{M}(W)^{-1}\|_2. \end{aligned}$$

From (2.3) and  $\mathcal{L}\mathcal{L}^T = \mathcal{M}(W)$ , we have

$$\mathcal{L}^T \mathcal{M}(W)^{-1} \mathcal{H} \mathcal{L}^{-T} = \mathcal{S}.$$

Hence,  $\mathcal{S}$  and  $\mathcal{M}(W)^{-1} \mathcal{H}$  have same eigenvalues.

By Lemma 2.1, the eigenvalues of  $\mathcal{S}$  are bounded within the two intervals

$$\left[ -1, \frac{1 - \sqrt{5}}{2} \right] \cup \left[ 1, \frac{1 + \sqrt{5}}{2} \right].$$

Hence all eigenvalues of  $\mathcal{S}^{-1}$  satisfy

$$|\lambda_i| \leq \frac{2}{\sqrt{5} - 1}, \quad i = 1, 2, \dots, l.$$

From (2.4), we obtain

$$\|\mathcal{H}^{-1}\|_2 \leq \frac{2}{\sqrt{5} - 1} \|\mathcal{M}(W)^{-1}\|_2.$$

Moreover, from (2.1), we have

$$\begin{aligned} \|\mathcal{M}(W)^{-1}\|_2 &\leq \max(\|M(W)^{-1}\|_2, \|(B^T M(W)^{-1} B)^{-1}\|_2) \\ &\leq \max(\|M(W)^{-1}\|_2, \|M(W)\|_2 \|(B^T B)^{-1}\|_2), \end{aligned}$$

where the last inequality uses

$$\begin{aligned}\lambda_{\min}(B^T M(W)^{-1} B) &= \min_{\substack{y \in \mathbb{R}^m \\ y \neq 0}} \frac{(M(W)^{-1} B y, B y)}{(B y, B y)} \frac{(B^T B y, y)}{(y, y)} \\ &\geq \lambda_{\min}(M(W)^{-1}) \lambda_{\min}(B^T B).\end{aligned}$$

□

Theorem 2.2 shows that an upper bound of the inverse  $\|\mathcal{H}^{-1}\|_2$  can be obtained by computing upper bounds for the norm of inverses of two symmetric positive definite matrices with sizes of  $n \times n$  and  $m \times m$ . When  $n$  and/or  $m$  are large, the number of its flops is much less than the methods working on the  $(n+m) \times (n+m)$  matrix  $\mathcal{H}$ . For example, the LU decomposition method for (1.1) requires  $O((n+m)^3)$  flops, but estimating  $\|M(W)^{-1}\|_\infty$  and  $\|(B^T B)^{-1}\|_\infty$  only requires  $O(n^3) + O(m^3)$  flops, which will save  $O(n^2 m + n m^2)$  flops computational cost. Moreover, since the two matrices are symmetric, we can replace  $\|\cdot\|_2$  by  $\|\cdot\|_\infty$  for the matrix norm in (2.2) and have

$$(2.5) \|u^* - u\|_2 \leq \frac{2}{\sqrt{5} - 1} \max(\|M(W)^{-1}\|_\infty, \|M(W)\|_\infty \|(B^T B)^{-1}\|_\infty) \|b - \mathcal{H}u\|_2.$$

In general, (2.5) is easier to implement than (2.2).

**3. Verification methods.** When we apply Theorem 2.2 and other verification methods to verify the accuracy of an approximate solution of (1.1) on a computer, it is necessary to consider rounding error. The IEEE 754 arithmetic standard [4] defines the rounding modes for double precision floating point numbers. Since Intel's CPU follows this standard, the rounding modes can be used on most PC's and workstations. We use rounding downwards, rounding upwards and rounding nearest to compute rigorous error bounds for (2.5). We also apply these rounding modes to the following three verification methods.

**Krawczyk method** [11, 13]

$$\begin{aligned}K(U) &:= u - \mathcal{R}(\mathcal{H}u - b) + (I - \mathcal{R}\mathcal{H})(U - u), \\ K(U) \subset \text{int}(U) &\Rightarrow u^* \in K(U) \Rightarrow \|u^* - u\|_\infty \leq \|\text{radius}(U)\|_\infty,\end{aligned}$$

where  $\mathcal{R}$  is an approximate inverse of  $\mathcal{H}$  and  $U$  is an interval vector whose center is  $u$ .

**LU decomposition method** [8]

Let  $\mathcal{L}\mathcal{U}$  be an approximate LU factorization of  $\mathcal{H}$ , that is,  $\mathcal{H} \approx \mathcal{L}\mathcal{U}$ .

$$\|u^* - u\|_\infty \leq \frac{\|\mathcal{U}^{-1} \mathcal{L}^{-1} (b - \mathcal{H}u)\|_\infty}{1 - \|\mathcal{U}^{-1} \mathcal{L}^{-1} \mathcal{H} - \mathcal{I}\|_\infty}$$

**Block component verification method** [2]

Let  $u^* = (x^*, y^*)$ ,  $r_1 = Ax + By - c$  and  $r_2 = B^T x - d$ .

$$\begin{aligned}\|u^* - u\|_\infty &\leq \max(\|x^* - x\|_\infty, \|y^* - y\|_\infty), \\ \|x^* - x\|_\infty &\leq \|A^{-1}\|_\infty (\|r_1\|_\infty + \|B\|_\infty \|y^* - y\|_\infty), \\ \|y^* - y\|_\infty &\leq \|A\|_\infty \|(B^T B)^{-1}\|_\infty \left( \|r_2\|_2 + \sqrt{\|BB^T\|_\infty \|A^{-1}\|_\infty} \|r_1\|_2 \right).\end{aligned}$$

**verifylss of INTLAB** [14]

$$\begin{aligned}U &= \text{verifylss}(\mathcal{H}, b) \\ \|u - u^*\|_\infty &\leq 2\|\text{radius}(U)\|_\infty, \quad \text{for } u \in U.\end{aligned}$$

Note that when  $A$  is singular, the block component verification method cannot be applied to (1.1) directly, but to the equivalent system

$$(3.1) \quad \begin{pmatrix} A + BWB^T & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c + BWd \\ d \end{pmatrix}.$$

Obviously, the error bounds depend on the choice of  $W$ . We tested the error bounds with various  $W$ . In this paper, we consider two types of choices:

$$W_1(\gamma) = \gamma / \|B^T B\|_2 I$$

and

$$W_2(\gamma) = \gamma (B^T B)^{-1}$$

if  $A$  is singular. We set  $W = 0$  if  $A$  is nonsingular.

**4. Numerical experiment.** The numerical testing was carried out on a IBM PC (3.0 GHz Pentium 4 processor, 1GB of memory) with the use of MATLAB 7.0 and INTLAB (Version 5.4) [12, 14]. We use the function `setround` in INTLAB [14] to compute the error bound. The function `setround` allows the rounding mode of the processor to be changed between round nearest (`setround(0)`), round down (`setround(-1)`), round up (`setround(1)`) and round towards zero (`setround(2)`). To compute  $\|M(W)^{-1}\|_\infty$ , we first use `setround(0)` to compute an approximate inverse  $R$  of  $M(W)$ . Next we use

$$\begin{aligned} S &= \text{intval}(R) \\ \beta &= \text{abss}(\text{norm}(S, \text{inf}) / (1 - \text{norm}(S * M(W) - I))) \end{aligned}$$

to get an upper bound  $\beta$  for  $\|M(W)^{-1}\|_\infty$  as

$$\|M(W)^{-1}\|_\infty \leq \frac{\|R\|_\infty}{1 - \|RM(W) - I\|_\infty} \leq \beta.$$

Similarly, we apply the function `setround` to (2.5) to get an upper bounds  $\Gamma$  by taking all possible effects of rounding errors into account such that

$$(4.1) \quad \|u^* - u\|_2 \leq \Gamma.$$

We compare the error bound (2.5) with the Krawczyk method, the LU decomposition method, the block component verification method and the function `verifylss` of INTLAB using examples from CUTER [6], optimal surface fitting [3, 10], mixed finite element discretization of Stokes equations [1] and image restoration [5, 15].

**Example 4.1: CUTER matrices.** We used two test problems, `genhs28` and `gouldqp3` from CUTER collection [7], which were used in [6].

The `genhs28` is a  $(n + m) \times (n + m)$  saddle point matrix, where  $A$  is an  $n \times n$  tridiagonal matrix with 2, 4, 2 along its superdiagonal, main diagonal and subdiagonal, respectively, except  $A_{1,1} = A_{n,n} = 2$ . The rank of  $A$  is  $n - 1$ .  $B$  is  $n \times m$  with values 1, 2, 3 along its main diagonal, first subdiagonal and second subdiagonal, respectively.

The `gouldqp3` is a  $(n + m) \times (n + m)$  saddle point matrix, where  $A$  is an  $n \times n$  matrix with  $\text{rank}(A) = n - 2$ .

We set the exact solution  $u^*$  and right hand side vector  $b$  as

$$u^* = (1, \dots, 1)^T, \quad b = \mathcal{H}u^*.$$

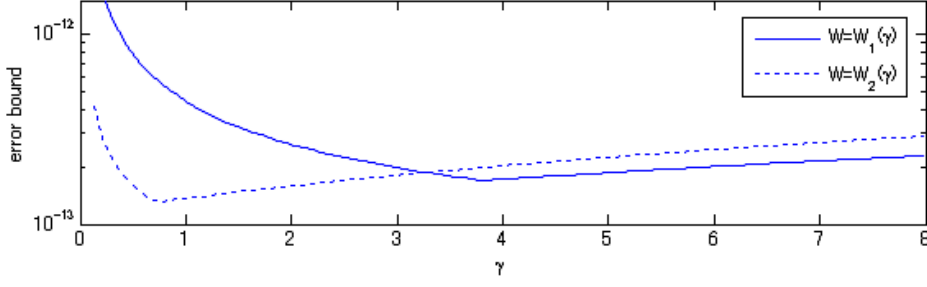


FIG. 4.1. Error bounds (2.5) for *genhs28* with different  $W$ .  
 $(n,m)=(500,498)$ ,  $\|A\|_2 \approx 8.0$ ,  $\|B^T B\|_2 \approx 36.0$ ,  $\|(B^T B)^{-1}\|_2 \approx 0.38$ .

TABLE 4.1  
The error bounds for Example 4.1, *genhs28*.  $\text{rank}(A)=n-1$ .

$(n, m)$		(10,8)	(500,498)	(1500,1498)	(3000,2998)
	$\text{cond}(\mathcal{H})$	40.232	40.461	40.461	40.460
	$\ u - u^*\ _\infty$	2.22e-16	2.22e-16	2.22e-16	9.99e-16
	$\ b - \mathcal{H}u\ _2$	2.46e-14	2.02e-13	3.50e-13	1.03e-12
(2.5)	$W_1(1)$	5.67e-13 [3.6e-15]	4.58e-12 [1.859]	7.95e-12 [48.093]	2.33e-11 [477.059]
	$W_2(1)$	1.46e-13 [7.1e-15]	1.41e-12 [1.983]	2.45e-12 [48.906]	7.21e-12 [1102.207]
Block component	$W_1(1)$	6.11e-10 [3.6e-15]	5.74e-09 [1.953]	9.95e-09 [89.373]	2.82e-08 [480.387]
	$W_2(1)$	3.96e-11 [1.1e-14]	3.82e-10 [2.078]	6.63e-10 [90.185]	1.88e-09 [1105.536]
LU		5.72e-16 [3.6e-15]	5.67e-16 [14.103]	5.67e-16 [385.535]	Fail <sup>1</sup>
Krawczyk		2.32e-15 [0.031]	2.32e-15 [6.625]	3.55e-15 [166.894]	Fail
verifylss		3.33e-16 [0.016]	3.33e-16 [4.795]	3.33e-16 [109.827]	Fail

Figure 1 shows the error bounds for *genhs28* with  $(n, m) = (500, 498)$  as  $\gamma$  changes from  $\min(\|A\|_2, 1/\|A\|_2)$  to  $\max(\|A\|_2, 1/\|A\|_2)$ . In Tables 1-2, we report numerical results with  $W_1(1)$  and  $W_2(1)$ .

**Example 4.2: Surface fitting problem.** Let  $\Omega \subset \mathbb{R}^2$  be a convex bounded domain,  $p_i = (p_i^1, p_i^2) \in \Omega$  be the measurement points and  $q_i$  be the corresponding real values, ( $i = 1 \cdots k$ ). We consider the following surface fitting problem [3, 10]

$$(4.2) \quad \min \sum_{i=1}^k (f(p_i) - q_i)^2 + \mu |f|_{H^2(\Omega)}^2,$$

over all functions  $f$  in the Sobolev space  $H^2(\Omega)$ . Here  $\mu$  is a fixed parameter.

<sup>1</sup>In Tables 1-5, "Fail" means out of the memory; [ ] shows CPU time(sec.).

TABLE 4.2  
The error bounds for Example 4.1, gouldqp3.  $\text{rank}(A) = n - 2$ .

$(n, m)$		(699,349)	(1999,999)	(2999,1499)
cond( $\mathcal{H}$ )		139.018	139.018	139.018
$\ u - u^*\ _\infty$		2.22e-16	2.22e-16	9.99e-16
$\ b - \mathcal{H}u\ _2$		8.29e-14	1.40e-13	3.80e-13
(2.5)	$W_1(1)$	5.63e-12 [2.406]	9.54e-12 [55.341]	2.58e-12 [194.385]
	$W_2(1)$	3.62e-12 [2.500]	6.13e-12 [81.605]	1.66e-12 [313.902]
Block component	$W_1(1)$	4.31e-09 [2.468]	7.30e-09 [55.951]	1.97e-08 [195.433]
	$W_2(1)$	1.94e-09 [2.546]	3.30e-09 [82.278]	8.90e-09 [314.996]
LU		4.75e-16 [19.420]	4.75e-16 [498.377]	Fail
Krawczyk		3.37e-15 [7.760]	3.37e-15 [179.915]	Fail
verifylss		3.33e-16 [5.968]	3.33e-16 [118.114]	Fail

We apply a finite element approximation with uniform triangular meshes to (4.2), and obtain a convex optimization problem in  $R^{4m+1}$  [3]

$$(4.3) \quad \begin{aligned} & \min \|Nx_1 + \mu e_k - q\|_2^2 + \mu(x_2^T G x_2 + x_3^T G x_3) \\ & \text{s.t. } Gx_1 = B_1 x_2 + B_2 x_3 \\ & \quad e_k^T N x_1 = 0, \end{aligned}$$

where  $N \in R^{k \times m}$ ,  $B_1, B_2, G \in R^{m \times m}$ . Here  $G$  is a symmetric positive semidefinite matrix. The problem (4.3) is equivalent to the following saddle point system

$$(4.4) \quad \left( \begin{array}{cc|cc} N^T N & & G & N^T e_k \\ & \mu G & -B_1^T & \\ \hline & & \mu G & -B_2^T \\ G & -B_1 & -B_2 & \\ e_k^T N & & & \end{array} \right) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 2N^T(q - \mu e_k) \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

where  $y_1$  and  $y_2$  are the Lagrange multipliers. In many applications,  $k < m$ , which results that the matrix  $N^T N$  is singular, i.e. the (1,1) block of the saddle point matrix in (4.4) is singular.

In this example, we use real data of average temperature 2006 in Aomori region from the Japan Meteorological Agency. We choose  $k = 47$ ,  $n = 3m + 1$ ,  $\text{rank}(A) = 2m + 47$ , and set  $\mu = 1.0e-5$  in (4.4). Numerical results of error estimate with various  $m, n$  are given in Table 4.3.

**Example 4.3: The Stokes equation.** We consider the saddle point system arising from the mixed finite element discretization of the stationary Stokes equation:

$$(4.5) \quad \begin{cases} -\nu \Delta u + \nabla p = \varphi & \text{in } \Omega, \\ -\text{div } u = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

TABLE 4.3  
The error bounds for Example 4.2, surface fitting problem.  $\text{rank}(A) = 2m + 47$ .

$(n, m)$		(766,255)	(1450,483)	(2350,783)	(3070,1023)
$\text{cond}(\mathcal{H})$		7.22e+05	1.31e+06	2.11e+06	2.56e+06
$\ b - \mathcal{H}u\ _2$		1.41e-14	1.61e-14	2.43e-14	2.54e-14
(2.5)	$W_1(1)$	4.46e-09 [2.749]	1.20e-08 [16.154]	3.38e-08 [66.213]	4.94e-08 [145.311]
	$W_2(1)$	4.85e-09 [2.812]	3.38e-09 [16.591]	9.25e-08 [68.119]	1.27e-08 [150.388]
Block component	$W_1(1)$	4.14e+01 [2.796]	5.36e+02 [16.310]	4.74e+03 [66.415]	1.04e+04 [145.951]
	$W_2(1)$	4.51e+01 [2.859]	1.02e+02 [16.763]	6.98e+02 [68.492]	1.27e+04 [151.122]
LU		8.46e-15 [19.422]	6.98e-15 [122.702]	3.51e-14 [514.185]	Fail
Krawczyk		1.43e-13 [7.243]	2.12e-13 [27.809]	4.76e-13 [198.123]	Fail
verifylss		2.93e-14 [5.047]	4.17e-14 [29.984]	7.84e-14 [124.562]	Fail

where  $\Omega = (0, 1) \times (0, 1)$ ,  $\partial\Omega$  is the boundary of  $\Omega$ ,  $\nu > 0$  is the kinematic viscosity coefficient,  $\varphi = (\varphi_1, \varphi_2)$  is a given force field,  $u : \Omega \rightarrow R^2$  is a velocity field and  $p : \Omega \rightarrow R$  is a pressure field. We apply a mixed finite element approximation with uniform triangular meshes, and obtain a saddle point linear system (1.1) where the velocity is approximated by the standard piecewise quadratic basis functions, and the pressure is approximated by piecewise linear basis functions.

In this example,  $A$  is nonsingular and  $A^{-1} \geq 0$ . We use Theorem 2 in [9] to get an upper bound of  $\|A^{-1}\|_\infty$ , i.e.

$$\|A^{-1}\|_\infty \leq \frac{\|\tilde{w}\|_\infty}{1 - \|s\|_\infty}$$

where  $\tilde{w}$  is an approximate solution of  $A\tilde{w} = e_n$ ,  $s = A\tilde{w} - e_n$ , and  $e_n = (1, 1, \dots, 1)^T \in R^n$ .

In this example,  $\|(B^T B)^{-1}\|_2 = O(h^{-2})$ , where  $h$  is the mesh size. To avoid using  $\|(B^T B)^{-1}\|_2$  for small  $h$ , we consider a preconditioned system. Let  $L$  be an  $m \times m$  nonsingular matrix such that  $L^T L \approx B^T B$ . Let

$$\mathcal{P} = \begin{pmatrix} I & \\ & L^{-1} \end{pmatrix}, \quad \tilde{\mathcal{H}} = \mathcal{P}^T \mathcal{H} \mathcal{P} = \begin{pmatrix} A & \tilde{B} \\ \tilde{B}^T & 0 \end{pmatrix}, \quad \tilde{b} = \mathcal{P}^T b,$$

$$\tilde{B} = BL^{-1}, \quad \tilde{M}(W) = A + \tilde{B}W\tilde{B}^T.$$

Applying Theorem 2.2 to the preconditioned system

$$\tilde{\mathcal{H}}\mathcal{P}^{-1}u = \tilde{b},$$

we obtain



$$\begin{aligned}
 (4.6) \quad & \|u - u^*\|_\infty \leq \|\mathcal{P}\|_\infty \|\mathcal{P}^{-1}(u^* - u)\|_\infty \leq \|\mathcal{P}\|_\infty \|\mathcal{P}^{-1}(u^* - u)\|_2 \\
 & \leq \frac{2\|\mathcal{P}\|_\infty}{\sqrt{5}-1} \max\left(\|\tilde{M}(W)^{-1}\|_2, \|\tilde{M}(W)\|_2 \|(\tilde{B}^T \tilde{B})^{-1}\|_2\right) \|\mathcal{P}^T(b - \mathcal{H}u)\|_2, \\
 & \leq \frac{2 \max(1, \|L^{-1}\|_\infty)}{\sqrt{5}-1} \max\left(\|\tilde{M}(W)^{-1}\|_\infty, \|\tilde{M}(W)\|_\infty \|L(B^T B)^{-1} L^T\|_\infty\right) \\
 & \quad \times \|\mathcal{P}^T(b - \mathcal{H}u)\|_2.
 \end{aligned}$$

We call (4.6) a preconditioned error bound. From  $\|L^{-1}\|_\infty \approx \sqrt{\|(B^T B)^{-1}\|_\infty} = O(h^{-1})$ , the preconditioned error bound is expected to be shaper than (2.5) for the Stokes equation. Similarly, we can get a preconditioned block component verification method as Method 2 in [2].

Numerical results of preconditioned error bounds for Example 4.3 with  $\nu = 1$  are given in Table 4.4.

TABLE 4.4  
The preconditioned error bounds for Example 4.3, Stokes equation.  $\text{rank}(A) = n$ .

$(n, m)$	(882,144)	(2738,400)	(9522,1296)	(20402,2704)
$\text{cond}(\mathcal{H})$	4.11e+05	1.21e+06	4.09e+06	8.66e+06
$\ b - \mathcal{H}u\ _2$	1.58e-15	2.56e-15	5.04e-15	7.32e-15
(4.7)	1.10e-11 [0.078]	9.24e-11 [1.001]	1.14e-09 [19.312]	5.11e-09 [149.383]
Preconditioned block component	7.24e-11 [0.078]	6.06e-10 [1.002]	7.46e-09 [19.295]	4.59e-08 [149.352]
LU	7.04e-16 [15.780]	1.53e-15 [430.490]	Fail	Fail
Krawczyk	2.00e-14 [2.516]	6.00e-14 [50.391]	Fail	Fail
verifylss	1.11e-16 [5.202]	2.22e-16 [124.224]	Fail	Fail

**Example 4.4: Image restoration.** Suppose the discretized scenes have  $p = p_1 \times p_2$  pixels. Let  $f \in R^p, g \in R^q$  be the underlying image and the observed image, respectively. Let  $H \in R^{q \times p}$  be the corresponding blurring matrix of block Toeplitz with Toeplitz blocks (BTTB). Restoration of  $f$  is an ill-conditioned problem. We consider the linear least squares problem with Tikhonov's regularization [5, 15]

$$(4.7) \quad \min_f \|Hf - g\|_2^2 + \alpha \|Df\|_2^2$$

where  $\alpha$  is a regularization parameter,  $D \in R^{(2p-p_1-p_2) \times p}$  is a regularization matrix of a first order finite difference operator

$$D = \begin{pmatrix} \Delta_{(p_1-1) \times p_1} \otimes I_{p_2 \times p_2} \\ I_{p_1 \times p_1} \otimes \Delta_{(p_2-1) \times p_2} \end{pmatrix} \quad \text{with} \quad \Delta = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & & 1 & -1 \end{pmatrix}$$

Problem (4.7) can be rewritten as a quadratic programming

$$(4.8) \quad \begin{aligned} \min \quad & \frac{1}{2}x^T Ax + c^T x \\ \text{s.t.} \quad & B^T x = 0 \end{aligned}$$

where

$$x = \begin{pmatrix} f \\ v \end{pmatrix}, \quad A = \begin{pmatrix} 2H^T H & 0 \\ 0 & \alpha I \end{pmatrix} \in R^{(3p-p_1-p_2) \times (3p-p_1-p_2)},$$

$$B^T = \begin{pmatrix} D & -I \end{pmatrix} \in R^{(2p-p_1-p_2) \times (3p-p_1-p_2)}, \quad c = \begin{pmatrix} -2H^T g \\ 0 \end{pmatrix}.$$

The optimal condition for (4.8) is a saddle point problem

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -c \\ 0 \end{pmatrix},$$

where  $y$  is the Lagrange multiplier vector for the constraints  $B^T x = 0$ . In this problem,  $A$  is a positive semidefinite matrix and  $B$  has full-column rank.

We generate an original image of the cameraman as shown in Figure 4.2 (a). The image is blurred by a Gaussian function

$$h(i, j) = e^{-(i^2+j^2)/18},$$

truncated such that the function has a support of  $7 \times 7$ , and then pixels are contaminated by Gaussian noise with the standard deviation of 0.05. The blurred and noisy image is shown in Figure 4.2 (b). We solve the saddle point problem to find a restored image, which is shown in Figure 4.2 (c). In the saddle point matrix,  $A$  has  $\text{rank}(A) = 2p - p_1 - p_2 + [p_1/2][p_2/2]$ . Here  $[\cdot]$  denotes the nearest integer.

Numerical results of error bounds for the restored image are given in Table 4.5.

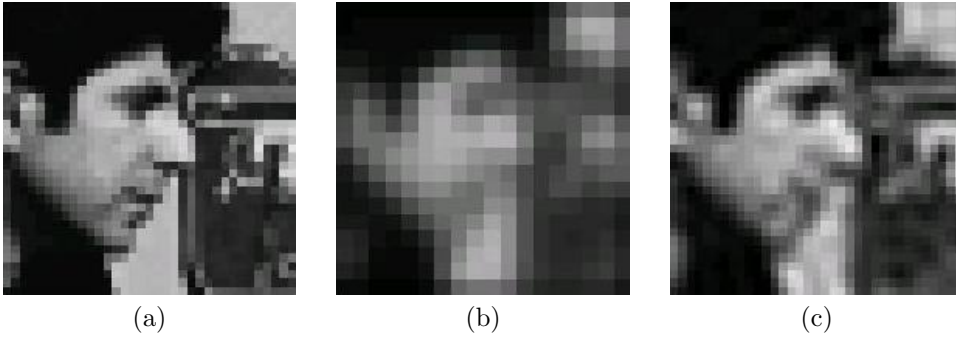


FIG. 4.2. (a) original image, (b) observed image, (c) restored image, PSNR=29.82 db.

To end this section, we use a  $3 \times 3$  block ill-conditioned saddle matrix to show that the error bound (2.5) is tight.

**Example 4.5.** We consider the following problem

$$\mathcal{H} = \begin{pmatrix} \epsilon I & 0 & 0 \\ 0 & 0 & B \\ 0 & B^T & 0 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},$$

TABLE 4.5  
The error bounds for Example 4.4, image restoration.

$(n, m)$		(1160,760)	(1825,1200)	(2465,1624)	(3400,2244)
cond( $\mathcal{H}$ )		3.15e+08	3.40e+08	3.67e+08	4.57e+08
$\ b - \mathcal{H}u\ _2$		2.13e-14	2.80e-14	3.39e-14	4.04e-11
(2.5)	$W_1(1)$	3.01e-06 [13.564]	4.33e-06 [50.798]	5.70e-06 [124.061]	6.96e-06 [353.168]
	$W_2(1)$	3.01e-06 [13.203]	4.32e-06 [48.687]	5.70e-06 [119.068]	6.96e-06 [320.720]
Block component	$W_1(1)$	5.42e+03 [13.798]	8.50e+03 [51.361]	1.23e+04 [125.123]	1.52e+04 [355.182]
	$W_2(1)$	1.11e+04 [13.203]	1.74e+04 [49.265]	2.51e+04 [120.161]	3.13e+04 [349.141]
LU		1.05e-09 [127.660]	1.56e-09 [481.395]	Fail	Fail
Krawczyk		$K(U) \not\subset U$	$K(U) \not\subset U$	Fail	Fail
verifylss		3.46e-14 [29.672]	3.91e-14 [112.698]	Fail	Fail

where  $b_1, b_2, b_3 \in R^m, I, B \in R^{m \times m}$ , and  $B$  is nonsingular. It is easy to find the inverse and the solution

$$\mathcal{H}^{-1} = \begin{pmatrix} \frac{1}{\epsilon}I & 0 & 0 \\ 0 & 0 & B^{-T} \\ 0 & B^{-1} & 0 \end{pmatrix}, \quad u^* = \begin{pmatrix} \frac{1}{\epsilon}b_1 \\ B^{-T}b_3 \\ B^{-1}b_2 \end{pmatrix}.$$

Consider  $0 < \epsilon \leq 1$  and  $\|B\|_\infty \geq 1$ . The condition number of  $\mathcal{H}$  satisfies

$$\|\mathcal{H}\|_\infty \|\mathcal{H}^{-1}\|_\infty \geq \epsilon^{-1} \|B\|_\infty.$$

When  $\epsilon \rightarrow 0$ , the condition number will go to  $\infty$ .

Using (2.2) or (2.5) with  $W = B^{-1}B^{-T}$ , we obtain

$$M(W) = \begin{pmatrix} \epsilon I & 0 \\ 0 & I \end{pmatrix}$$

and

$$(4.9) \quad \|u^* - u\|_2 \leq \frac{2}{\epsilon(\sqrt{5}-1)} \|b - \mathcal{H}u\|_2$$

for  $0 < \epsilon \leq 1/\|(B^T B)^{-1}\|_\infty$ . Furthermore, the equality holds in (4.9) when  $B = \frac{\epsilon(\sqrt{5}-1)}{2}I$  and  $u = (\frac{1}{\epsilon}b_1, u_2, u_3)$ .

**5. Final remark.** Using the algebraic analysis of a block diagonal preconditioner in [6], we proposed a fast verification method for saddle point linear systems where the  $(1,1)$  block may be singular. The method was implemented by using INTLAB [14] and taking all possible effects of rounding errors into account. Numerical results show that the method is efficient.

**Acknowledgements.** We are grateful to the two referees for their very helpful comments and suggestions. We wish to thank Prof. T. Yamamoto for carefully reading this paper and giving us helpful comments.

## REFERENCES

- [1] M. BENZI, G. H. GOLUB AND J. LIESEN, *Numerical solution of saddle point problems*, Acta Numerica, 14(2005), pp. 1-137.
- [2] X. CHEN AND K. HASHIMOTO, *Numerical validation of solutions of saddle point matrix equations*, Numer. Linear Algebra Appl., 10(2003), pp. 661-672.
- [3] X. CHEN AND T. KIMURA, *Finite element surface fitting for bridge management*, Inter. J. Information Technology & Decision Making, 5(2006), pp. 671-681.
- [4] ANSI/IEEE 754-1985, *Standard for Binary Floating-Point Arithmetic*, 1985.
- [5] H. FU, M.K. NG, M. NIKOLOVA AND J.L. BARLOW, *Efficient minimization methods of mixed  $\ell_2$ - $\ell_1$  and  $\ell_1$ - $\ell_1$  norms for image restoration*, SIAM J. Sci. Comput., 27(2006), pp. 1881-1902.
- [6] G. H. GOLUB, C. GREIF AND J. M. VARAH, *An algebraic analysis of a block diagonal preconditioner for saddle point system*, SIAM J. Matrix Anal. Appl., 27(2006), pp. 779-792.
- [7] N.I.M. GOULD, D. ORBAN AND PH.L. TOINT, *CUTEr, a Constrained and Unconstrained Testing Environment, revisited*, <http://hsl.rl.ac.uk/cuter-www/>, 2002.
- [8] S. OISHI AND S. M. RUMP, *Fast verification of solutions of matrix equations*, Numer. Math., 90(2002), pp. 755-773.
- [9] T. OGITA, S. OISHI AND Y. USHIRO, *Fast verification of solutions for sparse monotone matrix equations*, Computing, Supplement 15 (2001), pp.175-187.
- [10] S. ROBERTS, M. HELGLAND AND I. ALTAS, *Approximation of a thin plate spline smoother using continuous piecewise polynomial functions*, SIAM J. Numer. Anal., 41(2003), pp. 208-234.
- [11] S.M.RUMP, *Kleine Fehlerschranken bei Matrixproblemen*, Ph.D thesis, Universität Karlsruhe, 1980.
- [12] S.M. RUMP, *Verification methods for dense and sparse system of equations*, in J. Herzberger, editor, Topics in Validated Computations-Studies in Computational Mathematics, Elsevier, Amsterdam, 1994, pp.63-136.
- [13] S.M. RUMP, *Self-validating methods*, Linear Alg. Appl. 324(2001), pp. 3-13.
- [14] S.M. RUMP, *Interval computations with INTLAB*, Brazilian Electronic Journal on Mathematics of Computation, 1, 1999, <http://www.ti3.tu-harburg.de/publications/rump>
- [15] A. TIKHONOV AND V. ARSEININ, *Solution of Ill-Posed Problems*, Winston, Washington, DC, 1977.