SPARSE APPROXIMATE INVERSES AND TARGET MATRICES *

RUTH M. HOLLAND†, ANDY J. WATHEN‡, AND GARETH J. SHAW§

Abstract. If P has a prescribed sparsity and minimizes the Frobenius norm $||I - PA||_F$ it is called a sparse approximate inverse of A. It is well known that the computation of such a matrix P is via the solution of independent linear least squares problems for the rows separately (and therefore in parallel). In this paper we consider the choice of other norms, and introduce the idea of 'target' matrices. A target matrix, T, is readily inverted and thus forms part of a preconditioner when $||T - PA||$ is minimized over some appropriate sparse matrices P. The use of alternatives to the Frobenius norm which maintain parallelizability whilst discussed in early literature does not appear to have been exploited.

Key words. preconditioning, sparse approximate inverses, target matrices

1. Introduction. We consider the derivation of algebraic preconditioners for large, sparse, linear systems $Ax = b$. The construction of sparse approximate inverse preconditioners by

$$\min_{P \in S_P} ||I - PA|| \quad \left(\text{or} \quad \min_{P \in S_P} ||I - AP||\right) \tag{1.1}$$

has been considered in [Ben73], [KY86], [CDG92],[GH97],[GS98] where $||.||$ is widely taken to be the Frobenius norm $||.||_F$ and S_P denotes a set of matrices with prescribed sparsity pattern, i.e. for a set of indices $I_P \subseteq \{(i,j) : 1 \leq i,j \leq n\}$ we have that $S_P = \{P \in \mathbb{R}^{nxn} : P_{ij} = 0 \, for \, (i,j) \notin I_P\}$ where n is the dimension of A. For this norm the construction of P can be broken down so that each row (or column) is calculated by the solution of an independent small least squares problem, meaning that all rows can be computed in parallel. In the notation of the work we will introduce in this paper, we reinterpret (1.1) as constructing P such that the product PA (or AP) *targets* the identity. We need not restrict ourselves to the identity but can instead consider:

$$\min_{P \in S_P} ||T - PA|| \quad \left(\text{or} \quad \min_{P \in S_P} ||T - AP||\right)$$

i.e. find a preconditioner P such that PA (or AP) *targets* T (which we term the 'target' matrix), whereby our preconditioned system becomes $T^{-1}PA$. This approach requires that the action of T^{-1} is readily available. Easy parallelization is preserved when a target matrix, T, is employed.

Sparse approximate inverse preconditioners were introduced by Benson [Ben73], and developed to include a dynamic method of choosing S_P by Cosgrove, Diaz, and Griewank [CDG92], Grote and Huckle [GH97], and Gould and Scott [GS98]. The underlying idea was further developed to find P in factored form by Kolotilina and

*This work was supported by the Engineering and Physical Sciences Research Council, and NAG Ltd.

†Oxford University Computing Laboratory, Numerical Analysis Group, Wolfson Building, Parks Road, Oxford, OX1 3QD, U.K. (ruth@comlab.ox.ac.uk).

‡Oxford University Computing Laboratory, Numerical Analysis Group, Wolfson Building, Parks Road, Oxford, OX1 3QD, U.K. (wathen@comlab.ox.ac.uk).

§Schlumberger, Abingdon Technology Center, Lambourn Court, Abingdon Business Park, Abingdon OX14 1UJ, U.K. (Gareth.Shaw@slb.com).

Yeremin [KY93]. Computing approximate inverses using other methods has also received attention: Benzi and Tuma construct a factored form of approximate inverse by a method of A-biconjugation [BMT96], [BT98], [BMT99], whereas Chow and Saad use sparse-sparse iterations in their procedure [CS98]. We refer the survey [BT99] which compares these preconditioners.

In this paper we consider the use of target matrices and reconsider the use of alternative norms. In section 2 we will comment on the choice of norm, discussing Frobenius $H-$norms (generalized Frobenius norms) defined by

$$||B||_{F,H}^2 =< B, B >_{F,H} = trace\left(BHB^T\right) \tag{1.2}$$

where H is symmetric positive definite. These were introduced in this context by Kolotilina and Yeremin in [KY86] and described by Axelsson in [Axe96], p.322. In particular we will consider minimization in $||.||_{F,A^{-1}}$ which can be achieved without explicit knowledge of A^{-1} and requires n independent small linear solves. In sections 3 and 4 we will introduce two types of target matrix. Firstly a target with a specified sparsity pattern, and secondly a specific target matrix. We give results for the former in section 3 and for the latter in section 5 where we in particular show its use as an optimal preconditioner for low Peclet number advection-diffusion problems. Finally we draw some conclusions.

2. Frobenius $H-$Norms. In the Frobenius $H-$norm we construct our 'approximate inverse' (or more appropriately the multiplicative part of our preconditioner), P, and our target matrix, T, by minimizing $||T - PA||_{F,H}$ (as defined in (1.2)), which can be decomposed row-wise as follows.

$$\min ||T - PA||_{F,H}^2 = \sum_{j=1}^{n} \min ||t_j - p_j A||_{F,H}^2$$

where t_j and p_j represent the j^{th} rows of T and P respectively. Note that for $H = I$ this is the usual Frobenius norm. There are many possibilities for H, but an interesting one which we would like to highlight is the use of the inverse of A. Curiously we are able to minimize with respect to this norm without any explicit knowledge of the inverse itself, and moreover the construction of a preconditioner with respect to this norm is cheaper than using the basic Frobenius norm. Though identified by Kolotilina and Yeremin, and Axelsson this observation has not been widely exploited.

We show two versions of why the construction of the preconditioner in the Frobenius A^{-1}norm becomes so easy. The first is that used by Kolotilina and Yeremin. Consider

$$\min_{P \in S_P} ||T - PA||_{F,H} = \min_{P \in S_P} F_H(P) \tag{2.1}$$

where

$$F_H(P) = trace(THT^T) - \sum_{i,j} P_{ij}[(AHT^T)_{ji} + (AH^T T^T)_{ji}] + \sum_{i,j,k} P_{ij}(AHA^T)_{jk} P_{ik}$$

Thus we can find the minimum by finding the stationary points for each p_{ij}, $(i, j) \in I_P$. i.e. solve for $(i, j) \in I_P$:

$$\frac{\partial F_H(P)}{\partial P_{ij}} = -(TH^T A^T)_{ij} - (THA^T)_{ij} + (PAH^T A^T)_{ij} + (PAHA^T)_{ij} = 0 \tag{2.2}$$

For the case when $H = A^{-1}$, when A is symmetric positive definite, this gives us that:

$$T_{ij} = (PA)_{ij} \qquad \text{for}\,(i,j) \in I_P \tag{2.3}$$

i.e. $T_{*j} = P_{*k}A_{kj}$ for $(i,j) \in I_P$ where * represents the row containing the elements in I_P. Thus calculating a row of P is equivalent to solving a small, dense, symmetric square system which just contains the rows and columns of A which correspond to the sparsity pattern of the required row of P. This can be achieved by using a Cholesky factorization for example. Figure 2.1 highlights the elements of A which are needed to form this system.

FIG. 2.1. *Diagram showing how the rows and columns of A are selected to find the preconditioner with respect to the A^{-1}norm. The darkest entries are those of the required small square matrix.*

A second way of thinking about this problem is in terms of best approximations. When A is symmetric positive definite solving (2.1) can be considered as finding $P \in S_P$ such that PA is the 'best approximation' to T in $\|.\|_{F,A^{-1}}$ from the set $S_P A$, i.e.

$$\begin{aligned} < T - PA, VA >_{F,A^{-1}} &= 0 && \text{for}\, V \in S_P \\ < T - PA, V >_{F,I} &= 0 && \text{for}\, V \in S_P \\ < T, V >_{F,I} &= < PA, V >_{F,I} && \text{for}\, V \in S_P \end{aligned}$$

i.e. take $V = E_{ij}\, for\, (i,j) \in I_P$ where we use the 'delta' blocks E_{ij} as the basis functions, i.e. E_{ij} has one element in the (i,j) position equal to 1 and zeros elsewhere, and consider $P = \sum_{(r,s)\in I_P} p_{rs}E_{rs}$. So we obtain

$$\begin{aligned} < T, E_{ij} >_{F,I} &= \left\langle \sum_{(r,s)\in I_P} p_{rs}E_{rs}A, E_{ij} \right\rangle_{F,I} && \text{for}\,(i,j) \in I_P \\ T_{ij} &= \sum_{(r,s)\in I_P} p_{rs}\,(\delta_{ir}A_{sj}) && \text{for}\,(i,j) \in I_P \\ T_{ij} &= (PA)_{ij} && \text{for}\,(i,j) \in I_P \end{aligned}$$

which we find is identical to equation (2.3). This view is an adaptation of the work of Guillaume, Huard and Le Calvez from [GHLC03] where they consider block constant approximate inverses in a similar fashion.

The above are calculations necessary to produce a left preconditioner, and we will use this version in the remainder of the paper. However, it is just as straight forward to find a right preconditioner. To do this it is natural to consider the 'dual' Frobenius $H-$norm which is defined by

$$\|A\|_{F,H} = \left[trace\left(A^T H A\right)\right]^{1/2}.$$

The $\|.\|_{F,A^{-1}}$ norm can also be used when choosing the sparsity pattern dynamically, for details see [Hol03]. In this paper we restrict ourselves to choosing a sparsity pattern a priori, namely that of the transpose of A.

When we wish to solve a system which is not symmetric, our choice of H should not rely on A but on the symmetric part of A. Rather than using the A^{-1}norm we consider the $A^{-1} + A^{-T}$norm. Note that this choice of H will be positive definite for positive real A. Substituting into equation (2.2) we obtain

$$(T(A^{-1}A^T + I))_{ij} = (P(A + A^T))_{ij} \qquad \text{for}\,(i,j) \in I_P \tag{2.4}$$

which unfortunately includes A^{-1}. However, since we are constructing an approximate inverse $T^{-1}P \approx A^{-1}$ we can use $T^{-1}P$ in (2.4) to obtain

$$\begin{aligned}(T(T^{-1}PA^T + I))_{ij} &\approx (P(A + A^T))_{ij} \qquad &\text{for}\,(i,j) \in I_P \\ T_{ij} &\approx (PA)_{ij} \qquad &\text{for}\,(i,j) \in I_P\end{aligned}$$

We will use this form for the non-symmetric advection diffusion problems in section 5.

3. Using Target Matrices with Specified Sparsity Patterns. We consider two ways to use target matrices. In the first we merely choose a sparsity pattern for T in a similar manner as that in which a (usually different) sparsity pattern is chosen for P, and minimize over the possible choices for each. i.e.

$$\begin{aligned}\|T - PA\|^2 &= \min_{\hat{P}\in S_P} \min_{\hat{T}\in S_T} \|\hat{T} - \hat{P}A\|^2 \\ &= \sum_{j=1}^{n} \min_{p_j\in S_{p_j}} \min_{t_j\in S_{t_j}} \|t_j - p_jA\|^2\end{aligned}$$

In order for us to have a viable preconditioner we need to be able to cheaply solve systems involving T. Options available include for instance banded, triangular, and Hessenberg matrices.

In order not to obtain a trivial result one has to choose a normalization for T. As we consider this an extension of the approximate inverse technique we have normalized by setting the diagonal entries to be 1, though there are other options available. Then, where in the j^{th} row there is a non-zero entry in the i^{th} position of T (where $i \neq j$) we can set $t_j(i) = \sum_k p_j(k)a_{ki}$. That is we make $t_j - p_jA$ zero in these positions. Therefore to minimize over S_{t_j} and S_{p_j} we merely remove these columns from t_j and A and then solve the resulting smaller problem.

Unfortunately for this version of the target, the Frobenius A^{-1}norm is not applicable as it requires the solution of a square system initially and so removing columns would cause the system to become under-determined. However for the Frobenius norm we reduce the size of the least squares problems, see Figure 3.1. Then the entries of t_j can be formed via small dot products. Thus the construction of both a target matrix and P requires less work than that of P alone! As well as removing columns of A because of the introduction of the target matrix we can also remove any columns which are zero in the selected (dark) rows, let's call these 'zero rows'.

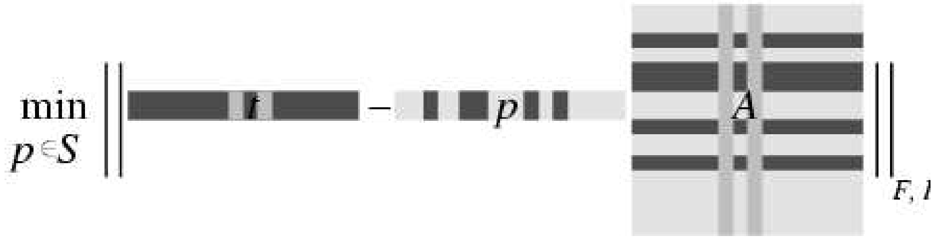

FIG. 3.1. *Diagram showing how the rows of A are selected (dark) to find the preconditioner with respect to the usual Frobenius norm, whereas the 2 identified columns are removed from the target problem as these correspond to non-zero (and non-diagonal) elements of T.*

Occasionally too many entries in P and T could lead to the problem of obtaining an under-determined system. This happens for row j if

$$dimension(A) - \#zero\ rows - nnz(t_j) + 1 \geq nnz(p_j).$$

This happens rarely and can be easily remedied, for instance by choosing to remove entries in p_j and/or t_j.

We will now show some results of iteration counts where such preconditioners have been used to solve the 5-point Laplacian problem. This problem is difficult for sparse approximate inverse preconditioners as the inverse of such an operator is not only dense but also exhibits slow decay of the entries away from the main diagonal, and we try to approximate it with a sparse banded matrix. We have given results (see Table 3.1) for no preconditioning, for the usual Frobenius norm sparse approximate inverse, and for a tridiagonal, lower triangular, and lower Hessenberg target in combination with a matrix P. In the latter cases we use only the corresponding non-zero entries of A to prescribe the sparsity pattern for T.

grid	no prec	$P, \|\|.\|\|_F$	P, tridiag. T	P, lower triang. T	P, lower H'berg T
8x8	22	13	12	13	10
16x16	43	24	21	23	18
32x32	82	46	39	42	33
64x64	153	84	70	84	63

TABLE 3.1
Iteration Counts of (full) GMRES *for the 5-point Laplacian Problem*

We see that using a tridiagonal or Hessenberg target can give us reasonable improvements, and as we have discussed these preconditioners are cheaper to construct than P alone. The improvement is not dramatic however, and we have not found real problems for which this is the case.

4. Using Specific Target Matrices. The second version of the target matrix is when we specify the exact matrix (as oppose to just the sparsity pattern). This idea was briefly mentioned in [CS98] where they suggested using a block diagonal matrix, although as we discussed previously, our only real restriction is that T should be easily inverted. One of our motivating ideas with specific target matrices is that we may have a problem consisting of several operators and for one of which we have a fast solver (like multigrid or fast Fourier transforms). We may then target this operator, and use P to 'move the problem closer' to this matrix. That is, we are using P to map the more difficult problem towards our target, T, which we can solve efficiently.

The construction of T and P is exactly equivalent to the construction of P alone in whichever norm is desired, with a row of T replacing the row of the identity in the small linear least squares problems (in the Frobenius norm), or the small linear solves (in the Frobenius A^{-1}norm).

For a partial differential equation boundary value problem, the equivalent operator theory of Manteuffel and Parter in [FMP90] and [MP90] indicates that the use of a preconditioner with the same boundary conditions is advantageous. This can be achieved if $S_T \subseteq S_A$, though the role of P makes this less clear.

5. Results. Let's consider the advection-diffusion equation.

$$\begin{aligned} -\epsilon\nabla^2 u + w.\nabla u &= f \quad in\ \Omega \\ u &= u^* \quad on\ \Gamma \end{aligned} \tag{5.1}$$

where w is some 'wind' function which is causing the advection. Clearly this system is naturally split into two distinct parts. Firstly we have the diffusion of u, i.e. a Laplacian system, and secondly the advection term. If the former quantity is significant it is sensible to target the Laplacian and to use a fast solver such as multigrid for the action of T^{-1}.

We consider two test problems taken from [ESW03]: $\Omega = (-1,1)\text{x}(-1,1)$, with bilinear finite elements on a uniform mesh of rectangles (Q1), and the streamline upwind Petrov-Galerkin formation [BH82]. In each case the wind function is scaled to have norm of $O(1)$ and so the value of epsilon in equation (5.1) represents the ratio of diffusion to advection in the problem.

Example 1 Constant wind.

$$w = \left(-\sin\frac{\pi}{6}, \cos\frac{\pi}{6}\right)$$

with boundary conditions

$$\begin{aligned} &u(x,-1) = 0,\ x \in (-1,0), \qquad u(x,-1) = 1,\ x \in (0,1) \qquad &u(x,1) = 1, \\ &u(-1,y) = 0, &u(1,y) = 1. \end{aligned}$$

which cause both internal and boundary layer formulation.

Example 2 Circular wind.

$$w = \left(2y(1-x^2), -2x(1-y^2)\right)$$

with boundary conditions

$$\begin{aligned} u(x,-1) &= 0, \qquad u(x,1) = 0, \\ u(-1,y) &= 0, \qquad u(1,y) = 1. \end{aligned}$$

which also has layers.

In each case we target the 5-point Laplacian matrix, and the sparsity pattern of P is that of A^T.

We solve iteratively the preconditioned system

$$T^{-1}PAx = T^{-1}Pb$$

We compare results using several versions of our preconditioner against the no preconditioning case. We consider preconditioning with:

- the usual approximate inverse in the Frobenius norm
- the usual approximate inverse in the Frobenius A^{-1}norm
- only a target ($P = I$)
- a target and a diagonal P (denoted D for clarity) in the Frobenius Norm
- a target and a P in the Frobenius norm
- a target and a P in the Frobenius A^{-1}norm

The number of iterations required for full GMRES to reach a stopping criterion of a relative residual of 10^{-6} are tabulated in Tables 5.1 and 5.2. Left preconditioning was used.

In each of the problems when epsilon is fairly large the target works very well, and we achieve a solution which is independent of the size of the mesh. As the advection operator becomes more important the 5-point Laplacian target is less effective, as we would expect, but even when epsilon is as small as 0.01 we are still apparently doing better than the $O(n^2)$ complexity that the usual approximate inverse preconditioners display.

Finally we show that the results are not dependent upon the type of elements we are using. For a circular wind problem and a random vector b we present analogous results to the above for

Example 3 linear elements on uniform mesh of triangles (P1)
Example 4 quadratic elements on uniform mesh of triangles (P2)
Example 5 biquadratic elements on uniform mesh of rectangles (Q2)

The corresponding full GMRES iteration counts are tabulated in Tables 5.3 - 5.5.

6. Conclusions. Target matrices in the context of sparse approximate inverse preconditioners have been described and demonstrated to be an effective method for an important class of partial differential equation problems. The efficiency inherent in the use of the Frobenius A^{-1}norm in this context has been highlighted.

Acknowledgements. We thank Howard Elman, Alison Ramage, and David Silvester for the use of their IFISS software which we used to provide the problems for examples 1 and 2. We also thank Daniel Loghin for the use of his codes from which we obtained examples 3 to 5.

REFERENCES

[Axe96] O. Axelsson. *Iterative Solution Methods.* Cambridge University Press, 1996.

[Ben73] M. W. Benson. Iterative solution of large scale linear systems. Master's thesis, Lakehead University, Thunder Bay, Canada, 1973.

[BH82] A. N. Brooks and T. J. R. Hughes. Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Comput. Methods Appl. Mech. Engrg.*, 32(1-3):199–259, 1982. FENOMECH '81, Part I (Stuttgart, 1981).

[BMT96] M. Benzi, C. D. Meyer, and M. Tůma. A sparse approximate inverse preconditioner for the conjugate gradient method. *SIAM J. Sci. Comput.*, 17(5):1135–1149, 1996.

[BMT99] M. Benzi, J. Marín, and M. Tůma. Parallel preconditioning with factorized sparse approximate inverses. In *Proceedings of the Ninth SIAM Conference on Parallel Processing for Scientific Computing 1999 (San Antonio, TX)*, page 5, Philadelphia, PA, 1999. SIAM.

[BT98] M. Benzi and M. Tůma. A sparse approximate inverse preconditioner for nonsymmetric linear systems. *SIAM J. Sci. Comput.*, 19(3):968–994 (electronic), 1998.

[BT99] M. Benzi and M. Tůma. A comparative study of sparse approximate inverse preconditioners. *Appl. Numer. Math.*, 30(2-3):305–340, 1999. Iterative methods and preconditioners (Berlin, 1997).

[CDG92] J. D. F. Cosgrove, J. C. Diaz, and A. Griewank. Approximate inverse preconditionings for sparse linear systems. *Intern. J. Computer Math*, 44:91–110, 1992.

ϵ	grid	no prec	$P, \|\|.\|\|_F$	$P, \|\|.\|\|_{F,A^{-1}}$	T	$T, D, \|\|.\|\|_F$	$T, P, \|\|.\|\|_F$	$T, P, \|\|.\|\|_{F,A^{-1}}$
1.00	8x8	16	9	9	12	10	6	5
	16x16	29	17	16	12	11	7	6
	32x32	58	33	31	12	11	7	7
	64x64	112	64	60	12	11	7	7
0.20	8x8	17	10	9	15	15	7	7
	16x16	31	18	16	16	16	11	11
	32x32	56	32	30	16	16	12	12
	64x64	108	61	57	16	16	13	13
0.10	8x8	17	10	8	19	19	8	8
	16x16	29	16	15	22	22	13	13
	32x32	53	30	28	22	22	18	18
	64x64	101	58	54	23	23	20	20
0.04	8x8	18	11	8	23	23	10	10
	16x16	28	17	13	32	31	14	15
	32x32	50	28	26	39	38	22	22
	64x64	94	53	49	39	38	33	33
0.02	8x8	19	12	9	25	25	11	11
	16x16	28	18	13	38	37	15	16
	32x32	49	27	24	49	48	23	25
	64x64	95	53	48	59	58	35	35
0.01	8x8	20	13	10	27	27	12	12
	16x16	30	19	15	43	42	17	19
	32x32	48	28	22	60	58	27	30
	64x64	92	49	44	75	74	41	43

TABLE 5.1
Iteration Counts of (full) GMRES *for Example 1*

ϵ	grid	no prec	$P, \|\|.\|\|_F$	$P, \|\|.\|\|_{F,A^{-1}}$	T	$T, D, \|\|.\|\|_F$	$T, P, \|\|.\|\|_F$	$T, P, \|\|.\|\|_{F,A^{-1}}$
1.00	8x8	18	10	10	12	11	5	5
	16x16	34	20	18	12	11	6	6
	32x32	66	38	36	12	10	7	6
	64x64	126	72	68	11	10	7	7
0.20	8x8	23	12	12	15	15	8	7
	16x16	45	25	23	16	15	10	10
	32x32	86	49	45	16	15	12	12
	64x64	164	94	87	16	16	12	12
0.10	8x8	28	14	14	19	18	9	9
	16x16	55	29	27	24	23	14	14
	32x32	106	59	54	23	23	18	18
	64x64	202	115	106	23	24	20	20
0.04	8x8	36	19	18	25	23	14	13
	16x16	76	38	34	34	30	19	18
	32x32	151	74	69	40	38	25	24
	64x64	286	154	143	42	40	32	34
0.02	8x8	43	22	22	31	30	17	17
	16x16	95	49	42	45	41	27	24
	32x32	196	91	81	58	51	37	32
	64x64	382	175	169	62	63	43	41
0.01	8x8	48	25	27	36	35	20	20
	16x16	117	63	52	60	55	38	33
	32x32	249	126	99	77	69	84	46
	64x64	498	218	195	92	85	93	56

TABLE 5.2
Iteration Counts of (full) GMRES *for Example 2*

ϵ	grid	no prec	$P, \|\|.\|\|_F$	$P, \|\|.\|\|_{F,A^{-1}}$	T	$T, D, \|\|.\|\|_F$	$T, P, \|\|.\|\|_F$	$T, P, \|\|.\|\|_{F,A^{-1}}$
1.00	8x8	20	11	11	6	4	4	4
	16x16	43	24	23	6	5	4	4
	32x32	87	48	45	6	4	4	4
	64x64	172	94	89	6	4	4	4
0.20	8x8	21	12	13	8	7	7	6
	16x16	50	27	26	8	8	7	7
	32x32	104	57	53	9	8	8	8
	64x64	207	113	107	9	8	8	8
0.10	8x8	25	13	14	12	11	8	8
	16x16	58	31	30	13	12	11	11
	32x32	120	64	60	13	13	12	12
	64x64	246	132	124	13	13	12	12
0.04	8x8	31	17	20	19	18	12	11
	16x16	76	39	36	26	24	17	17
	32x32	162	82	76	27	25	22	22
	64x64	344	176	165	28	26	25	24
0.02	8x8	34	20	24	26	23	14	14
	16x16	95	45	44	37	34	23	22
	32x32	208	99	92	43	41	32	32
	64x64	440	216	202	45	43	39	38
0.01	8x8	36	22	27	31	27	17	17
	16x16	117	56	53	53	48	29	29
	32x32	270	117	106	65	62	43	42
	64x64	573	251	235	72	68	56	56

TABLE 5.3
Iteration Counts of (full) GMRES *for Example 3 (P1)*

ϵ	grid	no prec	$P, \|\|.\|\|_F$	$P, \|\|.\|\|_{F,A^{-1}}$	T	$T, D, \|\|.\|\|_F$	$T, P, \|\|.\|\|_F$	$T, P, \|\|.\|\|_{F,A^{-1}}$
1.00	8x8	50	24	28	12	8	7	6
	16x16	100	50	61	12	9	7	7
	32x32	200	99	123	12	9	7	7
0.20	8x8	58	28	31	12	11	9	9
	16x16	121	59	70	13	11	10	10
	32x32	242	120	149	13	12	10	10
0.10	8x8	64	31	34	18	15	12	12
	16x16	137	67	81	18	16	14	14
	32x32	282	137	171	19	16	15	15
0.04	8x8	82	37	40	32	26	18	18
	16x16	185	84	100	34	28	24	23
	32x32	391	183	229	35	30	28	27
0.02	8x8	101	45	49	45	38	25	24
	16x16	236	94	117	52	43	32	30
	32x32	505	214	276	56	47	40	40
0.01	8x8	122	56	63	63	54	34	32
	16x16	304	115	148	79	65	44	44
	32x32	663	240	317	86	72	55	55

TABLE 5.4
Iteration Counts of (full) GMRES *for Example 4 (P2)*

ϵ	grid	no prec	$P, \|\|.\|\|_F$	$P, \|\|.\|\|_{F,A^{-1}}$	T	$T, D, \|\|.\|\|_F$	$T, P, \|\|.\|\|_F$	$T, P, \|\|.\|\|_{F,A^{-1}}$
1.00	8x8	40	18	18	13	11	5	5
	16x16	82	37	38	14	11	5	5
	32x32	167	77	78	14	11	6	5
0.20	8x8	47	22	21	14	13	7	7
	16x16	99	46	46	15	14	9	8
	32x32	202	94	96	15	15	9	9
0.10	8x8	56	25	24	19	17	11	10
	16x16	115	53	53	19	18	12	12
	32x32	239	109	111	20	18	13	13
0.04	8x8	76	32	32	33	27	17	16
	16x16	166	69	71	36	30	21	20
	32x32	336	148	153	36	33	24	24
0.02	8x8	94	40	41	47	39	23	22
	16x16	220	83	86	55	43	30	28
	32x32	461	180	193	58	50	37	37
0.01	8x8	116	50	52	64	55	32	29
	16x16	283	104	111	83	63	42	40
	32x32	620	214	234	92	71	54	52

TABLE 5.5
Iteration Counts of (full) GMRES *for Example 5 (Q2)*

[CS98] E. Chow and Y. Saad. Approximate inverse preconditioners via sparse-sparse iterations. *SIAM J. Sci. Comput.*, 19(3):995–1023 (electronic), 1998.

[ESW03] H. C. Elman, D. J. Silvester, and A. J. Wathen. Finite elements and fast iterative solvers. Manuscript, 2003.

[FMP90] V. Faber, T. A. Manteuffel, and S. V. Parter. On the theory of equivalent operators and application to the numerical solution of uniformly elliptic partial differential equations. *Adv. in Appl. Math.*, 11(2):109–163, 1990.

[GH97] M. J. Grote and T. Huckle. Parallel preconditioning with sparse approximate inverses. *SIAM J. Sci. Comput.*, 18(3):838–853, 1997.

[GHLC03] P. Guillaume, A. Huard, and C. Le Calvez. A block constant approximate inverse for preconditioning large linear systems. *SIAM J. Matrix Anal. Appl.*, 24(3):822–851 (electronic), 2003.

[GS98] N. I. M. Gould and J. A. Scott. Sparse approximate-inverse preconditioners using norm-minimization techniques. *SIAM J. Sci. Comput.*, 19(2):605–625 (electronic), 1998.

[Hol03] R. M. Holland. *Preconditioned Iterative Methods and Software.* DPhil thesis, Oxford University, 2003. Manuscript.

[KY86] L. Y. Kolotilina and A. Y. Yeremin. On a family of two-level preconditionings of the incomplete block factorization type. *Soviet J. Numer. Anal. Math. Modelling*, 1(4):293–320, 1986.

[KY93] L. Y. Kolotilina and A. Y. Yeremin. Factorized sparse approximate inverse preconditionings. I. Theory. *SIAM J. Matrix Anal. Appl.*, 14(1):45–58, 1993.

[MP90] T. A. Manteuffel and S. V. Parter. Preconditioning and boundary conditions. *SIAM J. Numer. Anal.*, 27(3):656–694, 1990.