Eidgenössische Technische Hochschule Zürich

Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Robust Parallel Smoothing for Multigrid Via Sparse Approximate Inverses

O. Bröker*, M.J. Grote, C. Mayer† and A. Reusken‡

Research Report No. 2000-13
November 2000

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

submitted to SIAM J. Sc. Comp.

*Departement Informatik, ETH Zürich, CH-8092 Zürich
†Fachbereich Mathematik, Universität Kaiserslautern, D-67653 Kaiserslautern
‡Institut für Geometrie und Praktische Mathematik, RWTH Aachen, D-52056 Aachen

Robust Parallel Smoothing for Multigrid Via Sparse Approximate Inverses

O. Bröker[§], M.J. Grote, C. Mayer[¶] and A. Reusken[‖]

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Research Report No. 2000-13 November 2000

Abstract

Sparse approximate inverses are considered as smoothers for multigrid. They are based on the SPAI-Algorithm (Grote and Huckle, 1997), which constructs a sparse approximate inverse M of a matrix A by minimizing $I - MA$ in the Frobenius norm. This yields a new hierarchy of smoothers: SPAI-0, SPAI-1, SPAI(ε). Advantages of SPAI smoothers over classical smoothers are inherent parallelism, possible local adaptivity and improved robustness. The simplest smoother, SPAI-0, is based on a diagonal matrix M. It is shown to satisfy the smoothing property for symmetric positive definite problems. Numerical experiments show that SPAI-0 smoothing is usually preferable to damped Jacobi smoothing. In more difficult situations, where the simpler SPAI-0 and SPAI-1 smoothers are not adequate, the SPAI(ε) smoother provides a natural procedure for improvement where needed. Numerical examples illustrate the usefulness of SPAI smoothing.

submitted to SIAM J. Sc. Comp.

[§]Departement Informatik, ETH Zürich, CH-8092 Zürich
[¶]Fachbereich Mathematik, Universität Kaiserslautern, D-67653 Kaiserslautern
[‖]Institut für Geometrie und Praktische Mathematik, RWTH Aachen, D-52056 Aachen

1 Introduction

Multigrid methods are efficient iterative solvers for large linear systems of equations, which result from the discretization of partial differential equations – see Brandt [8], Hackbusch [14, 15], Wesseling [25], and the references therein. They also yield efficient preconditioners when combined with Krylov subspace methods [7]. Any multigrid algorithm relies on the complementary interplay of smoothing and coarse grid correction. While the smoothing process aims at reducing the high-frequency error component, namely that which cannot be represented on coarser grids, the coarse grid correction solves for the low frequency error component, precisely that which is well represented on the coarser grid. The careful combination of both smoothing and coarse grid correction yields a multigrid iteration, which has a high convergence rate independent of the mesh size.

Standard smoothing techniques typically result from the application of a few steps of a basic iterative method. Here we shall consider smoothers that are based on sparse approximate inverses. Starting from the linear system

$$Ax = b, \tag{1}$$

we denote by M a sparse approximation of A^{-1}. Then, the corresponding basic iterative method is

$$x^{(k+1)} = x^{(k)} - M(Ax^{(k)} - b). \tag{2}$$

As the approximate inverse M is known explicitly, each iteration step requires only one additional $M \times v$ matrix-vector multiply; thus, it is easy to parallelize and cheap to evaluate, because M is sparse.

Recently, various algorithms have been proposed, all of which attempt to compute directly a sparse approximate inverse of A. Examples are the FSAI approach by Kolotilina and Yeremin [17], the MR algorithm by Chow and Saad [10], and the AINV approach by Benzi, Meyer, and Tuma [5]. Once computed, the approximate inverse M is applied as a preconditioner to the linear system (1) for use with a Krylov subspace iterative method. For a comparative study of various sparse approximate inverse preconditioners we refer to Benzi and Tuma [6]. By choosing an a priori sparsity pattern for M, the cost of computing M can be greatly reduced. Possible choices include powers of A or $A^\top A$, as suggested by Huckle [16] and Chow [11].

Approximate inverse techniques are also gaining in importance as smoothers for multigrid methods. First introduced by Benson and Frederickson [3, 4], they were shown to be effective on various difficult elliptic problems on unstructured grids by Tang and Wan [23]. Advantages of sparse approximate inverse smoothers over classical smoothers, such as damped Jacobi, Gauss-Seidel or ILU, are inherent parallelism, possible local adaptivity and improved robustness.

Here we shall consider sparse approximate inverse (SPAI) smoothers based on the SPAI-Algorithm by Grote and Huckle [13]. The SPAI-Algorithm computes an approximate inverse M explicitly by minimizing $I - MA$ in the Frobenius norm. Both

the computation of M and its application as a smoother are inherently parallel. Since an effective sparsity pattern of M is in general unknown a priori, the SPAI-Algorithm attempts to determine the most promising entries dynamically. This strategy has proved effective in generating preconditioners for many difficult and ill-conditioned problems (see Barnard, Bernardo, and Simon [1], Tang [22], and [13]). Moreover, it provides the means for adjusting the smoother locally and automatically, if necessary.

We shall consider the following hierarchy of sparse approximate inverse smoothers: SPAI-0, SPAI-1, and SPAI(ε). For SPAI-0 and SPAI-1 the sparsity pattern of M is fixed: M is diagonal for SPAI-0, whereas for SPAI-1 the sparsity pattern of M is that of A. For SPAI(ε) the sparsity pattern of M is determined automatically by the SPAI-Algorithm ([13]); the parameter ε controls the accuracy and the amount of fill-in of M.

Besides the SPAI smoothing operators, all other multigrid components, such as the prolongation, the restriction, and the coarse grid operators, result from standard choices. It is well-known that for certain classes of problems, such as convection-diffusion equations, a significant improvement in the efficiency of the multigrid solver can be obtained by using matrix-dependent prolongation and restriction operators (see [12, 19, 25, 27]). An interesting topic for future research is the combination of this new hierarchy of local and inherently parallel smoothers with algebraic multigrid techniques (see for instance [18, 20, 21, 24]).

In Section 2 we briefly review the SPAI-Algorithm and show how sparse approximate inverses are used as smoothers in multigrid. In Section 3 we prove that for SPAI-0 the smoothing property ([15]) holds under reasonable assumptions on the matrix A. More precisely, for A symmetric and positive definite, we prove that SPAI-0 satisfies the smoothing property, either if A is weakly diagonally dominant, or if A has at most seven nonzero off-diagonal entries per row. To our knowledge this is the first fairly general theoretical result on the smoothing property of iterative methods that are based on sparse approximate inverses. Previously Tang and Wan [23] analyzed the smoothing property of sparse approximate inverse smoothers for boundary value problems with constant coefficients on a two-dimensional regular grid. From a comparison of the SPAI-0 and damped Jacobi smoothers via numerical experiments, we conclude that the parameter-free SPAI-0 smoother is usually preferable to the damped Jacobi method. Finally, in Section 4, we present an extensive set of numerical experiments, which demonstrate the usefulness of SPAI smoothing.

2 Sparse approximate inverse smoothing

Starting from a standard multigrid setting, such as found in [14], ([15], Ch. 10), or [25], we recall some basic notions and briefly introduce relevant notation. We assume the following hierarchy of spaces,

$$X_\ell = \mathbb{R}^{n_\ell}, \qquad \ell = 0, 1, 2, \ldots, \quad n_0 < n_1 < n_2 < \ldots,$$

together with the prolongation and restriction operators

$$p: \ X_{\ell-1} \to X_\ell, \qquad r: \ X_\ell \to X_{\ell-1}, \quad \ell = 1, 2, \ldots \ .$$

To each space X_ℓ we associate a nonsingular operator,

$$A_\ell: \ X_\ell \to X_\ell.$$

We now wish to solve iteratively the linear system

$$A_{\ell_{\max}} x_{\ell_{\max}} = b_{\ell_{\max}}$$

by using a multigrid method. A multigrid iteration results from the recursive application of a two-grid method. A two-grid method on level ℓ consists of ν_1 presmoothing steps on level ℓ, a coarse grid correction on level $\ell - 1$, and ν_2 postsmoothing steps again on level ℓ. The corresponding error propagation is

$$e_\ell^{(m+1)} = [S_\ell^{\nu_2}(I - pA_{\ell-1}^{-1} r A_\ell) S_\ell^{\nu_1}]\, e_\ell^{(m)},$$

where S_ℓ denotes the iteration matrix of the smoother.

2.1 Classical smoothers

We shall limit the present discussion to the choice of the smoother. All other multigrid components, such as p, r, and $A_{\ell-1}$, follow from standard choices. If the smoother results from a consistent linear iterative method, the iteration matrix of the smoother, S_ℓ, can be written as

$$S_\ell = I - N_\ell A_\ell. \tag{3}$$

For instance, let $A = D+L+U$, with D the diagonal, L the lower triangular part, and U the upper triangular part of A. Then damped Jacobi smoothing corresponds to

$$S_\omega = I - \omega D^{-1} A, \tag{4}$$

whereas Gauss-Seidel smoothing corresponds to

$$S_{GS} = I - (D + L)^{-1} A. \tag{5}$$

In (4) the choice of ω must ensure good smoothing properties of the resulting damped Jacobi method. Yet the "optimal" value of ω is known only for certain model problems (see Sect. 3.2). In contrast, the Gauss-Seidel method is parameter-free and typically leads to improved smoothing over the damped Jacobi method. Unfortunately, the Gauss-Seidel method (5) is inherently sequential and therefore difficult to implement on a parallel architecture. Yet with an appropriate coloring of the unknowns (e.g., red-black ordering on a regular grid) it is sometimes possible to attain reasonable parallel efficiency with the Gauss-Seidel approach.

If neither damped Jacobi nor Gauss-Seidel leads to satisfactory smoothing, one can resort to more robust smoothers, such as the popular ILU smoothers based on the incomplete LU decomposition (ILU) of A_ℓ – see for instance [26]. Because each ILU smoothing step requires the solution of upper and lower triangular systems, it remains inherently sequential and difficult to implement in parallel. It is also difficult to improve the ILU smoother locally, say near the boundary or a singularity, without seriously affecting the sparsity of the LU factors.

2.2 SPAI smoothers

Most smoothers commonly used in multigrid methods, such as damped Jacobi, Gauss-Seidel, or ILU, have the form

$$x_\ell^{(k+1)} = x_\ell^{(k)} - W_\ell^{-1}(A_\ell x_\ell^{(k)} - b_\ell), \tag{6}$$

with W_ℓ a (sparse) approximation of A_ℓ; moreover, the computational cost of solving a linear system with matrix W_ℓ must be reasonable. In contrast, the SPAI smoothers lead to the iteration

$$x_\ell^{(k+1)} = x_\ell^{(k)} - M_\ell\,(A_\ell x_\ell^{(k)} - b_\ell), \tag{7}$$

where M_ℓ is sparse and explicitly known. Hence the iteration in (7) requires only matrix-vector multiplications and vector-vector additions, and no solution of a linear system; it is therefore easy to implement in a parallel environment.

To construct the sparse approximate inverse M of A, we shall minimize $I - MA$ in the Frobenius norm for a prescribed sparsity pattern of M – here we have dropped the index ℓ to simplify the notation. The Frobenius norm, denoted by $\|\cdot\|_F$, naturally leads to inherent parallelism because the rows m_k of M can be computed independently of one another. Indeed since

$$\|I - MA\|_F^2 = \sum_{k=1}^{n} \|e_k^\top - m_k A\|_2^2\ , \tag{8}$$

the solution of (8) separates into the n *independent least-squares problems* for the sparse (row) vectors m_k,

$$\min_{m_k} \|e_k^\top - m_k A\|_2, \qquad k = 1, \ldots, n. \tag{9}$$

Here e_k denotes the k-th unit vector. Because A and M are sparse these least-squares problems have small dimensions.

Since an effective sparsity pattern of M is usually unknown a priori, the original SPAI-Algorithm ([13]) begins with a diagonal pattern. Then the algorithm proceeds with augmenting the sparsity pattern of M to further reduce each residual $r_k = e_k^\top - m_k A$. The progressive reduction of the 2-norm of r_k involves two steps. First, the algorithm identifies a set of potential new candidates, based on the sparsity pattern of A and the current (sparse) residual r_k. Second, the algorithm selects

the most profitable entries, usually less than five entries, by computing for each candidate a cheap upper bound for the reduction in $\|r_k\|_2$. Once the new entries have been selected and added to m_k, the (small) least-squares problem (9) is solved again with the augmented set of indices. The algorithm proceeds until each row m_k of M satisfies

$$\|e_k^\top - m_k A\|_2 < \varepsilon. \tag{10}$$

Here ε is a tolerance set by the user, which controls the fill-in and the quality of the preconditioner M. A lower value of ε usually yields a more effective preconditioner, but the cost of computing $M = \text{SPAI}\,(\varepsilon)$ may become prohibitive; moreover, a denser M results in a higher cost per iteration in (7). The optimal value of ε minimizes the total time; it depends on the problem, the discretization, the desired accuracy, and the computer architecture. Further details about the original SPAI-Algorithm can be found in [13].

In addition to SPAI(ε), we shall also consider the following two greatly simplified SPAI smoothers with fixed sparsity patterns: SPAI-0, where M is diagonal, and SPAI-1, where the sparsity pattern of M is that of A. Both solve the least-squares problem (9), and thus minimize $\|I - MA\|_F$ for the sparsity pattern chosen a priori. This eliminates the search for the sparsity pattern of M, and thus greatly reduces the cost of computing the approximate inverse. The SPAI-1 smoother coincides with the SAI(0,1) smoother of Tang and Wan [23].

For SPAI-0, $M = \text{diag}(m_{kk})$ is diagonal and can be calculated directly:

$$m_{kk} = \frac{a_{kk}}{\|a_k\|_2^2}, \qquad 1 \le k \le n, \tag{11}$$

with a_k the k-th row of A. We note that M is always well-defined if A is non-singular. Unlike damped Jacobi, the SPAI-0 smoother is parameter-free.

To summarize, we shall consider the following hierarchy of SPAI smoothers, which all minimize $\|I - MA\|_F$ for a certain sparsity pattern of M.

SPAI-0: $M = \text{diag}(m_{kk})$ is diagonal, with m_{kk} given by (11).

SPAI-1: The sparsity pattern of M is that of A.

SPAI(ε): The sparsity pattern of M is determined automatically by the SPAI-Algorithm [13]. Then each row m_k satisfies (10) for a given ε.

We have found that in many situations, SPAI-0 and SPAI-1 yield ample smoothing. However, the added flexibility in providing an automatic criterion for improving the smoother via the SPAI-Algorithm remains very useful. Indeed, either SPAI-0 or SPAI-1 can be used as initial guess for SPAI(ε), and thus be locally improved upon where needed by reducing ε (see Section 4.2). For matrices with inherent (small) block structure, typical from the discretization of systems of partial differential equations, the Block-SPAI-Algorithm [2] greatly reduces the cost of computing M.

3 SPAI-0 smoothing

In this section we consider the simplest sparse approximate inverse smoother, SPAI-0. First, we shall show that SPAI-0 satisfies the smoothing property in two quite general situations. Second, we shall compare the two diagonal smoothers, SPAI-0 and damped Jacobi, via numerical experiments.

3.1 The smoothing property

From [14] and [15] we recall the following two conditions, which play a fundamental role in multigrid convergence theory:

1. The smoothing property ([15], Definition 10.6.3):

$$\|A_\ell S_\ell^\nu\|_2 \le \eta(\nu)\|A_\ell\|_2, \qquad \text{for all} \quad 0 \le \nu < \infty, \quad \ell \ge 1, \tag{12}$$

$$\eta(\nu) \text{ any function with } \lim_{\nu\to\infty} \eta(\nu) = 0.$$

2. The approximation property ([15], Section 10.6.3):

$$\|A_\ell - pA_{\ell-1}r\|_2 \le \frac{C_A}{\|A_\ell\|_2}, \qquad \text{for all} \quad \ell \ge 1. \tag{13}$$

Although we have stated these properties with respect to the Euclidean norm, other choices are possible. In general, the smoothing and approximation properties together imply convergence of the two-grid method and of the multigrid W-cycle, with a contraction number independent of the level number ℓ. Moreover, for symmetric positive definite problems, both conditions also imply multigrid V-cycle convergence independent of ℓ – see Hackbusch ([15], Sect. 10.6) for details.

The approximation property is independent of the smoother, S_ℓ; it depends only on the discretization $(A_\ell,\ A_{\ell-1})$, the prolongation operator p, and the restriction operator r. In [15] the approximation property is shown to hold for a large class of discrete elliptic boundary value problems. For symmetric positive definite problems the smoothing property usually holds for classical smoothers like damped Jacobi, (symmetric) Gauss-Seidel, and incomplete Cholesky. We shall now prove that the smoothing property (12) holds for SPAI-0 under reasonable assumptions on A_ℓ. To do so, we first recall (in a slightly simpler form) the following result for later reference.

Lemma 1 *(Lemma 10.7.4, [15]) Let A_ℓ and W_ℓ be symmetric and positive definite, and $S_\ell = I - W_\ell^{-1}A_\ell$. Assume that*

$$0 < A_\ell \le \Gamma W_\ell \quad \textit{for all} \quad \ell \ge 0 \quad \textit{with} \quad 0 < \Gamma < 2, \tag{14}$$

and that

$$\|W_\ell\|_2 \le C_W\|A_\ell\|_2, \qquad \forall \ell \ge 0. \tag{15}$$

Then S_ℓ satisfies the smoothing property (12), with

$$\eta(\nu) = C_W \max\{\eta_0(\nu), \Gamma|1-\Gamma|^\nu\}, \quad \eta_0(\nu) = \frac{1}{e\nu} + O(\nu^{-2}) \quad (\nu \to \infty) \tag{16}$$

In (14) and (15) both Γ and C_W must be independent of ℓ.

We shall now apply Lemma 1 to prove that SPAI-0 satisfies the smoothing property (12). To do so, we must show that W_ℓ satisfies (14) and (15), with $\Gamma < 2$. Here W_ℓ is the inverse of the diagonal approximate inverse defined in (11). Hence

$$W = \operatorname{diag}\left(\frac{\|a_i\|_2^2}{a_{ii}}\right), \tag{17}$$

where a_i denotes the i-th row of A – we have dropped the level index ℓ to simplify the notation. Since A is symmetric and positive definite, $a_{ii} > 0,\ 1 \le i \le n$, and thus W is positive definite.

Lemma 2 *Let W be given by (17). Furthermore, let p_i denote the number of nonzero off-diagonal entries in the i-th row of A, and assume that*

$$p \equiv \max_i p_i \le 7. \tag{18}$$

Then A satisfies $A \le \Gamma W$, with $\Gamma = (1 + \sqrt{1+p})/2 < 2$.

Proof:
We seek $\Gamma < 2$ such that $A \le \Gamma W$. First, we let $A = D - R$, with $D = \operatorname{diag}(A)$. Then

$$\begin{aligned} & \Gamma W - A \ge 0 \\ \iff \quad & \Gamma\left(D + \operatorname{diag}\left(\sum_{j\ne i}\frac{a_{ij}^2}{a_{ii}}\right)\right) - D + R \ge 0 \\ \iff \quad & (\Gamma-1)D + \Gamma\operatorname{diag}\left(\sum_{j\ne i}\frac{a_{ij}^2}{a_{ii}}\right) + R \ge 0. \end{aligned} \tag{19}$$

We note that the first two terms in (19) are diagonal matrices, while all off-diagonal entries are located in R. We now assume that $\Gamma \ge 1$, so that all entries on the main diagonal in (19) are non-negative. According to Gershgorin's theorem, for (19) to hold it is sufficient to have

$$\sum_{j\ne i}|a_{ij}| \le (\Gamma-1)a_{ii} + \Gamma\sum_{j\ne i}\frac{a_{ij}^2}{a_{ii}}, \qquad 1 \le i \le n. \tag{20}$$

Next, we divide (20) by a_{ii}, which yields the equivalent condition

$$\sum_{j\ne i}\beta_{ij} \le \Gamma - 1 + \Gamma\sum_{j\ne i}\beta_{ij}^2, \qquad 1 \le i \le n. \tag{21}$$

Here we have defined

$$\beta_{ij} = \frac{|a_{ij}|}{a_{ii}}.$$

Since p_i is the number of nonzero off-diagonal elements in row i of A, we conclude by Cauchy-Schwarz that

$$\sum_{j\neq i} \beta_{ij} \leq \sqrt{p_i} \sqrt{\sum_{j\neq i} \beta_{ij}^2} \leq \sqrt{p} \sqrt{\sum_{j\neq i} \beta_{ij}^2}, \qquad 1 \leq i \leq n.$$

Thus, for (21) to hold it is sufficient to have

$$\sqrt{p} \sqrt{\sum_{j\neq i} \beta_{ij}^2} \leq \Gamma - 1 + \Gamma \sum_{j\neq i} \beta_{ij}^2, \qquad 1 \leq i \leq n.$$

Therefore, since $x = \sum_{j\neq i} \beta_{ij}^2$ is real, non-negative, but otherwise arbitrary, it is sufficient to require that

$$\begin{aligned} & \sqrt{p}\,\sqrt{x} \leq \Gamma - 1 + \Gamma x, \qquad \forall x \in [0,\infty), \\ \iff \quad & \Gamma^2 x^2 + (2\Gamma(\Gamma-1) - p)x + (\Gamma-1)^2 \geq 0, \qquad \forall x \in [0,\infty), \\ (22) \qquad \iff \quad & (2\Gamma(\Gamma-1) - p)^2 - 4\Gamma^2(\Gamma-1)^2 \leq 0. \end{aligned}$$

The last inequality (22) is equivalent to

$$-4\Gamma(\Gamma-1)p + p^2 \leq 0,$$

which holds for $\Gamma = (1 + \sqrt{1+p})/2$. The assumption $p \leq 7$ yields $\Gamma < 2$.

□

Lemma 3 *Let W be given by (17), and assume that*

$$\max_i \sum_{j=1}^{n} \frac{a_{ij}^2}{a_{ii}^2} \leq \hat{C}. \tag{23}$$

Then W satisfies $\|W\|_2 \leq \hat{C}\,\|A\|_2$.

Proof:
The proof is immediate, since

$$\begin{aligned} \|W\|_2 &= \|\operatorname{diag}(a_{ii}) \operatorname{diag}\left(\frac{\|a_i\|_2^2}{a_{ii}^2}\right)\|_2 \\ &\leq \|\operatorname{diag}(a_{ii})\|_2 \|\operatorname{diag}\left(\frac{\|a_i\|_2^2}{a_{ii}^2}\right)\|_2 \\ &\leq \|A\|_2\,\hat{C}. \end{aligned}$$

□

From lemmas 1, 2, and 3 we now immediately conclude the following result.

Theorem 1 *Let A be symmetric positive definite, and let $S = I - MA$, with M the SPAI-0 preconditioner given by (11). Assume that the maximal number of nonzero off-diagonal entries in each row ($\equiv p$) is less than or equal to 7 (condition (18)), and that (23) holds. Then S satisfies the smoothing property (12), with $\eta(\nu)$ as in (16), $C_W = \hat{C}$, and $\Gamma = (1 + \sqrt{1+p})/2 < 2$.*

We remark that neither M-matrix properties nor diagonal dominance of A_ℓ are needed to show that the smoothing property holds for SPAI-0. In the context of a multigrid convergence analysis the constant $\hat{C}$ in (23) must be independent of the level number ℓ. Still, condition (23) is very mild and satisfied by most reasonable discretization schemes. Condition (18) is satisfied by standard second-order finite difference approximations of scalar elliptic boundary value problems in two or three space dimensions. It is also satisfied by linear finite element discretizations on a triangular mesh, if each node on the coarsest mesh has at most seven neighbors. This property is then transfered to all finer levels, if regular mesh refinement is used.

Next, we show that if A is weakly diagonally dominant, that is

$$\sum_{j \neq i} |a_{ij}| \leq |a_{ii}| \qquad \text{for all } i,$$

we may drop condition (18) and thus obtain another criterion for the smoothing property.

Theorem 2 *Let A be symmetric, positive definite, and weakly diagonally dominant. Furthermore, let $S = I - MA$, with M the SPAI-0 preconditioner given by (11), and assume that*

$$\sum_{j \neq i} \frac{a_{ij}^2}{a_{ii}^2} \geq C > 0. \tag{24}$$

Then S satisfies the smoothing property (12) with $\eta(\nu)$ as in (16), $C_W = 2$, and $\Gamma = 2/(1+C) < 2$.

Proof:
Again we seek $\Gamma < 2$ such that $A \leq \Gamma W$. To do so, we first follow the proof of lemma 2 until equation (21). Now, since A is weakly diagonally dominant, we have

$$\sum_{j \neq i} \beta_{ij} = \sum_{j \neq i} \frac{|a_{ij}|}{a_{ii}} \leq 1. \tag{25}$$

Hence for (21) to hold, it is sufficient to require

$$1 \leq \Gamma - 1 + \Gamma \sum_{j \neq i} \beta_{ij}^2, \qquad 1 \leq i \leq n,$$

which is equivalent to

$$\Gamma \geq \frac{2}{1 + \sum_{j \neq i} \beta_{ij}^2}, \qquad 1 \leq i \leq n.$$

Because of assumption (24) we can choose $\Gamma = 2/(1+C) < 2$.

To show that inequality (15) is indeed satisfied with $C_W = 2$, we first note that

$$\begin{aligned}\|W\|_2 = \|M^{-1}\|_2 &= \max_i \left(\frac{\|a_i\|_2^2}{a_{ii}} \right) \\ &= \max_i \left(a_{ii} + \frac{1}{a_{ii}} \sum_{j \neq i} a_{ij}^2 \right) \\ &\leq \max_i \left(a_{ii} + \frac{1}{a_{ii}} \left(\sum_{j \neq i} |a_{ij}| \right)^2 \right).\end{aligned}$$

The weak diagonal dominance of A then implies

$$\|W\|_2 \leq 2 \max_i a_{ii} \leq 2\|A\|_2.$$

□

Condition (24) is very mild and satisfied by most discretization schemes. It is not satisfied if a particular row i of A has the single entry a_{ii}. In that case, however, equation i is trivial and can be solved independently of the remaining equations.

In summary, we have shown for A symmetric positive definite that SPAI-0 satisfies the smoothing property, either if A is weakly diagonally dominant, or if the maximal number of nonzero off-diagonal entries per row is less than or equal to seven.

3.2 SPAI-0 versus Jacobi

Before we proceed with a comparison of the performance of these two diagonal smoothers via numerical experiments, we first point out a very special situation where SPAI-0 and damped Jacobi, with optimal relaxation parameter ω^*, lead to identical smoothers.

For the discrete Laplacian on a regular grid with periodic boundary conditions, the damping parameter ω^*, which is "optimal" with respect to smoothing, is known. Following the standard Fourier analysis in ([25], Section 7.3), we consider a regular d-dimensional equispaced mesh with n grid points in each dimension – for simplicity we assume n to be even. Then the eigenvalues of the discrete Laplacian with periodic boundary conditions are

$$\mu(\theta) = \frac{4}{h^2} \sum_{j=1}^{d} \sin^2 \left(\frac{\theta_j}{2} \right), \qquad \theta = (\theta_1, \theta_2, \ldots, \theta_d),$$

with $h = 1/n$ and $\theta_j \in \{-\pi + 2\pi h, -\pi + 4\pi h, \ldots, \pi\}$. Note that $\theta \in \prod_{j=1}^{d} [-\pi, \pi]$. The iteration matrix of the damped Jacobi method has eigenvalues

$$\lambda(\theta) = 1 - \frac{2\omega}{d} \sum_{j=1}^{d} \sin^2 \left(\frac{\theta_j}{2} \right).$$

We now concentrate on the high frequencies, which correspond to the subset

$$\bar{\Theta}_r = \prod_{j=1}^{d}[-\pi,\pi] \setminus \prod_{j=1}^{d}(-\frac{\pi}{2},\frac{\pi}{2}).$$

The amount of damping on the high-frequency components is measured by the *smoothing factor*,

$$\rho_{\rm sm}^{(\omega)} = \max_{\theta\in\bar{\Theta}_r} |\lambda(\theta)|.$$

It is independent of the mesh size but depends on ω. The optimal damping parameter, ω^*, is that for which the smoothing factor is minimal,

$$\rho_{\rm sm}^{(\omega^*)} = \min_{\omega} \rho_{\rm sm}^{(\omega)}.$$

To determine ω^* we first calculate $\rho_{\rm sm}^{(\omega)}$. For $\theta \in \bar{\Theta}_r$, the extreme values of $\sum_{j=1}^{d} \sin^2(\theta_j/2)$ are $1/2$ and d. Thus, we find that

$$\rho_{\rm sm}^{(\omega)} = \max\{|1-\frac{\omega}{d}|, |1-2\omega|\}.$$

The minimization of $\rho_{\rm sm}^{(\omega)}$ then yields the optimal damping parameter,

$$\omega^* = \frac{2d}{2d+1}. \tag{26}$$

Proposition 1 *Consider the discrete Laplacian with periodic boundary conditions in d space dimensions, which results from a standard second-order finite difference discretization on an equispaced grid. Then SPAI-0 and Jacobi smoothing, with optimal damping parameter ω^* given by (26), are identical.*

Proof:
Since for the discrete d-dimensional Laplacian we have $a_{ii} = (2d)h^{-2}$ and $\|a_i\|_2^2 = 2d(2d+1)h^{-4}$, for all $1 \le i \le n$, the approximate inverse M defined in (11) for SPAI-0 has the constant diagonal entry $m_{ii} = h^2/(2d+1)$. Therefore $M = \omega^* D^{-1}$ and the SPAI-0 and the damped Jacobi smoothers, with optimal damping parameter ω^* given in (26), coincide.

□

In this special situation, the parameter-free SPAI-0 smoother automatically yields a scaling of diag(A) which minimizes the smoothing factor; in that sense it is optimal.

Although both SPAI-0 and damped Jacobi yield diagonal smoothers, they typically differ, even with constant coefficients on an equispaced mesh. Indeed, the presence of boundary conditions modifies the rows of A_ℓ, which correspond to nodes on the boundary, and thus modifies SPAI-0 locally. We now compare quantitatively the performance of these two diagonal smoothers on the following class of model problems:

$$\begin{aligned} -\mathrm{div}(a(x,y)\nabla u(x,y)) &= 1, \qquad \text{in } \Omega \\ u(x,y) &= 0, \qquad \text{on } \partial\Omega. \end{aligned} \tag{27}$$

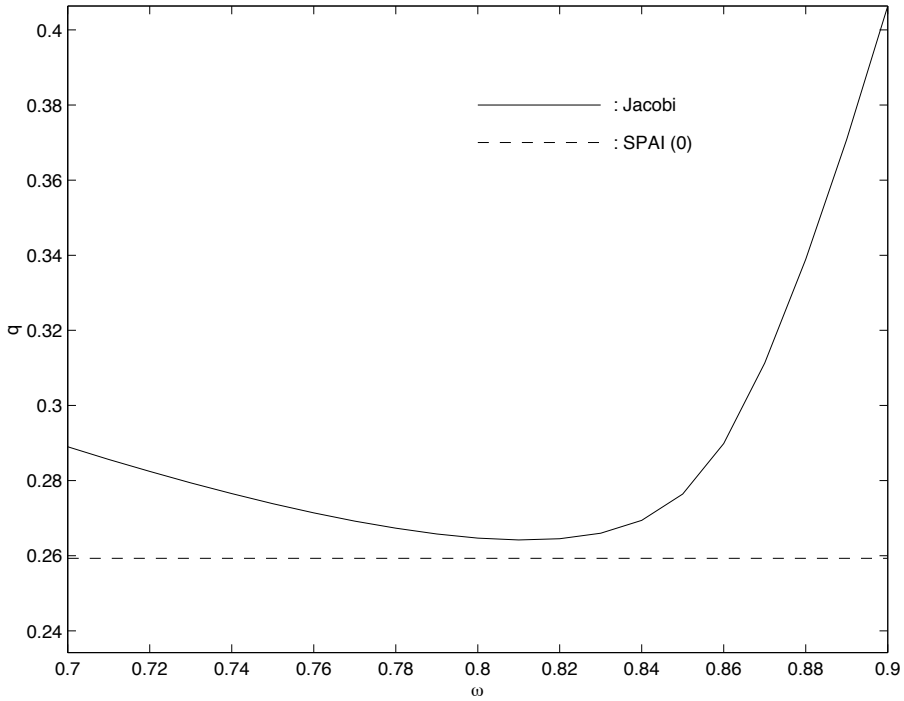

Figure 1: Comparison of the convergence rates q obtained with SPAI-0 and damped Jacobi for varying relaxation parameter ω – see (29) for the definition of q.

First, we choose $\Omega = (0,1) \times (0,1)$, set $a(x,y) \equiv 1$, and discretize the problem with continuous, piecewise linear finite elements on a triangular mesh. The various meshes are obtained successively by uniform refinement, starting from the coarsest mesh with a single unknown in the center of Ω. Here we use a multigrid V-cycle iteration with one pre- and one post-smoothing step ($\nu_1 = \nu_2 = 1$). We recall that SPAI-0 is parameter-free, whereas damped Jacobi contains the single relaxation parameter ω. In general the optimal damping parameter ω^* is unknown in advance and must be determined numerically in any given situation.

Is it possible for damped Jacobi to beat SPAI-0 by varying ω and thus determining the optimal value ω^* ? In Figure 1 we compare the convergence rate of SPAI-0 with that of Jacobi smoothing. The minimal convergence rate with damped Jacobi is attained with $\omega^* \simeq 0.81$, but the convergence rate obtained with SPAI-0 still lies slightly below it.

Next, we repeat the above numerical experiment, first for Ω the L-shaped domain, and second for discontinuous $a(x,y)$ with contrasts as high as 10^6. In all cases the overall picture remains the same: the convergence rate obtained with SPAI-0 smoothing consistently lies below that obtained with damped Jacobi smoothing for all values of ω. Although the improvement is usually small, and thus not really significant, it is remarkable because SPAI-0 is parameter-free.

In summary our numerical experiments with multigrid for 2D elliptic boundary value problems suggest that SPAI-0 is an attractive alternative to damped Jacobi.

Indeed SPAI-0 is parameter-free and typically leads to slightly better convergence rates than damped Jacobi.

4 Numerical results

To illustrate the usefulness and versatility of SPAI smoothing, we shall now consider various standard test problems. In all cases we use a regularly refined sequence of equispaced grids. The differential equation considered is discretized on the finest level with standard finite differences. The coarse grid operators are obtained via the Galerkin product formula, $A_{\ell-1} = r\, A_\ell\, p$, with $r = p^\top$ and p standard linear interpolation. We use a multigrid V-cycle iteration, with two pre- and two post-smoothing steps ($\nu_1 = \nu_2 = 2$) and $x_\ell^{(0)} = 0$ as initial guess. On the coarsest level ($\ell = 0$), we solve exactly for the single unknown remaining at the center of the domain. The iteration proceeds until the relative residual drops below the prescribed tolerance,

$$\frac{\|b - A_\ell\, x_\ell^{(m)}\|}{\|b\|} < 10^{-8}. \tag{28}$$

Then we calculate the average rate of convergence,

$$q = \left(\frac{\|b - A_\ell\, x_\ell^{(m)}\|}{\|b\|} \right)^{1/m}. \tag{29}$$

We recall that the ordering of the unknowns does not affect the SPAI-smoothing iteration (7), but that it does affect the Gauss-Seidel iteration. In all numerical examples we shall use lexicographic ordering of the unknowns.

Clearly, when comparing the performance of various smoothers, we cannot limit ourselves to comparing the number of multigrid iterations, but must also take into account the amount of arithmetic work due to the smoother. To do so, we calculate the total density ratio, ρ, of nonzero entries in M to those in A on all grid levels, $1 \le i \le \ell$, where smoothing is applied:

$$\rho = \frac{\sum_{i=1}^{\ell} \mathrm{nnz}(M_i)}{\sum_{i=1}^{\ell} \mathrm{nnz}(A_i)}. \tag{30}$$

Hence the additional amount of work due to the smoother is proportional to ρ. For a standard five-point stencil on a regular two-dimensional grid, $\rho_{\text{SPAI-0}} \simeq 1/5$, like damped Jacobi. Since $\rho_{\text{SPAI-1}} = 1$, the SPAI-1 smoother is about 67% times more expensive here than the SPAI-0 smoother. For SPAI(ε), the total density ratio, $\rho_{\text{SPAI}(\varepsilon)}$, depends on ε: it increases monotonically with decreasing ε. We remark that $\rho_{\text{SPAI}(\varepsilon)} < 1$, whenever SPAI($\varepsilon$) leads to a sparser approximate inverse than SPAI-1.

4.1 Poisson problem

We first consider the Poisson problem on the unit square with Dirichlet boundary conditions (27). In Table 1 we compare the convergence rates obtained with various smoothers. All SPAI smoothers lead to h-independent convergence rates. We observe a steady decrease of the convergence rate, q, for smaller values of ε, paralleled, of course, by an increase in ρ given in parenthesis. Note that SPAI-1 leads to a more effective but denser smoother than SPAI(ε) with $\varepsilon = 0.35$, yet the situation is reversed as ε decreases below 0.25.

Grid size	G-S	SPAI-0	SPAI-1	SPAI(0.35)	SPAI(0.25)
32×32	0.04	0.09	0.04	0.06 (0.7)	0.03 (1.5)
64×64	0.05	0.09	0.04	0.07 (0.7)	0.03 (1.5)
128×128	0.05	0.09	0.04	0.08 (0.7)	0.04 (1.5)

Table 1: Convergence rates q obtained with SPAI-0, SPAI-1, SPAI(ε), and Gauss-Seidel smoothing. The relative total density, ρ, defined in (30), is given in parenthesis. For SPAI-0, $\rho = 0.17$ and for SPAI-1, $\rho = 1$.

4.2 Locally anisotropic diffusion

We now consider the locally anisotropic diffusion problem

$$-\left(\nu(x,y)\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) = 1 \quad \text{in} \quad (0,1)\times(0,1), \tag{31}$$

with $u(x,y) = 0$ on the boundary. We set $\nu(x,y) = 1$ everywhere except inside the square $[1/4, 3/4] \times [1/4, 3/4]$, where $\nu(x,y) \equiv \nu$ is constant.

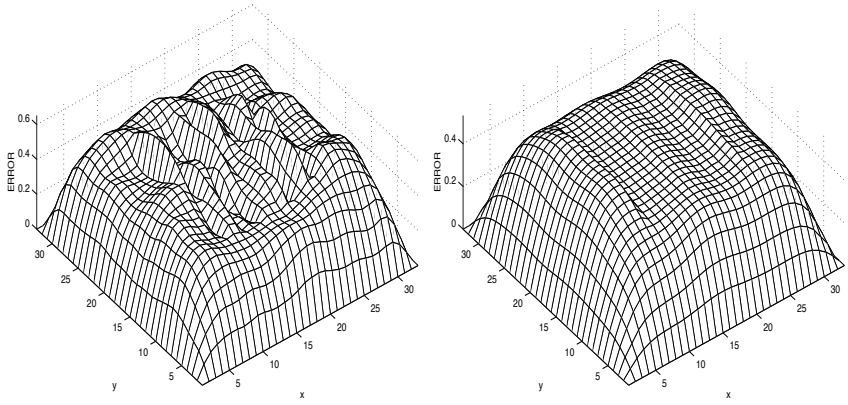

Figure 2: The error after five smoothing steps for the locally anisotropic diffusion problem with $\nu = 0.01$. Left: Gauss-Seidel; right: SPAI(0.25).

In Table 2 we compare the performance of Gauss-Seidel with SPAI smoothing for varying ν. For $\nu \leq 0.1$, the convergence rates for SPAI(0.4) lie consistently below those for Gauss-Seidel, while SPAI(0.25) leads to an even greater improvement. In particular, for $\nu = 0.01$ Gauss-Seidel smoothing barely converges, whereas SPAI(ε) with $\varepsilon \leq 0.4$ still yields acceptable convergence. Figure 2 shows that Gauss-Seidel is unable to smooth the error throughout Ω, mainly in the x-direction. In contrast, SPAI(ε) smoothing with $\varepsilon = 0.25$ results in a smooth error across the entire computational domain for $\nu = 0.01$. In Figure 3 we compare rows $e_j^\top A^{-1}$ and $e_j^\top M$, where M is computed with SPAI(0.25), for two unit basis vectors e_j. We consider e_j corresponding to the grid points $(1/2, 1/2)$ in the center of Ω, where $\nu = 0.01$, and $(1/8, 1/8)$ inside the surrounding region where $\nu = 1$. We observe how the approximate inverse, computed by the SPAI-Algorithm, captures the distinct local features of the true inverse. We recall that the sparsity pattern of M is not fixed a priori, but adapted automatically by the SPAI-Algorithm. For $\nu \leq 0.1$, the SPAI(0.4) smoother is not only more effective but also sparser than the SPAI-1 smoother. Hence the sophisticated search of the SPAI-Algorithm for an effective sparsity pattern of M clearly benefits the smoother.

Smoother	ν				
	1	10^{-1}	10^{-2}	10^{-3}	10^{-6}
Gauss–Seidel	0.05	0.42	0.89	0.97	0.98
SPAI-0	0.09	0.72	0.95	0.98	0.98
SPAI-1	0.04	0.37	0.89	0.97	0.98
SPAI(0.4)	0.12 (0.7)	0.16 (0.7)	0.81 (0.7)	0.95 (0.8)	0.97 (0.8)
SPAI(0.25)	0.04 (1.5)	0.07 (1.6)	0.37 (1.7)	0.75 (1.9)	0.87 (1.9)

Table 2: Locally anisotropic diffusion: convergence rates q for varying ν on a 128×128 grid. The relative density, ρ, is given in parenthesis.

These results demonstrate the usefulness of SPAI smoothing; they corroborate previous results obtained in [23], with $\nu = 0.01$ everywhere in Ω. Nevertheless, as the contrast in the anisotropy increases further, the convergence rates obtained with SPAI smoothing deteriorate as well. If the anisotropy occupies a small area of the domain of interest, say only within a boundary layer, further reducing ε enables to locally adjust the smoother and still reach acceptable convergence rates. But if the anisotropy is strong and present throughout the domain, the SPAI smoothers will become too dense, and thus too expensive.

Of course, various standard approaches combined with Gauss-Seidel smoothing, such as line relaxation or semi-coarsening, allow to circumvent some of these problems. However, these techniques are specifically designed for regular Cartesian grids in two space dimensions, when the anisotropy is usually constant and aligned with the grid. They are difficult to use on unstructured grids and become expensive in three dimensions. In contrast, the SPAI smoothers are not tied to a regular mesh or any special discretization; moreover, they do not encounter any particular hurdle in three dimensions.

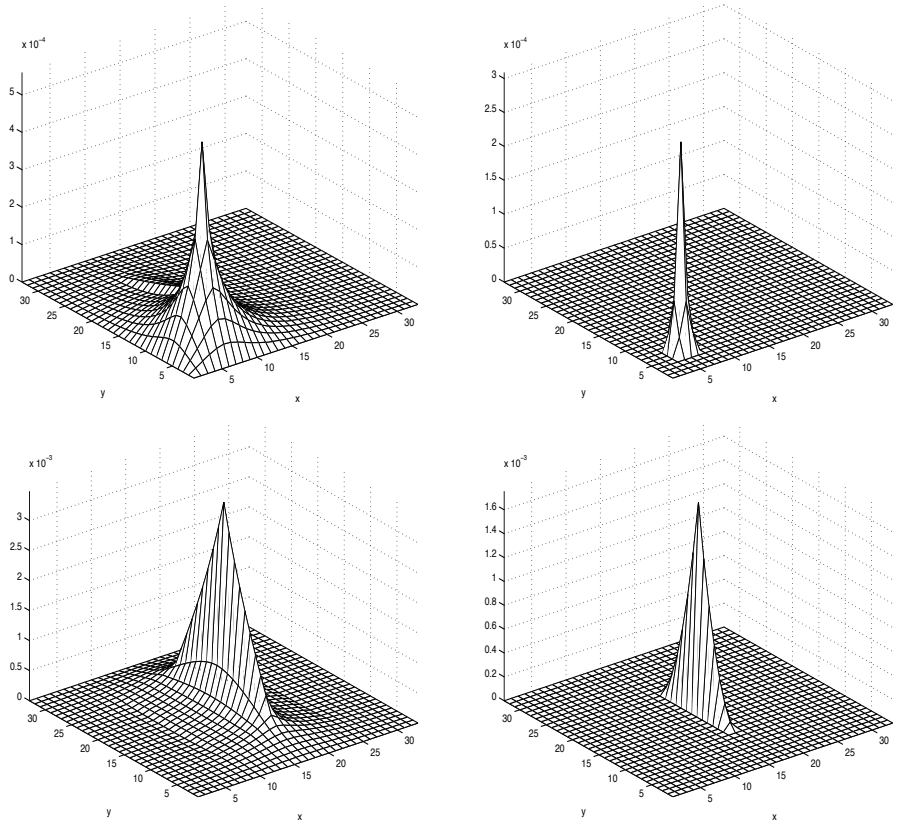

Figure 3: Comparison of a row of the true inverse, $e_j^\top A^{-1}$ (left), and of the approximate inverse, $e_j^\top M$ (right), for $\nu = 0.01$, where M is computed with SPAI(0.25). Top: e_j corresponds to the grid point $(1/8, 1/8)$; bottom: e_j corresponds to the grid point $(1/2, 1/2)$.

4.3 Convection-diffusion problems

We now consider

$$-\nu\Delta u(x,y) + \vec{v}(x,y)\cdot\nabla u(x,y) = 1, \quad \nu > 0, \tag{32}$$

in the unit square, where u vanishes on the boundary. Here u represents any scalar quantity advected by the flow field $\vec{v}$. For convection dominated flow, $\nu << h$, the linear systems cease to be symmetric and positive definite, so that these problems lie outside of classical multigrid theory. We use centered second-order finite differences for the diffusion, but discretize the convection with first-order upwinding to ensure numerical stability.

First, we consider a situation of unidirectional flow, with angle α from the x-axis, that is with constant flow direction $\vec{v}(x,y) = [\cos(\alpha), \sin(\alpha)]$. In Table 3 we compare the performance of Gauss-Seidel with SPAI smoothing on a regular 128×128 grid. For diffusion dominated flow, $\nu = 0.1$, the convergence rates obtained either with Gauss-Seidel or with SPAI smoothing are essentially independent on the flow direction.

Table 3: Constant flow direction with angle α from the x-axis: the convergence rates, q, obtained with different smoothers on a 128×128 grid, for the diffusion and the convection dominated cases, $\nu = 0.1$ and $\nu = 0.001$, respectively. The relative density, ρ, is given in parenthesis.

Problem	G-S$_G$	SPAI-0	SPAI-1	SPAI(0.35)	SPAI(0.25)
$\nu = 0.1, \alpha = 45°$	0.05	0.11	0.05	0.08 (0.8)	0.04 (1.6)
$\nu = 0.1, \alpha = 225°$	0.05	0.11	0.05	0.08 (0.8)	0.04 (1.6)
$\nu = 0.001, \alpha = 45°$	†	0.98	0.98	0.06 (1.7)	0.02 (2.2)
$\nu = 0.001, \alpha = 225°$	†	0.98	0.98	0.06 (1.7)	0.02 (2.2)

For convection dominated flow, the multigrid iteration with Gauss-Seidel smoothing does not converge, when the coarse grid operators are computed via Galerkin projection. To obtain a convergent scheme, one needs to compute the coarse grid operators via discrete coarse grid approximation, that is by discretizing (32) explicitly on all grid levels ([25], pp. 79). In that situation it is well-known that the convergence with Gauss-Seidel smoothing strongly depends on the ordering of the unknowns: the mere reversal of the flow direction (or equivalently the ordering of the unknowns) results in a large increase in the convergence rate, from $q = 0.17$ to $q = 0.55$; the contrast becomes even more striking for smaller values of ν. Indeed, when the flow direction is aligned with the ordering of the unknowns, the problem degenerates into a quasi-lower triangular system as $\nu \to 0$. In that situation, the Gauss-Seidel inverse, $(L + D)^{-1}$ essentially yields A^{-1} and thus results in rapid convergence. In contrast, the SPAI smoothers are not affected by the ordering of the unknowns, and thus yield identical results in Table 3 for both $\alpha = 45°$ and $\alpha = 225°$. For $\nu = 0.001$ the performance of SPAI-1 is poor, while SPAI(ε) with $\varepsilon \leq 0.35$ yields excellent convergence rates at little extra cost.

Next, we consider a situation of rotating flow, where u solves (32) with $\vec{v}(x, y) = [y - 1/2, 1/2 - x]$. In this special situation, shown in Figure 4, it is generally impossible to reorder the unknowns so that the entire system becomes lower triangular for vanishing ν. As a consequence, the multigrid iteration with Gauss-Seidel smoothing usually diverges for small ν. Convergence can be attained with symmetric Gauss-Seidel smoothing, that is by repeated application of the Gauss-Seidel iterations in natural and reverse ordering of the unknowns [25]; this approach does not generalize easily to unstructured grids. In contrast, SPAI-1 and SPAI(ε) lead to convergent multigrid iterations regardless of the flow pattern or the ordering of the unknowns, and without modifying additional components of the multigrid iteration.

In Table 4 we observe that all SPAI smoothers yield h-independent convergence rates in the diffusion dominated case, with $\nu = 0.1$. Although the convergence rate essentially doubles from SPAI-1 to SPAI-0 smoothing, the density ratio ρ drops from 1 to 0.17, which reduces the amount of work in applying the smoother. For SPAI(0.35) the convergence rate lies between those obtained with SPAI-0 and SPAI-1, and so does the relative density $\rho = 0.7$.

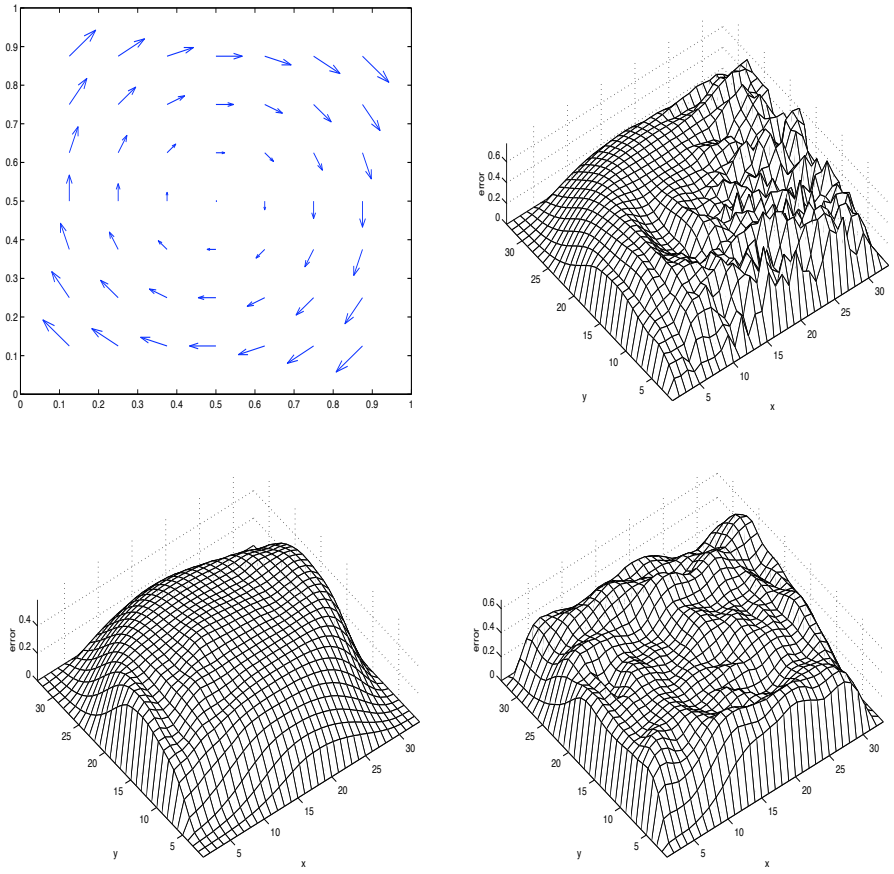

Figure 4: Rotating flow, $\nu = 10^{-3}$, 32×32 grid: flow field $\vec{v}(x, y)$ (upper left); the effect of three smoothing steps applied to an initially random error: Gauss-Seidel smoothing (upper right), ILU(0) smoothing (lower left), SPAI-1 smoothing (lower right). Note that Gauss-Seidel is unable to smooth the error throughout the domain.

Finally, we set $\nu = 10^{-3}$ and thus consider a convection dominated rotating flow, that is $\nu << h$. Both the convergence rates and the relative densities ρ are shown in Table 5 for different smoothers. Neither Gauss-Seidel nor SPAI-0 smoothing lead to a convergent multigrid iteration. Indeed Gauss-Seidel is unable to smooth the error throughout the domain due to the absence of a dominant "unidirectional" flow direction, as shown in Figure 4. Again we note that convergence may be achieved by using symmetric Gauss-Seidel smoothing [25]. Our attempt to use ILU(0) on the 128×128 grid eventually failed because of numerical instability encountered in the computation of the incomplete LU-decomposition – alternative variants of ILU would probably work [26].

Table 5 illustrates the typical behavior of SPAI smoothing versus ε. It shows that reducing ε in the SPAI Algorithm produces a steady reduction in the convergence rate. Hence a greater reduction of $\|I - MA\|_F$ typically yields an improved smoother. Of course, as ε decreases, both the work in computing and in applying the smoother M rapidly increase, so that the optimal value of ε with the smallest total time

Table 4: Diffusion dominated case, $\nu = 0.1$: the convergence rates q obtained with various smoothers. For SPAI(0.35) the relative density is $\rho = 0.7$.

	Gauss-Seidel	SPAI-0	SPAI-1	SPAI(0.35)
64×64	0.04	0.10	0.04	0.07
128×128	0.05	0.10	0.04	0.08
256×256	0.05	0.10	0.04	0.08

probably lies between 0.2 and 0.5 for this problem. For SPAI(0.4) both q and ρ remain essentially constant as the mesh is refined. As we compare the performance of SPAI-1 with that of SPAI(0.4), we remark that both the convergence rate q and the total density ratio ρ are reduced. Therefore the sophisticated search of the original SPAI-Algorithm ([13]) benefits the smoother by selecting an effective sparsity pattern for M; clearly, the increase in the cost of computing M is worthwhile when memory resources are critical. It may even pay off in reducing total time, since fewer nonzero entries in M results in a cheaper smoother. Again these considerations are problem and architecture dependent; hence the importance of providing the user with a simple but effective way to tune the algorithm, here by adjusting the value of ε.

Table 5: Convection dominated case, $\nu = 10^{-3}$: the convergence rates q obtained with various smoothers. The values of ρ are given in parenthesis.

	SPAI-1	SPAI (0.5)	SPAI (0.4)	SPAI (0.3)	SPAI (0.2)
128×128	0.61	0.67 (0.4)	0.42 (0.6)	0.22 (1.4)	0.09 (3.6)
256×256	0.68	0.75 (0.2)	0.45 (0.6)	0.31 (1.3)	0.12 (3.2)

Table 6: Robust convergence with SPAI (0.3) smoothing with respect to ν on a 128×128 grid: both the convergence rate q and the relative density ρ are shown.

	$\nu = 1$	$\nu = 0.1$	$\nu = 0.01$	$\nu = 10^{-3}$	$\nu = 10^{-4}$	$\nu = 10^{-5}$	$\nu = 10^{-6}$
q	0.07	0.07	0.05	0.22	0.73	0.74	0.75
ρ	0.86	0.85	0.92	1.41	2.11	2.29	2.31

What if we decrease ν while keeping the grid spacing fixed ? Do we obtain a robust multigrid algorithm (in the sense of [26]) ? In Table 6 we show the convergence rate and density obtained with SPAI (0.3) on the 128×128 grid for varying ν. Although both q and ρ vary throughout the range of values for ν, they remain bounded as $\nu \to 0$.

5 Conclusion

Our numerical results show that sparse approximate inverses based on minimizing the Frobenius norm are an attractive alternative to classical Jacobi or Gauss-Seidel smoothing. For symmetric positive definite problems, SPAI-1 typically behaves like Gauss-Seidel, whereas SPAI-0, which is parameter-free, usually has a slight edge over damped Jacobi with optimal relaxation parameter. Moreover, our proof of the smoothing property for SPAI-0 applies also in situations, where that of Jacobi smoothing cannot be shown. For convection dominated flow problems, such as rotating flow, the ordering independence of SPAI-1 leads to a more robust smoother than Gauss-Seidel. In situations, where neither SPAI-0 nor SPAI-1 suffice, SPAI(ε) automatically improves the smoother locally where needed. Both the computation and the application of the SPAI smoothers are inherently parallel.

Nevertheless, our results also show the limitations of SPAI smoothing in difficult situations, such as strong anisotropy, where the lack of isotropic smoothing needs to be compensated by appropriate prolongation and restriction operators. It is very interesting to combine this new hierarchy of local and inherently parallel smoothers with matrix-dependent coarsening strategies, such as found in algebraic multigrid [21]. The first two authors are currently pursuing these issues and will report on them elsewhere [9] in the near future.

The C/MPI version of the SPAI-Algorithm ([1, 2]) with Matlab and PETSc interfaces can be downloaded from the following address:
`http://www.sam.math.ethz.ch/~grote/spai/`.

References

[1] S. T. Barnard, L. M. Bernardo, and H. D. Simon, *An MPI implementation of the SPAI preconditioner on the T3E*, Intl. J. High Perf. Comput.

Appl. 13, 1999, pp. 107–128.

[2] S. T. Barnard and M. J. Grote, *A block version of the SPAI preconditioner*, in Proc. of the 9th SIAM conference on Parallel Processing for Scientific Computing, held in San Antonio, TX, March 1999.

[3] M. W. Benson and P. O. Frederickson, *Iterative solution of large sparse linear systems arising in certain multidimensional approximation problems*, Utilitas Mathematica 22, 1982, pp. 127–140.

[4] M. W. Benson, *Frequency domain behavior of a set of parallel multigrid smoothing operators*, Intern. J. Computer Math., 36 (1990), pp. 77–88.

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

[6] M. Benzi and M. Tuma, *A comparative study of sparse approximate inverse preconditioners*, Appl. Num. Math. 30, 1999, pp. 305–340.

[7] J. Bramble, J. Pasciak, and J. Xu, *Parallel multilevel preconditioners*, Math. Comp. 31, 1990, pp. 333–390.

[8] A. Brandt, *Multi-level adaptive solution to boundary-value problems*, Math. Comp. 31, pp. 333–390, 1977.

[9] O. Bröker and M. J. Grote, *Parallel sparse approximate inverse smoothers for geometric and algebraic multigrid*, Appl. Num. Math., in preparation.

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

[11] E. Chow, *A priori sparsity patterns for parallel sparse approximate inverse preconditioners*, SIAM J. Sc. Comput. 21, 2000, pp. 1804–1822.

[12] T. Grauschopf, M. Griebel, H. Regler, *Additive multilevel-preconditioners based on bilinear interpolation, matrix-dependent geometric coarsening and algebraic multigrid coarsening for second order elliptic PDESs*, Appl. Numer. Math. 23, 1997, pp. 63–96.

[13] M. J. Grote and T. Huckle, *Parallel preconditioning with sparse approximate inverses*, SIAM J. Sci. Comput., 18, 1997, pp. 838–853.

[14] W. Hackbusch, *Multi-grid Methods and Applications*, Springer-Verlag, Berlin, 1985.

[15] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, Springer-Verlag, New York, 1994.

[16] T. Huckle, *Approximate sparsity patterns for the inverse of a matrix and preconditioning*, Appl. Num. Math. 30, pp. 291–303, 1999.

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

[18] A. Krechel and K. Stüben, *Operator dependent interpolation in algebraic multigrid*, In: Multigrid Methods V (eds. W. Hackbusch and G. Wittum), Lecture Notes in Computational Science and Engineering 3, 1998, pp. 189–211.

[19] A. Reusken, *On a robust multigrid solver*, Computing 56, 1996, pp. 303–322.

[20] A. Reusken, *On the approximate cyclic reduction preconditioner*, SIAM J. Scientific Comp. 21 , 2000, pp. 565–590.

[21] J. W. Ruge and K. Stüben, *Algebraic multigrid*, in Multigrid Methods, S. F. McCormick, ed., SIAM, Philadelphia, PA, 1987, pp. 73–130.

[22] W.-P. Tang, *Toward an effective approximate inverse preconditioner*, SIAM J. Matrix Anal. Appl. 20, 1999, pp. 970–986.

[23] W.-P. Tang and W. L. Wan, *Sparse approximate inverse smoother for multi-grid*, SIAM J. Matrix Anal. Appl. 21, 2000, pp. 1236–1252.

[24] P. Vanek, J. Mandel and M. Brezina, *Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems*, Computing 56, 1996, pp. 179–196.

[25] P. Wesseling, *An Introduction to Multigrid Methods*, John Wiley, Chichester, UK, 1992.

[26] G. Wittum, *On the robustness of ILU-smoothing*, SIAM J. Sci. Stat. Comput. 10, 1989, pp. 699–717.

[27] de Zeeuw, P. M., *Matrix prolongations and restrictions in a black-box multigrid solver*, J. Comp. and Appl. Math. 33, 1990, pp. 1–27.

Research Reports

No.	Authors	Title
00-13	O. Bröker, M.J. Grote, C. Mayer, A. Reusken	Robust Parallel Smoothing for Multigrid Via Sparse Approximate Inverses
00-12	C. Lasser, A. Toselli	An overlapping domain decomposition preconditioner for a class of discontinuous Galerkin approximations of advection-diffusion problems
00-11	J. Liesen, M. Rozložník, Z. Strakoš	On Convergence and Implementation of Minimal Residual Krylov Subspace Methods for Unsymmetric Linear Systems
00-10	C. Schwab, O. Sterz	A scalar boundary integrodifferential equation for eddy current problems using an impedance boundary condition
00-09	M.H. Gutknecht, M. Rozložník	By how much can residual minimization accelerate the convergence of orthogonal error methods?
00-08	M. Rozložník, V. Simoncini	Short-term recurrences for indefinite preconditioning of saddle point problems
00-07	P. Houston, C. Schwab, E. Süli	Discontinuous *hp*-Finite Element Methods for Advection-Diffusion Problems
00-06	W.P. Petersen	Estimation of Weak Lensing Parameters by Stochastic Integration
00-05	M.H. Gutknecht	A Matrix Interpretation of the Extended Euclidean Algorithm
00-04	M.J. Grote	Nonreflecting Boundary Conditions for Time Dependent Wave Propagation
00-03	M.H. Gutknecht	On Lanczos-type methods for Wilson fermions
00-02	R. Sperb, R. Strebel	An alternative to Ewald sums. Part 3: Implementation and results
00-01	T. Werder, K. Gerdes, D. Schötzau, C. Schwab	*hp* Discontinuous Galerkin Time Stepping for Parabolic Problems
99-26	J. Waldvogel	Jost Bürgi and the Discovery of the Logarithms
99-25	H. Brunner, Q. Hu, Q. Lin	Geometric meshes in collocation methods for Volterra integral equations with proportional time delays
99-24	D. Schötzau, Schwab	An *hp* a-priori error analysis of the DG time-stepping method for initial value problems
99-23	R. Sperb	Optimal sub- or supersolutions in reaction-diffusion problems