Symmetric Positive Definite Based Preconditioners For Discrete Convection-diffusion Problems

Mohammad Sameer Sunhaloo
School of Business Informatics and Software Engineering
University of Technology Mauritius
La Tour Koenig, Mauritius.

Muddun Bhuruth
Department of Mathematics
University of Mauritius
Reduit, Mauritius.

Abstract — We experimentally examine the performance of preconditioners based on entries of the symmetric positive definite part and small subspace solvers for linear system of equations obtained from the high-order compact discretization of convection-diffusion equations. Numerical results are described to illustrate that the preconditioned GMRES algorithm converges in a reasonable number of iterations.

Keywords: preconditioner, convection-diffusion equation, high-order compact discretization.

1.0 Introduction

An important aspect in the numerical solution of convection diffusion problems is the use of efficient iterative methods for solving the large sparse nonsymmetric linear systems. GMRES and the BiCSGTAB algorithm [9] are two widely used nonsymmetric iterative solvers. In order to obtain fast convergence, some form of preconditioning of the linear system is necessary. Incomplete LU factorizations [7] and sparse approximate inverses [3, 5] are two popular preconditioning strategies. For HOC schemes, such preconditioners have been considered by Zhang [12] and Gopaul and Bhuruth [2].

Xu and Cai [11] proposed a different preconditioning approach for second-order elliptic equations with first derivative terms where the coefficients of the lower order derivatives are not too large. For such systems, the matrix A corresponding to the discretised equations is dominated by its symmetric positive definite part. Xu and Cai [11] motivate their preconditioning approach by mentioning that the symmetric positive definite (SPD) part governs the equation very well on the propagation of the high frequency modes but poorly on the low frequencies and they employ an exact solver for the nonsymmetric system on a coarser mesh to smooth out the low-frequency modes. An optimal preconditioner is then constructed using the small subspace solver and a preconditioner for the SPD operator.

In this paper we carry out an experimental study of the efficiency of SPD-based preconditioners for HOC linear systems. In §**2**, we briefly describe high-order compact (HOC) discretizations for two-dimensional convection-diffusion problems . In §**3**, we give a brief account of the relevant theory following the work by [10] and [11] and in §**4** we discuss the applications to HOC linear systems.

2.0 High-Order Compact Discretizations

We consider the two-dimensional constant-coefficient convection-diffusion problem

$$-\epsilon\Delta u(x,\, y) + \mathrm{w}\cdot\nabla u(x,\, y) = f(x,\, y), \tag{1}$$

in the unit square $\Omega = (0,\, 1)^2)$ with Dirichlet boundary conditions $u(x,\, y) = g(x,\, y)$ on $\partial\Omega$. In (1), $\epsilon > 0$ is the diffusivity parameter, w is the convective velocity field and $f(x,\, y)$ is the source term.

To derive the HOC schemes for constant coefficients two dimensional problems of the form (1), we take the velocity field as $\mathrm{w} = (p,\, q)$ in (1), where p and q are known positive constants. Central difference approximations on a uniform mesh with mesh size $h = 1/(n+1)$ in both directions, the representative difference equation at an interior node $(i,\, j)$ is

$$-\epsilon\left(\delta_x^2 + \delta_y^2\right)u_{ij} + \left(p\delta_x + q\delta_y\right)u_{ij} = f_{ij} + T_{ij}, \tag{2}$$

with associated truncation error given by

$$T_{ij} = \frac{h^2}{12}\left[2\left(p\frac{\partial^3 u}{\partial x^3} + q\frac{\partial^3 u}{\partial y^3}\right) - \epsilon\left(\frac{\partial^4 u}{\partial x^4} + \frac{\partial^4 u}{\partial y^4}\right)\right]_{ij} + \mathcal{O}(h^4). \tag{3}$$

Expressing the leading terms in T_{ij} in terms of the lower-order derivatives, we obtain the HOC scheme [4, 6, 8]

$$\begin{aligned} &-\left(\epsilon + \frac{p^2h^2}{12\epsilon}\right)\delta_x^2 u_{ij} - \left(\epsilon + \frac{q^2h^2}{12\epsilon}\right)\delta_y^2 u_{ij} + p\delta_x u_{ij} + q\delta_y u_{ij} \\ &\qquad + \frac{h^2}{6}\left(q\delta_x^2\delta_y + p\delta_x\delta_y^2 - \frac{pq}{\epsilon}\delta_x\delta_y - \epsilon\delta_x^2\delta_y^2\right)u_{ij} \\ &\qquad\qquad = f_{ij} + \frac{h^2}{12}\left(\delta_x^2 + \delta_y^2 - \frac{p}{\epsilon}\delta_x - \frac{q}{\epsilon}\delta_y\right)f_{ij}. \end{aligned} \tag{4}$$

Considering a natural rowwise ordering of the unknowns u_{ij} at the nodes (x_i, y_j), the scheme (4) leads to a coefficient matrix A which has the block-tridiagonal structure

$$A = \text{blocktridiagonal}[L,\, K,\, M],$$

with the matrices L, K and M given by

$$\begin{aligned} L &= \text{tridiagonal}\left[-\left(\frac{\epsilon}{h^2} + \frac{p+q}{2h} + \frac{pq}{4\epsilon}\right),\, \frac{4\epsilon}{h^2} + \frac{2q}{h} + \frac{q^2}{2\epsilon},\, -\left(\frac{\epsilon}{h^2} - \frac{p-q}{2h} - \frac{pq}{4\epsilon}\right)\right], \\ K &= \text{tridiagonal}\left[-\left(\frac{4\epsilon}{h^2} + \frac{2p}{h} + \frac{p^2}{2\epsilon}\right),\, \frac{20\epsilon}{h^2} + \frac{p^2}{\epsilon} + \frac{q^2}{\epsilon},\, -\left(\frac{4\epsilon}{h^2} - \frac{2p}{h} + \frac{p^2}{2\epsilon}\right)\right], \\ M &= \text{tridiagonal}\left[-\left(\frac{\epsilon}{h^2} + \frac{p-q}{2h} - \frac{pq}{4\epsilon}\right),\, \frac{4\epsilon}{h^2} - \frac{2q}{h} + \frac{q^2}{2\epsilon},\, -\left(\frac{\epsilon}{h^2} - \frac{p+q}{2h} + \frac{pq}{4\epsilon}\right)\right]. \end{aligned} \tag{5}$$

3.0 Iterative Methods by SPD and Small Subspace Solvers

Suppose that $A = A_1 + A_2$ where A_1 is symmetric positive definite with respect to the usual inner product in $\mathbb{R}^N$ and let $\|\cdot\|_{A_1}$ denote the norm induced by A_1 on $\mathbb{R}^N$. Starting with an initial guess $\mathrm{u}^{(0)}$, suppose that the following iteration

$$\mathrm{u}^{(m+1)} = u^{(m)} + B\left(\mathrm{b} - A\mathrm{u}^{(m)}\right), \quad m \geq 0, \tag{6}$$

is used to solve the corresponding HOC system. A sufficient condition for the convergence of the iteration (6) is that

$$\rho = \|I - BA\|_{A_1} < 1,$$

and in this case, the number ρ is the convergence factor of the iteration.

Let $\mathcal{V}_0$ be a subspace of $\mathbb{R}^N$ such that for any given $\mathrm{u} \in \mathbb{R}^N$, there exists a unique $\hat{\mathrm{u}}_0 \in \mathcal{V}_0$ such that

$$(A\hat{\mathrm{u}}_0,\, \mathrm{v}) = (A\mathrm{u},\, \mathrm{v}) \;\; \forall\, \mathrm{v} \in \mathcal{V}_0. \tag{7}$$

We note that (7) defines an operator $P_0 : \mathbb{R}^N \mapsto \mathcal{V}_0$ such that $P_0\mathrm{u} = \hat{\mathrm{u}}_0$. Also define $A_0 : \mathcal{V}_0 \mapsto \mathcal{V}_0$ and a projection $Q_0 : \mathbb{R}^N \mapsto \mathcal{V}_0$ by

$$(A_0\hat{\mathrm{u}}_0,\, \mathrm{v}_0) = (A\hat{\mathrm{u}}_0,\, \mathrm{v}_0)\; \forall \hat{\mathrm{u}}_0,\, \mathrm{v}_0 \in \mathcal{V}_0,$$

and

$$(Q_0\mathrm{u},\, \mathrm{v}_0) = (\mathrm{u},\, \mathrm{v}_0)\; \forall \mathrm{u} \in \mathbb{R}^N,\, \mathrm{v}_0 \in \mathcal{V}_0,$$

respectively. We then have

$$A_0 P_0 = Q_0 A. \tag{8}$$

Now consider the iteration given by

$$\mathrm{v}^{(i+1)} = \mathrm{v}^{(i)} + B_1\left(\mathrm{g} - A_1\mathrm{v}^{(i)}\right),\; i \geq 0, \quad \mathrm{v}^{(0)} \quad \text{given}, \tag{9}$$

for the SPD system

$$A_1\mathrm{v} = \mathrm{g}.$$

The iterative algorithm for solving

$$Au = f \tag{10}$$

consists of two steps: a first step is the correction on the coarse space $\mathcal{V}_0$ and the second step consists of smoothing by the SPD matrix A_1.

Algorithm: Given $\mathrm{u}^{(0)} \in \mathbb{R}^N$

(1). Find the exact solution on $\mathcal{V}_0$ of
$A_0\hat{\mathrm{u}}_0 = Q_0(\mathrm{f} - A\mathrm{u}^{(m)}),\; m \geq 0$

(2). Let $\mathrm{g} = \mathrm{f} - A(\mathrm{u}^k + \hat{\mathrm{u}}_0)$ and apply p steps of iteration (9) to the system $A_1\mathrm{v} = \mathrm{g}$ to get an improved solution $\mathrm{v}^{(p)}$.

(3). $\mathrm{u}^{(m+1)} = \mathrm{u}^{(m)} + \hat{\mathrm{u}}_0 + \mathrm{v}^{(p)},\; m \geq 0.$

The convergence of the above algorithm depends on the following parameter

$$\varrho_0 = \sup_{\mathrm{u},\mathrm{v}\in\mathbb{R}^N} \frac{(A_2\,(I-P_0)\,\mathrm{u},\,\mathrm{v})}{\|\mathrm{u}\|_{A_1}\|\mathrm{v}\|_{A_1}}. \tag{11}$$

In [11], Xu and Cai assumed that that $\varrho_0 < 1$ and let $E = I - BA$. Then they obtained

$$\|E\|_{A_1} \leq \sigma, \tag{12}$$

where

$$\sigma = \varrho^p + \varsigma_0, \quad \varrho = \|I - B_1 A_1\|_{A_1}, \quad \text{and} \quad \varsigma_0 = \frac{4\varrho_0}{1-\varrho_0}.$$

A consequence of (12) is that the iterative algorithm for the nonsymmetric problem is always convergent as long as the iterative algorithm for the SPD system is convergent.

3.1 Preconditioners for GMRES

The above approach for designing iterative methods can be used to derive preconditioners for Krylov subspace methods such as GMRES. The preconditioning strategy proposed in [11] is based on the following assumptions:

1. there exists a constant c_1 such that
$$\|\mathrm{u}\|_2 \leq c_1\|\mathrm{u}\|_{A_1}, \forall \mathrm{u}\in\mathbb{R}^N,$$
2. that A_2 satisfies
$$(A_2\mathrm{u},\,\mathrm{v}) \leq c_2\|\mathrm{u}\|_2\|\mathrm{v}\|_{A_1},\ \forall\,\mathrm{u},\,\mathrm{v}\in\mathbb{R}^N.$$

The preconditioner for A then takes the form

$$M = A_0^{-1}Q_0 + \tau B_1, \tag{13}$$

where B_1 is a given symmetric positive definite preconditioner for A_1 and τ is a positive constant. Combining (8) and (13), yields

$$MA = P_0 + \tau B_1 A. \tag{14}$$

Let

$$\kappa_0 = \sup_{\mathrm{u}\in\mathbb{R}^N} \frac{\|(I-P_0)\mathrm{v}\|_2}{\|\mathrm{v}\|_{A_1}}.$$

Now, let λ_0 and λ_1 be the minimal and maximal eigenvalues of B_1A_1, respectively. Xu and Cai proved in [11] that there exists positive constants ϵ, α, τ and μ, depending on λ_0 and λ_1 such that, for the preconditioner given by (13), and if $\varrho_0 \leq \epsilon$,

$$\|MA\mathrm{u}\|_{A_1} \leq \mu\|\mathrm{u}\|_{A_1} \quad \forall\,\mathrm{u}\in\mathbb{R}^N, \tag{15}$$

and

$$(MA\mathrm{u},\,\mathrm{u})_{A_1} \geq \alpha\,(\mathrm{u},\,\mathrm{u})_{A_1}, \quad \forall\,\mathrm{u}\in\mathbb{R}^N. \tag{16}$$

From (15) and (16), it can be shown that the GMRES convergence rate estimate is given by

$$\frac{\|\mathrm{r}_m\|_{A_1}}{\|\mathrm{r}_0\|_{A_1}} \leq \left(1 - \frac{\alpha^2}{\mu^2}\right)^{m/2}. \tag{17}$$

The implication of (17) is that if the condition number of B_1A_1 is uniformly bounded, then the GMRES method will have a good convergence rate.

4.0 Applications To HOC Systems

We describe some numerical experiments using the preconditioner (13) for GMRES. We consider problems of the form (10) arising from HOC approximations of elliptic equations. We consider different preconditioners for the symmetric part of A. Specifically, we consider Jacobi preconditioning (JAC),Incomplete Choleski factorization (ICHOL) and a sparse approximate inverse [1] for the symmetric part (SAI).

Briefly, the SAI preconditioner is derived as follows. Consider the decomposition of the inverse of the SPD matrix A_1 in the form $A_1^{-1} = ZDZ^{\mathrm{T}}$, where Z is an unit upper triangular matrix and D a diagonal matrix. A sparse approximate inverse of A_1 is obtained by enforcing sparsity in Z. This is achieved by neglecting elements whose absolute value falls below a prescribed drop tolerance, *droptol*.

We consider the case when $\mathrm{w} = (p(x,\, y),\, q(x,\, y))^T$ in (1). The function $f(x, y)$ is chosen such that the exact solution is

$$u(x,\, y) = e^{xy} sin(\pi x) sin(\pi y).$$

The initial guess for the preconditioned GMRES method is always a vector of zeros. For the SAI and ICHOL preconditioners, we choose a drop tolerance of *droptol*=10^{-3} and the stopping criteria is

$$\|\mathrm{r}_i\|_{A_1}/\|\mathrm{r}_0\|_{A_1} < 10^{-6}.$$

In the tables, we give the number of iterations required to satisfy the stopping criteria. A ‡ indicates that the iterative algorithm diverges or requires more than 50 iterations to converge.

Example 4.1 *We choose* $(p,\, q)^T = (10,\, 10)$ *and we discretise on a* 24×24 *grid. The coarse mesh spacing is* $H = 1/4$.

We choose the SPD matrix A_1 such that A_1 is the fourth-order discretisation of the Poisson equation with Dirichlet boundary conditions. Table 1 shows the number of iterations required for convergence when the SPD part is preconditioned using ICHOL and SAI respectively. We note that a choice of $\tau = 5$ does not lead to a convergent algorithm. On the other hand, we notice that, the same number of iterations are required when τ is varied. This shows the robustness of the algorithm for the choice of τ. However, a remarkable difference from the numerical results of Xu and Cai for piecewise linear finite element approximation of the same problem is that the values of τ for the HOC discretisation is relatively large whereas in the former case, the values of τ are less than 1.

Table 1: Number of iterations for $h = 1/24$ and $H = 1/4$.

τ	5	10	15	20	25	30	35
SAI	‡	9	9	9	9	9	9
ICHOL	‡	9	9	9	9	9	9

We now fix $\tau = 20$ and consider the effect of varying the magnitude of the convection term on the performance of the algorithm. This choice of τ is justified from our numerical experiments as it keeps the number of GMRES iterations low and in a sense this value keeps an optimal balance between both terms in the right-hand side of (14). We choose a 32×32 grid and we consider in this case the Jacobi (JAC) preconditioning for the SPD part.

Table 2: Number of iterations for $h = 1/32$, $H = 1/4$ and $\tau = 20$.

$(p,\, q)^{\mathrm{T}}$	(100, 0)	(100, 100)	(200, 100)
JAC	35	46	41
SAI	20	23	23
ICHOL	20	23	23

We note that as the convection term gets larger, the iteration counts for the diagonal preconditioning increases. However, no such increase in the number of iterations is observed for the other two preconditioners.

We now investigate the effect of varying the coarse mesh size H on the performance of the algorithm. We choose ICHOL as the preconditioner for the SPD part. The results in Table 3 indicate that there is practically no difference in the number of iterations and the amount of work required for convergence when the coarse mesh size is decreased. However, we note an increase in the iteration counts and the amount of work required as convection term decomes larger.

Table 3: Iteration and Flops counts.

	$H^{-1} = 4$		$H^{-1} = 8$		$H^{-1} = 16$	
	(10,0)	(100,50)	(10,0)	(100,50)	(10,0)	(100,50)
Iter.	8	23	8	23	8	23
Flops	3.06E7	8.89E7	3.06E7	8.89E7	3.07E7	8.92E7

Example 4.2 *We consider the case when* $p(x,\, y) = 10x(1-y)$ *and* $q(x,\, y) = -10x(1-y^2)$. *We show results for the SAI and ICHOL preconditioners for problems on a* 32×32 *grid.*

We notice that although both preconditioners require the same number of iterations for convergence, the cost for the SAI preconditioner is higher. We again find that there is no effect on the choice of the coarse mesh size on the number of iteration counts.

Table 4: Number of iterations for $p(x, y) = 10x(1-y)$ and $q(x, y) = -10x(1-y^2)$.

	$H^{-1} = 8$		$H^{-1} = 16$	
	SAI	ICHOL	SAI	ICHOL
Iter.	6	6	6	6
Flops	1.03E9	5.47E7	1.03E9	5.48E7

We consider another variable coefficient problem with $p(x, y) = 5(1 + x^2)$ and $q(x, y) = 10y$. The results are shown in Table 5. Similar conclusions as for the above problem can be reached in this case also.

Table 5: Number of iterations for $p(x, y) = 5(1 + x^2)$ and $q(x, y) = 10y$ on a 32×32 grid.

	$H^{-1} = 8$		$H^{-1} = 16$	
	SAI	ICHOL	SAI	ICHOL
Iter.	8	8	8	8
Flops	1.05E9	6.34E7	1.06E9	6.35E7

Example 4.3 *In the next example, we compare the performance of the algorithm when the SPD part arises from the discretisation of the Laplacian, denoted by A_Δ, with the case when the SPD part of A is given by*

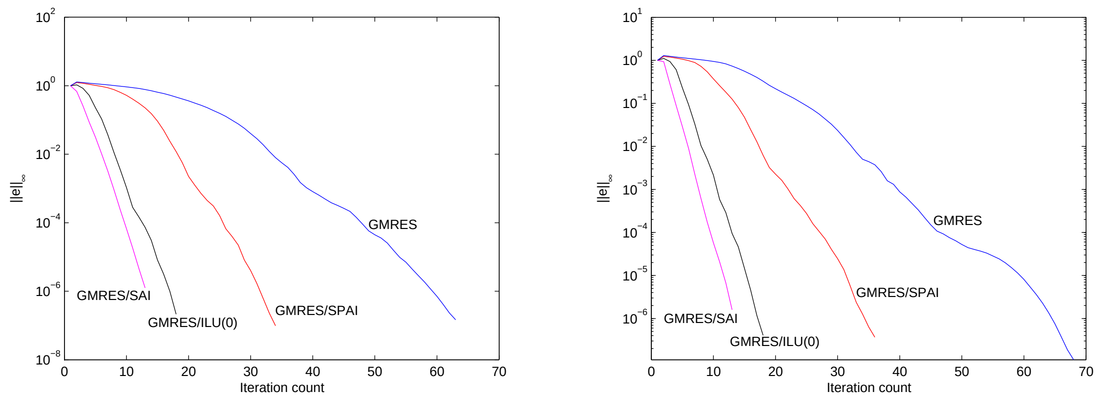
$$A_1 = \frac{1}{2}\left(A + A^{\mathrm{T}}\right).$$

In Table 6, we give the results for the case $(p, q)^T = (100, 10)$ and $(p, q)^T = (200, 10)$ on a 32×32 grid.

Table 6: Number of iterations for $A_1 = A_\Delta$ and $A_1 = \frac{1}{2}\left(A + A^{\mathrm{T}}\right)$ on a 32×32 grid.

	$(p, q)^T = (100, 10)$		$(p, q)^T = (200, 10)$	
	$A_1 = A_\Delta$	$A_1 = \frac{1}{2}\left(A + A^{\mathrm{T}}\right)$	$A_1 = A_\Delta$	$A_1 = \frac{1}{2}\left(A + A^{\mathrm{T}}\right)$
Iter.	21	22	20	21
Flops	8.11E7	8.79E7	7.72E7	8.40E7

In Figure 1, we plot the the L_∞ norm of error vector against iteration count for the case $h^{-1} = 32$ and $H^{-1} = 8$.



(a) $p = q = 1/\sqrt{2}$ and $\nu = 0.1$

(b) $p = 5(1 + x^2)$, $q = 10y$ and $\nu = 1$.

Figure 1: L_∞ norm of error vector against iteration count for the case $h^{-1} = 32$ and $H^{-1} = 8$.

5.0 Concluding Remarks

A common procedure for determining a preconditioner for a nonsymmetric matrix is based on the entries of the matrix itself. In this paper, we have investigated the performance of preconditioners based on entries of the SPD part and small subspace solvers. Numerical experiments indicate that for linear systems arising from HOC approximations of convection-diffusion problems, the preconditioned GMRES algorithm converges in a reasonable number of iterations. The convergence analysis of the preconditioned algorithm is based on a few algebraic assumptions. Although,the numerical experiments indicate convergence for various problems, we have not addressed here these theoretical issues. We believe that it is reasonable to assume that these assumptions hold here, but a theoretical justification is necessary. We are presently carrying out a study in this direction. We are also studying the quality of the preconditioned matrix in terms of the field of values and pseudospectra along the lines discussed in[2]. Another point of interest is some adaptive procedure for choosing the balancing parameter τ and a comparison of the preconditioning method with other methods such as ILU and SAI.

References

[1] M. Benzi, C. D. Meyer, and M. Tuma. A sparse approximate inverse preconditioner for the conjugate gradient method. *SIAM J. Sci. Comp.*, 17:1135–1149, 1996.

[2] M. Bhuruth, M. K. Jain, and A. Gopaul. Preconditioned iterative methods for the nine-point approximation to the convection-diffusion equation. *J. Comp. Appl. Math.*, 138:73–92, 2002.

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

[4] M. M. Gupta, R. Manohar, and J. W. Stephenson. A single cell high order difference scheme for the convection diffusion equation with variable coefficients. *Int. J. Numer. Meth. Fluids*, 4:641–651, 1984.

[5] T. K. Huckle. Efficient computation of sparse approximate inverses. *Numer. Lin. Alg. Appl.*, 5:57–71, 1998.

[6] M. K. Jain, R. K. Jain, and R. K. Mohanty. Fourth order difference method for elliptic equations with nonlinear first derivative terms. *Numer. Meth. Partial Diff. Eq.*, 5:87–95, 1989.

[7] J. A. Meijerink and H. A. van der Vorst. An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix. *Math. Comp.*, 31:148–162, 1977.

[8] W. F. Spotz and G. F. Carey. Iterative and parallel performance of high-order compact systems. *SIAM J. Sci. Comput.*, 19:1–14, 1998.

[9] H. A. van der Vorst. Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of non-symmetric linear systems. *SIAM J. Sci. Comput.*, 13:631–644, 1992.

[10] J. Xu. A new class of iterative methods for nonselfadjoint or indefinite problems. *SIAM J. Numer. Anal.*, 29(2):303–319, April 1992.

[11] J. Xu and X-C Cai. A preconditioned GMRES method for nonsymmetric or indefinite problems. *Math. Comp.*, 59:311–319, 1992.

[12] J. Zhang. Preconditioned iterative methods and finite difference schemes for convection-diffusion. *Appl. Math. Comp.*, 109:11–30, 2000.