European Conference on Computational Fluid Dynamics
ECCOMAS CFD 2006
P. Wesseling, E. Oñate and J. Périaux (Eds)

DOMAIN DECOMPOSITION METHODS AND DEFLATED KRYLOV SUBSPACE ITERATIONS

Reinhard Nabben*, Cornelis Vuik†

* TU Berlin,
Institut für Mathematik, MA 3-3
Strasse des 17. Juni 136,
D-10623 Berlin,
Germany
e-mail: nabben@math.tu-berlin.de

†Delft University of Technology,
Faculty Faculty of Electrical Engineering, Mathematics and Computer Science,
Delft Institute of Applied Mathematics, P.O. Box 5031, 2600 GA Delft,
The Netherlands
e-mail: c.vuik@tudelft.nl

Key words: Deflation, balancing and additive coarse grid correction

Abstract. *The balancing Neumann-Neumann (BNN) and the additive coarse grid correction (BPS) preconditioner are fast and successful preconditioners within domain decomposition methods for solving partial differential equations. For certain elliptic problems these preconditioners lead to condition numbers which are independent of the mesh sizes and are independent of jumps in the coefficients (BNN). Here we give an algebraic formulation of these preconditioners. This formulation allows a comparison with another solution or preconditioning technique - the deflation technique.*

By giving a detailed introduction into the deflation technique we establish analogies between the balancing, the additive coarse grid correction and the deflation technique.

We prove that the effective condition number of the deflated preconditioned system is always, i.e. for all deflation vectors and all restrictions and prolongations, below the condition number of the system preconditioned by the balancing preconditioner and the coarse grid correction preconditoner. This implies that the conjugate gradient method applied to the deflated preconditioned system is expected to converge always faster than the conjugate gradient method applied to the system preconditioned by the coarse grid correction or balancing. Moreover, we prove that the A-norm of the errors of the iterates built by the deflation preconditioner is always below the A-norm of the errors of the iterates built by the balancing preconditioner.

Numerical results for porous media flows emphasize the theoretical results.

1 Introduction

In 1952 Hestenes and Stiefel [9] introduced the Conjugate Gradient method (cg method) to solve large linear systems of equations

$$Ax = b$$

whose coefficient matrices A are sparse and symmetric positive definite. The convergence rate of the Conjugate Gradient method is bounded as a function of the condition number of the system matrix to which it is applied. Let us denote the i-th eigenvalue in non-decreasing order by $\lambda_i(A)$ or simply by λ_i when it is clear to which matrix we are referring. After k iterations of the conjugate gradient method, the error is bounded by (cf. [8], Thm. 10.2.6)

$$\|x - x_k\|_A \leq 2 \|x - x_0\|_A \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k, \tag{1}$$

where $\kappa = \kappa(A) = \lambda_n/\lambda_1$ is the spectral condition number of A and the A-norm of x is given by $\|x\|_A = (x^T A x)^{1/2}$.

If the condition number of A is large it is advisable to solve, instead, a preconditioned system $M^{-1}Ax = M^{-1}b$, where the symmetric positive definite preconditioner M is chosen such that $M^{-1}A$ has a more clustered spectrum or a smaller condition number than that of A. Furthermore, systems $Mz = r$ must be cheap to solve relative to the improvement it provides in convergence rate.

Today the design and analysis of preconditioners for the cg method are in the main focus whenever a linear system with symmetric positive definite coefficient matrix need to be solved. Even fast solvers, like multigrid or domain decomposition methods, are used as preconditioners. However, there are just a few theoretical comparisons of different preconditioners.

Here we consider three different preconditioning techniques the additive coarse grid correction, the balancing and the deflation preconditioner.

We show that the condition number of the system matrix preconditioned by the deflation method is always below the condition number of the system matrix preconditioned by the additive coarse grid correction. Moreover, we show that the condition number of the system matrix preconditioned by the deflation method is always below the condition number of the system matrix preconditioned by the balancing preconditioner.

We also establish a direct comparison between the condition numbers of the coarse grid correction preconditioner and the balancing preconditioner. We are able to show that the effective condition number of the system preconditioned by the balancing preconditioner is less or equal to the condition number of the system preconditioned by the additive coarse grid correction method. Moreover, we show that the A-norm of the errors of the iterates built by the deflation preconditioner is always below the A-norm of the errors of the iterates built by the balancing preconditioner.

2 Deflation

The deflation technique has been exploited by several authors. Among them are Nicolaides [19], Morgan [15], Kolotilina [11], and Saad, Yeung, Ehrel, and Guyomarc'h [23]. There are also many different ways to describe the deflation technique. We prefer the following one.

We define the projection P_D by

$$P_D = I - AZ(Z^TAZ)^{-1}Z^T, \quad Z \in \mathbb{R}^{n\times r}, \tag{2}$$

where the column space of Z is the deflation subspace, i.e. the space to be projected out of the residual, and I is the identity matrix of appropriate size.

We assume that $r \ll n$ and that Z has rank r. Under this assumption $E \equiv Z^TAZ$ may be easily computed and factored and is symmetric positive definite. Since $x = (I - P_D^T)x + P_D^Tx$ and because

$$(I - P_D^T)x = Z(Z^TAZ)^{-1}Z^TAx = ZE^{-1}Z^Tb \tag{3}$$

can be immediately computed, we need only to compute P_D^Tx. In light of the identity $AP_D^T = P_DA$, we can solve the deflated system

$$P_DA\tilde{x} = P_Db \tag{4}$$

for $\tilde{x}$ using the Conjugate Gradient method, premultiply this by P_D^T and add it to (3).

Obviously (4) is singular. But a positive semidefinite system can be solved by the cg method as long as the right-hand side is consistent (i.e. as long as $b = Ax$ for some x) [10]. This is certainly true for (4), where the same projection is applied to both sides of the nonsingular system. The rate of convergence of the cg method can be described with the *effective* condition number

$$\kappa_{\text{eff}}(C) = \frac{\lambda_n}{\lambda_{r+1}}.$$

which replaces the condition number in (1).

The deflated system can also be solved by using a symmetric positive definite preconditioner M^{-1},

$$M^{-1}P_DA\tilde{x} = M^{-1}P_Db. \tag{5}$$

3 Comparison of deflation and additive coarse grid correction

In this section we compare the deflation preconditioner with a well-known coarse grid correction preconditioner of the form

$$P_C = I + ZE^{-1}Z^T \tag{6}$$

and in the preconditioned case

$$P_{CM^{-1}} = M^{-1} + ZE^{-1}Z^T. \tag{7}$$

In the multigrid or domain decomposition language the matrices Z and Z^T are known as restriction and prolongation or interpolation operator. Moreover, the matrix $E = Z^T AZ$ is the Galerkin operator.

The above coarse grid correction preconditioner belongs to the class of additive Schwarz preconditioner. It is called the two level additive Schwarz preconditioner. If used in domain decomposition methods, typically, M^{-1} is the sum of the local (exact or inexact) solves in each domain. To speed up convergence a coarse grid correction $ZE^{-1}Z^T$ is added.

These methods are introduced by Bramble, Paschiak and Schatz [1] and Dryja and Widlund [2, 3, 4]. They show that under mild conditions the convergence rate of the PCG method is independent of the grid sizes.

For more details about this preconditioner we refer to the books of Toselli and Widlund [25], Quarteroni and Valli [22], and Smith, Bjørstad and Gropp [24]. A more abstract analysis of this preconditioner is given by Padiy, Axelsson and Polman [20], recently. To make the condition number of $P_{CM^{-1}}A$ smaller Padiy, Axelsson and Polman used a parameter $\sigma > 0$ and considered

$$P_{CM^{-1}}(\sigma) = M^{-1} + \sigma ZE^{-1}Z^T. \tag{8}$$

Note that in our abstract formulation Z is just an arbitrary rectangular but full rank matrix, $Z \in \mathbb{R}^{n\times r}$.

We compare this preconditioner to the corresponding deflated preconditioner

$$M^{-1}P_D. \tag{9}$$

Then, we obtain the following theorem.

Theorem 3.1 *Let $A \in \mathbb{R}^{n\times n}$ and $M \in \mathbb{R}^{n\times n}$ be symmetric positive definite. Let $Z \in \mathbb{R}^{n\times r}$ with rank $Z = r$. Then*

$$\lambda_n(M^{-1}P_DA) \leq \lambda_n(P_{CM^{-1}}(\sigma)A), \tag{10}$$

$$\lambda_{r+1}(M^{-1}P_DA) \geq \lambda_1(P_{CM^{-1}}(\sigma)A). \tag{11}$$

Proof: See the proof of Theorem 2.11 in [17]. □

In this theorem we prove that the effective condition number of the deflated preconditioned system $M^{-1}P_DA$ is always below the condition number of the system preconditioned by the coarse grid correction $P_{CM^{-1}}(\sigma)A$. This implies that we expect for all matrices $Z \in \mathbb{R}^{n\times r}$ and all positive definite preconditioners M^{-1} that the conjugate gradient method applied to the deflated preconditioned system will converge always faster than the conjugate gradient method applied to the system preconditioned by the coarse grid correction. Finally, we note that the memory requirements and the amount of work per iteration of both preconditioners are comparable. This implies, that in practice it is better to use deflation instead of the additive coarse grid correction.

4 Comparison of deflation and balancing

In this section we compare the preconditioned deflation operator to the balancing preconditioner proposed by Mandel [12, 13] and analyzed by Widlund et al. [5, 21, 25]. As the FETI algorithm [6, 7] the balancing Neumann-Neumann preconditioner is one of the domain decomposition methods that have been most carefully implemented and severely tested on the very largest existing parallel computer systems.

Applied to some specific symmetric positive definite problems the balancing Neumann-Neumann preconditioner leads to moderately growing condition numbers if the size of the systems increases [24]. Moreover, if an appropriate scaling is used, the condition numbers are independent of jumps in the coefficients in the matrices [24].

In our notation the balancing preconditioner is given by

$$P_B = (I - ZE^{-1}Z^TA)M^{-1}(I - AZE^{-1}Z^T) + ZE^{-1}Z^T, \tag{12}$$

where $Z \in \mathbb{R}^{n\times r}$, $E = Z^TAZ$ and M is a symmetric positive definite matrix. Note that P_B is symmetric and positive definite. For more details we refer to [12] and the books [22, 24, 25].

As a first comparison of both preconditioners we observe that the balancing preconditioner needs per iteration 3 matrix vector products and the coarse grid operator is used 2 times. This makes the balancing preconditioner per iteration more expensive than the deflation approach. However, if an optimal implementation of the balancing preconditioner is used (see e.g. [25]), the amount of work per iteration is the same.

Theorem 4.1 *Suppose that the spectrum of $M^{-1}P_DA$ is given by:*

$$\mathit{spectrum}(M^{-1}P_DA) = \{0, \ldots, 0, \mu_{r+1}, \ldots, \mu_n\},$$

then

$$\mathit{spectrum}(P_BA) = \{1, \ldots, 1, \mu_{r+1}, \ldots, \mu_n\}.$$

Proof: See the proof of Theorem 2.8 in [18]. □

Thus both preconditioners lead to almost the same spectra with the same clustering. The zero eigenvalues of the deflation preconditioned system are replaced by eigenvalues which are one if the balancing preconditioner is used. It follows from Theorem 4.1 that

$$\kappa(P_BA) \geq \kappa_{eff}(M^{-1}P_DA)$$

so the convergence bound based on the effective condition number implies that preconditioned deflated CG never converges slower than CG preconditioned by the balancing preconditioner. This combined with memory requirements and work per iteration counts suggested that it is better to use the deflation preconditioner instead of the balancing preconditioner.

5 Comparison of additive coarse grid correction and balancing

We start this section with a well-known result which relates the error of the i-th iterate of the preconditioned cg method with the effective condition number with respect to x and x_o.

Theorem 5.1 *Let A and M be symmetric positive definite. Let $(\tilde{\lambda}_i, \tilde{y}_i)$ be the eigenpairs of $M^{\frac{1}{2}}AM^{\frac{1}{2}}$. For $x_0 \in \mathbb{R}$ let*

$$M^{-\frac{1}{2}}x - M^{-\frac{1}{2}}x_0 = \sum_{j=1}^{n} \tilde{\gamma}_j \tilde{y}_j.$$

Define

$$\begin{aligned} \tilde{\alpha} &:= min\{\tilde{\lambda}_j | \tilde{\gamma}_j \neq 0\}, \\ \tilde{\beta} &:= max\{\tilde{\lambda}_j | \tilde{\gamma}_j \neq 0\}, \\ \kappa(MA, x - x_0) &:= \frac{\tilde{\beta}}{\tilde{\alpha}}. \end{aligned}$$

If the PCG method is applied to solve $Ax = b$ with starting vector x_0 and preconditioner M the i-th iterate satisfies

$$||x - x_i||_A \leq 2 \left\{ \frac{\sqrt{\kappa(MA, x - x_0)} - 1}{\sqrt{\kappa(MA, x - x_0)} + 1} \right\}^i ||x - x_0||_A.$$

Proof: See Corollary 3.3 in [16].
We then have

Theorem 5.2 *Let A be symmetric positive definite. Let the preconditioners P_B and P_{CM} be defined as in* (12) *and* (8)*, where $\sigma = 1$. With $x_{0,B} = ZE^{-1}Z^Tb$ we obtain*

$$\kappa(P_B A, x - x_{0,B}) \leq \kappa(P_{CM} A).$$

Proof: See Theorem 3.4 in [16].

From Theorem 5.2 we conclude that the balancing preconditioner with starting vector $x_{0,B} = ZE^{-1}Z^Tb$ is asymptotically a better preconditioner than the coarse grid correction preconditioner. Hence, we expect a faster convergence of the PCG method if the balancing preconditioner is used.

In order to make a more detailed comparison of the deflation operator and the balancing preconditioner for general projection vectors we compare the error vectors.

With respect to the approximation using preconditioned CG combined with deflation, we note that $x = (I - P_D^T)x + P_D^T x = ZE^{-1}Z^Tb + P_D^T x$. So after k iterations of preconditioned CG applied to $AP_D^T x = P_D Ax = P_D b$ we get the approximation $\tilde{x}_{k,D}$. The approximation $x_{k,D}$ of the solution vector x is then given by $x_{k,D} = ZE^{-1}Z^Tb + P_D^T \tilde{x}_{k,D}$.

Theorem 5.3 *The iterates $x_{k,D}$ and $x_{k,B}$ of the CG-method with start vector zero and preconditioned by the deflation preconditioner and the balancing preconditioner, respectively, satisfy*

$$\|x - x_{k,D}\|_A \leq \|x - x_{k,B}\|_A.$$

Proof: See the proof of Theorem 3.4 in [16]. □

6 Numerical experiments

In all our numerical experiments, the multiplication $y = E^{-1}b$ is done by solving y directly from $Ey = b$, where E is decomposed in its Cholesky factor. In this section, coarse grid correction is abbreviated as CGC. The choice of the boundary conditions is such that all problems have as exact solution the vector with components equal to 1. In order to make the convergence behavior representative for general problems we chose a random vector as starting solution, in stead of the zero starting vector.

We simulate a porous media oil flow in a 3-dimensional layered geometry, where the layers vary in thickness and orientation (see Figure 1 and 2 for a 4 layer problem). The fluid pressure and permeability are denoted by p and σ_p, respectively. The pressure p satisfies the equation:

$$-\text{div}(\sigma_p \nabla p) = 0 \text{ on } \Omega, \tag{13}$$

with boundary conditions

$$p = 1 \text{ on } \partial\Omega^D \text{ (Dirichlet) and } \frac{\partial p}{\partial n} = 0 \text{ on } \partial\Omega^N \text{ (Neumann)},$$

where $\partial\Omega = \partial\Omega^D \cup \partial\Omega^N$. In this problem $\partial\Omega^D$ is the top boundary of the domain. Figure 1 shows a part of the earth's crust. The depth of this part varies between 3 and 6 kilometers, whereas horizontally its dimensions are 40×60 kilometers. The upper layer is a mixture of sandstone and shale and has a permeability of 10^{-4}. Below this layer, shale and sandstone layers are present with permeabilities of 10^{-7} and 10 respectively. An incomplete Cholesky factorization with no fill in is used as preconditioner [14]. We consider a problem with 9 layers: 5 sandstone layers are separated by 4 shale layers. Due to the Dirichlet boundary condition at the top the preconditioned matrix has 4 small eigenvalues. We use 4 *physical* projection vectors [26] and stop if $\|r_k\|_2 \leq 10^{-5}$. Trilinear hexahedral elements are used and the total number of gridpoints is equal to 148185.

The number of iterations and the CPU time of ICCG are given in Table 1. From this table it appears that the coarse grid correction preconditioner takes more iterations than the deflation preconditioner, which makes the required CPU time longer. It appears that the norm of the residuals of the deflation and balancing preconditioners are the same. Due to extra work per iteration, the balancing preconditioner costs more CPU time. The computations are performed on an AMD Athlon, 1.4 GHz processor with 256 Mb of RAM. The code is compiled with FORTRAN g77 on LINUX.

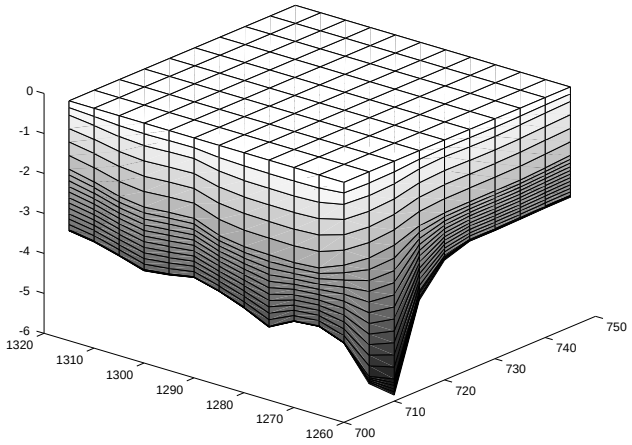

Figure 1: The geometry of an oil flow problem with 4 layers

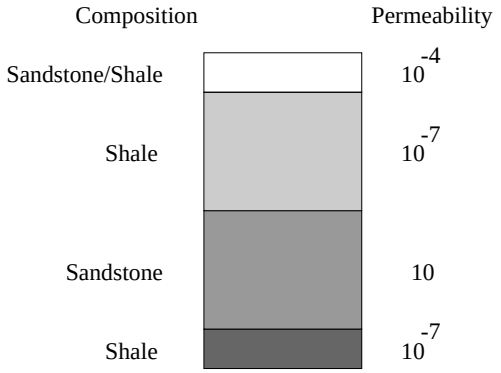

Figure 2: Permeabilities for each layer

method	deflation	CGC	balancing
iterations	36	47	36
CPU time in seconds	6.3	8.2	9.8

Table 1: The results for the ICCG method combined with various preconditioners applied to the oil flow problem

7 Conclusion

We compared different preconditioners for the cg method, namely the balancing, the additive coarse grid correction and the deflation technique. The balancing preconditioner and the additive coarse grid correction preconditioner are used in domain decomposition methods (BNN- and BPS-method). We showed that the effective condition number of the deflated preconditioned system is always, i.e. for all deflation vectors and all restrictions and prolongations, below the condition number of the system preconditioned by the balancing preconditioner and the coarse grid correction preconditoner. This implies that the conjugate gradient method applied to the deflated preconditioned system is expected to converge always faster than the conjugate gradient method applied to the system preconditioned by the coarse grid correction or balancing. Moreover, we proved that the A-norm of the errors of the iterates built by the deflation preconditioner is always below the A-norm of the errors of the iterates built by the balancing preconditioner.

REFERENCES

[1] J.H. Bramble, J.E. Pasciak, and A.H. Schatz. The construction of preconditioners for elliptic problems by substructuring. I. *Math. Comput.*, 47:103–134, 1986.

[2] M. Dryja. An additive Schwarz algorithm for two- and three dimensional finite element elliptic problems. In *Domain Decomposition methods*, pages 168–172. SIAM Philadelphia PA, 1989.

[3] M. Dryja and O.B. Widlund. Towards a unified theory of domain decomposition algorithms for elliptic problems. In *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 3–21. SIAM Philadelphia PA, 1990.

[4] M. Dryja and O.B. Widlund. Multilevel additive methods for elliptic finite element problems in three dimensions. In *Parallel Algorithms for PDEs*, pages 58–69. Vieweg Braunschweig Germany, 1991.

[5] M. Dryja and O.B. Widlund. Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite element problems. *Comm. Pure Appl. Math.*, 48:121–155, 1995.

[6] C. Farhat and F.-X. Roux. A method of finite element tearing and interconnecting and its parallel solution algorithm. *Int. J. Numer. Methods Eng.*, 32:1205–1227, 1991.

[7] C. Farhat and F.-X. Roux. Implicit parallel processing in structural mechanics. *Comput. Mech. Adv.*, 2:1–124, 1994.

[8] G. H. Golub and C. F. van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, third edition, 1996.

[9] M.R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Stand.*, 49:409–436, 1952.

[10] E .F. Kaasschieter. Preconditioned Conjugate Gradients for solving singular systems. *J. Comput. Appl. Math.*, 24:265–275, 1988.

[11] L.Yu. Kolotilina. Preconditioning of systems of linear algebraic equations by means of twofold deflation. I. Theory. *J. Math. Sci.*, 89:1652–1689, 1998.

[12] J. Mandel. Balancing domain decomposition. *Commun. Appl. Numer. Meth.*, 9:233–241, 1993.

[13] J. Mandel and M. Brezina. Balancing domain decomposition for problems with large jumps in coefficients. *Mathematics of Computation*, 216:1387–1401, 1996.

[14] 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. *Mathematics of Computation*, 31:148–162, 1977.

[15] R. B. Morgan. A restarted GMRES method augmented with eigenvectors. *SIAM J. Matrix Anal. Appl.*, 16(4):1154–1171, 1995.

[16] R. Nabben and C. Vuik. A comparison of deflation, balancing and additive coarse grid correction. *submitted*.

[17] R. Nabben and C. Vuik. A comparison of deflation and coarse grid correction applied to porous media flow. *SIAM J. Numer. Anal.*, 42:1631–1647, 2004.

[18] R. Nabben and C. Vuik. A comparison of deflation and the balancing Neumann-Neumann preconditioner. *SIAM J. Sci. Comput.*, 27:1742–1759, 2006.

[19] R. A. Nicolaides. Deflation of Conjugate Gradients with applications to boundary value problems. *SIAM J. Numer. Anal.*, 24(2):355–365, 1987.

[20] A. Padiy, O. Axelsson, and B. Polman. Generalized augmented matrix preconditioning approach and its application to iterative solution of ill-conditioned algebraic systems. *SIAM J. Matrix Anal. Appl.*, 22:793–818, 2000.

[21] L.F. Pavarino and O.B. Widlund. Balancing Neumann-Neumann methods for incompressible Stokes equations. *Comm. Pure Appl. Math.*, 55:302–335, 2002.

[22] A. Quarteroni and A. Valli. *Domain Decomposition Methods for Partial Differential Equations*. Oxford Science Publications, Oxford, 1999.

[23] Y. Saad, M. Yeung, J. Ehrel, and F. Guyomarc'h. A deflated version of the Conjugate Gradient algorithm. *SIAM Journal on Scientific Computing*, 21:1909–1926, 2000.

[24] B. Smith, P. Bjørstad, and W. Gropp. *Domain Decomposition*. Cambridge University Press, Cambridge, 1996.

[25] A. Toselli and W. Widlund. *Domain Decomposition Methods*. Computational Mathematics, Vol. 34. Springer, Berlin, 2005.

[26] C. Vuik, A. Segal, L. El Yaakoubi, and E. Dufour. A comparison of various deflation vectors applied to elliptic problems with discontinuous coefficients. *Applied Numerical Mathematics*, 41:219–233, 2002.