Preconditioners for regularized saddle point problems with an application for heterogeneous Darcy flow problems

Owe Axelsson*, Radim Blaheta†, Petr Byczanski‡, János Karátson§ and Bashir Ahmad¶

April 14, 2014

Abstract

Saddle point problems arise in the modelling of many important practical problems. Preconditioners for the corresponding matrices on block triangular form, based on coupled inner-outer iteration methods are analyzed and applied to a Darcy flow problem, possibly with strong heterogeneity and to non-symmetric saddle point problems. Using proper regularized forms of the given matrix and its preconditioner it is shown that, for large values of the regularization parameters, the eigenvalues cluster about one or two points on the real axis and that eigenvalue bounds do not depend on this variation. Therefore, just two outer iterations can suffice. To solve the inner iteration systems various preconditioners are used.

1 Introduction

Saddle point problems arise in various physical problems, such as fluid flow problems, where a pair of variables are coupled via some differential operators. Other problems, like second or fourth order elliptic problems, can be reformulated as a coupled system by the introduction of a new variable, typically the gradient of the potential function. This can give a higher order of accuracy for the new variable and additionally, as we shall see, in some cases enable a more efficient solution method for instance for strongly variable coefficient problems. Such mixed variable formulations of second order problems has also the favorable properly that mass is conserved locally. Frequently the problems are highly ill-conditioned due to some coefficients, such as a diffusion coefficient, taking widely varying values in the given domain of definition. The construction

*Institute of Geonics AS CR, IT4 Innovations, Ostrava, The Czech Republic
†Institute of Geonics AS CR, IT4 Innovations, Ostrava, The Czech Republic
‡Institute of Geonics AS CR, IT4 Innovations, Ostrava, The Czech Republic
§Department of Applied Analysis, ELTE University, Budapest, Hungary; karatson@cs.elte.hu, Supported by the European Union and co-financed by the European Social Fund (grant agreement no. TAMOP 4.2.1./B-09/1/KMR-2010-0003) and by the Hungarian Research Grant OTKA No.K 67819.
¶King Abdulaziz University, Jeddah, Saudi Arabia ,

of robust and efficient preconditioners for solving such problems is of a primary concern in this paper. We show also how non–symmetric problems, including problems with singular or indefinite pivot (1.1) block matrix, can be preconditioned. In general, it is most efficient to solve saddle point matrix equations by iterative solution methods, see e.g. [1] – [5].

In order to enable a fast and robust solution method, saddle point problems must be properly preconditioned. Commonly used methods (see e.g. [6] – [11]) are based on approximate block factorization. They lead to a form where a Schur complement matrix which appears must be approximated. This may cause complications since the Schur complement is a full matrix and evaluation of the corresponding residuals may need many inner iterations for the evaluation of actions of arising inverses of ill–conditioned matrices and slow rates of convergence. See, however [12] for a discussion of the use of elementwise constructed Schur complements. In this paper we show that certain preconditioners on block tridiagonal form are essentially free of Schur complement matrices, and can be very efficient. The methods involve some regularized form of the given matrix or of its preconditioner and are based on coupled inner outer iteration methods. The idea is to use an efficient regularization method and block matrix factorization which gives very few, typically only 2–4 outer iterations. Thereby the block diagonal matrices which arise in the preconditioner are solved by inner iterations. The regularization used leads also to better conditioned inner systems which can enable faster and more robust solutions than if the reduced Schur complement matrix is solved.

Other arguments for using a Schur complement free preconditioner can be found in [13]. The inner iteration need also efficient preconditioners. The construction of such is a topic by itself and will only be shortly commented on in this paper.

We show first conditions for non-singularity of the given matrix and how it can be regularized if the conditions are not met, or if the matrix is nearly singular. Two preconditioners on block triangular form, one for the unregularized and one for the regularized matrix are then presented. As the regularization parameter increases, we show that clustering of the eigenvalues of the preconditioned matrix occurs about just one or two points on the real axis. This hold also for non-symmetric saddle point problems. This result is shown algebraically for the given finite element matrices and then, in an alternative way, shown to hold by using relations between the corresponding operator pairs, namely for the preconditioning and the given operators. This result shows a mesh-independent rate of convergence. The clustering of eigenvalues implies that after some initial stage a fast, superlinear rate of convergence takes place.

Similar clustering results have appeared previously in various publications, such as in [4], [15], [16]. The major contribution of the present paper is to show these results using short proofs as well as proofs based on operator pair settings. The results for the regularized matrices and the applications are new.

The remainder of the paper is composed as follows. In Section 2 we show conditions for the non-singularity of the given matrix in saddle point form and introduce a regularization of it if necessary. Then the corresponding preconditioned system is analysed. The regularization used to handle singular systems corresponds actually to the case where the so called LBB (inf–sup)stability condition is violated, see e.g. [17] for an earlier presentation of this approach. The preconditioner involves in general two, but similar, parameters where one is

used for the regularized form of the pivot(1,1) block matrix and the inverse of one to make the (2,2) block non-zero, but not too big, as it is a perturbation of the given matrix. As the parameters increase, the eigenvalues cluster about one or two points, which in the limit gives just two or three outer iterations for the method.

In Section 3 the corresponding results for the differential operator pairs are presented. The choice of the weight matrix used in the regularization to get a better conditioned pivot block matrix is discussed here. This choice is problem dependent and can be crucial for the efficiency of the method. The results show when mesh-independence and compact perturbation properties i.e., the preconditioned matrix is a compact perturbation of identity, hold. Furthermore, it is shown that the eigenvalue bounds for the preconditioned matrix depend little on the variation of coefficients in the differential operator and, hence, are efficient for heterogeneous material problems, such as can occur for Darcy flow problems.

The solution of the regularized form of the pivot matrix can take place by use of inner iterations. It is then crucial for the efficiency of the whole method to balance the number of inner and outer iterations. This topic is discussed in Section 4. Section 5 contains some numerical tests which involve heterogeneous material coefficient problems as well as advection dominated problem and illustrate the practical bearings of the methods. We end the paper with some concluding remarks.

Unless otherwise stated, we will denote by $\lambda_{min}(A)$ and $\lambda_{max}(A)$ the minimal and maximal eigenvalues, respectively, of a symmetric matrix A. We use the notations $\mathcal{R}(A)$ and $\mathcal{N}(A)$ respectively, for the range and nullspace of a matrix A. Further rank $(A) = \dim \mathcal{R}(A)$.

2 Efficient preconditioning

Consider a given, possibly nonsymmetric, real-valued saddle point matrix in the form

$$\mathcal{M} = \begin{bmatrix} M & B^T \\ C & 0 \end{bmatrix},$$

where M is a square matrix of order $n \times n$ and B, C have orders $m \times n$, where $m \leq n$. Here M may be indefinite but we assume that its symmetric part has at least some positive eigenvalues.

First, we give assumptions to ensure that $\mathcal{M}$ is nonsingular and then, if these conditions are not met, we consider a regularization of $\mathcal{M}$ that becomes nonsingular. In both cases we present efficient block matrix preconditioners, for which a strong clustering of the eigenvalues of the preconditioned matrix takes place.

2.1 Preconditioning for a nonsingular saddle point matrix

The construction of the preconditioner will be based on the following assumption. Recall that $\text{rank}(A) = \dim \mathcal{R}(A)$. If A has order $m \times n, m \leq n$, then $\text{rank}(A) \leq m$ and $\dim \mathcal{N}(A) = n - \text{rank}(A) \geq n - m$.

Assumptions 2.1.

(i) There exists a nonsingular matrix W such that

$$M_1 := M + B^T W^{-1} C$$

is nonsingular. (Note that this implies in particular that $\mathcal{N}(C) \cap \mathcal{N}(M) = \{0\}$, $\mathcal{N}(B) \cap \mathcal{N}(M^T) = \{0\}$.)

(ii) The matrices B and C have full rank ($= m$).

Lemma 2.1 *The matrix $\mathcal{M}$ is nonsingular if and only if the conditions in Assumption 2.1 hold .*

PROOF. Singularity of $\mathcal{M}$ is equivalent to singularity of $\mathcal{M}^T$. If

$$\mathcal{M} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \text{i.e.} \qquad \begin{cases} Mx + B^T y &= 0 \\ Cx &= 0. \end{cases} \tag{2.1}$$

then, if $x = 0$ and $B^T y = 0$, (2.1) has a nontrivial solution $x = 0, y \neq 0$ if B is rank deficient. A similar conclusion holds if C is rank deficient.

It follows from (2.1) that $B^T W^{-1} C x = 0$ and thus, $(M + B^T W^{-1} C)x + B^T y = M_1 x + B^T y = 0$. Hence, if $M_1 x = 0, x \neq 0$, then there exists vectors $x \neq 0, y = 0$ as nontrivial solutions to (2.1). Hence, the conditions of Assumption 2.1 must hold if $\mathcal{M}$ is nonsingular.

To prove the sufficiency of the conditions, it follows from (2.1) that $B^T W^{-1} C x = 0$, thus

$$(M + B^T W^{-1} C)x + B^T y = M_1 x + B^T y = 0, \text{ or } x = -M_1^{-1} B^T y,$$

so $0 = Cx = -C M_1^{-1} B^T y$. By the assumptions of full rank, the matrix $C M_1^{-1} B^T$, of order $m \times m$, is nonsingular. Hence $y = 0$ and thus also $x = 0$, i.e. (2.1) has only the trivial solution.

■

The following alternate form of Assumption 2.1 (i) holds (cf. [5]). Let $U = [\mathbf{u}_1, \ldots, \mathbf{u}_{n-m}]$, $V = [\mathbf{v}_1, \ldots, \mathbf{v}_{n-m}]$ where $\{\mathbf{u}_i\}, \{\mathbf{v}_i\}$ are linearly independent vectors in $\mathcal{N}(B)$ and $\mathcal{N}(C)$, respectively. It holds:

Corollary 2.2 *$\mathcal{M}$ is nonsingular if and only if B and C have full rank and the matrix $U^T M V$ is nonsingular.*

PROOF. It holds $U^T M_1 V = U^T M V + (BU)^T W^{-1} C V = U^T M V$, which means that the nonsingularity of M_1 can be replaced with that of $U^T M V$.

■

In practical applications the assumption on M_1 is normally more useful.

To find efficient preconditioners to $\mathcal{M}$, consider first a preconditioner for a nonsingular matrix $\mathcal{M}$ in the form

$$\mathcal{B} := \begin{bmatrix} M_r & 2B^T \\ 0 & -W_r \end{bmatrix},$$

where W_r is a given nonsingular matrix, depending on a parameter $r > 0$, and

$$M_r := M + B^T W_r^{-1} C.$$

This preconditioner with the factor 2 in the (1,2) block turns out to be efficient in clustering the eigenvalues of the preconditioned matrix, see e.g. [15, 16, 18] for this choice. The factor will not be included when we consider preconditioning of the regularized matrix M_r in the next subsection.

Theorem 2.3 *Let Assumptions 2.1 hold. Then the preconditioned matrix*

$$\mathcal{B}^{-1}\mathcal{M} = \begin{bmatrix} M_r & 2B^T \\ 0 & -W_r \end{bmatrix}^{-1} \begin{bmatrix} M & B^T \\ C & 0 \end{bmatrix}$$

has eigenvalues equal to unity if $x \in \mathcal{N}(C)$ or $x \in \mathcal{N}(B)$. Likewise, if $x \in \mathcal{N}(M)$, i.e. $x \notin \mathcal{N}(C) \cup \mathcal{N}(B)$ then $\lambda = 1$. If $W_r = \frac{1}{r}W$, where W is a given nonsingular matrix, then the eigenvalues equal unity of multiplicity at least $n - m + q$, where $q \leq m$ is the dimension of $\mathcal{N}(M)$. The remaining eigenvalues satisfy $\lambda = \lambda_r \to 1$ as $r \to \infty$. If $C = B$ and $M = M^T$, then

$$\frac{r}{r + \nu_{max}} \leq \lambda \leq \frac{r}{r + \nu_{min}}, \tag{2.2}$$

where ν_{min} and ν_{max} are the extreme eigenvalues of $Mx = \nu B^T W^{-1} Bx$, $x \notin \mathcal{N}(B)$ and $r > |\nu_{min}|$ is chosen.

Proof The corresponding generalized eigenvalue problem takes the form

$$\lambda \begin{bmatrix} M_r & 2B^T \\ 0 & -W_r \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} M & B^T \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad \|x\| + \|y\| \neq 0\,. \tag{2.3}$$

This implies

$$\lambda y = -W_r^{-1} C x\,.$$

Since both block matrices in (2.3) are nonsingular, it follows that $\lambda \neq 0$. Hence $y = -\frac{1}{\lambda} W_r^{-1} C x$, and from the first equation of (2.3) it follows that

$$\lambda(Mx + B^T W_r^{-1} C x) - 2B^T W_r^{-1} C x = Mx - \frac{1}{\lambda}\, B^T W_r^{-1} C x,$$

or

$$\lambda(\lambda - 1)\, Mx = -(\lambda - 1)^2\, B^T W_r^{-1} C x. \tag{2.4}$$

It follows that if $x \in \mathcal{N}(C)$ or, by taking the matrix transpose, if $x \in \mathcal{N}(B)$, then $\lambda = 1$. Likewise, if $x \in \mathcal{N}(M)$, i.e. $x \notin \mathcal{N}(C) \bigcup \mathcal{N}(B)$ then $\lambda = 1$. This is the first part of the Theorem.

For $\lambda \neq 1$, it follows from (2.4) that

$$\lambda\, Mx = -(\lambda - 1)\, B^T W_r^{-1} C x$$

or

$$\lambda\, M_r x = B^T W_r^{-1} C x. \tag{2.5}$$

If $W_r := \frac{1}{r}\, W$ $(r > 0)$, where W is a given nonsingular matrix, then (2.5) takes the form

$$\lambda\, (M + r B^T W^{-1} C) x = r\, B^T W^{-1} C x.$$

It is seen that the above implies that for the eigenvalues $\lambda \neq 1$ one has $\lambda = \lambda_r \to 1$ as $r \to \infty$, i.e., all eigenvalues cluster at the unit value.

More precisely, the eigenvalues λ in (2.5) satisfy $\lambda = -(\lambda - 1)\mu r$, that is

$$\lambda = \frac{\mu r}{1 + \mu r}, \tag{2.6}$$

where μ is an eigenvalue of

$$\mu M x = B^T W^{-1} C x. \tag{2.7}$$

(Note that, since $\lambda \neq 1$ by assumption, it holds $x \notin \mathcal{N}(M) \bigcup \mathcal{N}(B) \bigcup \mathcal{N}(C)$, so $\mu \neq 0$.) If the latter eigenvalues are μ_i ($i = 1, \ldots, n$, with multiplicity) and nonzero, then the ith eigenvalue in (2.5) satisfies

$$\lambda_i = \frac{\mu_i r}{1 + \mu_i r} \to 1 \quad \text{as} \quad r \to \infty$$

for all $i = 1, \ldots, n$.

In the special case when $\mathcal{M}$ is symmetric, i.e. $C = B$ and $M = M^T$, then all eigenvalues λ of the generalized eigenvalue problem (2.3) are real. Let us rewrite (2.7) as

$$M x = \nu B^T W^{-1} B x, \tag{2.8}$$

where $\mu = 1/\nu$. Then (2.6) becomes

$$\lambda = \frac{r}{r + \nu}. \tag{2.9}$$

Let ν_{min} and ν_{max} be the extreme eigenvalues of (2.8), i.e. $\nu_{min} \leq \nu \leq \nu_{max}$ for all eigenvalues ν. We choose r such that $r > |\nu_{min}|$, then $\lambda > 0$ for all λ. Moreover, it follows from (2.9) that

$$\frac{r}{r + \nu_{max}} \leq \lambda \leq \frac{r}{r + \nu_{min}}.$$

■

In the rest of this section we consider the following case.

Example 2.1. Let $\mathcal{M}$ be symmetric, i.e. $C = B$ and $M = M^T$, moreover, M is symmetric and positive definite (SPD). We make two choices (a), (b) of W.

(a) We first choose the matrix W as

$$W := BB^T.$$

Proposition 2.4 *If $W := BB^T$ then $B^T W^{-1} B x = x_1$, where $x_1 := P_{\mathcal{R}(B^T)} x$ denotes the orthogonal projection of x onto the range of B^T.*

Proof. Any $x \in \mathbf{R}^n$ can be written as $x = x_1 + x_2$, where $x_1 \in \mathcal{R}(B^T)$ and $x_2 \in \mathcal{R}(B^T)^\perp = \mathcal{N}(B)$. Then $x_1 = B^T y$ for some y and $B x_2 = 0$, hence

$$B^T W^{-1} B x = B^T W^{-1} B x_1 = B^T W^{-1} B B^T y = B^T W^{-1} W y = B^T y = x_1. \quad ■$$

Let us consider the eigenvalue problem (2.8) for this choice of W:

$$Mx = \nu B^T W^{-1} Bx, \quad \text{i.e.} \quad Mx = \nu P_{\mathcal{R}(B^T)} x. \tag{2.10}$$

Let us denote the extreme eigenvalues ν by ν_{min} and ν_{max} and further, the extreme eigenvalues of M by $\lambda_{min}(M)$ and $\lambda_{max}(M)$.

Proposition 2.5 *We have* $\lambda_{min}(M) \le \nu_{min}, \quad \nu_{max} \le \lambda_{max}(M)$.

PROOF. Let x be a solution of (2.10), and write $x = x_1 + x_2$, where $x_1 \in \mathcal{R}(B^T) = \mathcal{N}(B)^\perp$ and $x_2 \in \mathcal{N}(B)$. Then $Mx = \nu x_1$, hence $x_1 + x_2 = x = \nu M^{-1} x_1$. An inner product with x_1 yields

$$\|x_1\|^2 = \nu \, \langle M^{-1} x_1, x_1 \rangle.$$

The ratio of $\|x_1\|^2$ and $\langle M^{-1} x_1, x_1 \rangle$ lies between $1/\lambda_{max}(M^{-1}) = \lambda_{min}(M)$ and $1/\lambda_{min}(M^{-1}) = \lambda_{max}(M)$, hence so does ν. ∎

Then (2.2) yields

Corollary 2.6 *If $\mathcal{M}$ is symmetric, $W := BB^T$ and $r > |\lambda_{min}(M)|$, then the eigenvalues λ of (2.3) satisfy*

$$\frac{r}{r + \lambda_{max}(M)} \le \lambda \le \frac{r}{r + \lambda_{min}(M)}.$$

(b) We now choose the matrix W more generally as

$$W := BG^{-1}B^T,$$

where G is spd.

Proposition 2.7 *If $W := BG^{-1}B^T$, then $B^T W^{-1} Bx = Gx_1$, where x_1 denotes the G-orthogonal projection of x to the range of $G^{-1}B^T$.*

PROOF. Similar as above. It follows easily that $\mathcal{R}(G^{-1}B^T)$ is G-orthogonal to $\mathcal{N}(B)$, hence $x = x_1 + x_2$, where $x_1 \in \mathcal{R}(G^{-1}B^T)$ and $Bx_2 = 0$. Then $x_1 = G^{-1}B^T y$ for some y and

$$B^T W^{-1} Bx = B^T W^{-1} Bx_1 = B^T W^{-1} BG^{-1} B^T y = B^T W^{-1} Wy = B^T y = Gx_1.$$

∎

Let us now consider the eigenvalue problem

$$Mx = \nu B^T W^{-1} Bx \equiv \nu G x_1. \tag{2.11}$$

Proposition 2.8 *We have* $\lambda_{min}(G^{-1}M) \le \nu_{min}, \quad \nu_{max} \le \lambda_{max}(G^{-1}M)$.

PROOF. From (2.11) follows $x = \nu M^{-1} G x_1$, i.e. $x_1 + x_2 = \nu M^{-1} G x_1$, where the G-orthogonal decomposition is used such that $x_1 \in \mathcal{R}(G^{-1}B^T)$ and $Bx_2 = 0$. Multiplying by Gx_1, we obtain

$$\langle Gx_1, x_1 \rangle = \nu \langle M^{-1} G x_1, G x_1 \rangle,$$

i.e.

$$\nu = \frac{\|x_1\|_G^2}{\langle M^{-1}Gx_1, x_1\rangle_G}.$$

The range of this Rayleigh quotient lies between the extreme eigenvalues independently of the inner product, hence ν lies between the reciprocals of the extreme eigenvalues of $M^{-1}G$, i.e. between the extreme eigenvalues of $G^{-1}M$. ■

Corollary 2.9 *If $\mathcal{M}$ is symmetric, $W := BG^{-1}B^T$ and $r > |\lambda_{min}(G^{-1}M)|$, then the eigenvalues $\lambda \neq 1$ of (2.3) satisfy*

$$\frac{r}{r + \lambda_{max}(G^{-1}M)} \leq \lambda \leq \frac{r}{r + \lambda_{min}(G^{-1}M)}.$$

2.2 Preconditioning for a regularized saddle point matrix

The matrix $M_r = M + rB^TW^{-1}C$, which appeared in the previous section, motivates to introduce the following regularized saddle point matrix

$$\widetilde{\mathcal{M}}_r = \begin{bmatrix} M_r & B^T \\ C & 0 \end{bmatrix}.$$

Note that the regularized and unregularized systems

$$\widetilde{\mathcal{M}}_r \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \widetilde{f} \\ g \end{bmatrix}, \ \widetilde{f} = f + rB^Tg, \quad \text{and} \quad \mathcal{M} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix},$$

have the same solution (unique if $\mathcal{M}$ is nonsingular). This is the familiar augmented Lagrangian method, see e.g.[9]. The major problem here is the approximation of the Schur complement matrix $CM_r^{-1}B^T$. It can be shown that its inverse equals $rW^{-1} + (CM^{-1}B^T)^{-1}$ of M is nonsingular and C, B have full rank. Hence for large values of r it is less important how the unregularized Schur complement matrix $CM^{-1}B^T$ is approximated. However, depending on the choice of W, applying actions of $\widetilde{\mathcal{M}}_r$ and solutions of systems with M_r can be costly. Instead, we can use such a regularized form, where M_r is replaced by M_{r_0}, in the preconditioner to $\mathcal{M}$. Here $M_{r_0} = M + r_0B^TW^{-1}C$ and $r_0 < r$, where r_0 is an additional method parameter. Then we need to solve systems with the better conditioned matrix $\mathcal{M}_{r_0}$ instead of with $\mathcal{M}_r$ when applying the preconditioner.

In this paper we choose another regularization where the Schur complement $CM_r^{-1}B^T$ does not arise, which can also handle the case where $\mathcal{M}$ is singular due to a rank-deficient matrix B. It arises when we replace the zero $(2,2)$ block in the matrix, and its preconditioner, with $-W_r = -\frac{1}{r}W$, where W is regular. This means that the matrix $\mathcal{M}$ is perturbed with the matrix $\begin{bmatrix} 0 & 0 \\ 0 & -W_r \end{bmatrix}$. We assume that $r >> 1$ takes sufficiently large values so that the corresponding perturbation of the solution is negligible or, otherwise, we can use some steps of a defect-correction method to correct for this perturbation. The arising perturbed matrix will be denoted by $\mathcal{M}_r$. As preconditioner to $\mathcal{M}_r$ we use now B_r in (2.13) or B_{r_0}.

The regularized matrix $\mathcal{M}_r$ can be factorized as

$$\mathcal{M}_r = \begin{bmatrix} M & B^T \\ C & -W_r \end{bmatrix} = \begin{bmatrix} M_r & B^T \\ 0 & -W_r \end{bmatrix} \begin{bmatrix} I_1 & 0 \\ -W_r^{-1}C & I_2 \end{bmatrix}. \tag{2.12}$$

It follows that $\mathcal{M}_r$ is nonsingular if and only if M_r is nonsingular.
We make therefore the following assumption.

Assumption 2.2. The matrix $M_r = M + rB^TW^{-1}C$ is nonsingular for all $r \geq r_0$ for some $r_0 > 0$.

Clearly M_r is nonsingular if M is nonsingular on the subspace (V_0) where $B^TW^{-1}C$ is singular. This subspace contains $\mathcal{N}(B) \bigcup \mathcal{N}(C)$ and includes also vectors due to possible rank deficiency of B and/or C. Correspondingly, $B^TW^{-1}C$ must be nonsingular on $\mathcal{N}(M)$, i.e. $\mathcal{N}(M) \bigcap V_0 = \{0\}$.
As preconditioner to $\mathcal{M}_r$ we take

$$\mathcal{B}_r = \begin{bmatrix} M_r & B^T \\ 0 & -W_r \end{bmatrix}. \tag{2.13}$$

Since, as is seen from (2.12),

$$\mathcal{B}_r^{-1}\mathcal{M}_r = \begin{bmatrix} I_1 & 0 \\ -W_r^{-1}C & I_2 \end{bmatrix},$$

it is obvious that $p_2(\mathcal{B}_r^{-1}\mathcal{M}_r) = 0$, where $p_2(\lambda) = (\lambda - 1)^2$ and, correspondingly, the preconditioned generalized conjugate gradient method (see [11],[34]) converges in just two steps.
We state the result in the next proposition.

Proposition 2.10 *Assume that $M_r = M + rB^TW^{-1}C$ is nonsingular. Then the preconditioned matrix*

$$\mathcal{B}_r^{-1}\mathcal{M}_r = \begin{bmatrix} M_r & B^T \\ 0 & -W_r \end{bmatrix}^{-1} \begin{bmatrix} M & B^T \\ C & -W_r \end{bmatrix} = \begin{bmatrix} I_1 & 0 \\ -W_r^{-1}C & I_2 \end{bmatrix}$$

has eigenvalues $\lambda = 1$ and GMRES or a generalized conjugate GCG gradient method converges in just two iterations.

Clearly, the above preconditioner is nothing but the left part of the exact block matrix LU factorization of $\mathcal{M}_r$.
In practice, we solve the arising systems with M_r and W_r by inner iterations. Depending on the size of the relative stopping criteria used for the inner iterations, there may occur a few more iterations. For discussions of related issues, see [19] and for the flexible GMRES method, see [20]. For simplicity, we consider the case where just systems with M_r are solved by inner iterations, but systems with W_r are assumed to be solved exactly, or to a negligible small error. Let then

$$\widetilde{B}_r^{-1} = \begin{bmatrix} \widetilde{M}_r & B^T \\ 0 & -W_r \end{bmatrix}^{-1} = \begin{bmatrix} \widetilde{M}_r^{-1} & \widetilde{M}_r^{-1}B^TW_r^{-1} \\ 0 & -W_r^{-1} \end{bmatrix}$$

be the corresponding multiplicative preconditioner, where $\widetilde{M}_r^{-1}$ denotes an approximate action of M_r^{-1}, defined via an inner iteration process. Note that there will occur only one action of $\widetilde{M}_r^{-1}$ for each outer iteration process. It holds then

$$\begin{aligned}\widetilde{B}_r^{-1}\mathcal{M}_r &= \begin{bmatrix} \widetilde{M}_r^{-1} & \widetilde{M}_r^{-1}B^TW_r^{-1} \\ 0 & -W_r^{-1} \end{bmatrix}\begin{bmatrix} M_r - B^TW_r^{-1}C & B^T \\ C & -W_r \end{bmatrix} \\ &= \begin{bmatrix} \widetilde{M}_r^{-1}M_r & 0 \\ -W_r^{-1}C & I_2 \end{bmatrix}\end{aligned}$$

so

$$\widetilde{B}_r^{-1}\mathcal{M}_r - \begin{bmatrix} I_1 & 0 \\ 0 & I_2 \end{bmatrix} = \begin{bmatrix} \widetilde{M}_r^{-1}M_r - I_1 & 0 \\ -W_r^{-1}C & 0 \end{bmatrix}$$

Clearly, the perturbation $\widetilde{M}_r^{-1}M_r - I_1$ of the unit eigenvalues can be arbitrarily small, $\|\widetilde{M}_r^{-1}M_r - I_1\| \leq \varepsilon$, for some $\varepsilon > 0$ by making a sufficiently large number of inner iterations.

Since the matrix M_r can be ill–conditioned for large values of r, we may use a smaller value, r_0 for this part of the regularization. The preconditioner to $\mathcal{M}_r$ will then be taken as

$$\mathcal{B}_{r_0} := \begin{bmatrix} M_{r_0} & B^T \\ 0 & -W_r \end{bmatrix},$$

It holds,

Theorem 2.11 *Assume that $M_r = M + rB^TW^{-1}C$ is nonsingular for all $r \geq r_0$ for some $r_0 > 0$. Then the preconditioned matrix*

$$\mathcal{B}_{r_0}^{-1}\mathcal{M}_r = \begin{bmatrix} M_{r_0} & B^T \\ 0 & -W_r \end{bmatrix}^{-1}\begin{bmatrix} M & B^T \\ C & -W_r \end{bmatrix}$$

has eigenvalues $\lambda = 1$ of multiplicity at least $n - m$. If $r \to \infty$ and $r_0 = \nu_0 r$ for some fixed $0 < \nu_0 < 1$, then the remaining eigenvalues cluster about ν_0. More precisely, if $C = B$ and M is symmetric, then if $r_0 > \mu_0$ the eigenvalues are contained in the interval

$$\left[\frac{r+\mu_1}{r_0+\mu_1}, \frac{r-\mu_0}{r_0-\mu_0}\right],$$

where μ_0, μ_1 are the extreme eigenvalues of $Mx = \mu B^TW^{-1}Bx$, $\quad x \notin \mathcal{N}(C) \cup \mathcal{N}(B)$.

Proof The corresponding generalized eigenvalue problem takes the form

$$\lambda\begin{bmatrix} M_{r_0} & B^T \\ 0 & -W_r \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} M & B^T \\ C & -W_r \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix},$$

or

$$(\lambda - 1)\begin{bmatrix} M_{r_0} & B^T \\ 0 & -W_r \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -r_0B^TW^{-1}Cx \\ Cx \end{bmatrix}.$$

It follows that $\lambda = 1$ for any eigenvector $\begin{bmatrix} x \\ y \end{bmatrix}$ in the form $x \in \mathcal{N}(C)$, $y \in \mathbf{R}^m$.

Since the preconditioner is nonsingular by assumption, it follows for any $x \notin \mathcal{N}(C)$ that

$$(\lambda - 1)y = -W_r^{-1}Cx = -rW^{-1}Cx\,,$$

i.e. $y = -\frac{r}{\lambda - 1}W^{-1}Cx$. Hence

$$(\lambda - 1)M_{r_0}x = (r - r_0)B^T W^{-1}Cx \quad (x \notin \mathcal{N}(C)).$$

Let μ be an eigenvalue of

$$Mx = \mu B^T W^{-1}Cx\,, \quad x \notin \mathcal{N}(C) \cup \mathcal{N}(B).$$

It follows that $(\lambda - 1)(\mu + r_0) = r - r_0$ or

$$\lambda = \frac{r + \mu}{r_0 + \mu}\,.$$

If $C = B$ and M is symmetric, then the problem

$$Mx = \mu B^T W^{-1}Bx\,, \quad x \notin \mathcal{N}(B)$$

has real eigenvalues,

$$-\mu_0 \le \mu \le \mu_1.$$

Therefore if $r_0 > \mu_0$, the eigenvalues λ are real and bounded as

$$1 \le \frac{r + \mu_1}{r_0 + \mu_1} \le \lambda \le \frac{r - \mu_0}{r_0 - \mu_0}$$

As $r > r_0 \to \infty$, the corresponding eigenvalues $\lambda \ne 1$ satisfy $\lambda \approx \frac{r}{r_0}$. It is advisable to let $r_0 = \nu_0 r$ for some (not very small) value $0 < \nu_0 < 1$. Then the eigenvalues cluster about ν_0. ■

Remark 2.12 *For applications of the action of the inverse of the preconditioner $\mathcal{B}_{r_0}$ we must solve inner systems with matrix $\mathcal{M}_{r_0}$. This matrix gets increasingly ill–conditioned for large values of r_0. However, as follows from Theorem 2.11, its spectrum is clustered in two tight intervals on the real axis. It is known, see e.g. [21], that using an spd preconditioner, in this case the conjugate gradient method converges essentially as if there were just one eigenvalue interval, with a correspondingly smaller condition number.*

Clearly, rounding errors may cause problems, and lead to effectively more spread–out eigenvalues but with the use of a few steps of a defect–correction method, where the residuals get increasingly smaller, one can avoid the need to use of higher precision than, say, double precision.

3 An application

Saddle point problems arise for instance when one uses mixed variable formulations of elliptic problems, see e.g. [2], [22].

An important problem of mixed form, which is similar to a mixed formulation of a diffusion equation, arises for flows in porous media, modelled by the Darcy flow equation,

$$\mathbf{u} = -\mathcal{K}\nabla p\,, \qquad \operatorname{div}\mathbf{u} = f \quad \text{in} \quad \Omega \tag{3.1}$$

where $\mathbf{u} \cdot \mathbf{n} = q_1$ on $\partial\Omega_N$, $p = q_2$ on $\partial\Omega_D$. Here Ω is a bounded domain in $\mathcal{R}^n$. Further $\mathbf{u}$ denotes the primal (vector) variable in the form of velocities, and p is the dual (scalar) variable corresponding to the pressure variable. The outward unit normal is given on the boundary $\partial\Omega_N$.

The weak formulation of problem (3.1) is

$$\begin{aligned} \int_\Omega \mathcal{K}^{-1}\mathbf{u} \cdot \mathbf{v} - \int_\Omega (\operatorname{div} \mathbf{v})p = 0 \qquad & \forall \mathbf{v} \in H(\operatorname{div})_{0,N}(\Omega), \\ \int_\Omega (\operatorname{div} \mathbf{u})q = \int_\Omega fq \qquad & \forall q \in L^2(\Omega), \end{aligned} \tag{3.2}$$

where $H(\operatorname{div})_{0,N}(\Omega) := \{\mathbf{v} \in H(\operatorname{div})(\Omega) : \ \mathbf{v} \cdot \mathbf{n} = 0 \text{ on } \partial\Omega_N\}$.

The pressure variable can be given on $\partial\Omega_D$. Possibly $\partial\Omega_D$ is an empty set, $\partial\Omega_D = \emptyset$, and the compatibility condition

$$\int_\Omega f dx = \int_{\partial\Omega_N} q_1 ds$$

must then hold. In this case, the pressure p is determined only up to an additive constant. The permeability coefficient matrix $\mathcal{K} = [\mathcal{K}_{ij}]$ in assumed to be symmetric, positive definite, and bounded in Ω, i.e., the eigenvalues of $\mathcal{K}(x)$ lie between uniform bounds

$$0 < \lambda_{min}(\mathcal{K}) < \lambda_{max}(\mathcal{K})$$

independent of x. In heterogeneous media, as is generally the case for porous media, $\mathcal{K}$ can vary a great deal between narrow small regions, i.e. $\lambda_{min}(\mathcal{K}) << \lambda_{max}(\mathcal{K})$. Therefore the reduced, elliptic equation

$$-\operatorname{div}\,(\mathcal{K}\nabla p) = f$$

can be extremely ill–conditioned and is even singular it we have pure Neumann boundary conditions.

As mixed methods are more insensitive to the variation in the coefficient term, which has also been pointed out in [23], this is one of the reasons why we keep the coupled form. Another reason is that mixed methods preserve mass locally, which property is of crucial importance, e.g. in modelling interfaces in groundwater flow over a porous media, see e.g. [24].

Furthermore, Darcy flow problems are normally coupled with other equations, such as in poro–elasticity problems, where the elasticity equations for the deformation of a porous medium is coupled to the Darcy flow velocity variable, or in fluid flow over such a media where the Stokes equations are coupled with the Darcy flow. In such problems it is important to have accurate approximation of the velocity and to satisfy mass conservation properties. Hence, the mixed formulation must be kept. See [14] and references therein for such applications.

The FEM discretization of this problem leads to a stiffness matrix of the saddle point form

$$\mathcal{M}_h = \begin{bmatrix} M_h & B_h^T \\ B_h & 0 \end{bmatrix},$$

where M_h and B_h are the Gramian matrices corresponding to the bilinear forms

$$a(u,v) = \int_\Omega \mathcal{K}^{-1}\mathbf{u}\cdot\mathbf{v}\,, \qquad b(u,p) = -\int_\Omega (\operatorname{div}\mathbf{u})p$$

arising in (3.2).

The problem can be reduced to homogeneous boundary conditions in a standard way. Here M_h is the discrete analogue (i.e. projection) of the operator $\mathbf{u}\to\mathcal{K}^{-1}\mathbf{u}$. Further, B_h is the discrete analogue of the operator

$$-\operatorname{div} : H(\operatorname{div})(\Omega) \supset\!\!\to L^2(\Omega)$$

defined on the subspace

$$D(-\operatorname{div}) := H(\operatorname{div})_{0,N}(\Omega) \subset H(\operatorname{div})(\Omega),$$

and B_h^T is the discrete analogue of the operator

$$\nabla : L^2(\Omega) \supset\!\!\to L^2(\Omega)^n,$$

defined on the subspace

$$D(\nabla) := \{p\in H^1(\Omega) :\ p = 0 \text{ on } \partial\Omega_D\} \subset L^2(\Omega).$$

We note that if $p\in L^2(\Omega)$ is not in the above $D(\nabla)$, then ∇p can be defined only in weak sense, see Remark 3.2 later.

3.1 Some estimates for the corresponding operators

For completeness, we state first some basic results, which in one or other form are familiar in operator theory.

Proposition 3.1 *The operator ∇ coincides with the adjoint of $-\operatorname{div}$ on its domain.*

Proof. For any $\mathbf{v}\in D(-\operatorname{div})$ and $p\in D(\nabla)$

$$0 = \int_{\partial\Omega_D} (p\,\mathbf{v}\cdot\mathbf{n})\,d\sigma + \int_{\partial\Omega_N} (p\,\mathbf{v}\cdot\mathbf{n})\,d\sigma = \int_{\partial\Omega} (p\,\mathbf{v}\cdot\mathbf{n})\,d\sigma$$

$$= \int_\Omega \operatorname{div}(p\,\mathbf{v})\,dx = \int_\Omega \nabla p\cdot\mathbf{v}\,dx + \int_\Omega p\,(\operatorname{div}\mathbf{v})\,dx,$$

i.e.

$$\langle\nabla p,\mathbf{v}\rangle_{L^2(\Omega)^n} = \langle p, -\operatorname{div}\mathbf{v}\rangle_{L^2(\Omega)}. \qquad \blacksquare \tag{3.3}$$

Remark 3.2 Relation (3.3) can also be used as the definition of the weak gradient ∇p as an element in the dual of $H(\operatorname{div})(\Omega)$ when $p\in L^2(\Omega)$ in general.

In particular, if $\mathbf{v}\in\ker\operatorname{div}$, i.e. $\operatorname{div}\mathbf{v} = 0$, then $\langle\nabla p,\mathbf{v}\rangle_{L^2(\Omega)^n} = 0$ for all $p\in D(\nabla)$, that is, $\mathbf{v}$ is orthogonal to the range of the operator ∇.

Corollary 3.3 *The range $\mathcal{R}(\nabla)$ is orthogonal to* $\ker\operatorname{div}$ *in $L^2(\Omega)^n$.*

Our goal is to apply a preconditioner based on Example 2.1. First, on the operator level, the role of W is proposed to be played by the operator $-\Delta$, defined on

$$D(-\Delta) := \{p \in H^1(\Omega) : \ p = 0 \text{ on } \partial\Omega_D, \ \ \nabla p \cdot \mathbf{n} = 0 \text{ on } \partial\Omega_N\}.$$

Then one can derive an analogue of (2.10), such that M is replaced by the operator G, where $G\mathbf{u} := \mathcal{K}^{-1}\mathbf{u}$, and B is replaced by the operator $-\mathrm{div}$. Namely, if λ is an eigenvalue of the operator equation corresponding to (2.3), then $\lambda = \frac{r}{r+\nu}$, where ν comes from the generalized eigenvalue problem

$$\mathcal{K}^{-1}\mathbf{u} = \nu\, \nabla(-\Delta)^{-1}(-\mathrm{div}\,\mathbf{u}) \qquad (\mathbf{u} \in D(\mathrm{div}),\ \mathbf{u} \neq 0). \tag{3.4}$$

Proposition 3.4 *The operator $\nabla(-\Delta)^{-1}(-\mathrm{div})$ satisfies*

$$\nabla(-\Delta)^{-1}(-\mathrm{div}\,\mathbf{u}) = P_{R(\nabla)}\mathbf{u} \qquad (\mathbf{u} \in D(-\mathrm{div})),$$

where $P_{\mathcal{R}(\nabla)}$ denotes the orthogonal projection to the range of ∇.

Proof. Denote the l.h.s. by $\mathbf{z} := \nabla(-\Delta)^{-1}(-\mathrm{div}\,\mathbf{u})$. Then $\mathbf{z} = \nabla w$, where $w \in H^1(\Omega)$ is the weak solution of problem

$$\begin{cases} -\Delta w = -\mathrm{div}\,\mathbf{u} \\ w_{\,|\partial\Omega_D} = 0, \qquad \nabla w \cdot \mathbf{n}_{\,|\partial\Omega_N} = 0\,. \end{cases}$$

Here $\mathbf{u} \in D(-\mathrm{div})$, further, $\mathbf{z} \in H(\mathrm{div})(\Omega)$ (since $\mathrm{div}\,\mathbf{z} = \Delta w$ exists in $L^2(\Omega)$) and $\mathbf{z}\cdot\mathbf{n} = \nabla w \cdot \mathbf{n} = 0$ on $\partial\Omega_N$, hence also $\mathbf{z} \in D(-\mathrm{div})$. Thus $\mathbf{u}-\mathbf{z} \in D(-\mathrm{div})$. Further,

$$\mathrm{div}\,(\mathbf{u}-\mathbf{z}) = \mathrm{div}\,\mathbf{u} - \Delta w = 0,$$

hence $\mathbf{u}-\mathbf{z} \in \ker(\mathrm{div})$, and thus $\mathbf{u}-\mathbf{z} \perp R(\nabla)$ by Corollary 3.3. Here $\mathbf{z} = \nabla w$ and $w \in D(\nabla)$, hence $\mathbf{z} \in R(\nabla)$. Since, $\mathbf{u} = \mathbf{z} + (\mathbf{u}-\mathbf{z})$, where $\mathbf{z} \in R(\nabla)$ and $\mathbf{u}-\mathbf{z} \perp R(\nabla)$, this means that $\mathbf{z}$ is the orthogonal projection of $\mathbf{u}$ to $R(\nabla)$. ■

Then an analogue of Proposition 2.5 can be formulated:

Proposition 3.5 *The generalized eigenvalues μ in (3.4) satisfy*

$$1/\lambda_{max}(\mathcal{K}) \leq \mu \leq 1/\lambda_{min}(\mathcal{K}),$$

where $0 < \lambda_{min}(\mathcal{K}) < \lambda_{max}(\mathcal{K})$ are the best uniform bounds for the eigenvalues of $\mathcal{K}(x)$.

Proof. Equation (3.4) and Proposition 3.4 imply

$$\mathcal{K}^{-1}\mathbf{u} = \mu\mathbf{z},$$

where $\mathbf{z}$ denotes the orthogonal projection of $\mathbf{u}$ to the range of ∇, i.e. $\mathbf{u} = \mathbf{z}+\widetilde{\mathbf{z}}$, where $\mathbf{z} \in R(\nabla)$ and $\widetilde{\mathbf{z}} := \mathbf{u}-\mathbf{z} \perp R(\nabla)$. Then

$$\mathbf{z} + \widetilde{\mathbf{z}} = \mathbf{u} = \mu\,\mathcal{K}\mathbf{z}.$$

Multiplying by $\mathbf{z}$, integrating and using the L^2-orthogonality of $\mathbf{z}$ and $\widetilde{\mathbf{z}}$,

$$\int_\Omega |\mathbf{z}|^2 = \mu \int_\Omega \mathcal{K}\mathbf{z}\cdot\mathbf{z}.$$

Here

$$\lambda_{min}(\mathcal{K}) \int_\Omega |\mathbf{z}|^2 \le \int_\Omega \mathcal{K}\mathbf{z}\cdot\mathbf{z} \le \lambda_{max}(\mathcal{K}) \int_\Omega |\mathbf{z}|^2,$$

which implies the desired inequalities. ∎

Consequently, by the relation $\lambda = \frac{r}{r+\mu}$,

$$\frac{r}{r+1/\lambda_{min}(\mathcal{K})} \le \lambda \le \frac{r}{r+1/\lambda_{max}(\mathcal{K})}.$$

3.2 Eigenvalue bounds for the discretized system

Consider now the discrete system. We shall verify that the above bound holds in this case also, independently of h. The preconditioner to $\mathcal{M}_h$ is

$$\mathcal{B} := \begin{bmatrix} M_r & 2B_h^T \\ 0 & -W_r \end{bmatrix}, \qquad \text{where } W_r = \frac{1}{r} W^{(h)},\ W^{(h)} := B_h G_h^{-1} B_h^T .$$

Here G_h denotes the unweighted mass matrix, whereas M_h denotes the weighted mass matrix using $\mathcal{K}^{-1}$ as in $a(u,v)$. Then

$$W^{(h)} = -\Delta_h,$$

i.e., corresponds to the discrete Laplacian. There are many fast solvers available for this matrix. Therefore the solution of systems with W_r is not costly. We note here that, on the contrary, the Schur complement $S_h = B_h M_h^{-1} B^T$, of the original system contains the ill–conditioned matrix M_h, depending on the variable coefficient matrix $\mathcal{K}$. Note also that in a functional space framework, a regularization with $\frac{1}{r}\nabla\nabla\cdot\mathbf{u}$ corresponds to a penalization in $L_2(\Omega)$–norm whereas the term $\frac{1}{r}\nabla\triangle^{-1}\nabla\cdot\mathbf{u}$ penalizes in a weaker norm, corresponding to the dual space $H^{-1} = (H_0^1(\Omega))'$.

Proposition 3.6 *We have* $\lambda_{min}(G_h^{-1}M_h) \ge 1/\lambda_{max}(\mathcal{K})$, $\lambda_{max}(G_h^{-1}M_h) \le 1/\lambda_{min}(\mathcal{K})$.

Proof. For all $\mathbf{u} \in V_h$ we have

$$\langle G_h\mathbf{c},\mathbf{c}\rangle = \int_\Omega \mathbf{u}^2, \qquad \langle M_h\mathbf{c},\mathbf{c}\rangle = \int_\Omega \mathcal{K}^{-1}\mathbf{u}\cdot\mathbf{u}$$

where $\mathbf{c}$ is the coordinate vector of $\mathbf{u}$ in V_h. Hence

$$(1/\lambda_{max}(\mathcal{K}))\,\langle G_h\mathbf{c},\mathbf{c}\rangle \ \le\ \langle M_h\mathbf{c},\mathbf{c}\rangle \le (1/\lambda_{min}(\mathcal{K}))\,\langle G_h\mathbf{c},\mathbf{c}\rangle$$

for all $\mathbf{c}$, which implies the desired inequalities. ∎

Let us consider the eigenvalues of $\mathcal{B}^{-1}\mathcal{M}_h$, i.e. the solutions of the generalized eigenvalue problem

$$\lambda \begin{bmatrix} M_r & 2B_h^T \\ 0 & -W_r \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} M_h & B_h^T \\ B_h & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad \|x\| + \|y\| \ne 0\,. \tag{3.5}$$

Proposition 2.8 and Corollary 2.9 then yield

Corollary 3.7 *We have* $\nu_{min} \geq 1/\lambda_{max}(\mathcal{K})$, $\nu_{max} \leq 1/\lambda_{min}(\mathcal{K})$, *and thus*

$$\frac{r}{r + 1/\lambda_{min}(\mathcal{K})} \leq \lambda \leq \frac{r}{r + 1/\lambda_{max}(\mathcal{K})}.$$

Note that here $\lambda \leq 1$, i.e. we must only ensure a uniform lower bound of λ. The above formula yields

Corollary 3.8 *If* r *is chosen to satisfy*

$$r \geq \frac{c}{\lambda_{min}(\mathcal{K})}$$

for some constant $c > 0$, *then* $\frac{c}{1+c} \leq \lambda \leq 1$, *i.e.* λ *is bounded independently of the variation of* $\mathcal{K}$.

It follows that, by the proposed regularization, the matrices involved in the preconditioner and the effect of the preconditioner depend little on the variation of the permeability coefficient matrix $\mathcal{K}$. This means that the method can robustly handle heterogeneous problems such as arising in Darcy flow problems for underground water flows see, e.g., [24] for further discussions of such problems.

Remark 3.9 *An interesting problem where the off–diagonal block matrices* C, B *are different is presented in [25]. Here the saddle point structure arises from a fictitious domain approach to solve elliptic boundary value problems, where the boundary conditions on the auxiliary boundary outside the given domain are enforced by certain control variables.*

Another problem with $C \neq B$ arises in a mixed variable formulation of a convection – diffusion problem

$$-\nabla \cdot A\nabla u - \mathbf{b}^T \nabla u + cu = f \quad \text{in } \Omega$$

with proper boundary conditions, when one rewrites it as a first order system

$$\begin{aligned} A^{-1}\underline{\sigma} + \nabla u &= 0 \\ \nabla \cdot \underline{\sigma} + \mathbf{b}^T A^{-1}\underline{\sigma} + cu &= f \end{aligned}$$

by introducing the flux $\underline{\sigma} = -A\nabla u$, see [26]. This problem is treated as one of the test problems in Section 5.

4 Inner and outer iterations

To solve systems with matrix $\mathcal{M} = \begin{bmatrix} M & B^T \\ C & 0 \end{bmatrix}$ arising in Darcy flow problems with strongly heterogeneous coefficients, it is less efficient to use a method based on the Schur complement matrix, $CM^{-1}B^T$. Such a system must be solved by iteration but, since M is highly ill–conditioned, it is difficult to construct a good preconditioner for it. Furthermore, each iteration requires the computation of residuals for it, which implies that one must use inner iterations to solve for the arising systems with M. Instead, in this paper we have based the methods on regularized forms of $\mathcal{M}$.

In the previous sections two preconditioners for saddle point matrices have been analyzed. For nonsingular matrices $\mathcal{M} = \begin{bmatrix} M & B^T \\ C & 0 \end{bmatrix}$, the preconditioner $\mathcal{B} := \begin{bmatrix} M_r & 2B^T \\ 0 & -W_r \end{bmatrix}$ can be used and, for large values of the regularization parameter r, it results in a strong clustering of all the eigenvalues around unity. For possibly singular matrices $\mathcal{M}$, a regularization can be done by the subtraction of a small perturbation in the $(2,2)$ block to form the matrix $\mathcal{M}_r := \begin{bmatrix} M & B^T \\ C & -W_r \end{bmatrix}$, which is nonsingular. This matrix can be factorized in the form

$$\mathcal{M}_r = \begin{bmatrix} M_{r_0} & B^T \\ 0 & -W_r \end{bmatrix} \begin{bmatrix} M_{r_0}^{-1}M_r & 0 \\ -W_r^{-1}C & I_2 \end{bmatrix} = \mathcal{B}_{r_0,\,r} \begin{bmatrix} M_{r_0}^{-1}M_r & 0 \\ -W_r^{-1}C & I_2 \end{bmatrix}$$

where $M_{r_0} = M + r_0 B^T W^{-1} C$ and $r_0 < r$. Here the first factor $\mathcal{B}_{r_0,\,r}$ can be used as preconditioner which, for large values of r_0, results in a clustering of the eigenvalues about the points r/r_0 and 1. Hence if we let $r = r_0$, then all eigenvalues cluster around unity for this method too. In that case, the preconditioned matrix $\mathcal{B}_{r_0,\,r}^{-1}\mathcal{M}_r$ equals $\begin{bmatrix} I_1 & 0 \\ -W_r^{-1}C & I_2 \end{bmatrix}$, so the use of a generalized conjugate gradient method like GCG or GMRES – in exact arithmetic – results in just two iterations. If $\mathcal{B}_{r_0,\,r}$ is applied to $\mathcal{M}$, then at any rate, even if $r_0 < r$ in this method as in the first case, the strong clustering results in an iteration which rapidly goes over in a superlinear rate of convergence.

Both of the above preconditioners involve matrices in the form $M_r = M + rB^T W^{-1} C$ and $W_r = \frac{1}{r}W$. In general, the corresponding linear systems that arise during each outer iteration step, are best solved by use of CG or GCG inner iterations. To preserve the good behaviour of the outer iteration method, the inner iterations should be solved fairly accurately, so that the outer iteration preconditioner behaves nearly as if it involved the corresponding exact matrices. The choice of preconditioners for the inner iteration systems is problem dependent. For the Darcy flow problem, presented in Section 3, we have seen that $W = -\Delta$ is an efficient choice, since it results in a matrix M_r which is spectrally equivalent to a mass matrix $\hat{M}$. Therefore $\hat{M}$ can be used as a preconditioner for M_r. For W we can use any of many available efficient preconditioners for the discrete Laplacian. As it turns out none of the corresponding inner iterations however are fully independent of the heterogeneity in the form of a strongly varying coefficient matrix $\mathcal{K}$. For other, more robust solvers, see [14].

As for the choice of the regularization parameters, in the first method r need not be very large, but in the second preconditioning method it must be large to make the perturbation of the $(2,2)$ block in $\mathcal{M}$ by W_r small. One can choose r_0 somewhat smaller to make M_{r_0} less ill-conditioned and the inner iterations fewer.

One can also approximate M_{r_0} with some of the following matrices

$$\tilde{M}_{r_0} = mI + r_0 P, \text{ where } P = B^T(BB^T)^{-1}B, \tag{4.1}$$

$$\tilde{M}_{r_0} = \sum R_k^T M_{r_0,k}^{-1} R_0. \tag{4.2}$$

The preconditioner (4.1) can be efficiently implemented by using the fact that P is a projection. Since $P^2 = P$ it follows that $(m+r_0)Pu = Pg$ and $m(I-P)u =$

$(I-P)g$ so a system $(mI+r_0P)u=g$ has the solution

$$u = Pu + (I-P)u = \frac{1}{m+r_0}Pg + \frac{1}{m}(I-P)g = \frac{r_0}{m+r_0}\left(\frac{1}{m}(I-P) + \frac{1}{r_0}I\right)g$$

which is easily computable to the cost of one application of P. The value m can e.g. be taken as some average, say, a some convex combination of the harmonic and geometric averages of diagonal elements of M.

The accuracy of this preconditioner follows from the next result.

Proposition 4.1 *Let P be a projection matrix and let $\frac{r_0}{m+r_0}(mI+r_0P)$ be a preconditioner to $\tilde{M}+r_0P$ where $m>0$. Then for the preconditioned matrix it holds $\frac{m+r_0}{r_0}(mI+r_0P)^{-1}(\tilde{M}+r_0P) = I + (I-P)\left(\frac{1}{m}\tilde{M}-I\right) + \frac{1}{r_0}\tilde{M}$*

Proof . It holds

$$\begin{aligned}
&\frac{m+r_0}{r_0}(mI+r_0P)^{-1}(\tilde{M}+r_0P) \\
&\quad = \left(\frac{1}{m}(I-P)+\frac{1}{r_0}I\right)(\tilde{M}+r_0P) = \left(\frac{1}{m}(I-P)\tilde{M}+\frac{1}{r_0}\tilde{M}+P\right) \\
&\quad = I + (I-P)\left(\frac{1}{m}\tilde{M}-I\right)+\frac{1}{r_0}\tilde{M}\,.
\end{aligned}$$

■

It follows from Proposition 4.1 that for large values of r_0 the preconditioned matrix approaches

$$I + (I-P)\left(\frac{1}{m}\tilde{M}-I\right),$$

which implies that the preconditioner (4.1) is accurate if $\frac{1}{m}\tilde{M}-I$ takes vectors essentially into the projection space of P that is, in our application, for vectors in the orthogonal complement of the divergence free space of functions.

The Schwarz type preconditioner (4.2) has been previously investigated in [16].

It is possible to combine the methods (4.1) and (4.2), i.e. use a very coarse mesh and use method (4.1) for each arising submatrix. This may enable the computation of more accurate average coefficients values m and, additionally, gives the opportunity to utilize parallel computations.

5 Numerical tests

To test the performance of the preconditioners we consider two problems

(i) A Darcy flow type problems.

(ii) An advection–diffusion transport type problem.

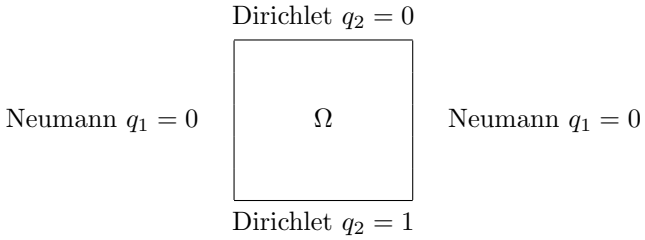

Figure 1: A unit square domain Ω and boundary conditions.

5.1 Darcy flow

We consider the Darcy flow type problem (3.1) in a unit square Ω with Dirichlet and homogeneous Neumann boundary conditions as given in Figure 5.1. We assume isotropic case $\mathcal{K} = kI$ with a scalar coefficient k which is stochastically generated. For such a problem a standard finite element discretization will lead to an ill–conditioned matrix and multigrid methods, for instance, will be less efficient. Furthermore, this system is singular when we have pure Neumann boundary conditions. In addition one is frequently more interested in computing the velocity vector $\underline{v} = -k\nabla p$ accurately then the pressure variable. For this reason we use the mixed variable formulation of (5.1). Furthermore, in important practical applications such as in porous media flow, the Darcy flow equations are coupled with either elasticity equations for the deformation of the material or with a Stokes equation for a fluid velocity, see e.g. [27], where the pressure is the coupling variable and it is not possible to solve the pressure separately.

The following choices of k have been made:

(1) Ω is divided into $h^{-1} \times h^{-1}$ square elements, where $k = k(x)$ is assumed to be constant,
(2) for any square element we first generate k_1, so that $\log k_1$ has normal distribution $N(0, 1)$ with the mean value 0 and variance 1,
(3) for $\sigma = 0$, 2 and 4, we define $k = k_1^\sigma$, which provide $\log k = \sigma \log k_1$ with normal distribution $N(0, \sigma^2)$ and the coefficient jumps with contrast $k_{\max}/k_{\min} = 1,\ \sim 0.5 \cdot 10^3$, and $\sim 0.275 \cdot 10^6$ for $\sigma = 0$, 2, 4, respectively.

The discretization of the corresponding mixed variable form is done with the triangular Thomas–Raviart elements of lowest order, $P_1 - P_0$ for u and p, respectively, see e.g. [28]. A uniform triangulation is used with rectangular elements which have two sides of length $h = 1/100$.

The FEM matrix $\mathcal{M}$ and the preconditioner $\mathcal{B}_r$ then take the form

$$\mathcal{M} = \begin{bmatrix} M & B^T \\ B & 0 \end{bmatrix}, \quad \mathcal{B}_r = \begin{bmatrix} M_r & B^T \\ 0 & -W_r \end{bmatrix}, \tag{5.1}$$

where the blocks M and B are described in Section 3. The matrix M is symmetric, positive definite with eigenvalues which can be bounded independently of h but not independently of the variation of the coefficient k. Due to the boundary conditions imposed and the inf-sup condition, the matrix B has full rank. The block matrices M_{r_0}, W_r in the preconditioner to $\mathcal{B}$ are defined as $M_{r_0} = M + r_0 B^T W_r^{-1} B$ and $W_r = \frac{1}{r} W$, respectively. Two choices, $W = I$ and

	r=	1	10^2	10^4	10^6	10^8
$\sigma = 1$	$W = I$	207	25	7	4	4
	$W = BB^T$	5	3	3	2	4
$\sigma = 2$	$W = I$	241	27	8	4	4
	$W = BB^T$	6	4	3	2	4
$\sigma = 3$	$W = I$	434	33	8	4	4
	$W = BB^T$	8	4	3	2	4
$\sigma = 4$	$W = I$	diverg.	53	10	4	4
	$W = BB^T$	12	4	4	2	4

Table 1: Number of outer iterations for GMRES with $\mathcal{B}_r$ preconditioner and different values of parameters: r-regularization and σ-coefficient jumps. Here $r_0 = r$.

	$r_0 = 1$	$r_0 = 10$	$r_0 = 10^2$	$r_0 = 10^3$	$r_0 = 10^5$	$r_0 = 10^6$
$W = I$	121	54	21	10	5	4
$W = BB^T$	4	3	3	2	2	2

Table 2: Number of outer iterations for GMRES with $\mathcal{B}_{r_0}$ preconditioner and different values of r_0. In all cases $r = 10^6$ and $\sigma = 2$.

$W = BB^T$ are tested.

The FE system $\mathcal{M}w = f$ arising from the discretization of the model problem is solved by the GMRES method [4] with the nonsymmetric matrix $\mathcal{B}_r$ as preconditioner. Thereby we use $\varepsilon = 10^{-6}$ as a relative stopping criteria.

The first set of experiments (Table 1, Table 2) shows the efficiency of the preconditioner $\mathcal{B}_r$ and the dependence of the number of iterations on the regularization parameters r and r_0 and the variation in coefficient k determined by the parameter σ. The inner systems with matrices M_r and W are here solved accurately.

It can be seen that the choice of $W = BB^T$ provides a more robust preconditioner, especially for smaller values of r. The number of iterations reach the asymptotical value 2 in some cases. The number of these (outer) iterations for large values of r is independent on coefficient heterogeneity and oscillations (σ) and for smaller values of r depends only very moderately on heterogeneity and oscillations of the coefficient k.

inner		$\sigma = 0$		$\sigma = 2$		$\sigma = 4$	
accur. ε_0	precnd.	$r_0 = 10^6$	$r_0 = 10^3$	$r_0 = 10^6$	$r_0 = 10^3$	$r_0 = 10^6$	$r_0 = 10^3$
$10e-6$	(4.2)	4(8)	15(30)	5(157)	19(531)	6(372)	24(808)
$10e-4$	(4.2)	4(8)	15(30)	6(131)	20(377)	6(226)	24(490)
$10e-2$	(4.2)	4(8)	15(30)	19(145)	21(198)	22(273)	29(298)

Table 3: Total number of outer and inner iterations (inner iterations are in brackets, for solving the model problem with $W = I$. The inner problem with matrix $M + r_0 B^T B$ is solved by the Schwarz preconditioner (4.2), with domain decomposition into 4 subdomains with overlap $2h$.

inner		$\sigma = 0$		$\sigma = 2$		$\sigma = 4$	
accur. ε_0	precnd.	$r_0 = 10^6$	$r_0 = 10^3$	$r_0 = 10^6$	$r_0 = 10^3$	$r_0 = 10^6$	$r_0 = 10^3$
$10e-6$	(4.0)	2(6)	3(7)	4(515)	4(487)	4(74h)	4(70h)
$10e-6$	(4.1)	2(6)	3(6)	3(305)	3(293)	4(53h)	4(71h)
$10e-6$	(4.2)	2(40)	3(24)	2(82)	3(102)	2(150)	4(239)
$10e-4$	(4.0)	2(6)	3(5)	14(555)	12(430)	13(89h)	13(74h)
$10e-4$	(4.1)	3(6)	3(5)	6(422)	6(365)	11(91h)	10(55h)
$10e-4$	(4.2)	7(57)	9(15)	8(231)	10(211)	8(406)	8(322)
$10e-2$	(4.0)	4(6)	3(5)	18(635)	18(523)	20(92h)	20(82h)
$10e-2$	(4.1)	6(7)	3(3)	10(419)	11(336)	19(94h)	18(65h)
$10e-2$	(4.2)	8(24)	7(21)	-	-	-	-

Table 4: Total number of outer and inner iterations (inner iterations are in brackets), 1h$\sim$ 1 hundred iterations) for solving the model problem with $W = BB^T$. The inner problem with matrix $M + r_0 B^T W^{-1} B$ is solved by the preconditioner (4.1), and the Schwarz preconditioner (4.2) with domain decomposition into 4 subdomains with overlap $2h$.

The next Table 2 shows the influence of taking smaller r_0 in the (1,1) block. We can see that the smaller values of r_0 influence the outer iterations for $W = BB^T$ only to a minor extent but deteriorates the method for $W = I$.

In practice a direct solution of the systems with M_{r_0} can be costly and in the case $W = BB^T$ even impossible. Thus in the second set of experiments, we use inner iterations for solving the systems for the block M_{r_0}, whereas the system with the (Laplacian) matrix W is solved by a direct method. Tables 3 and 4 show the influence of inaccurate solution of the systems with M_{r_0} , which means the solution by inner CG iterations [4], [11] with a relative stopping parameter $\varepsilon_0 = 10^{-2}$, 10^{-4} and 10^{-6}, respectively.

Tables 3 and 4 show dependence of the numbers of outer and inner iterations on the choice of regularization parameters, influence of inaccurate solution of systems with M_{r_0} and influence of heterogeneity and variation in coefficient k. As expected, the number of outer iterations decrease for smaller ε_0. Mostly the total number of inner iterations decreases also with ε_0 when $W = BB^T$. The number of inner iterations are very large for $W = I$, see Table 3. In this case we should need a better preconditioner like the Schwarz-type one (4.2). As also the theoretical results in Section 3 shows for $W = BB^T$ the efficiency of inner iterations is good, in particular for the projection type preconditioner (4.2) in the case of smaller oscillations in the coefficient k. For the Schwartz type preconditioner, which however is more costly, the results are good for all values of σ. A further improvement might be possible by choosing a proper matrix G in proposition 2.8 to let $W = BG^{-1}B^T$. Such a choice could be a weighted mass matrix $G(K)$ depending on the heterogeneous coefficients k. This was also discussed in Section 3.2.

In Table 4 has also results for the preconditioner (4.0) $\hat{M} + r_0 I_1$ been included. It is seen that its behaviour is similar to that of preconditioner (4.1). Since it is cheaper to implement and apply, it is preferable.

5.2 Advection-diffusion problem

We test also the block preconditioners for solving a nonsymmetric saddle point system which is a modification of (5.1). For this purpose, let us consider an advection–diffusion transport problem (see [31]) written in the form

$$\underline{\xi} = -\delta \nabla c \quad \text{div } \underline{\xi} - \mathbf{b} \cdot \frac{1}{\delta} \underline{\xi} = \varphi \quad \text{in } \Omega \,.$$

Here $\delta > 0$ is the diffussion–dispersion coefficient, c is the unknown concentration of a species, $\underline{\xi}$ is the diffusive flux, $\mathbf{b}$ is the velocity in the transport advective field, and φ is the volume source term. In the test problem, we assume that Ω is the unit square, $\delta \in \{10^{-2}, 10^{-4}, 10^{-6}\}$, the velocity vector is constant, $\mathbf{b} = (2,1)^T$ and $\varphi = 0$. The boundary conditions are given as

$$\begin{aligned} (\underline{\xi} + \mathbf{b}c) \cdot \mathbf{n} &= g \text{ on } \Gamma_{in} = \{x \in \partial\Omega : \mathbf{b} \cdot \mathbf{n}(x) < 0\} \,, \\ \frac{1}{\delta} \underline{\xi} \cdot \mathbf{n} &= 0 \text{ on } \Gamma_{out} = \{x \in \partial\Omega : \mathbf{b} \cdot \mathbf{n}(x) \geq 0\} \,. \end{aligned}$$

Here $\mathbf{n}$ is the unit outer normal, Γ_{in} and Γ_{out} are the inlet and outlet parts of the boundary, $g(x) = 1$ for $x \in \Gamma_{in}$, with $x_2 = 0$ and $g(x) = 0$ for $x \in \Gamma_{in}$, with $x_1 = 0$.

The mixed formulation of the above problem with unknowns $\underline{\xi} \in H^2(\text{div}, \Omega)$, $\underline{\xi} \cdot \mathbf{n} = 0$ on Γ_{out}, $c \in L_2(\Omega)$ and subsequent discretization by the lowest order Raviart–Thomas finite elements (see [32]) on the same grid with $h = 1/100$ as in Subsection 5.1 provides the FEM matrix

$$\mathcal{M} = \begin{bmatrix} M & B^T \\ C & 0 \end{bmatrix} \tag{5.2}$$

where M, B are the same as in Subsection 5.1, $C = B + M_u$ and M_u is the matrix arising from discretization of the bilinear form m_u ,

$$m_u(\underline{\eta}, \gamma) = \int_\Omega \frac{1}{\delta} \underline{\eta}\, \gamma \, dx \quad \text{for } \underline{\eta} \in H^2(\text{div}, \Omega),\ \gamma \in L_2(\Omega) \,.$$

The tested preconditioners have the form

$$\mathcal{B} = \begin{bmatrix} M_{r_0} & B^T \\ 0 & -W_r \end{bmatrix},\ M_{r_0} = M + r_0 B^T W^{-1} C, \quad W_r = \frac{1}{r} W$$

where $W = I$ or $W = BB^T$.

From Table 5, we can again see that only 2 outer Krylov type method iterations are required when approaching the limit $r \to \infty$ as the preconditioned system approaches a matrix with quadratic minimal polynomial. From Table 6, it is seen that the choice $W = BB^T$ is efficient only if the diffusion part of the problem is not too much suppressed by the advection.

In the case, when the convection part C dominates over B, it is more efficient to use the regularized matrix $\mathcal{M}_r$ with (2,2) block $-W_r = \frac{1}{r} I$ as a preconditioner. Due to the factorization (2.12), the application of this preconditioner

	r=	1	10^2	10^4	10^6	10^8
$\delta = 10^{-2}$	$W = I$	$\succ$99	20	6	4	2
	$W = BB^T$	8	4	3	2	2
$\delta = 10^{-4}$	$W = I$	$\succ$99	68	8	4	2
	$W = BB^T$	$\succ$99	40	6	4	3
$\delta = 10^{-6}$	$W = I$	$\succ$99	86	8	4	2
	$W = BB^T$	$\succ$99	$\succ$99	46	8	4

Table 5: Number of outer iterations, for the advection–diffusion problem, $r_0 = r$, $\succ$99 means that the accuracy $\varepsilon = 10^{-6}$ is not reached within 99 iterations.

	$r_0 = 1$	$r_0 = 10$	$r_0 = 10^2$	$r_0 = 10^3$	$r_0 = 10^5$	$r_0 = 10^6$
$W = I$	$\succ$99	$\succ$99	63	15	5	4
$W = BB^T$	$\succ$99	$\succ$99	30	11	5	4

Table 6: $r = 10^6$ and $\delta = 10^{-4}$, $r = 10^6$.

$\mathcal{B}_{II}$ reduces to the solution of two triangular systems

$$(\mathcal{B}_{II})^{-1}\, w = \begin{bmatrix} M & B^T \\ C & -W_r \end{bmatrix}^{-1} w = \begin{bmatrix} I_1 & 0 \\ -rC & I_2 \end{bmatrix}^{-1} \begin{bmatrix} M_r & B^T \\ 0 & -r^{-1} I_2 \end{bmatrix}^{-1} w. \tag{5.3}$$

As above, it is also possible to use M_{r_0} instead of M_r in (5.3), which provides

$$(\mathcal{B}_{II})^{-1}\, w = \begin{bmatrix} I_1 & 0 \\ rC & I_2 \end{bmatrix} \mathcal{B}^{-1} w. \tag{5.4}$$

and therefore the difference in the computational effort between application of $\mathcal{B}$ and $\mathcal{B}_{II}$ is very small. Furthermore,

$$(B_{II})^{-1}\mathcal{M} = \begin{bmatrix} I_1 & 0 \\ 0 & I_2 \end{bmatrix} + \begin{bmatrix} 0 & M_r^{-1}B^T \\ 0 & rCM_r^{-1}B^T - I_2 \end{bmatrix} \;\to\; \begin{bmatrix} I_1 & 0 \\ 0 & I_2 \end{bmatrix},$$

as $r \to \infty$, since $M_r^{-1}B^T = \frac{1}{r}(\frac{1}{r}M + B^TC)^{-1}B^T \to 0$ and $rCM_r^{-1}B^T = C(\frac{1}{r}M + B^TC)^{-1}B^T \to I_2$ because $(B^TC)^{-1}$ exists and $C(B^TC)^{-1}B^TC = C$ implies $C(B^TC)^{-1}B^T = I_2$ due to full rank of the matrices C and B. The efficiency of the $\mathcal{B}_{II}$ preconditioner is therefore high as the preconditioned system converges to identity with $r \to \infty$. The use of $\mathcal{B}_{II}$ for preconditioning the system (5.2) is illustrated in Tables 7 and 8.

	r=	1	10^2	10^4	10^6	10^8
$\delta = 10^{-2}$	$W = I$	78	10	3	2	1
$\delta = 10^{-4}$	$W = I$	$\succ$99	34	4	2	1
$\delta = 10^{-6}$	$W = I$	$\succ$99	43	4	2	1

Table 7: Number of outer iterations, $r_0 = r$.

	$r_0 = 1$	$r_0 = 10$	$r_0 = 10^2$	$r_0 = 10^3$	$r_0 = 10^5$	$r_0 = 10^6$
$W = I$	$\succ 99$	$\succ 99$	$\succ 99$	35	6	2

Table 8: Number of outer iterations, $r = 10^6$ and $\delta = 10^{-4}$.

6 Conclusions

Special block matrix preconditioners for original and regularized forms of saddle point matrix have been analysed and tested numerically on a Darcy flow problem with strongly heterogeneous coefficients and advection–diffusion transport problems with dominating advection. The main focus is on the outer iterations but also various inner iteration preconditioners for the regularized (1,1) pivot block matrix have been applied. The arising Schur complement matrix is approximated simply by a scalar multiple of the identity matrix or by a discrete isotropic Laplacian. For the most efficient versions of the regularization only very few outer iterations are needed, which is in full accordance with the presented theory. Moreover, the number of outer iterations is independent of degree of heterogeneity involved in the material coefficient. The number of inner iterations, however, is sensitive to the choice of preconditioner to (1,1) block and heterogeneity. Good choices are Schwarz type preconditioner (4.2) in both of the cases $W = I$ and $W = BB^T$, cf. [16], and also the preconditioner (4.1) in the case of $W = BB^T$. For the advection problem, $W = I$ is preferable.

The main conclusion is that the number of outer iterations is always small and the block matrix preconditioning is robust. The number of inner iterations increases with increasing variance of the heterogeneous coefficients but the number of inner iterations is still quite acceptable for method (4.2) and also for (4.1), if we let $W = BB^T$. That this topic can be quite involved follows from various papers on the solution of heterogeneous material problems, such as [24] and [30]. Another promising solution method could be an algebraic multigrid method, see e.g. [33] and [34].

An alternative method would be to eliminate the velocity variable and solve the arising Schur complement system with the matrix $BM_r^{-1}B^T$. Here, however, to get sufficiently small residuals, one must solve the inner system with M_r more accurately and more frequently than needed for our block matrix approach, which would increase the number of inner iterations even further.

References

[1] E.G. D'yakhonov, Iterative methods with saddle operators. *Dokl. Akad. Nauk SSSR*, 292(1987), pp. 1037–1041.

[2] O. Axelsson, Preconditioning of indefinite problems by regularization. *SIAM J. Numer. Anal.*, 16(1979), pp. 58–69.

[3] R.E. Bank, B. D. Welfert, H. Yserentant, A class of iterative methods for solving saddle point problems. *Numer. Math.*, 56(1990), pp. 645–666.

[4] Y. Saad, *Iterative Methods for Sparse Linear Systems*. PS Publishing Company, Boston, 1996.

[5] W. Zulehner, Analysis of iterative methods for saddle point problems: a unified approach. *Math. Comp.*, 71(2002), pp. 479–505.

[6] A. Klawonn, Block-triangular preconditioners for saddle point problems with a penalty term. *SIAM J. Sci. Comp.*, 19(1998), pp. 172–184.

[7] O. Axelsson, M. Neytcheva, Preconditioning methods for linear systems arising in constrained optimization problems. *Numer. Linear Algebra Appl.*, 10(2003), pp. 3–31.

[8] G.H. Golub, C. Greif, On solving block–structured indefinite linear systems. *SIAM J. Sci. Comp.*, 24(2003), pp. 2076-2092.

[9] M. Benzi, G.H. Golub, J. Liesen, Numerical solution of saddle point problems. *Acta Numerica* 14(2005), pp. 1-137.

[10] V. Simoncini, Block triangular preconditioners for symmetric saddle–point problems. *Appl. Numer. Math.*, 49(2004), pp. 63–80.

[11] O. Axelsson, *Iterative Solution Methods*. Cambridge University Press, 1994.

[12] O. Axelsson, R. Blaheta, M. Neytcheva, Preconditioning for boundary value problems using elementwise Schur complements. *SIAM J. Matrix Anal. Appl.*, 31(2009), pp. 767–789.

[13] C. Greif, D. Schötzau, Preconditioners for saddle point problems with highly singular (1–1) blocks. *Electronic Transactions on Numerical Analysis*, 22(2006), pp. 114–121.

[14] O. Axelsson, R. Blaheta, P. Byczanski, Stable discretization and efficient preconditioners of poroelasticity problems for arising saddle point type matrices. J. Computing and Visualisation in Science, 15 (2012) 191–207.

[15] O. Axelsson, M. Neytcheva, A general approach to analyse preconditioners for two–by–two block matrices. *Numer. Linear Algebra Appl.*, (2011). Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nla.830.

[16] O. Axelsson, R. Blaheta, Preconditioning of matrices partitioned in two by two block form: Eigenvalue estimates and Schwarz DD for mixed FEM. *Numer. Linear Algebra Appl.* 17(2010), pp. 787–810.

[17] O. Axelsson, M. Neytcheva, Eigenvalue estimates for preconditioned saddle point matrices. *Numer. Linear Algebra Appl.* 13(2006), pp. 339–360.

[18] Z.-H. Cao, A note on spectrum distribution of constraint preconditioned generalized saddle point matrices. *Numer. Linear Algebra Appl.* 16(2009), pp. 503–516.

[19] O. Axelsson, P.S. Vassilevski, A black box generalized conjugate gradient solver with inner iterations and variable–step preconditioning. *SIAM J. Matrix Anal. Appl.* 12(1991), pp. 625-644.

[20] Y. Saad, A flexible inner-outer preconditioned GMRES-algorithm. *SIAM J. Scientific Computing*, 14(1993), pp. 461-469.

[21] O. Axelsson, A.V. Barker. Finite Element Solution of Boundary Value Problems: Theory and Computation. *Classics in Applied Mathematics*, 35, *SIAM*, Philadelphia, 2001.

[22] P.S. Vassilevski, R.D. Lazarov, Preconditioning mixed finite element saddle–point elliptic problems. *Numer. Linear Algebra Appl.*, 3(1996), pp. 1–20.

[23] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods. Springer Verlag, 1991.

[24] M.F. Wheeler, T. Wildey, G. Xue, Efficient algorithms for multiscale modelling in porous media. *Numer. Linear Algebra Appl.*, 17(2010), pp. 771–786.

[25] J. Haslinger, T. Kozubek, R. Kučera, A. Peichl, Projected Schur complement method for solving non–symmetric systems arising from a smooth fictitious domain approach. *Numer. Linear Algebra Appl.*, 14(2007), pp.713–739.

[26] A.I. Pehlivanov, G.F. Carey, P.S. Vassilevski, Least–squares mixed finite element methods for non–selfadjoint elliptic problems: I. Error estimates. *Numer. Math.*, 72(1996), pp. 501–522.

[27] W.L. Layton, F. Schieweck, I. Yotov, Coupling fluid flow with porous media flow. *SIAM J. Numer. Anal.*, 40(2003),pp. 2195–2218.

[28] C.E. Powell, S. Silvester, Optimal preconditioning for Raviart–Thomas mixed formulation of second–order elliptic problems. *SIAM J. Matrix Anal.*, 25(2004), pp. 718–738.

[29] C. C. Paige, M. A. Saunders, Solutions of sparse indefinite systems of linear equations. *SIAM J. Numerical Analysis*, 12(1975), 617–629.

[30] J. Galvis, Y. Efendiev, Domain decomposition preconditioners for multiscale flows in high–contrast media. *SIAM J. Multiscale Model. Simul.*, 8(2010), pp. 1461–1483.

[31] L. El Alaoni, A. Ern, E. Burman, A priori and a posteriori analysis of nonconforming finite elements with face penalty for advection–diffusion equations. *IMA J. Numer. Anal. (2005).*

[32] J. Douglas Jr., J.E. Roberts, Mixed finite element methods for second order elliptic problems. *Math. Appl. Comput.* 1(1982), pp.91–103.

[33] Y. Notay, AGMG software and documentation.
See http://homepages.ulb.ac.be/∼ ynotay/AGMG

[34] P.S. Vassilevski, *Multilevel Block Factorization Preconditioners. Matrix–based Analysis and Algorithms for Solving Finite Element Equations,* Springer, New York, 2008, 514 p.