ISSN 1360-1725

Parameter-free H(div) preconditioning for mixed finite element formulation of diffusion problems

Catherine E. Powell

Numerical Analysis Report No. 453

July 2004

Manchester Centre for Computational Mathematics
Numerical Analysis Reports

DEPARTMENTS OF MATHEMATICS

Reports available from:	And over the World-Wide Web from URLs
Department of Mathematics University of Manchester Manchester M13 9PL England	http://www.ma.man.ac.uk/MCCM/MCCM.html ftp://ftp.ma.man.ac.uk/pub/narep

PARAMETER-FREE H(DIV) PRECONDITIONING FOR MIXED FINITE ELEMENT FORMULATION OF DIFFUSION PROBLEMS

CATHERINE E. POWELL *

Abstract. In this paper, we implement the lowest-order Raviart-Thomas scheme for generalised diffusion problems and analyse the robustness of so-called $H(div)$ preconditioning for the resulting saddle-point systems. Properties of the exact version of the suggested block-diagonal preconditioner were discussed in some previous work, 'Optimal preconditioning for Raviart-Thomas mixed formulation of second-order elliptic problems', SIAM J. Matrix Anal. Appl., 25, 3(2003). Here, new bounds are established for the eigenvalue spectrum of the preconditioned system matrix when a single V-cycle of any available multigrid scheme is used to approximately invert the matrix corresponding to the discrete weighted $H(div)$ operator.

Key words: saddle-point problems, variable coefficients, mixed finite elements, Raviart-Thomas, $H(div)$, minimum residual method, preconditioning.

AMS subject classifications: 65F10, 65N22, 65M60, 65M55

1. Introduction.

Let Ω be a polygon in $I\!R^d$, with boundary $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$. We consider scalar, second-order elliptic problems of the form,

$$\begin{aligned} -\nabla \cdot \mathcal{A}\nabla p &= f \quad \text{in } \Omega, \\ p &= g \quad \text{on } \partial\Omega_D, \\ \mathcal{A}\nabla p \cdot \vec{n} &= 0 \quad \text{on } \partial\Omega_N, \end{aligned} \tag{1.1}$$

where $\partial\Omega_D \neq \emptyset$ and $\mathcal{A} = \mathcal{A}(\vec{x})$ is a $d \times d$ bounded, symmetric and uniformly positive definite matrix-valued function. Specifically, we assume that there exist positive constants γ and Γ with $0 < \gamma \leq \Gamma$ such that,

$$\gamma(\vec{v}, \vec{v}) \leq (\mathcal{A}^{-1}\vec{v}, \vec{v}) \leq \Gamma(\vec{v}, \vec{v}), \tag{1.2}$$

for every $\vec{v} : \Omega \to I\!R^d$, where $(\cdot, \cdot)$ denotes the standard L^2 inner-product. Substituting $\vec{u} = \mathcal{A}\nabla p$ in (1.1) yields the mixed first-order system,

$$\begin{aligned} \mathcal{A}^{-1}\vec{u} - \nabla p &= 0, \\ \nabla \cdot \vec{u} &= -f \quad \text{in } \Omega, \\ p &= g \quad \text{on } \partial\Omega_D, \\ \vec{u} \cdot \vec{n} &= 0 \quad \text{on } \partial\Omega_N. \end{aligned} \tag{1.3}$$

Boundary-value problems of this type occur, for example, in mathematical models of fluid flow in porous media (see [24], [11]). The macroscopic flow of groundwater in a porous medium was first shown, by Henri Darcy in the nineteenth century, to satisfy the relation $\vec{u} = -k\mu^{-1}\nabla P_R$, where $\vec{u}$ denotes fluid discharge or velocity, P_R is the 'residual pressure', k is the permeability coefficient of the medium and μ is the fluid viscosity. Coupling Darcy's law with the mass conservation law $\nabla \cdot \vec{u} = 0$, yields precisely the first

*Department of Mathematics, University of Manchester Institute of Science and Technology, Manchester M60 1QD, UK.cp@fire.ma.umist.ac.uk.

equation of (1.1) with $f = 0$, $\mathcal{A} = \frac{k}{\mu}\mathcal{I}$ and $p := P_R$. To fix ideas, then, we will call p and $\vec{u} = \mathcal{A}\nabla p$ the 'pressure' and 'velocity' solutions, respectively. Mixed finite element methods are favoured when $\vec{u}$ is the variable of interest since post-processing primal pressure solutions leads to loss of accuracy. Crucially, mixed methods also conserve mass locally.

If the entries of the coefficient tensor $\mathcal{A}$ have the same value at all points in the domain considered, we describe the medium as homogeneous. In practise, flow domains are comprised of different media with spatially varying porosity volumes, leading to heterogeneous problems with discontinuous $\mathcal{A}$. If $\mathcal{A}$ depends on the direction of flow, then we say that the medium is anisotropic. This occurs in stratified media with alternating layers. The robustness of numerical solution schemes for (1.3), with respect $\mathcal{A}$, is of critical importance in modelling Darcy flow.

It is well known (see [1], [3], [19]) that (1.3) may be formulated as a saddle-point problem in two distinct pairs of function spaces. In this paper, we work in the standard test spaces,

$$V = H_{0,N}(div;\Omega) = \{\vec{v} \in H(div;\Omega) \mid \vec{v}\cdot\vec{n} = 0 \text{ on } \partial\Omega_N\},$$

and $W = L^2(\Omega)$. Here, $L^2(\Omega)$ is the usual space of square-integrable scalar functions and we use the notation, $H(div;\Omega) = \left\{\vec{v} \in L^2(\Omega)^d \mid \nabla\cdot\vec{v} \in L^2(\Omega)\right\}$ to denote the Hilbert space of vector functions endowed with the inner-product,

$$(\vec{u},\vec{v})_{div} = (\vec{u},\vec{v}) + (\nabla\cdot\vec{u},\nabla\cdot\vec{v}), \tag{1.4}$$

and the induced norm $\|\cdot\|_{div}$. Multiplying by arbitrary $\vec{v} \in V$ and $w \in W$ in (1.3) and integrating by parts, yields the continuous variational problem,

find $(\vec{u},p) \in V \times W$ satisfying,

$$\begin{array}{rcll} (\mathcal{A}^{-1}\vec{u},\vec{v}) + (p,\nabla\cdot\vec{v}) &=& < g,\vec{v}\cdot\vec{n} > & \forall\,\vec{v}\in V, \\ (w,\nabla\cdot\vec{u}) &=& -(f,w) & \forall\,w\in W, \end{array} \tag{1.5}$$

where $\langle g,\vec{v}\cdot\vec{n}\rangle = \int_{\partial\Omega_D} g\,\vec{v}\cdot\vec{n}\,ds$. The data requirements are $f \in L^2(\Omega)$ and $g \in H^{\frac{1}{2}}(\partial\Omega_D)$. That a unique solution exists, follows from the stability theory of Brezzi and Babuška (see [6], [23]) and condition (1.2). To compute a solution numerically, we choose finite-dimensional subspaces $V_h \subset V$ and $W_h \subset W$, consisting of piecewise polynomials defined on a partition T_h of Ω into finite elements, and solve the discrete variational problem,

find $(\vec{u_h},p_h) \in V_h \times W_h$ satisfying,

$$\begin{array}{rcll} (\mathcal{A}^{-1}\vec{u_h},\vec{v_h}) + (p_h,\nabla\cdot\vec{v_h}) &=& \langle g,\vec{v_h}\cdot\vec{n}\rangle & \forall\,\vec{v_h}\in V_h, \\ (w_h,\nabla\cdot\vec{u_h}) &=& -(f,w_h) & \forall\,w_h\in W_h. \end{array} \tag{1.6}$$

To ensure stability (see [6]) of the approximation, V_h and W_h must be chosen so that there exist constants α_h and β_h with $\alpha_h > \alpha_* > 0$, $\beta_h > \beta_* > 0$, for some α_* and β_*

independent of the representative discretisation parameter, h, satisfying,

$$(\mathcal{A}^{-1}\vec{v_h}, \vec{v_h}) \geq \alpha_h \parallel \vec{v_h} \parallel_V^2, \quad \forall \vec{v_h} \in Z_h, \tag{1.7}$$

$$\sup_{\vec{v_h} \in V_h} \frac{(\nabla \cdot \vec{v_h}, w_h)}{\parallel \vec{v_h} \parallel_V} \geq \beta_h \parallel w_h \parallel_W \quad \forall w_h \in W_h, \tag{1.8}$$

with $Z_h = \{\vec{v_h} \in V_h \mid (\nabla \cdot \vec{v_h}, w_h) = 0 \quad \forall w_h \in W_h\}$.

We focus attention on the lowest-order Raviart-Thomas method (see [22].) Hence, let T_h be a partition of Ω into triangular ($d = 2$) or tetrahedral ($d = 3$) finite elements with $h = \max_{T \in T_h} h_T$ where h_T denotes the diameter of an element T. Define $V_h = \{\vec{v} \in H(div;\Omega) \mid \vec{v} \mid_T \in (P_0^d(T) + [x,y]P_0(T)) \quad \forall T \in T_h\}$ and $W_h = \{w \in L^2(\Omega) \mid w \mid_T \in P_0(T) \quad \forall T \in T_h\}$, where $P_0(T)$ denotes the set of polynomials of degree zero defined on element T. Note that to be conforming in $H(div;\Omega)$, the velocity degrees of freedom are the normal components at the centres of the edges ($d = 2$) or faces ($d = 3$). The piecewise constant pressure solution is sampled at element centroids. Condition (1.8) is established in [22] and since, crucially, $\nabla \cdot V_h \subset W_h$, (1.7) is satisfied if (1.2) holds. The lowest-order quadrilateral elements can also be used (see [22], [19].) Higher order schemes could be used but for discontinuous coefficients we cannot expect high solution regularity.

Choosing bases $V_h = \mathtt{span}\{\vec{\varphi}_i\}_{i=1}^n$ and $W_h = \mathtt{span}\{\phi_j\}_{j=1}^m$, yields,

$$\underbrace{\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}}_{C} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{g} \\ \underline{f} \end{pmatrix}, \tag{1.9}$$

with matrices $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m \times n}$ defined via,

$$A_{i,j} = \left(\mathcal{A}^{-1}\vec{\varphi}_i, \vec{\varphi}_j\right) \quad i,j = 1:n,$$
$$B_{k,j} = (\nabla \cdot \vec{\varphi}_j, \phi_k) \quad k = 1:m,\ j = 1:n,$$

and vectors $g_i = \langle g, \vec{\varphi}_i \cdot \vec{n}\rangle$, $f_k = -(f, \phi_k)$ for $i = 1:n$ and $k = 1:m$. Saddle-point systems such as (1.9) are amenable to solution by the Krylov subspace method MINRES (see [18]). However, C is extremely ill-conditioned (see [25] and [27]) and the method will converge poorly unless a preconditioner P is available such that the eigenvalues of $P^{-1}C$ are bounded independently of h and $\mathcal{A}$.

The minimum residual scheme, or MINRES, is our method of choice since it is a parameter-free iteration with an optimal approximation property. Specifically, the iterate produced at the kth step minimises the Euclidean norm of the residual error. Surprisingly, it does not seem to be widely used. An alternative is Uzawa's method which, in its standard form, is the algorithm shown in Fig. 1.1. It requires the action of the inverse of the matrix A at each step. Since this is infeasible for general coefficient tensors, modern variants use an approximation, leading to nested iteration. The choice of the relaxation parameter α is crucial for convergence. Optimal values are known (see [12]) but require the solution of a costly eigenvalue problem.

```
Given initial guess p^(0),
for  n = 1,2,...,          until convergence:
       Solve:              A u^(n+1) = g - B^T p^(n)
       Update:             p^(n+1) = p^(n) + α (B u^(n+1) - f)
end
```

FIG. 1.1. *Standard Uzawa algorithm*

Another deficiency is the need to tune the parameter α and stopping tolerances for the inner-iteration (e.g. see [25].) A second popular approach is the so-called augmented Lagrangian method (see [12, Ch.1].) This refers to the application of Uzawa's algorithm to the modified saddle-point problem,

$$\begin{pmatrix} A + rB^T B & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{g} + rB^T \underline{f} \\ \underline{f} \end{pmatrix}, \tag{1.10}$$

which has the same solution as (1.9) for any $r \in \mathbb{R}$. Choosing the relaxation parameter $\alpha = r$, where r is large, yields arbitrarily fast convergence. However, the method has a serious flaw. The condition number $\kappa\left(A + rB^T B\right)$ increases with r (see Proposition 2.3 in [12]), crippling the inner-iteration. Preconditioners for the inner-solve have been suggested (e.g. see [14]) but optimal values of α and r have not been discussed. Parameter tuning is required. The authors of [29] solve (1.10) with MINRES and propose a preconditioning scheme similar to the one we will use for the original system (1.9). However, convergence of their scheme is highly sensitive to the choice of two artificial parameters for which optimal values are not established.

Arnold, Falk and Winther observe in [1] and [2] that for unit coefficients $\mathcal{A} = \mathcal{I}$, C in (1.9) also has the same mapping properties as the matrix operator,

$$\mathcal{P} = \begin{pmatrix} \mathcal{H} & 0 \\ 0 & \mathcal{I} \end{pmatrix}, \tag{1.11}$$

where $\mathcal{H} : H(div) \times H(div) \to I\!R$ is the PDE operator defined via $(\mathcal{H}\vec{u}, \vec{v}) = (\vec{u}, \vec{v})_{div}$. $\mathcal{H}$ gives rise to a matrix similar to the one occurring in the penalty method and augmented Lagrangian methods and can be approximated by any of the multilevel schemes suggested for those formulations (see [7], [28], [10], [14]). A simpler method is presented by Arnold et al. in the form of a standard multigrid V-cycle, with additive Schwarz smoothing. The resulting approximation is optimal for weighted $H(div)$ operators $\mathcal{H}_{\mathcal{A}}$, defined via,

$$(\mathcal{H}_{\mathcal{A}}\vec{u}, \vec{v}) = (\vec{u}, \vec{v})_{div,\mathcal{A}} = \left(\mathcal{A}^{-1}\vec{u}, \vec{v}\right) + (\nabla \cdot \vec{u}, \nabla \cdot \vec{v}), \tag{1.12}$$

in the special case of quasi-uniform tetrahedral meshes and coefficients of the form $\mathcal{A} = \rho\mathcal{I}$ where $\rho \in \mathbb{R}_+$. However, incorporating an $H(div)$ solver into a block-diagonal preconditioner for (1.9) has not been fully explored. Until now, there has been no systematic analysis of the eigenvalue spectrum of the Raviart-Thomas saddle-point matrix preconditioned by a matrix corresponding to an operator of the form (1.11) without artificial parameters. Note also that theory concerning the quality of multilevel approximations to operators $\mathcal{H}_{\mathcal{A}}$ with anisotropic and/or variable ρ seems to be limited.

1.1. Overview. This paper is an extension of some work reported in [19], where we established a uniform eigenvalue bound for the saddle-point matrix in (1.9) preconditioned by a matrix of the form (1.11). That analysis focused on the exact inversion of the matrix corresponding to the operator $\mathcal{H}$. In the next section, we review some properties of the parameter-free preconditioning scheme studied in [19]. In section 3, we consider a practical implementation, the key building block for which is any fast solver for a weighted $H(div)$ operator and diagonal scaling for a pressure mass matrix. A new eigenvalue bound is established and shown to be tight. In section 4, numerical experiments assess the impact of a range of different coefficients on that bound. We reiterate our earlier message that the existence of an optimal multilevel solver for the weighted $H(div)$ operator does not necessarily lead to a good block-diagonal preconditioner for (1.9) when arbitrary diffusion coefficients are prescribed.

2. Exact $\mathbf{H(div)}$ preconditioning.

The stability theory of Brezzi motivates the use of a particular type of preconditioning scheme. Conditions (1.7) and (1.8) essentially ensure (see [5], and Proposition 2.1 in [4]), that the linear operator $\mathcal{C}$ corresponding to the saddle-point matrix in (1.9) defines an isomorphism from $H(div;\Omega)\times L^2(\Omega)$ onto it's dual space, and furthermore that $\|\ \mathcal{C}^{-1}\ \|$ is bounded in the natural norm defined on $H(div;\Omega)\times L^2(\Omega)$. Hence, an operator $\mathcal{P}$ with the same mapping properties as $\mathcal{C}$, such that $\|\ \mathcal{P}^{-1}\ \|$ is bounded in the same norm, will yield a good preconditioner for $\mathcal{C}$.

Following [19], we consider preconditioning (1.9) with a block-diagonal matrix,

$$P=\begin{pmatrix} P_1 & 0\\ 0 & P_2\end{pmatrix}, \tag{2.1}$$

with symmetric and positive definite blocks chosen to represent norms in which the finite element method is stable. The most intuitive (but not the only) choice is,

$$\underline{v}^T P_1 \underline{v} = \|\ \vec{v}_h\ \|_{div}^2 \quad \forall\, \vec{v}_h\in V_h, \tag{2.2}$$

$$\underline{w}^T P_2 \underline{w} = \|\ w_h\ \|_0^2 \quad \forall\, w_h\in W_h, \tag{2.3}$$

where $\underline{v}$ and $\underline{w}$ are the vectors of coefficients corresponding to the expansion of $\vec{v_h}$ and w_h in the basis sets. Constructing the unweighted velocity mass matrix $A_I\in\mathbb{R}^{n\times n}$ and $D\in\mathbb{R}^{n\times n}$ via,

$$A_{I\,ij}=(\vec{\varphi}_i,\vec{\varphi}_j)\,, \qquad i,j=1:n, \tag{2.4}$$

$$D_{ij}=(\nabla\cdot\vec{\varphi}_i,\nabla\cdot\vec{\varphi}_j)\,, \qquad i,j=1:n, \tag{2.5}$$

leads to $P_1=A_I+D$ since $\|\ \vec{v}_h\ \|_{div}^2=\|\ \vec{v}_h\ \|_0^2+\|\ \nabla\cdot\vec{v}_h\ \|_0^2=\ \underline{v}^T\left(A_{\mathcal{I}}+D\right)\underline{v}$. Defining the pressure mass matrix N via,

$$N_{rs}=(\phi_r,\phi_s) \qquad r,s=1:m, \tag{2.6}$$

yields $P_2 = N$. It is important to note that although N represents an identity *operator* with respect to the L^2 inner-product, it is not an identity *matrix*. In fact, N is a diagonal matrix which supplies h-dependent scaling.

This preconditioning scheme was recommended by Arnold et al. in [1]. In [19] we established the relation $D = BN^{-1}B^T$ and the following eigenvalue bounds.

LEMMA 2.1. *The $(n+m)$ eigenvalues of the generalised eigenvalue problem,*

$$\begin{pmatrix} A_I & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \sigma \begin{pmatrix} A_I + D & 0 \\ 0 & N \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix}, \tag{2.7}$$

arising in the Raviart-Thomas approximation of (1.6) are bounded by constants independent of h and lie in the intervals $\left[-1, -\beta_h^2\right] \cup [1]$, where β_h is the discrete inf-sup constant in (1.8).

When $\mathcal{A} = \mathcal{I}$, choosing P_1 and P_2 to represent the norms for which stability holds, leads to an h-optimal eigenvalue bound. To efficiently precondition a problem with a general coefficient tensor, however, P must also supply scaling with respect to $\mathcal{A}$. Hence, we apply,

$$P = \begin{pmatrix} A + D & 0 \\ 0 & N \end{pmatrix}, \tag{2.8}$$

so that the leading diagonal block represents the weighted norm $\| \cdot \|_{div,\mathcal{A}}$.

LEMMA 2.2. *If T_h is quasi-uniform, the $(n+m)$ eigenvalues $\{\sigma_i\}_{i=1}^{n+m}$ of,*

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \sigma \begin{pmatrix} A + D & 0 \\ 0 & N \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix}, \tag{2.9}$$

arising in the Raviart-Thomas approximation of (1.6), lie in the union of the intervals,

$$\left(-1, -\left(\frac{c\mu_{min}}{|T|_{min} + \mu_{min}}\right)\right] \cup [1], \tag{2.10}$$

where μ_{min} is the minimum eigenvalue of the Schur complement matrix $BA^{-1}B^T$, $\mid T \mid_{min}$ is the volume of the smallest element in T_h and c is a constant independent of h and $\mathcal{A}$.

Proof. See Lemma 2.3 in [19] □

To summarise, for quasi-uniform meshes, (2.8) is h-optimal but may not be $\mathcal{A}$-optimal. Lemma 2.2 in [19] suggests that small coefficients, or large values of Γ in (1.2) can cause μ_{min}, the minimum eigenvalue of $BA^{-1}B^T$, to be small. Lemma 2.2 above suggests that MINRES convergence will not be efficient in such cases. However, the dependence of μ_{min} on the coefficient tensor is not straightforward, as demonstrated in [19]. For instance, anisotropic diagonal tensors with small entries on the diagonal can be accommodated if quadrilateral meshes are used. If there are portions of the domain where all the entries of the coefficient tensor are close to zero, then μ_{min} is small on any type of mesh and the preconditioning scheme is not efficient. For non quasi-uniform meshes, the bound established in Lemma 2.2 also holds. However, it is too pessimistic in general.

3. Practical $H(div)$ preconditioning. Let V be *any* symmetric and positive definite approximation to the matrix $H = A + D$ arising in the discretisation of the weighted $H(div)$ inner-product (1.12). To achieve optimal work complexity, we have in mind here that V is one V-cycle of a multigrid routine. Possible candidates include any of the schemes proposed by the authors of [1], [2], [14], [15] and [28], generalised to include variable, anisotropic and discontinuous diffusion coefficients. However, any available approximation can be substituted in what follows, provided that for the chosen V, there exist positive constants θ and Θ, satisfying,

$$\theta \le \frac{\underline{u}^T H \underline{u}}{\underline{u}^T V \underline{u}} \le \Theta \le 1 \quad \forall \underline{u} \in I\!R^n \backslash \{\underline{0}\}. \tag{3.1}$$

The condition $\Theta \le 1$ is not restrictive since the chosen V can always be rescaled. It is purely to simplify presentation in the sequel. Now consider,

$$P = \begin{pmatrix} V & 0 \\ 0 & N \end{pmatrix}. \tag{3.2}$$

3.1. Eigenvalue bounds. Below, we establish new bounds for the eigenvalues of the saddle-point matrix in (1.9) preconditioned by (3.2). In section 4 we perform numerical computations with a particular V and compute values of θ.

LEMMA 3.1. *The eigenvalues $\{\sigma_i\}_{i=1}^n$ of $D\underline{u} = \sigma H \underline{u}$, lie in the interval $[0, 1)$.*

Proof. Recall that $H = A + D = A + B^T N^{-1} B$. Since H is positive definite and D is semi-positive definite, $\sigma \ge 0$. If $\underline{u} \in null(B)$, $\sigma = 0$ by positive definiteness of A. The dimension of $null(B)$ is $n - m$, so there are $n - m$ zero eigenvalues. If $\underline{u} \notin null(B)$, $\sigma > 0$ and we have $(1 - \sigma)\, \underline{u}^T D \underline{u} = \sigma \underline{u}^T A \underline{u} > 0$. Since $\underline{u}^T D \underline{u} > 0$ in this case, we must have $1 - \sigma > 0$ and $\sigma < 1$. ☐

To simplify notation, let

$$a = \left(\frac{c\mu_{min}}{\mid T \mid_{min} + \mu_{min}} \right), \tag{3.3}$$

so that the bound (2.10) in Lemma 2.2 reads $[-1, -a] \cup [\,1\,]$.

THEOREM 1. *The $(n+m)$ eigenvalues $\{\lambda_i\}_{i=1}^{n+m}$ of the generalised eigenvalue problem,*

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \lambda \begin{pmatrix} V & 0 \\ 0 & N \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix}, \tag{3.4}$$

arising in the Raviart-Thomas approximation of (1.6), lie in the union of the intervals,

$$\left[-1\, , \frac{1}{2} \left(\theta \,(1 - a) - \sqrt{\theta^2 \,(a - 1)^2 + 4a\theta} \right) \right] \cup [\theta, 1], \tag{3.5}$$

where θ is the positive constant satisfying (3.1) and a is the positive constant defined in (3.3).

Proof. First, suppose that $\lambda > 0$. The eigenvalues $\{\lambda_i\}_{i=1}^{m+n}$ satisfy,

$$A\underline{u} + B^T \underline{p} = \lambda V \underline{u}, \qquad B\underline{u} = \lambda N \underline{p}.$$

Eliminating $\underline{p}$ yields,

$$\begin{aligned}\lambda A\underline{u} + B^T N^{-1} B\underline{u} &= \lambda^2 V\underline{u},\\ \lambda (A+D)\,\underline{u} + (1-\lambda)\, D\underline{u} &= \lambda^2 V\underline{u},\\ \lambda H\underline{u} + (1-\lambda)\, D\underline{u} &= \lambda^2 V\underline{u}.\end{aligned}$$

Thus,

$$(1-\lambda)\,\underline{u}^T D\underline{u} \;=\; \lambda^2 \underline{u}^T V\underline{u} - \lambda \underline{u}^T H\underline{u}, \tag{3.6}$$

and since, by assumption (3.1),

$$\underline{u}^T H\underline{u} \;\leq\; \underline{u}^T V\underline{u} \;\leq\; \frac{1}{\theta}\underline{u}^T H\underline{u}, \tag{3.7}$$

it follows that $(1-\lambda)\,\underline{u}^T D\underline{u} \geq \left(\lambda^2 - \lambda\right)\underline{u}^T H\underline{u}$. Applying Lemma 3.1, and noting that H is positive definite, we obtain,

$$(1-\lambda)\,\underline{u}^T H\underline{u} \geq \left(\lambda^2 - \lambda\right)\underline{u}^T H\underline{u},$$

and so $1 \geq \lambda$. From (3.6) and (3.7), it also follows that,

$$(1-\lambda)\,\underline{u}^T D\underline{u} \leq \left(\lambda^2\theta^{-1} - \lambda\right)\underline{u}^T H\underline{u}.$$

Since $0 \leq \underline{u}^T D\underline{u}$ and we have established $0 < \lambda \leq 1$, it follows that,

$$0 \leq \lambda\left(\lambda\theta^{-1} - 1\right)\underline{u}^T H\underline{u}.$$

Hence $\lambda \geq \theta$, and the bound for the positive eigenvalues is proved.

Now assume that $\lambda < 0$. Eliminating $\underline{u}$ yields,

$$B\,(A - \lambda V)^{-1}\, B^T \underline{p} \;=\; -\lambda N\underline{p}, \tag{3.8}$$

an implicit equation for λ. Note that since $\lambda < 0$, the matrix $(A - \lambda V)$ is positive definite. Now, the values of λ satisfying (3.8) are the eigenvalues of the matrix,

$$\begin{aligned}N^{-\frac{1}{2}} B\,(\lambda V - A)^{-1}\, B^T N^{-\frac{1}{2}} &= N^{-\frac{1}{2}} B\,(\lambda V - H + D)^{-1}\, B^T N^{-\frac{1}{2}}\\ &= N^{-\frac{1}{2}} B\left(\lambda V - H + B^T N^{-1} B\right)^{-1} B^T N^{-\frac{1}{2}}\\ &= N^{-\frac{1}{2}} B Y^{-\frac{1}{2}}\left(I + Y^{-\frac{1}{2}} B^T N^{-1} B Y^{-\frac{1}{2}}\right)^{-1} Y^{-\frac{1}{2}} B^T N^{-\frac{1}{2}}\\ &= X\left(I + X^T X\right)^{-1} X^T,\end{aligned}$$

where, here, $X = N^{-\frac{1}{2}} B Y^{-\frac{1}{2}}$ and $Y = \lambda V - H$. Applying the same arguments as in Lemma 3.1 in [29], and applying the Sherman-Morrison-Woodbury formula, the eigenvalues $\{\lambda_i\}$ we are seeking in (3.8) are the values:

$$\lambda_i \;=\; \frac{\sigma_i}{1+\sigma_i}, \tag{3.9}$$

where each σ_i is an eigenvalue of,

$$B\left(\lambda_i V - H\right)^{-1} B^T \underline{p} = \sigma N \underline{p}. \tag{3.10}$$

We now can obtain a bound for these values by exploiting the spectral equivalence of H and V defined by (3.1). Note that bounds for the eigenvalues of (3.8) cannot be obtained without the above manipulation, since we have no readily available information about the spectral equivalence of the leading diagonal blocks A and V.

Consider, first, the eigenvalues $\{\mu\}$ of,

$$\left(\lambda V - H\right)^{-1} \underline{u} = \mu \, H^{-1} \underline{u}. \tag{3.11}$$

Since $\lambda < 0$, the matrix $(\lambda V - H)$ is negative definite. Since H^{-1} is positive definite, the values of μ are negative. Rearranging gives,

$$\frac{\lambda}{\left(\frac{1}{\mu} + 1\right)} = \frac{\underline{u}^T H \underline{u}}{\underline{u}^T V \underline{u}}.$$

Combining this with (3.1) we obtain,

$$\theta \le \frac{\lambda \mu}{(\mu + 1)} \le 1.$$

Recalling that $\theta > 0$, $\mu < 0$ and $\lambda < 0$, we find that,

$$\mu \in \left[\frac{1}{\lambda - 1}, \frac{\theta}{\lambda - \theta}\right]. \tag{3.12}$$

Now, combining (3.11) and (3.12), we obtain $\forall \, \underline{p} \in I\!R^m \backslash \{\underline{0}\}$,

$$\frac{1}{\lambda - 1} \le \frac{\underline{p}^T B \left(\lambda V - H\right)^{-1} B^T \underline{p}}{\underline{p}^T B H^{-1} B^T \underline{p}} \le \frac{\theta}{\lambda - \theta},$$

and so,

$$\left(\frac{1}{\lambda - 1}\right) \frac{\underline{p}^T B H^{-1} B^T \underline{p}}{\underline{p}^T N \underline{p}} \le \frac{\underline{p}^T B \left(\lambda V - H\right)^{-1} B^T \underline{p}}{\underline{p}^T N \underline{p}} \le \left(\frac{\theta}{\lambda - \theta}\right) \frac{\underline{p}^T B H^{-1} B^T \underline{p}}{\underline{p}^T N \underline{p}}.$$

In the proof of Lemma 2.2 above, we obtained the bound,

$$a \le \frac{\underline{p}^T B H^{-1} B^T \underline{p}}{\underline{p}^T N \underline{p}} \le 1 \quad \forall \, \underline{p} \in I\!R^m \backslash \{\underline{0}\},$$

with $a > 0$ defined in (3.3). Since the bounds for the eigenvalues in (3.12) are negative, we obtain,

$$\left(\frac{1}{\lambda - 1}\right) \le \frac{\underline{p}^T B \left(\lambda V - H\right)^{-1} B^T \underline{p}}{\underline{p}^T N \underline{p}} \le \left(\frac{a\theta}{\lambda - \theta}\right).$$

For any $\lambda < 0$, the eigenvalues $\{\sigma\}$ of (3.10) therefore satisfy,

$$\frac{1}{\lambda - 1} \le \sigma \le \frac{a\theta}{\lambda - \theta} < 0,$$

and also,

$$\frac{1}{1+\frac{1}{\lambda-1}} \geq \frac{1}{1+\sigma} \geq \frac{1}{1+\frac{a\theta}{\lambda-\theta}} > 0.$$

Hence, we obtain,

$$\frac{\frac{1}{\lambda-1}}{1+\frac{1}{\lambda-1}} \leq \frac{\sigma}{1+\sigma} \leq \frac{\frac{a\theta}{\lambda-\theta}}{1+\frac{a\theta}{\lambda-\theta}},$$

and so, using (3.9), the eigenvalues λ of (3.8) satisfy,

$$\frac{1}{\lambda} \leq \lambda \leq \frac{a\theta}{\lambda+\theta\,(a-1)}.$$

Finally, solving for λ in,

$$1 \geq \lambda^2, \qquad \lambda^2+\lambda\theta\,(a-1)-a\theta \geq 0,$$

yields, since $\lambda < 0$,

$$-1 \leq \lambda \leq \frac{1}{2}\left(\theta\,(1-a)-\sqrt{\theta^2\,(a-1)^2+4a\theta}\right)$$

□

Remark 1. *Notice that when $\theta = 1$, we recover the eigenvalue bound (2.10) for the ideal preconditioner.*

Using Theorem 1, we can now clearly see that the $\mathcal{A}$-optimality of the preconditioner (3.2) is determined by that of the ideal preconditioner (2.8). If a in (3.3) decays to zero due to the small magnitude of μ_{min} which is caused by some types of diffusion coefficients then, asymptotically,

$$\lim_{a\to 0}\frac{1}{2}\left(\theta\,(1-a)-\sqrt{\theta^2\,(a-1)^2+4a\theta}\right) = \frac{1}{2}\left(\theta-\sqrt{\theta^2}\right) = 0.$$

Thus, even if the chosen multigrid approximation to H is both h-optimal and $\mathcal{A}$-optimal, the eigenvalue bound (3.5) deteriorates. On the other hand, if the coefficient term is sufficiently large so that $a \to 1$, we obtain, asymptotically,

$$\lim_{a\to 1}\frac{1}{2}\left(\theta\,(1-a)-\sqrt{\theta^2\,(a-1)^2+4a\theta}\right) = -\frac{1}{2}\sqrt{4\theta} = -\sqrt{\theta}.$$

Hence, for large coefficients and hence large μ_{min}, the eigenvalue bound (3.5) takes the form,

$$\left(-1\,,\,-\sqrt{\theta}\right] \cup [\theta\,,\,1],$$

and the efficiency of the preconditioning scheme is completely determined by the properties of the multigrid approximation.

4. Numerical results. To illustrate that the above theory is tight, we now report on MINRES convergence for a range of diffusion problems in two dimensions, using the preconditioner (3.2) with V chosen to be one V-cycle of the Arnold-Falk-Winther multigrid scheme (see [1], [2]) with additive Schwarz smoothing. This particular multigrid method is restricted to quasi-uniform meshes and to triangles. In addition note that θ and Θ have only been proved to be independent of the coefficient if $\mathcal{A} = \rho\mathcal{I}$ where $\rho \in \mathbb{R}_+$. For simplicity, all of the experiments below are assembled on square domains with uniform meshes of triangles. Examples with more general triangulations can be treated in exactly the same way choosing V to be a V-cycle of any other available optimal scheme. The iteration is terminated after k steps if the residual error $\underline{r}^{(k)} = \underline{b} - C\underline{x}^{(k)}$ satisfies,

$$\frac{\| \underline{r}^{(k)} \|_{P^{-1}}}{\| \underline{r}^{(0)} \|_{P^{-1}}} \leq 10^{-6}, \tag{4.1}$$

where $\| \underline{r} \|_{P^{-1}} = \underline{r}^T P^{-1} \underline{r}$. All experiments are performed in `Matlab 6.0` with a zero initial guess $\underline{x}^{(0)}$. Iteration counts reported in parentheses refer to unpreconditioned experiments. The symbol $(*)$ indicates that more than 500 iterations are required to satisfy (4.1).

Example 1 Let us begin with the most trivial example, choosing the coefficient $\mathcal{A} = \mathcal{I}$. Consider the domain $\Omega = [0,1] \times [0,1]$ together with the boundary condition $p = 0$ on $\partial\Omega$, and $f = 2(x^2 - x + y^2 - y)$, so that we obtain the analytical solution $p = x(x-1)y(y-1)$ for the pressure. The minimum and maximum eigenvalues of the corresponding matrix $V^{-1}H$, using the method of Arnold et al., are listed in Table 4.1. They confirm numerically that the multigrid approximation is h-optimal. In Table 4.2 we compare the observed eigenvalues of the preconditioned saddle-point system with the bounds in (3.5). MINRES iteration counts are listed in Table 4.3.

h	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$
θ	0.5938	0.4595	0.4273
Θ	1	1	1

TABLE 4.1
Eigenvalues of $V^{-1}H$, unit coefficients

h	*bounds*	*observed*
$\frac{1}{4}$	$[-0.9983, -0.7381] \cup [0.5938, 1]$	$[-0.9879, -0.8507] \cup [0.5943, 1]$
$\frac{1}{8}$	$[-0.9996, -0.6504] \cup [0.4595, 1]$	$[-0.9972, -0.8438] \cup [0.4598, 1]$
$\frac{1}{16}$	$[-0.9999, -0.6503] \cup [0.4273, 1]$	$[-0.9994, -0.8481] \cup [0.4273, 1]$

TABLE 4.2
Theoretical bounds and observed eigenvalues, unit coefficients

h	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$
	17	18	18	18	18
	(81)	(183)	(367)	(*)	(*)

TABLE 4.3
MINRES *iterations, Example 1*

Example 2 Next, introduce a jump in the coefficient and set $\mathcal{A} = \alpha\mathcal{I}$ in one quadrant of Ω so that $\mu_{min} \to 0$ in (3.3) as $\alpha \to 0$ (see [19].) Values of θ are listed in Table 4.4. Now we see that the chosen multigrid approximation to $\mathcal{H}_{\mathcal{A}}$ is both $\mathcal{A}$-optimal and h-optimal; $\Theta = 1$ in all cases. The negative eigenvalues of the preconditioned saddle-point system, for $\alpha = 10^{-3}$ and $\alpha = 10^{-6}$ are listed in Tables 4.5–4.6.

h α	10^6	10^3	10^0	10^{-3}	10^{-6}
$\frac{1}{8}$	0.4725	0.4725	0.4595	0.3513	0.3491
$\frac{1}{16}$	0.4326	0.4326	0.4273	0.3319	0.3299

TABLE 4.4
Values of θ, discontinuous $\mathcal{A}$, single jump

h	*bounds*	*observed*
$\frac{1}{4}$	$[-0.9982, -0.0695]$	$[-0.9864, -0.0766]$
$\frac{1}{8}$	$[-0.9996, -0.0665]$	$[-0.9969, -0.0741]$
$\frac{1}{16}$	$[-0.9999, -0.0653]$	$[-0.9993, -0.0733]$

TABLE 4.5
Theoretical bounds and observed negative eigenvalues, $\alpha = 10^{-3}$

h	*bounds*	*observed*
$\frac{1}{4}$	$[-0.9982, -0.8327e-4]$	$[-0.9864, -0.8328e-4]$
$\frac{1}{8}$	$[-0.9996, -0.8021e-4]$	$[-0.9972, -0.8022e-4]$
$\frac{1}{16}$	$[-0.9999, -0.7927e-4]$	$[-0.9994, -0.7928e-4]$

TABLE 4.6
Theoretical bounds and observed negative eigenvalues, $\alpha = 10^{-6}$

We observe that the right-hand bound for the negative eigenvalues is tighter as the jump parameter decreases. The eigenvalues of the preconditioned saddle-point system, for a fixed mesh and varying α are plotted in Fig. 4.1. The scale on the y-axis corresponds to values of $\alpha \in \left[10^{-6}, 10^{6}\right]$. Each line of the plot depicts the eigenvalues for a different value of α. It is clear that for small values $\alpha < 1$, MINRES convergence will deteriorate. Iteration counts obtained with the ideal preconditioner are listed in columns 2–4 of Table 4.7. Counts for the multigrid preconditioner are given in columns 5–7. As Theorem 1 predicts, the multigrid preconditioner exhibits exactly the same asymptotic behaviour as the ideal version as $\alpha \to 0$. It is important to realise here that the deterioration has nothing to do with the chosen multigrid routine. It performs optimally, with respect to all parameters, as demonstrated in Table 4.4. In this simple case, the deterioration can be corrected by a simple rescaling of the coefficients, as discussed in [19], so that the minimum value of any coefficient is unity. However, in complicated Darcy-flow models, this may not be possible.

$\alpha \quad h$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
10^6	5	5	5	17	18	18
10^3	5	5	5	17	18	18
10^0	5	5	5	17	18	18
10^{-1}	6	6	6	18	20	20
10^{-2}	9	9	9	22	24	25
10^{-3}	17	19	19	41	45	47
10^{-4}	28	42	48	71	119	130
10^{-5}	35	70	107	106	220	342
10^{-6}	42	92	162	131	328	574

TABLE 4.7
MINRES *iterations,ideal preconditioner (left) and with multigrid (right)*

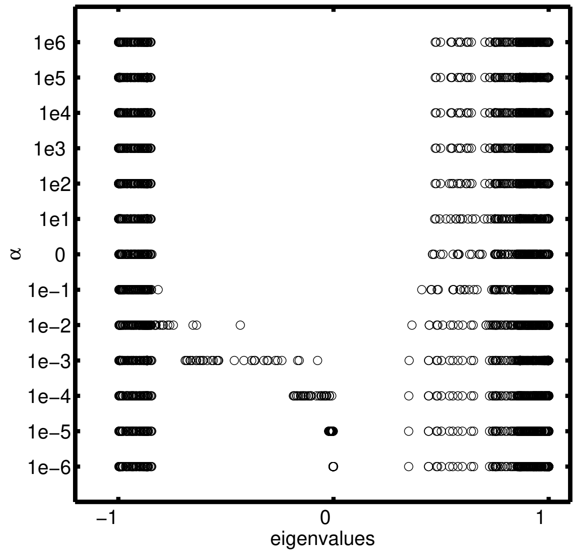

FIG. 4.1. *Eigenvalues of preconditioned saddle-point system,* $h = \frac{1}{16}, \alpha \in [10^{-6}, 10^{6}]$

Example 3 For a more challenging discontinuous coefficient example, we can consider so-called 'Kellogg problems' (see Kellogg [16]). Here, the coefficient is prescribed in a 2×2 checker-board fashion on $\Omega = [-1,1] \times [-1,1]$ such that $\mathcal{A} = a_1 \mathcal{I}$ in the first and third quadrants of Ω and $\mathcal{A} = a_2\mathcal{I}$ in the second and fourth quadrants. With $f = 0$, the analytical 'pressure' solution, in polar co-ordinates, is $p(r, \theta) = r^\gamma \mu(\theta)$, where,

$$\mu(\theta) = \begin{cases} \cos((\frac{\pi}{2} - \sigma)\gamma) \cdot \cos((\theta - \frac{\pi}{2} + \rho)\gamma) & 0 \le \theta \le \frac{\pi}{2}, \\ \cos(\rho\gamma) \cdot \cos((\theta - \pi + \sigma)\gamma) & \frac{\pi}{2} \le \theta \le \pi, \\ \cos(\sigma\gamma) \cdot \cos((\theta - \pi - \rho)\gamma) & \pi \le \theta \le \frac{3\pi}{2}, \\ \cos((\frac{\pi}{2} - \rho)\gamma) \cdot \cos((\theta - \frac{3\pi}{2} - \sigma)\gamma) & \frac{3\pi}{2} \le \theta \le 2\pi, \end{cases}$$

and the constants σ, ρ and γ are chosen to satisfy the non-linear relations,

$$\begin{aligned} R = \frac{a_1}{a_2} &= -\tan((\frac{\pi}{2} - \sigma)\gamma) \cdot \cot(\rho\sigma), \\ R^{-1} &= -\tan(\rho\sigma) \cdot \cot(\sigma\gamma), \\ R &= -\tan(\sigma\gamma) \cdot \cot((\frac{\pi}{2} - \rho)\gamma), \\ 0 < \gamma &< 2, \\ \max\{0, \pi\gamma - \pi\} &< 2\gamma\rho < \min\{\pi\gamma, \pi\} \\ \max\{0, \pi - \pi\gamma\} &< -2\gamma\sigma < \min\{\pi, 2\pi - \pi\gamma\}. \end{aligned}$$

The test problem we consider is one suggested by Morin et al. in [17]. Setting $\gamma = 0.1$. and solving for the various constants yields,

$$R = \frac{a_1}{a_2} \approx 161.4476387, \quad \rho = \frac{\pi}{4}, \quad \sigma \approx -14.922565105.$$

We then set $a_1 = R$ and $a_2 = 1$. The interpolant of the numerical pressure solution on a uniform mesh is shown in Fig. 4.2.

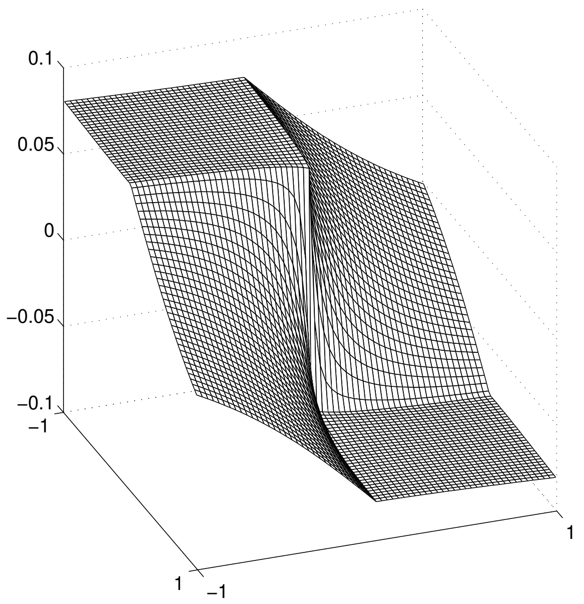

FIG. 4.2. *Numerical solution to Kellogg problem, $h = \frac{1}{32}$*

Iteration counts are listed in Table 4.8. For $h = \frac{1}{8}$, we obtain $\theta \approx 0.1538$ and $\Theta = 1$, yielding the eigenvalue bound, $[-0.9999, -0.7568] \cup [0.1538, 1]$. The observed values lie in the interval $[-0.9998, -0.8431] \cup [0.1541, 1]$.

h	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$
	23	25	26	26	26
	(*)	(*)	(*)	(*)	(*)

TABLE 4.8

MINRES *iterations, exact preconditioner*

Once again, we observe numerically that the multigrid approximation to the weighted $H(div)$ operator is h-optimal.

Example 4 Next, we consider a couple of variable coefficient tensors. First, choose the diagonal coefficient tensor,

$$\mathcal{A} = \begin{pmatrix} \frac{1}{1+1000(x^2+y^2)} & 0 \\ 0 & \frac{1}{1+1000(x^2+y^2)} \end{pmatrix}, \tag{4.2}$$

with $\Omega = [0,1] \times [0,1]$. In this case, we obtain $\kappa(V^{-1}H) \approx 1.95$. Iteration counts are listed in Table 4.9. Once again, the multigrid approximation is h-optimal and so therefore is the preconditioner (3.2). However, the iteration count rises significantly since the minimum eigenvalue μ_{min} of $BA^{-1}B^T$ is smaller here than in the other examples.

h	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$
	45	59	67	69	69
	(*)	(*)	(*)	(*)	(*)

TABLE 4.9

MINRES *iterations, Example 4*

Example 5 Next, choose the full tensor,

$$\mathcal{A} = \begin{pmatrix} 1 + 4\left(x^2+y^2\right) & 3xy \\ 3xy & 1 + 11\left(x^2+y^2\right) \end{pmatrix}. \tag{4.3}$$

In this case, we obtain $\kappa(V^{-1}H) = \theta^{-1} \approx 3.84$. Iteration counts for the full saddle-point problem are listed in Table 4.10. The eigenvalues of $\mathcal{A}$ in this example are fairly benign. The eigenvalue μ_{min} in (3.3) is bounded away from zero and MINRES iteration is extremely efficient. Note that it is not the structure of $\mathcal{A}$ that influences the iteration count.

Example 6 Finally, consider an anisotropic test problem with $\mathcal{A} = \mathtt{diag}(\alpha, 1)$. Condition numbers for $V^{-1}H$ are listed in Table 4.11; they show that the suggested multigrid scheme is totally unsuitable as an approximation to the $H(div)$ operator with anisotropic weighting. $\Theta = 1$ in all cases but θ deteriorates with α. MINRES convergence is extremely poor, even for mild anisotropies.

h	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$
	21	22	22	22	22
	(322)	(*)	(*)	(*)	(*)

TABLE 4.10
MINRES *iterations, Example 5*

h α	10^3	10^2	10^1	10^0	10^{-1}	10^{-2}	10^{-3}
$\frac{1}{8}$	8.99e2	9.11e1	1.27e1	2.18e0	0.97e1	5.95e1	1.98e2
$\frac{1}{16}$	9.61e2	9.74e1	1.09e1	2.34e0	1.75e1	8.54e1	4.30e2

TABLE 4.11
Condition number of $V^{-1}H$, Example 6

5. Concluding remarks. In this paper, we considered a block-diagonal parameter free preconditioning scheme for the linear systems (1.9) arising in the lowest-order Raviart-Thomas discretisation of generalised diffusion problems. Motivated by standard stability theory, and discussion in [1] and [19], the diagonal blocks are a weighted $H(div)$ matrix and the pressure mass matrix. New bounds are established for the eigenvalue spectrum of the preconditioned saddle-point matrix when the leading block of the exact preconditioner is replaced with a suitable approximation. In numerical experiments, we demonstrated the impact of general coefficient tensors on the performance of a particular multilevel approximation and on the theoretical eigenvalue bound.

We identified two distinct reasons for poor MINRES convergence. If the exact version of the preconditioner is not $\mathcal{A}$-optimal, which is likely in flow modelling, then neither is the suggested practical preconditioner (3.2). Iteration may only be efficient if scaling with respect to the minimum coefficient, is applied. This is true independently of the approximation properties of the chosen multilevel scheme for the $H(div)$ operator. In addition, the particular multigrid approximation used in our numerical tests failed in anisotropic test cases. As far as we are aware, the existence of optimal multigrid solvers for $H(div)$ operators with anisotropic and more general weighting is an open question. Our suspicion is that coefficient dependent transfer operators should be employed, rather than pure injection.

The key message for practitioners is that given an arbitrary coefficient tensor, the existence of an optimal multigrid solver for the weighted $H(div)$ operator does not guarantee the existence of a good block-diagonal preconditioner for the mixed system (1.9). A second block-diagonal preconditioning scheme, which can be motivated by an alternative stability theory discussed in [26], is presented in [20]. It consists of diagonal scaling for the weighted velocity mass matrix and a V-cycle of an algebraic multigrid routine applied to a sparse generalised diffusion operator. For certain types of diffusion coefficients, it may be more efficient.

Acknowledgements The author would like to thank Doug Arnold for providing an initial `matlab` implementation of his multigrid code and Manel Cascon and Ricardo Nochetto for providing the motivation to study the Kellogg problem in Example 3.

REFERENCES

[1] D.N. Arnold, R.S. Falk and R. Winther, Multigrid in $H(div)$ and $H(curl)$, *Numer. Math.,* 85(2000), pp.197–218.

[2] D.N. Arnold, R.S. Falk and R. Winther, Preconditioning in $H(div)$ and applications, *Math. Comp.,* 66, 219(1997), pp.957–984.

[3] D. Braess, Finite Elements: Theory, Fast Solvers and Applications in Solid Mechanics, Cambridge University Press, 1992.

[4] F.Brezzi, On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers, *RAIRO Anal. Numér.,* 8(1974), pp.129–151.

[5] F. Brezzi and K.J. Bathe, A discourse on the stability conditions for mixed finite element formulations, *Comp. Methods. Appl. Mech. Engrg.,* 82(1990), pp.27–57, North-Holland.

[6] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[7] Z. Cai, C.I. Goldstein and J.E. Pasciak, Multilevel iteration for mixed finite element systems with penalty, *SIAM J. Sci. Comput.,* 14(5), pp.1072–1088, 1993.

[8] K.A. Cliffe, I.G. Graham, R. Scheichl, L. Stals, Parallel computation of flow in heterogeneous media modelled by mixed finite elements, *J. Comput. Physics,* 164, 2(2000).

[9] R.E. Ewing and J. Wang, Analysis of the Schwarz algorithm for mixed finite elements methods, *M^2AN, Math. Model. Numer. Anal.,* 26, 6(1992) pp.739–756.

[10] R.E. Ewing and J. Wang, Analysis of multilevel decomposition iterative methods for mixed finite element methods, *M^2AN, Math. Model. Numer. Anal.,* 28, 4(1994) pp.377–398.

[11] R.E. Ewing and M.F. Wheeler, Computational aspects of mixed finite element methods. In: Scientific Computing, R. Stepleman (Ed), *IMACS*, North-Holland, pp.163–172, 1983.

[12] M. Fortin and R. Glowinski, Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems, *Stud. Math. Appl.,* 15, North-Holland, Amsterdam, 1983.

[13] A. Greenbaum, *Iterative Methods for Solving Linear Systems,* SIAM, Philadelphia, 1997.

[14] R. Hiptmair, Multigrid methods for $H(div)$ in three dimensions, *Electron. Trans. Numer. Anal.,* 6(1997) pp.133–152.

[15] R. Hiptmair, B. Wohlmuth, T. Scheikofer, Multilevel preconditioned augmented Lagrangian techniques for 2nd order mixed problems, *Computing,* 57, pp.25–48, 1996.

[16] R.B. Kellogg, On the Poisson equation with intersecting interfaces, *Appl., Anal.,* 4, pp.101–129, 1975.

[17] P. Morin, R.H. Nochetto and K.G. Siebert, Convergence of adaptive finite element methods, *SIAM Rev.,* pp.1–28, 2002.

[18] C.C. Paige and M.A. Saunders, Solution of sparse indefinite systems of linear equations, *SIAM J. Numer. Anal.,* 12(1975), pp.617–629.

[19] C.E. Powell and D. Silvester, Optimal preconditioning for Raviart-Thomas mixed formulation of second-order elliptic problems, *SIAM J. Matrix Anal. Appl.,* 25, 3(2003) pp.718–738.

[20] C.E. Powell and D. Silvester, Black-box preconditioning for mixed formulation of self-adjoint elliptic pdes, *Lect. Notes Comput. Sci. Eng.,* 35(2002), pp.268–284.

[21] A. Quarteroni and A. Valli, Numerical Approximation of Partial Differential Equations, Springer-Verlag, 1994.

[22] P.A. Raviart and J.M. Thomas, A mixed finite element method for second order elliptic problems, *Mathematical Aspects of the Finite Element Method, Lect. Notes in Math.*

606, Springer-Verlag, New York, 1977.

[23] J.E. Roberts and J.M. Thomas, Mixed and hybrid methods, in *Handbook of Numerical Analysis, Vol II: Finite Element Methods (Part 1),* P.G. Ciarlet and J.L. Lions, eds, North-Holland, 1991, pp. 523–633.

[24] T.F. Russell and M.F. Wheeler, Finite element and finite difference methods for continuous flows in porous media. In: Mathematics of Reservoir Simulation, R.E. Ewing (Ed.), pp.35–106, 1983.

[25] T. Rusten and R. Winther, A preconditioned iterative method for saddlepoint problems, *SIAM J. Matrix. Anal. Appl.,* 13, 3(1992), pp.887–904.

[26] T. Rusten, P.S. Vassilevski and R. Winther, Interior preconditioners for mixed finite element approximations of elliptic problems, *Math. Comp.,* 65, 214(1996) pp.447–466.

[27] R. Scheichl, Iterative Solution of Saddle-Point Problems Using Divergence-free Finite Elements with Applications to Groundwater Flow, *PhD. Thesis,* University of Bath, 2000.

[28] P.S. Vassilevski and J. Wang, Multilevel iterative methods for mixed finite element discretizations of elliptic problems, *Numer. Math.*, 63(1992) pp.503–520.

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