An Element-Based Preconditioner for Mixed Finite Element Problems

Tyrone Rees and Michael Wathen

An element-based preconditioner for mixed finite element problems [*]

Tyrone Rees[†] and Michael Wathen[†]

Abstract. We introduce a new and generic approximation to Schur complements arising from inf-sup stable mixed finite element discretizations of self-adjoint multi-physics problems. The approximation exploits the discretization mesh by forming local, or element, Schur complements of an appropriate system, and projecting them back to the global degrees of freedom. The resulting Schur complement approximation is sparse, has low construction cost (with the same order of operations as assembling a general finite element matrix), and can be solved using off-the-shelf techniques, such as multigrid. Using results from saddle point theory, we give conditions such that this approximation is spectrally equivalent to the global Schur complement. We present several numerical results to demonstrate the viability of this approach on a range of applications. Interestingly, numerical results show that the method gives an effective approximation to the non-symmetric Schur complement from the steady state Navier-Stokes equations.

1. Introduction. We seek $(u, p) \in \mathcal{X} \times \mathcal{M}$ that solves the saddle point problem

$$(1.1) \quad \begin{aligned} a(u, v) + b(v, p) &= \langle f, v \rangle \quad \forall v \in \mathcal{X}, \\ b(u, q) &= \langle g, q \rangle \quad \forall q \in \mathcal{M}, \end{aligned}$$

where $\mathcal{X}$ and $\mathcal{M}$ are two Hilbert spaces, $a(\cdot, \cdot) : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ and $b(\cdot, \cdot) : \mathcal{X} \times \mathcal{M} \to \mathbb{R}$ are two bounded bilinear forms, and $(f, g) \in \mathcal{X}' \times \mathcal{M}'$, where $\mathcal{X}'$ and $\mathcal{M}'$ are the dual spaces of $\mathcal{X}$ and $\mathcal{M}$, respectively. This is a well-posed problem provided that it satisfies certain conditions, given in Section 4.1. Such problems arise in fields as diverse as constrained optimization, constrained least-squares, fluid dynamics, electomagnetics, elasticity, optimal control, and many others (see, for example, the references within the survey of Benzi, Golub and Liesen [7]).

We focus on the case where (1.1) comes from the solution of a partial differential equation (PDE). We discretize using the Finite Element Method, and in particular, by choosing finite dimensional spaces $\mathcal{X}_h \subset \mathcal{X}$ and $\mathcal{M}_h \subset \mathcal{M}$. This entails overlaying the domain with a grid and using this to define local elements, giving an elemental structure. We select basis functions that have support only on a small number of neighbouring elements, which are usually chosen to be unity at one mesh point and vanish at the others. We describe the requirements for such a discretization to be well-defined in Section 4.1.

Suppose we have such a pair of spaces defined by the basis functions $\phi_i \in \mathcal{X}_h$, $\psi_i \in \mathcal{M}_h$; we assume they share the same elements, but the basis functions may differ. We look for an approximation $(u_h, p_h) = \left(\sum_i \mathbf{u}_i \phi_i, \sum_j \mathbf{p}_j \psi_j \right)$, where we find the coefficents of each basis

---

[†]STFC Rutherford Appleton Laboratory, Chilton, Didcot, UK (tyrone.rees@stfc.ac.uk, michael.wathen@stfc.ac.uk)

element by solving a linear system of the form

$$(1.2) \qquad \underbrace{\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}}_{\mathcal{A}} \underbrace{\begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix}}_{\mathbf{b}},$$

where $A_{i,j} = a(\phi_i, \phi_j) \in \mathbb{R}^{n \times n}$ and $B_{i,j} = b(\phi_j, \psi_i) \in \mathbb{R}^{m \times n}$ (see, e.g., [11, 17] for more detail).

We can decompose the matrices in (1.2) into a sum of σ elemental matrices. On the eth element, let n_e and m_e be the number of local degrees of freedom in $\mathcal{X}_h$ and $\mathcal{M}_h$, respectively (see Figure 1 for an example). Let the small dense matrices $A_e \in \mathbb{R}^{n_e \times n_e}$ and $B_e \in \mathbb{R}^{m_e \times n_e}$ be the element equivalents of A and B.
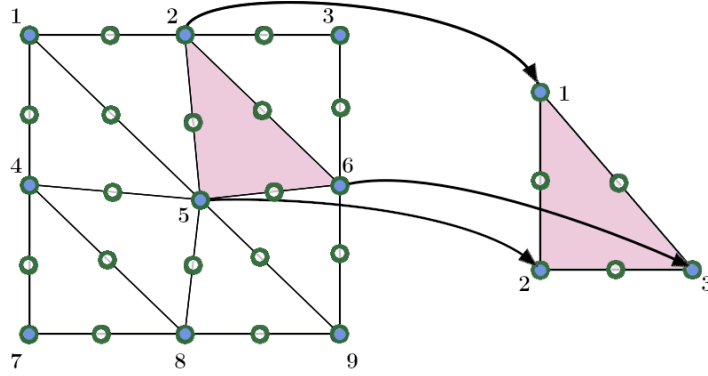

Figure 1: Global (left) and local (right) meshes for a P2-P1 triangulation of a domain. Here $n = 25$, $m = 9$, $n_e = 6$, $m_e = 3$, and $\sigma = 8$.

We introduce Boolean matrices $L_e \in \mathbb{R}^{n_e \times n}$ and $N_e \in \mathbb{R}^{m_e \times m}$ that map the local orderings to the global orderings. Thus, we write

$$(1.3) \qquad A = \sum_e L_e^T A_e L_e := L^T \widehat{A_e} L, \quad B = \sum_e N_e^T B_e L_e := N^T \widehat{B_e} L.$$

where $L^T = [L_1^T, \ldots L_\sigma^T]$, $N^T = [N_1^T, \ldots N_\sigma^T]$, $\widehat{A_e} = \texttt{blkdiag}(A_e)$, and $\widehat{B_e} = \texttt{blkdiag}(B_e)$.

As we will outline in Section 3.1, preconditioners based on Schur complement approximations work well for systems of the form (1.2). In this paper we describe a sparse approximation to the Schur complement based on local contributions by element matrices of another saddle point problem:

$$(1.4) \qquad \hat{S}_{\text{dual}} = \sum_e N_e^T C_e P_e^{-1} C_e^T N_e \quad \text{and} \quad \hat{S}_{\text{primal}} = \sum_e L_e^T (H_e + C_e^T D_e^{-1} C_e) L_e,$$

One of the main aims of the paper is to identify appropriate choices of matrices C_e, P_e, H_e and D_e to give effective preconditioners. In (1.4), we form the Schur complement on an element and then map the local matrix into a global matrix by using the standard assembly

process. The idea of using elements to define preconditioners is not new, and we outline their development in the literature in Section 2.

In Section 3, we review preconditioning for saddle point systems, in particular looking at (ideal) Schur complement preconditioners and natural norm, or Riesz-map, preconditioners. We further explore the link between these methods in Section 4, where we extend existing results in the literature. In Section 5, we prove our main result that, under certain conditions, we can derive matrices $\hat{S}_{\text{dual}}$ and $\hat{S}_{\text{primal}}$ that are spectrally equivalent to the Schur complements of the original system (1.2) or, correspondingly, are spectrally equivalent to a natural norm preconditioner.

One potential use of the class of preconditioners we describe here is to have a *black-box* method to derive optimal preconditioners without requiring any knowledge of the analytic formulation of the underlying system. We apply such preconditioners to three well-known examples of saddle-point systems, Stokes equation (Section 6.1), Maxwell's equation (Section 6.2) and the Navier-Stokes equation (Section 6.3). Finally, in Section 7 we draw conclusions and point out potential avenues for future work.

2. Related work. Solving a local system on an element or, more generally, on a patch of elements has a long history in the Computational Fluid dynamics community, either as a solution method in its own right, or (especially effectively) as a smoother for multigrid. For Navier-Stokes, Local Multilevel Pressure Schur Complement approaches solve for the local Schur complement, and then use a (block) simple iteration (such as Jacobi or Gauss-Seidel) to give a solution to the global problem [59, Section 2.4]. Replacing the (1,1) block of the local system by a diagonal matrix, and using this to form an approximation to the local Schur complement gives the well-known method proposed by Vanka [61]. Similar element-wise ideas have also been applied by Schörberl [52] when developing multilevel projection methods for nearly incompressible materials, which have been further utilized by Benzi and Olshanskii [8] to develop robust smoothing operations for multigrid. More recently, PCPATCH [19, 20] derives effective relaxation methods based on specific "patches", or local contributions from elements.

The solution of a linear system by solving local sub-problems is, of course, a cornerstone of domain decomposition methods. In particular, we highlight that local Schur complements have been used in the development of balancing Neumann-Neumann methods for saddle point systems [14]; for an overview of related methods see, e.g. Toselli and Widlund [58, Chapter 9] and the references therein. More recently, there has been interest in multilevel preconditioners based on local subdomains or elements. For example, in the context of overlapping domain decomposition, GENEO [54, 55] is designed specifically for heterogeneities within the variational form by incorporating local generalized eigenvalue problems on overlaps to define the coarsening.

More closely aligned to the methods we propose are the class of element-based methods inspired by the work of Hughes, Levit and Winget [28] and Ortiz, Pinksy and Taylor [45], who introduced element-by-element (EBE) methods for solving self-adjoint PDEs. Such methods approximate the sum in (1.3) by a product. Nour-Omid and Parlett [44] show that we may apply the EBE method as a preconditioner for conjugate gradients [26], and Wathen [62] and van Gijzen [60] give an analysis of such methods for symmetric and non-symmetric problems,

respectively. Gustafsson and Lindskog [25] also exploit local contributions by developing a preconditioner based on global structure preserving Cholesky factorizations on elements.

For symmetric positive definite systems, Kraus [31, 32] described the use of local (dual) Schur complements to build a preconditioner which they call the additive Schur complement approximation (ASCA), Kraus and co-authors [33] extend the work on ASCA to introduce the of idea of auxiliary space multigrid (ASMG) methods which they utilize the coarse-grid operator from ASCA. Axelsson, Baheta and Neytcheva [4] give descriptive eigenvalue bounds for $\hat{S}_{\text{dual}}^{-1} S_{\text{dual}}$, where $S_{\text{dual}} = BA^{-1}B^T$ and $\hat{S}_{\text{dual}}^{-1}$ is defined in (1.4) with $P_e = A_e$ and $C_e = B_e$, which rely only on the Cauchy-Bunyakowski-Schwarz constant. Neytcheva and co-authors [41, 42, 43] apply this idea to non-positive definite systems, and Neytcheva [41] gives a bound on $\|\hat{S}_{\text{dual}}^{-1} S_{\text{dual}}\|$ based on the norms of relations involving the constituent blocks. In that work, Neytcheva also shows numerically that this is a viable method for solving Stokes equation and the Oseen equation.

3. Iterative solution of saddle point systems. The convergence of a Krylov subspace method applied to (1.2) is generally unsatisfactory unless paired with a suitable preconditioner, $\mathcal{P}$, which has two, competing, properties: a solve with $\mathcal{P}$ must be cheap, and the Krylov method applied to $\mathcal{P}^{-1}\mathcal{A}$ must converge more quickly. We might satisfy the latter condition if, for example, the eigenvalues of $\mathcal{P}^{-1}\mathcal{A}$ are in clusters away from the origin. We recommend the surveys by Wathen [64] and Benzi, Golub, and Liesen [7] for a comprehensive overview of the state-of-the-art in preconditioning techniques in general (in the former), and for saddle point matrices in particular (in the latter).

In the case where the leading matrix A in (1.2) is symmetric, the saddle point system $\mathcal{A}$ is symmetric but indefinite, and MINRES [46] is the method of choice. MINRES requires a symmetric positive definite preconditioner, and convergence is exactly described by the eigenvalues of the preconditioned system $\mathcal{P}^{-1}\mathcal{A}$ (see, e.g., [17, Section 4.1]).

3.1. Schur complement preconditioners. An idea that has proved successful for solving systems of the form (1.2) is to build a preconditioner that exploits an approximation to the Schur complement. The ideal preconditioners are

$$\mathcal{P}_{\text{dual}} = \begin{bmatrix} A & 0 \\ 0 & BA^{-1}B^T \end{bmatrix}, \mathcal{P}_{\text{primal}} = \begin{bmatrix} A + B^T W^{-1} B & 0 \\ 0 & W \end{bmatrix}.$$

These involve the negative of the *dual Schur complement*, $BA^{-1}B^T$, and the *primal Schur complement*, $A + B^T W^{-1} B$. Often the choice between using $\mathcal{P}_{\text{dual}}$ and $\mathcal{P}_{\text{primal}}$ depends on the invertibility of the leading block, A. Murphy, Golub and Wathen [39] and Ipsen [29] show that the preconditioned matrix, $\mathcal{P}_{\text{dual}}^{-1}\mathcal{A}$, has three distinct eigenvalues. Thus, MINRES would converge in precisely three iterations. For the primal Schur complement, Greif and Schötzau [22] give bounds for the eigenvalues of $\mathcal{P}_{\text{primal}}^{-1}\mathcal{A}$. For the specific case where the dimension of the null space of leading block A is m (the dimension of the dual variables), Greif and Schötzau [23] show that the preconditioned matrix has two distinct eigenvalues.

Although in one sense these preconditioners are ideal, as convergence is rapid, forming and solving with these is not practical as the primal and dual Schur complements are typically dense. However, they give us something to aim for in a practical preconditioner.

3.2. Operator preconditioning for saddle point problems. Another framework for solving systems of the form (1.2) is operator preconditioning. Given a Hilbert space $\mathcal{V}$, the Riesz map is a mapping $\tau : \mathcal{V}' \to \mathcal{V}$ such that, for any $r \in \mathcal{V}'$,

$$(\tau r, v)_{\mathcal{V}} := \langle r, v \rangle \quad \forall v \in \mathcal{V}.$$

The continuous saddle point system (1.1) is an operator $\mathcal{L} : \mathcal{X} \times \mathcal{M} \to \mathcal{X}' \times \mathcal{M}'$, which returns functions outside of the space $(\mathcal{X}, \mathcal{M})$. Consider the Riesz map

$$(3.1) \qquad \tau = \begin{bmatrix} \chi & 0 \\ 0 & \mu \end{bmatrix},$$

where $\chi : \mathcal{X}' \to \mathcal{X}$ and $\mu : \mathcal{M}' \to \mathcal{M}$ are the Riesz maps associated with the spaces $\mathcal{X}$ and $\mathcal{M}$, respectively. Applying τ to (1.1) gives the transformed equation

$$\tau \mathcal{L}(u, p) = \tau(f, g),$$

where $\tau(f, g) \in \mathcal{X} \times \mathcal{M}$ and $\tau \mathcal{K} : \mathcal{X} \times \mathcal{M} \to \mathcal{X} \times \mathcal{M}$. We now have a reformuation of (1.1) posed entirely in the space $\mathcal{X} \times \mathcal{M}$, and the map τ is therefore analogous to applying a preconditioner to the operator $\mathcal{L}$. Furthermore, we can build a practical preconditioner by choosing an inner product which is sufficiently close to the bilinear forms associated with these spaces, yet is numerically tractable. We refer the reader to the monograph by Malek and Strakos [40], for example, for more detail.

In some sense, therefore, block diagonal preconditioners are natural for saddle point problems of the form (1.1). We refer the reader to the article by Mardal and Winther [37], which describes in detail such preconditioners. However, their derivation requires detailed knowledge of the problem, and may be non-trivial, especially in real-world applications.

It is well known that there is a strong link between operator preconditioning and Schur complement preconditioning. Preconditioners for a range of problems from mixed finite elements [23, 49, 53, 63] were developed as approximations to the primal or dual Schur complement, but can be thought of as a finite dimensional analogue to τ.

4. The relationship between natural norm and Schur complement preconditioners.

Pestana and Wathen [47] describe the link between the dual Schur complement $(BP^{-1}B^T)$ and the Riesz map, μ, on the secondary variables for a specific choice of P, and we give an alternative proof of this result below. Zulehner [66], presents similar results to [47] but considering the underlying operators instead of the discretized system. We also derive an analogous link between the primal Schur complement $(A + B^T D^{-1} B)$ and Riesz map, χ, on the primary variables for a specific choice of D. Again, Zulehner [66] presents similar arguments based operators. These relationships are central to the theory that we develop for practical element preconditioners in Section 5.

4.1. Saddle point theory. We assume that the operators $a(\cdot, \cdot)$ and $b(\cdot, \cdot)$ are bounded, satisfying

$$(4.1) \qquad |a(u, v)| \le \Gamma_a^* \|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}} \quad \forall u, v \in \mathcal{X}$$

$$(4.2) \qquad |b(u, p)| \le \Gamma_b^* \|u\|_{\mathcal{X}} \|p\|_{\mathcal{M}} \quad \forall u \in \mathcal{X}, p \in \mathcal{M},$$

for some positive constants Γ_a^*, Γ_b^*. We define the space

$$\mathcal{V} := \{v \in \mathcal{X} : b(v, p) = 0 \ \forall \ p \in \mathcal{M}\}.$$

Brezzi's splitting theorem [12] tells us that the mapping in (1.1) defines an isomorphism if and only if the bilinear forms satisfy the following conditions:

    1. the bilinear form $a(\cdot, \cdot)$ is $\mathcal{V}-$elliptic, i.e.

$$(4.3) \qquad \qquad \exists \alpha_* > 0 \text{ s.t. } a(v, v) \geq \alpha_* \|v\|_{\mathcal{X}}^2 \quad \forall v \in \mathcal{V}.$$

    2. The bilinear form $b(\cdot, \cdot)$ satisfies the inf-sup condition:

$$(4.4) \qquad \qquad \exists \beta_* > 0 \text{ s.t. } \inf_{p \in \mathcal{M}} \sup_{u \in \mathcal{X}} \frac{b(u, p)}{\|u\|_{\mathcal{X}} \|p\|_{\mathcal{M}}} \geq \beta_*.$$

When we discretize problems of the form (1.1), we need to be careful in the choice of approximation spaces. It is not necessarily true that the finite dimensional problem will satisfy the equivalent inf-sup condition. In particular, we cannot choose the spaces $\mathcal{X}_h$ and $\mathcal{M}_h$ independently; see, for example, Brezzi and Fortin [13, Chapter 2]. If the spaces are complimentary, we say they satisfy the Ladyshenskaja-Babuška-Brezzi (LLB) condition, or that they are inf-sup stable. The discrete analogue to (4.4) is therefore

$$(4.5) \qquad \qquad \inf_{p_h \in \mathcal{M}_h} \sup_{u_h \in \mathcal{X}_h} \frac{b(u_h, p_h)}{\|u_h\|_{\mathcal{X}_h} \|p_h\|_{\mathcal{M}_h}} \geq \beta.$$

For a stable discretization, the inf-sup constant, β, is independent of the discretization parameter (mesh size).

There is an affinity between the continuous operators and their matrix representations, and we can write the relations above in terms of matrices. In finite dimensions, (4.1) and (4.2) become

$$(4.6) \qquad \qquad |\mathbf{u}^T A \mathbf{v}| \leq \Gamma_a \|\mathbf{u}\|_X \|\mathbf{v}\|_X \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$$

$$(4.7) \qquad \qquad |\mathbf{p}^T B \mathbf{u}| \leq \Gamma_b \|\mathbf{u}\|_X \|\mathbf{p}\|_M \quad \forall \mathbf{u} \in \mathbb{R}^n, \mathbf{p} \in \mathbb{R}^m$$

for positive constants Γ_a and Γ_b, where

$$(4.8) \qquad \qquad X_{i,j} = (\phi_i, \phi_j)_{\mathcal{X}_h} \quad \text{and} \quad M_{i,j} = (\psi_i, \psi_j)_{\mathcal{M}_h},$$

are the matrices that define the natural inner products of the primary and secondary spaces, respectively. Similarly, the matrix representation of the inf-sup condition (4.5) is

$$(4.9) \qquad \qquad \sup_{\mathbf{u}} \frac{(B\mathbf{u}, \mathbf{p})}{\|\mathbf{u}\|_X} \geq \beta \|\mathbf{p}\|_M \quad \forall \mathbf{p} \in \mathbb{R}^m.$$

Following the theory in Braess [11, Chapter 4], we associate the mapping $\mathbb{B} : \mathcal{X}_h \to \mathcal{M}_h'$ with $(\mathbb{B}u_h, p_h)_{\mathcal{M}_h} = b(u_h, p_h)$ for all $p_h \in \mathcal{M}_h$, and its adjoint mapping $\mathbb{B}' : \mathcal{M}_h \to \mathcal{X}_h'$ with $(\mathbb{B}'q_h, v_h)_{\mathcal{M}_h} = b(v_h, q_h)$ for all $v_h \in \mathcal{X}_h$. The following lemma gives five alternative statements of the inf-sup condition (4.9).

Lemma 4.1. *Let $\mathcal{V}_h$ be the finite dimensional analogue of $\mathcal{V}$. The following statements are equivalent:*

1. *There exits a constant β that satisfies (4.9).*
2. *The operator $\mathbb{B}: \mathcal{V}_h^\perp \to \mathcal{M}_h'$ is an isomorphism, and*

$$\|\mathbb{B}v_h\|_{\mathcal{M}_h} \geq \beta \, \|v_h\|_{\mathcal{X}_h} \qquad \forall v_h \in \mathcal{V}_h^\perp.$$

3. *For all $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v} \in \mathrm{null}(B)^\perp$,*

$$\|B\mathbf{v}\|_{M^{-1}} \geq \beta \, \|\mathbf{v}\|_X,$$

4. *The operator $\mathbb{B}': \mathcal{M}_h \to \mathcal{V}_h^0 \subset \mathcal{X}_h'$ is an isomorphism, and*

$$\left\|\mathbb{B}'p_h\right\|_{\mathcal{X}_h} \geq \beta \, \|p_h\|_{\mathcal{M}_h} \qquad \forall p_h \in \mathcal{M}_h.$$

5. *For all $\mathbf{p} \in \mathbb{R}^m$,*

$$\left\|B^T\mathbf{p}\right\|_{X^{-1}} \geq \beta \, \|\mathbf{p}\|_M.$$

Proof. For the equivalence of 1, 2 and 4, see Braess [11, Lemma 4.2].

We now show $2 \iff 3$. First note that $\|v_h\|_{\mathcal{X}_h}^2 = \mathbf{v}^T X \mathbf{v}$. Since $\mathbb{B}v_h \in \mathcal{M}_h$, $\mathbb{B}v_h = \sum_j \mathbf{q}_j \psi_j$ for some coefficients $\mathbf{q} \neq \mathbf{0}$. Then

$$\|\mathbb{B}v_h\|_{\mathcal{M}_h} = \sum_i \sum_j \mathbf{v}_i \mathbf{q}_j \left(\mathbb{B}\phi_i, \psi_j\right)_{\mathcal{M}_h} = \sum_i \sum_j \mathbf{v}_i \mathbf{q}_j b(\phi_i, \psi_j) = \mathbf{q}^T B \mathbf{v}.$$

Furthermore, we have

$$(B\mathbf{v})_j = \sum_i \mathbf{v}_i (\mathbb{B}\phi_i, \psi_j)_{\mathcal{M}_h} = (\mathbb{B}v_h, \psi_j)_{\mathcal{M}_h} = \sum_k \mathbf{q}_k \left(\psi_k, \psi_j\right)_{\mathcal{M}_h} = (M\mathbf{q})_j,$$

and so $\mathbf{q} = M^{-1}B\mathbf{v}$, which shows the equivalence of 2 and 3.

A similar argument gives $4 \iff 5$. ∎

We will use these results below to show relationships between a Schur complement and the natural norm for such problems.

4.2. Dual Schur complements and the natural norm. The following theorem formalises the relationship between the dual Schur complement and the matrix associated with the natural norm of $\mathcal{M}_h$ for saddle point systems in general. We refer the reader to Pestana and Wathen [47, Section 3] for an alternative proof of this result.

Theorem 4.2. *Suppose we have an inf-sup stable discretization of a saddle point problem of the form (1.1) where $a(\cdot, \cdot)$ is $\mathcal{V}_h-$elliptic. Let λ satisfy the generalized eigenvalue problem*

$$BX^{-1}B^T\mathbf{p} = \lambda M\mathbf{p}.$$

Then λ is independent of h, and lies in the range $[\beta^2, \Gamma_b^2]$.

Proof. Consider the generalized Rayleigh quotient

$$\frac{\mathbf{p}^T B X^{-1} B^T \mathbf{p}}{\mathbf{p}^T M \mathbf{p}}.$$

We can bound this quantity from below by β^2 using Condition 5 of Lemma 4.1, and from above by Γ_b^2 using (4.7) with $\mathbf{u} = X^{-1} B^T \mathbf{p}$. ∎

Remark 4.3. *From Theorem 4.2, if A can be identified with X, then the dual Schur complement of the leading block of $\mathcal{A}$ is spectrally equivalent to the matrix M.*

4.3. Primal Schur complements and the natural norm. Zulehner [66] examines the relationship between the operator associated primal Schur complement $(\mathbb{A} + \mathbb{B}'\mathbb{M}^{-1}\mathbb{B})$ and the operator associated natural norm preconditioner on primary space ($\mathbb{X}$). However, we are unaware of a result in the literature that links the discrete primal Schur complement to the natural norm in a way analogous to Theorem 4.2. Below we prove results for the cases when A is symmetric positive definite (Theorem 4.4) and maximally rank deficient (Theorem 4.5).

Theorem 4.4. *Suppose that the bilinear form $a(\cdot, \cdot)$ is elliptic with ellipticity constant α_x, and the associated matrix A is symmetric. For an inf-sup stable discretization, the generalized eigenvalues satisfying*

$$(A + B^T M^{-1} B)\mathbf{x} = \lambda X \mathbf{x}$$

are such that $\lambda \in [\alpha_x, \Gamma]$, where $\Gamma = \Gamma_a + \Gamma_b^2$.

Proof. For the lower bound, we have

$$\mathbf{x}^T (A + B^T M^{-1} B)\mathbf{x} = \mathbf{x}^T A \mathbf{x} + \mathbf{x}^T B^T M^{-1} B \mathbf{x} \geq \mathbf{x}^T A \mathbf{x} \geq \alpha_x \mathbf{x}^T X \mathbf{x}.$$

For the upper bound, first note that

$$\mathbf{x}^T B^T M^{-1} B \mathbf{x} \leq \Gamma_b \|\mathbf{x}\|_X \|M^{-1} B \mathbf{x}\|_M = \Gamma_b (\mathbf{x}^T X \mathbf{x})^{1/2} (\mathbf{x}^T B^T M^{-1} B \mathbf{x})^{1/2},$$

where we have used (4.7). Therefore using this result, together with (4.6), we obtain

$$\mathbf{x}^T (A + B^T M^{-1} B)\mathbf{x} = \mathbf{x}^T A \mathbf{x} + \mathbf{x}^T B^T M^{-1} B \mathbf{x} \leq \Gamma_a \mathbf{x}^T X \mathbf{x} + \Gamma_b^2 \mathbf{x}^T X \mathbf{x},$$

which gives the required result. ∎

Theorem 4.5. *Suppose that the bilinear form is $\mathcal{V}_h$-elliptic with ellipticity constant α, and the matrix A associated with the bilinear form $a(\cdot, \cdot)$ is symmetric and positive semi-definite with nullity $n - m$. For an inf-sup stable discretization, the generalized eigenvalues satisfying*

$$(A + B^T M^{-1} B)\mathbf{x} = \lambda X \mathbf{x}$$

are such that $\lambda \in [\gamma, \Gamma]$, where $\gamma = \frac{1}{2} \min(\alpha, \beta^2)$ and $\Gamma = \Gamma_a + \Gamma_b^2$.

Proof. We can decompose $\mathbf{x} = \mathbf{y} + \mathbf{z}$ where $\mathbf{y} \in \text{null}(B)$ and $\mathbf{z} \in \text{null}(B)^{\perp}$. Since A has nullity $n - m$, and $a(\cdot, \cdot)$ is $\mathcal{V}_h$−elliptic, we must have that $A\mathbf{z} = \mathbf{0}$.

We have that

$$\mathbf{x}^T(A + B^T M^{-1}B)\mathbf{x} = \mathbf{y}^T A\mathbf{y} + \mathbf{z}^T B^T M^{-1}B\mathbf{z}$$
$$\geq \alpha\mathbf{y}^T X\mathbf{y} + \beta^2\mathbf{z}^T X\mathbf{z}$$
$$\geq \min(\alpha, \beta^2)(\mathbf{y}^T X\mathbf{y} + \mathbf{z}^T X\mathbf{z}),$$

using Lemma 4.1. Since $(\mathbf{y} - \mathbf{z})^T X(\mathbf{y} - \mathbf{z}) \geq 0$, we have that $\mathbf{y}^T X\mathbf{y} + \mathbf{z}^T X\mathbf{z} \geq 2\mathbf{y}^T X\mathbf{z}$. Therefore

$$\mathbf{x}^T(A + B^T M^{-1}B)\mathbf{x} \geq \min(\alpha, \beta^2)\left(\frac{1}{2}\left(\mathbf{y}^T X\mathbf{y} + \mathbf{z}^T X\mathbf{z}\right) + \frac{1}{2}\left(\mathbf{y}^T X\mathbf{y} + \mathbf{z}^T X\mathbf{z}\right)\right)$$
$$\geq \min(\alpha, \beta^2)\frac{1}{2}\left(\mathbf{y}^T X\mathbf{y} + 2\mathbf{y}^T X\mathbf{z} + \mathbf{z}^T X\mathbf{z}\right)$$
$$= \gamma\,\mathbf{x}^T X\mathbf{x},$$

from which we obtain the lower bound. The upper bound follows the proof of Theorem 4.4.■

Theorems 4.4 and 4.5 provide the motivation for the "natural" choice for the weighting matrix $D = M$. From this point onwards we consider the *primal* Schur complement to be $A + B^T M^{-1}B$.

Remark 4.6. *If* $\mathrm{null}(B)$ *and* $\mathrm{null}(B)^\perp$ *are* $X-$*orthogonal subspaces, then it is straightforward to adapt the proof of Theorem 4.5 to show that minimum eigenvalue is* $\min(\alpha, \beta^2)$. *This is the case in, for example, Maxwell's equations [23, Section 2.2].*

4.4. Schur complement relations. An interesting consequence of Theorems 4.2, 4.4 and 4.5 is that primal or dual Schur complements for different problems that share a primal or dual space, respectively, are spectrally equivalent. We formalize this statement in the following.

Theorem 4.7. *Suppose that you have two inf-sup stable discretizations on the mixed spaces* $\mathcal{U}_h \times \mathcal{L}_h$ *and* $\mathcal{Z}_h \times \mathcal{L}_h$ *where the corresponding saddle-point matrices are:*

$$(4.10) \qquad \begin{pmatrix} K_1 & B_1^T \\ B_1 & 0 \end{pmatrix} \quad and \quad \begin{pmatrix} K_2 & B_2^T \\ B_2 & 0 \end{pmatrix},$$

respectively. If U *and* Z *define the matrices associated with the natural inner products of* $\mathcal{U}_h$ *and* $\mathcal{Z}_h$, *then the eigenvalues,* λ, *of the generalized eigenvalue problem*

$$(4.11) \qquad B_1 U^{-1}B_1^T\mathbf{p} = \lambda B_2 Z^{-1}B_2^T\mathbf{p},$$

are bounded from above and away from zero independently of h, *and therefore the dual Schur complements in* (4.11) *are spectrally equivalent.*

Proof. Let us define L as the matrix associated with the natural inner product of $\mathcal{L}_h$. Then from Theorem 4.2, both $B_1 U^{-1}B_1^T$ and $B_2 Z^{-1}B_2^T$ are spectrally equivalent to L. Thus, $B_1 U^{-1}B_1^T$ and $B_2 Z^{-1}B_2^T$ are spectrally equivalent. ■

In a similar fashion to Theorem 4.7, suppose the saddle point systems (4.10) share a primal space, so that they are posed on the mixed spaces $\mathcal{W}_h \times \mathcal{Q}_h$ and $\mathcal{W}_h \times \mathcal{T}_h$. Then it is possible to show that $K_1 + B_1^T Q^{-1} B_1$ is spectrally equivalent to $K_2 + B_2^T T^{-1} B_2$, where Q and T define the matrices associated with the natural inner products of $\mathcal{Q}_h$ and $\mathcal{T}_h$, respectively. These results give intuition as to why the approximations we describe in Section 6 are effective.

5. Element Schur complement preconditioners.

We now turn our attention to the element Schur complement approximations defined in (1.4). In Section 5.1 we define the term elementally inf-sup stable, and show spectral equivalence of element Schur complements and relevant Gram matrices. We apply these results in Section 5.2 to describe element-based preconditioners for saddle point problems that may not be elementally inf-sup stable, and discuss how these may be practically applied in Section 5.3.

5.1. Element saddle-point system.

In Section 4.1 we looked at the conditions under which the global saddle point system is well-defined. Here we look at the problem on an element.

Definition 5.1 (Elementally inf-sup stable). *Given a domain, Ω, with a triangulation $\{T_i\}$, the discrete saddle point system, find $(u_h, p_h) \in \mathcal{P}_h \times \mathcal{D}_h$ such that*

$$
\begin{aligned}
h(u, v) + c(v, p) &= \langle f, v \rangle \quad \forall v \in P_h, \\
c(u, q) &= \langle g, q \rangle \quad \forall q \in D_h,
\end{aligned}
\tag{5.1}
$$

is said to be elementally inf-sup stable if, for all elements in the triangulation, the elemental saddle point systems

$$
\begin{bmatrix} H_e & C_e^T \\ C_e & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}_e \\ \mathbf{p}_e \end{bmatrix} = \begin{bmatrix} \mathbf{f}_e \\ \mathbf{g}_e \end{bmatrix}.
$$

are inf-sup stable.

Note that it is not necessarily true that if a mixed finite element discretization is inf-sup stable on the whole domain then it is elementally inf-sup stable. For example, it may be the case that there exists a macroelement (a subset of the domain) for which the discretization is inf-sup stable, but individual elements are not.

Lemma 5.2. *Consider an elementally inf-sup stable discretization, as in Definition 5.1. Let P_e and D_e be the elemental Gram matrices of the spaces P_h and D_h, respectively. Further, let $\hat{S}_{\mathrm{dual}}$ denote the global matrix assembled from the element matrices $C_e P_e^{-1} C_e^T$ and D be the assembled Gram matrix of the space $\mathcal{D}_h$. The generalized eigenvalues satisfying:*

$$
\hat{S}_{\mathrm{dual}} \mathbf{x} = \lambda D \mathbf{x}
$$

are bounded within some finite region $[\gamma_d^2, \Gamma_d^2]$ independently of h.

Proof. By Theorem 4.2 and Definition 5.1, there exist constants γ_e^2, Γ_e^2 independent of h such that

$$
\gamma_e^2 \leq \frac{\mathbf{x}^T C_e P_e^{-1} C_e^T \mathbf{x}}{\mathbf{x}^T D_e \mathbf{x}} \leq \Gamma_e^2.
$$

$\blacksquare$

We can write $\hat{S}_{\text{dual}} = N^T C_e P_e^{-1} C_e^T N$ and $D = N^T D_e N$ for a boolean matrix N. Therefore the result holds with $\gamma_d^2 = \min_e \gamma_e^2$ and $\Gamma_d^2 = \max_e \Gamma_e^2$.

Lemma 5.3. *Consider an elementally inf-sup stable discretization, as in Definition 5.1, where the underlying bilinear forms of (5.1) satisfies the requirements of Theorem 4.4 or 4.5. Let P_e and D_e be the elemental Gram matrices of the spaces P_h and D_h, respectively. Define $\hat{S}_{\text{primal}}$ to be the global matrix to be assembled from the element matrices $H_e + C_e^T D_e^{-1} C_e$ and P, which is the Gram matrix of the space $\mathcal{P}_h$, is assembled from the elemental matrices P_e. The generalized eigenvalues satisfying:*

$$\hat{S}_{\text{primal}} \mathbf{x} = \lambda P \mathbf{x}$$

are bounded within some finite region $[\gamma_p^2, \Gamma_p^2]$ independently of h.

Proof. Similar to Lemma 5.2. ∎

5.2. Relationship between element and global Schur complements. In Section 5.1, we showed that, for an elementally inf-sup stable problem, the assembled element Schur complements are spectrally equivalent to the appropriate Gram matrix. If either the primal or dual space in Definition 5.1 coincides with that of the saddle point problem in (1.2), we can further show spectral equivalence of the element Schur complements in Lemmas 5.2 and 5.3 and the global Schur complements presented in Section 4. We give the details via the following two theorems.

Theorem 5.4. *Consider a discete saddle point system (1.2) on $\mathcal{X}_h \times \mathcal{M}_h$ that satisfies the conditions of Theorem 4.2. Suppose further that we have an elementally inf-sup stable discretization of a saddle point problem on the mixed space $\mathcal{P}_h \times \mathcal{M}_h$. If $\hat{S}_{\text{dual}} = N^T C_e P_e^{-1} C_e^T N$, then generalized eigenvalues satisfying*

$$BX^{-1}B^T \mathbf{x} = \lambda \hat{S}_{\text{dual}} \mathbf{x}$$

satisfy $\lambda \in [\beta^2/\Gamma_d^2, \Gamma_b^2/\gamma_d^2]$, which is independent of the mesh size.

Proof. From Lemma 5.2 and Theorem 4.2, both $\hat{S}_{\text{dual}}$ and $BX^{-1}B^T$ are spectrally equivalent to M. Thus, $\hat{S}_{\text{dual}}$ and $BX^{-1}B^T$ are spectrally equivalent, as required. ∎

Theorem 5.5. *Consider a discrete saddle point system (1.2) on $\mathcal{X}_h \times \mathcal{M}_h$ that satisfies the conditions of Theorem 4.4 or 4.5. Suppose further that we have an elementally inf-sup stable discretization on the mixed space $\mathcal{X}_h \times \mathcal{D}_h$. If $\hat{S}_{\text{primal}} = L^T (H_e + C_e^T D_e^{-1} C_e) L$, then the generalized eigenvalues satisfying*

$$(A + B^T M^{-1} B) \mathbf{x} = \lambda \hat{S}_{\text{primal}} \mathbf{x}$$

are bounded away from zero, with $\lambda \in [\alpha_x/\widehat{\Gamma}_p^2, \Gamma/\widehat{\gamma}_p^2]$ in the elliptic case, and $\lambda \in [\gamma/\widehat{\Gamma}_p^2, \Gamma/\widehat{\gamma}_p^2]$ in the maximally rank deficient case.

Proof. Similar to Theorem 5.4. ∎

Theorems 5.4 and 5.5 show spectral equivalence between the element and global Schur complements for both the dual and primal cases. The existence of a natural norm preconditioner is vital in the proofs, but explicit knowledge of it is not required to form a practical

preconditioner. Again, we highlight the interesting observation that the elemental matrices that make up $\hat{S}_{\text{primal}}$ or $\hat{S}_{\text{dual}}$ need not be those that would be present in the primal or dual Schur complements of (1.2), all that is required is that the problems share a primal or dual space, respectively.

5.3. Solving systems associated with the element Schur complements. An important part of preconditioners is the availability of techniques to solve systems associated with them. It is clear from the construction of the primal and dual element Schur complements that the matrices are sparse. For the global primal Schur complement $(A + B^T D^{-1} B)$, it is possible to choose the weighting matrix D such that the Schur complement matrix is sparse. However, this does not necessarily mean that there are efficient preconditioners for such systems. Since the element Schur complements we proposed are spectrally equivalent to operators that define inner products, then off-the-shelf multigrid methods designed for Riesz-map preconditioners should perform well. Specifically, PCPATCH [19] and the auxiliary space multigrid from [27] are designed for Riesz-map operators arising from $H(\text{curl}, \Omega)$ and $H(\text{div}, \Omega)$ discretizations, whilst standard algebraic multigrid methods work well for Riesz-map operators arising from $H^1(\Omega)$ and $L^2(\Omega)$ discretizations.

6. Examples. In this section, we present numerical results for both symmetric and non-symmetric problems, using `Firedrake` [50] with `PETSc` [5, 6] and `PETSc4PY` [15] as the solver interface. We have also used `FEniCS` [1, 34] to generate the eigenvalue tables that depict the theoretical bounds produced in Section 5. For systems involving simple H^1 elliptic type operators, we use an algebraic multigrid solver from `HYPRE` [18], and for more complex operators, arising from $H(\text{curl})$ discretizations, we use the sparse direct solver `MUMPS` [2, 3]. We set the absolute and relative tolerance of the Krylov subspace solver to be 10^{-6} and 10^{-8}, respectively. Table 1 introduces the notation used in the results tables. We note that the timings are for the total time to solve the model, including the assembly of both the linear system and the preconditioners, as well as the linear solve.

Column label	Linear model	Nonlinear model
DoF	Total degrees of freedom (system size)	Total degrees of freedom (system size)
Time	Total time to solve the linear system (including assembly of the linear system and preconditioner, as well as the linear solve time)	Total time to solve the nonlinear system (including assembly of all linear systems and preconditioners, as well as linear solve time at each nonlinear iteration)
Iteration	Total number of linear iterations	Total number of nonlinear iterations/average number of linear iterations

Table 1: Notation used in the results tables

6.1. Stokes Flow. The Stokes equations describe viscous incompressible flow over some bounded, sufficiently regular, domain $\Omega \subset \mathbb{R}^d$:

$$\text{(6.1)} \qquad -\nu\nabla^2\vec{u} + \nabla p = \vec{f} \qquad \text{in } \Omega$$

$$\text{(6.2)} \qquad \nabla \cdot \vec{u} = 0 \qquad \text{in } \Omega,$$

where $\vec{u}$, p and ν are the fluid velocity, fluid pressure and the viscosity, respectively. For a detailed description of the problem, see, e.g., Temam [57, Chapter 1] or Elman, Silvester and Wathen [17, Chapter 3]. We seek a weak solution $(\vec{u}, p) \in H^1_E(\Omega)^d \times L^2(\Omega)$ such that

$$\begin{array}{rcll} \nu(\nabla\vec{u}, \nabla\vec{v})_\Omega - (p, \nabla\cdot\vec{v})_\Omega &=& (\vec{f}, \vec{v})_\Omega, & \forall\vec{v} \in H^1_{E_0}(\Omega)^d, \\ -(q, \nabla\cdot\vec{u})_\Omega &=& 0, & \forall q \in L^2(\Omega), \end{array}$$

where $H^1_E(\Omega)^d$ and $H^1_{E_0}(\Omega)^d$ are subsets of $H^1(\Omega)^d$ that satisfy the required boundary conditions. We consider the classical test problem known as the *leaky* cavity driven flow, see Elman, Silvester and Wathen [17, Example 3.1.3] for full details.

Upon discretization, we obtain a saddle point system of the form (1.2), where A is a discrete vector Laplacian scaled by the viscosity, ν, and B is a fluid divergence operator.

Theory. For a $H^1_E(\Omega)^d \times L^2(\Omega)$ discretization of the Stokes problem, the operator preconditioner [37, Example 3.1] takes the form

$$\text{(6.3)} \qquad \begin{bmatrix} \nu A & 0 \\ 0 & \frac{1}{\nu}Q \end{bmatrix},$$

where Q is the pressure mass matrix. In the language of (4.8), X and M are identified with A and Q, respectively. Silvester and Wathen [53, 63] showed that the mass matrix, Q, is spectrally equivalent to the dual Schur Complement, $BA^{-1}B^T$.

The elemental mass matrices, Q_e, are invertible, and so we can define the primal element Schur complement preconditioner as

$$\text{(6.4)} \qquad \begin{bmatrix} \nu L^T(A_e + B_e^T Q_e^{-1} B_e)L & 0 \\ 0 & \frac{1}{\nu}Q \end{bmatrix}.$$

When it comes to the dual element Schur complement, however, the elemental matrices A_e are singular, and so we cannot form a dual Schur complement in the same way. One way to get around this issue would be to add a small shift, taking $X_e = A_e + \epsilon T_e$, where T_e is the element mass matrix of the velocity space. Provided ϵ is sufficiently small, the conditions (4.3) and (4.4) will be satisfied. The preconditioner in this case will be

$$\text{(6.5)} \qquad \begin{bmatrix} \nu A & 0 \\ 0 & \frac{1}{\nu}N^T B_e(A_e + \epsilon T_e)^{-1} B_e^T N \end{bmatrix}.$$

The results in Section 4.4 show that the matrices of the element Schur complement need not arise from the same problem. The preconditioner in (6.5) can also be derived from a PDE with bilinear form $a(\vec{u}, \vec{v}) = (\nabla\vec{u}, \nabla\vec{v}) + \epsilon(\vec{u}, \vec{v})$; see, e.g., Mardal and Winther [37, Example 7.5].

Alternatively, consider the $H(\text{div}, \Omega) \times L_2(\Omega)$ formulation of the mixed Laplacian. From [49], the bilinear form for the constraint matrix, c, and the natural inner product, p, of $H(\text{div}, \Omega)$ are defined as:

$$
\begin{aligned}
c(\vec{\tau}, u) &= (\nabla \cdot \vec{\tau}, u) & \text{where } \vec{\tau} \in H(\text{div}, \Omega), u \in L^2(\Omega) \\
p(\vec{\tau}, \vec{\sigma}) &= (\nabla \cdot \vec{\tau}, \nabla \cdot \vec{\sigma}) + (\vec{\tau}, \vec{\sigma}) & \text{where } \vec{\tau}, \vec{\sigma} \in H(\text{div}, \Omega).
\end{aligned}
\tag{6.6}
$$

Thus, the matrix C and P are associated with the bilinear forms $c(\vec{\tau}, u)$ and $p(\vec{\tau}, \vec{\sigma})$ from (6.6), respectively. The dual space from the mixed Laplacian in [49] is the same as the Stokes discretization, and thus, by Theorem 4.7 $BA^{-1}B^T$ and $CP^{-1}C^T$ are spectrally equivalent. From Theorem 5.4, $N^T C_e P_e^{-1} C_e^T N$, is spectrally equivalent to the Stokes Schur complement, $BA^{-1}B^T$, where C_e and P_e are the element matrices of C and P, respectively. The preconditioner here will be given by

$$
\begin{bmatrix} \nu A & 0 \\ 0 & \frac{1}{\nu} N^T C_e P_e^{-1} C_e^T N \end{bmatrix}.
\tag{6.7}
$$

To verify the quality of the preconditioners (6.4), (6.5) and (6.7) we consider the generalized eigenvalue problems from Theorems 5.4 and 5.5. We compute the minimum and maximum eigenvalues coming from the eigenvalue problems

$$
\begin{aligned}
N^T C_e P_e^{-1} C_e^T N x = \lambda_1 (BA^{-1}B^T)x, \quad N^T B_e (A_e + \epsilon T_e)^{-1} B_e^T N x = \lambda_2 (BA^{-1}B^T)x, \\
L^T (A_e + B_e^T Q_e^{-1} B_e) L x = \lambda_3 (A + B^T Q^{-1} B)x,
\end{aligned}
\tag{6.8}
$$

and present these in Table 2. From the table, we can see that for all three generalized eigenvalue problems the minimum and maximum eigenvalues are bounded away from zero and are bounded from above, and thus, the approximations (6.4), (6.5) and (6.7) are spectrally equivalent to the corresponding global Schur complement.

# cells	$\hat{S}_1 x = \lambda_1 BA^{-1}B^T x$		$\hat{S}_2 x = \lambda_2 BA^{-1}B^T x$		$\hat{S}_3 x = \lambda_3 (A + B^T Q^{-1} B)x$	
	$\min_e(\lambda_1)$	$\max_e(\lambda_1)$	$\min_e(\lambda_2)$	$\max_e(\lambda_2)$	$\min_e(\lambda_3)$	$\max_e(\lambda_3)$
8	0.134775	0.906035	0.081566	0.589784	0.506419	1.0
32	0.135293	0.993885	0.079786	0.625184	0.500296	1.0
128	0.134122	0.999828	0.078566	0.653874	0.500000	1.0
512	0.133646	1.000029	0.077982	0.663387	0.500000	1.0

Table 2: Minimum and maximum eigenvalues of the element Schur complements with $S_1 = N^T C_e P_e^{-1} C_e^T N$, $S_2 = N^T B_e (A_e + \epsilon T_e)^{-1} B_e^T N$ and $S_3 = L^T (A_e + B_e^T Q_e^{-1} B_e) L$

While Table 2 gives us information about how closely the element Schur complements approximate the relevant global Schur complement, it does not give a complete picture of the quality of the preconditioner for the whole system. Figure 2 shows the generalized eigenvalues

$$
\mathcal{A}x = \lambda \mathcal{P}_* x,
$$

where $\mathcal{A}$ is the discretized Stokes equation and $\mathcal{P}_*$ is one of the three preconditioners (6.4), (6.5) and (6.7). Aside from the eigenvalue at zero, that corresponds to the known null-space of the Stokes problem [17, Chapter 3.3], the eigenvalues are bounded away from zero and strongly clustered, which will give fast convergence for MINRES. For reference, a block diagonal preconditioner with the pressure mass matrix in the (2,2) block gives eigenvalues that are indistinguishable to those for $\mathcal{P}_1$ to the eye.
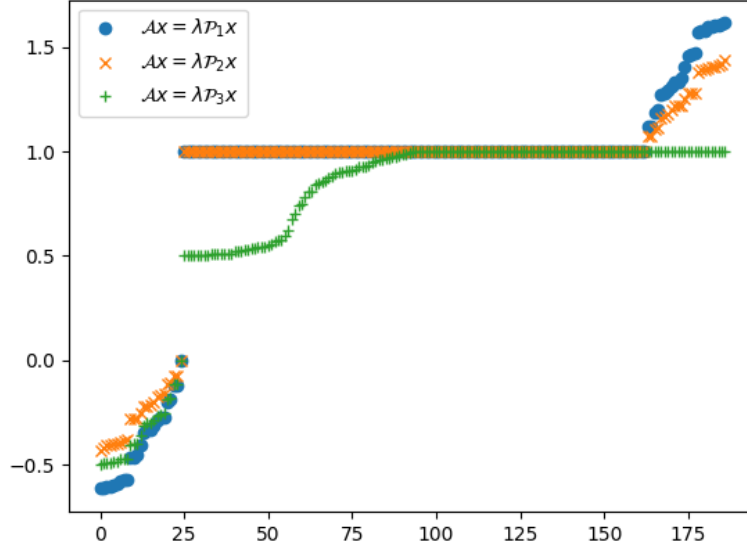

Figure 2: Eigenvalues of generalized eigenvalue problem with total 187 degrees of freedom, where $\mathcal{A}$ is the discretized Stokes problem and $\mathcal{P}_1$, $\mathcal{P}_2$ and $\mathcal{P}_3$ are given in (6.7), (6.5) and (6.4), respectively.

Numerical Results. In the following we use the (inf-sup stable) P2-P1 (Taylor-Hood) element [56], although the methods can be applied to any stable element pair. Table 3 give the two and three dimensional timing and iteration results for a constant Reynolds number, $\nu = 1e\text{-}3$. The table shows that for both the two and three dimensional results that the element Schur complement preconditioners perform the same as the natural norm preconditioners. It appears that all iterations numbers increase slightly with respect to the mesh size but are all comparable on a single mesh. The timings for primal element Schur complement slightly higher than the other preconditioners. This is likely to the construction of the algebraic multigrid solver for the leading block of (6.4). The timing results appear to be linear with respect to the problem size.

15

	Dual precond in (6.7)		Dual-ε precond in (6.5)		Primal precond in (6.4)		Natural norm precond in (6.3)	
DoF	Time	Iteration	Time	Iteration	Time	Iteration	Time	Iteration
9,539	0.4	43	0.4	42	0.3	42	0.3	41
37,507	0.8	43	1.0	43	0.8	45	0.7	41
148,739	3.6	45	3.8	47	4.9	45	3.8	43
592,387	11.8	51	14.2	47	15.6	48	13.8	51
2,364,419	53.1	57	55.7	53	52.6	45	61.1	55
9,447,427	214.7	55	200.0	50	245.1	51	212.1	53

(a) Two dimensional results

	Dual precond in (6.7)		Dual-ε precond in (6.5)		Primal precond in (6.4)		Natural norm precond in (6.3)	
DoF	Time	Iteration	Time	Iteration	Time	Iteration	Time	Iteration
402	0.4	49	0.2	43	0.2	51	0.2	40
2,312	0.6	59	0.3	57	0.3	57	0.2	55
15,468	1.8	64	1.7	63	2.3	60	1.7	59
112,724	18.8	66	18.6	66	26.5	63	17.3	62
859,812	190.3	68	186.7	68	270.7	66	181.7	66

(b) Three dimensional results

Table 3: Iteration and timing results for the two-dimensional Stokes *leaky* cavity driven flow problem with $\nu = $ 1e-3 and $\varepsilon = $1e-6.

6.2. Mixed formulation of Maxwell's equations. We now consider the time-harmonic Maxwell equation in mixed form [23, 24, 38]. The continuous problem is given as follows:

(6.9)
$$\begin{aligned} \frac{1}{\mathrm{Re}_m}\nabla \times \nabla \times \mathbf{b} + \nabla r &= \mathbf{f} \quad \text{in } \Omega \\ \nabla \cdot \mathbf{b} &= 0 \quad \text{in } \Omega, \end{aligned}$$

where $\mathbf{b}$ is the magnetic field, r is the Lagrange multiplier associated with the divergence constraint on the magnetic field and Re_m is the magnetic Reynolds number. For this problem, we set the forcing terms and Dirichlet boundary conditions corresponding to the exact solution

$$\vec{u} = \begin{bmatrix} \exp(x)\cos(y) \\ \exp(x)\sin(y) \end{bmatrix} \quad \text{and} \quad p = xy.$$

The standard weak form is: find $(\vec{b}, r) \in H(\mathrm{curl}, \Omega) \times H_E^1(\Omega)$ such that

$$\begin{aligned} \frac{1}{\mathrm{Re}_m}(\nabla \times \vec{b}, \nabla \times \vec{c})_\Omega + (\vec{c}, \nabla r)_\Omega &= (\mathbf{f}, \vec{c})_\Omega, \quad \forall \vec{c} \in H(\mathrm{curl}, \Omega), \\ (\vec{b}, \nabla s)_\Omega &= 0, \quad \forall s \in H_{E_0}^1(\Omega). \end{aligned}$$

In the mixed Maxwell case, the matrices A and B in (1.2) are the discrete curl-curl operator and magnetic divergence operator, which we denote as $\frac{1}{\mathrm{Re}_m}K$ and J, respectively.

Theory. As with the previous example, the operator preconditioner preconditioner for a $H(\text{curl}, \Omega) \times H_E^1(\Omega)$ discretization of the mixed Maxwell problem (see Mardal and Winther [37, Example 7.4]) is given by:

$$
(6.10) \qquad \begin{bmatrix} \frac{1}{\text{Re}_m} K + R & 0 \\ 0 & U \end{bmatrix},
$$

where R and U are the $H(\text{curl}, \Omega)$ mass matrix and scalar Laplacian operator, respectively. In the language of (4.8), X and M are identified with $\frac{1}{\text{Re}_m} K + R$ and U, respectively. Greif and Schötzau [23], showed that if you chose the weighting matrix in the the primal Schur complement to be the scalar Laplacian then $\frac{1}{\text{Re}_m} K + J^T U^{-1} J$ is spectrally equivalent to $\frac{1}{\text{Re}_m} K + R$, and thus, the proposed approximate Schur complement preconditioner in [23] is the same as that in (6.10).

Since the elemental matrix $\frac{1}{\text{Re}_m} K_e + R_e$, which when assembled form $\frac{1}{\text{Re}_m} K + R$, is invertible, then the dual element Schur complement preconditioner is

$$
(6.11) \qquad \begin{bmatrix} \frac{1}{\text{Re}_m} K + R & 0 \\ 0 & \text{Re}_m N^T J_e \left(\frac{1}{\text{Re}_m} K_e + R_e \right)^{-1} J_e^T N \end{bmatrix}.
$$

Similarly to the Stokes, the elemental matrix, U_e, of U is singular so we cannot form the primal Schur complement in the same was as (6.11). Thus, in an analogous fashion we introduce a small perturbation, ϵ, and consider the preconditioner:

$$
(6.12) \qquad \begin{bmatrix} \frac{1}{\text{Re}_m} K + L^T D_e^T (U_e + \epsilon Z_e)^{-1} D_e L & 0 \\ 0 & U \end{bmatrix},
$$

where Z_e is the H^1 mass matrix.

Alternatively, consider the $H(\text{curl}, \Omega) \times H(\text{div}, \Omega)$ formulation of the first order Maxwell's equations. From [38, 48], the bilinear form for the constraint matrix, c, the leading block matrix, h, and the natural inner product, d, of $H(\text{div}, \Omega)$ are defined as:

$$
(6.13) \qquad
\begin{aligned}
c(\vec{\tau}, u) &= (\nabla \times \vec{u}, \vec{\tau}) & &\text{where } \vec{u} \in H(\text{curl}, \Omega), \vec{\tau} \in H(\text{div}, \Omega) \\
h(\vec{u}, \vec{v}) &= (\vec{u}, \vec{v}) & &\text{where } \vec{u}, \vec{v} \in H(\text{curl}, \Omega), \\
d(\vec{\tau}, \vec{\sigma}) &= (\nabla \cdot \vec{\tau}, \nabla \cdot \vec{\sigma}) + (\vec{\tau}, \vec{\sigma}) & &\text{where } \vec{\tau}, \vec{\sigma} \in H(\text{div}, \Omega).
\end{aligned}
$$

Thus, the matrix C, H and D are associated with the bilinear forms $c(\vec{\tau}, u)$, $h(\vec{u}, \vec{v})$ and $d(\vec{\tau}, \vec{\sigma})$ from (6.13), respectively. Here the primal space of the first order Maxwell problem is the same as the mixed Maxwell problem we consider in (6.9), and thus, by results in Section 4.4 $K + B^T U^{-1} B$ and $H + C^T D^{-1} C$ are spectrally equivalent. Let C_e, H_e and P_e be the elemental matrices of C, H and D, respectively, then by Lemma 5.3 we consider the preconditioner

$$
(6.14) \qquad \begin{bmatrix} L^T \left(\frac{1}{\text{Re}_m} H_e + C_e^T D_e^{-1} C_e \right) L & 0 \\ 0 & U \end{bmatrix},
$$

since $L^T \left(\frac{1}{\text{Re}_m} H_e + C_e^T D_e^{-1} C_e \right) L$ is spectrally equivalent to $\frac{1}{\text{Re}_m} K + R$, the matrix associated with the natural inner product of $H(\text{curl}, \Omega)$.

As was the case for Stokes, the behavior of the preconditioners (6.11), (6.12) and (6.14) depends on the quality of approximations in Theorems 5.4 and 5.5. Again, we compute the minimum and maximum eigenvalues coming from the eigenvalue problems

$$N^T J_e \left(K_e + R_e \right)^{-1} J_e^T N x = \lambda_1 J (K + R)^{-1} J^T x,$$
$$L^T \left(H_e + C_e^T D_e^{-1} C_e \right) L x = \lambda_2 \left(K + J^T U^{-1} J \right) x,$$
$$L^T \left(K_e + J_e^T \left(U_e + \epsilon Z_e \right)^{-1} J_e \right) L x = \lambda_3 \left(K + J^T U^{-1} J \right) x,$$

and present these in Table 4. For simplicity, we have taken Re_m to be 1. The table shows the minimum and maximum eigenvalues are bounded away from zero and are bounded from above for these generalized eigenvalue problems, and thus, again, the element Schur complements are spectrally equivalent to the corresponding global Schur complement.

# cells	$\hat{S}_1 x = \lambda_1 J(K+R)^{-1} J^T x$		$\hat{S}_2 x = \lambda_2 (K + J^T U^{-1} J) x$		$\hat{S}_3 x = \lambda_3 (K + J^T U^{-1} J) x$	
	$\min_e(\lambda_1)$	$\max_e(\lambda_1)$	$\min_e(\lambda_2)$	$\max_e(\lambda_2)$	$\min_e(\lambda_3)$	$\max_e(\lambda_3)$
8	1.0	1.0	0.473142	1.0	0.911392	1.0
32	1.0	1.0	0.475186	1.0	0.909101	1.0
128	1.0	1.0	0.475714	1.0	0.908287	1.0
512	1.0	1.0	0.475847	1.0	0.908073	1.0

Table 4: Minimum and maximum eigenvalues of the element Schur complements with $S_1 = N^T J_e \left(K_e + R_e \right)^{-1} J_e^T N$, $S_2 = L^T \left(K_e + J_e^T \left(U_e + \epsilon Z_e \right)^{-1} J_e \right) L$ and $S_3 = L^T (H_e + C_e^T D_e^{-1} C_e) L$

Similarly to in the previous section, Figure 3 shows the generalized eigenvalues of the whole system for preconditioners (6.11), (6.12) and (6.14). The eigenvalues are again seen to be bounded away from zero and strongly clustered. Interestingly, the eigenvalues of the system preconditioned with $\mathcal{P}_3$ seem to asymptote to 0.5, suggesting this will perform slightly better than the others if an optimal Krylov method, such as MINRES, is used.

Numerical results. Table 5 gives the timing and iteration results for this example. Similarly with the previous example, the element Schur complement preconditioners exhibit iterations almost identical to the natural norm preconditioner. Interestingly, the iteration results for the primal Schur complement preconditioner in (6.14) are slightly lower than the others, apart from the final mesh level. This might be caused but the clustering of the eigenvalues around a half in Figure 3.

The timings do not scale linearly in this case, which is due to our use of a direct solver to apply the preconditioner. Since we are applying the preconditioner with a direct solver we do not provide three dimensional results for this example.
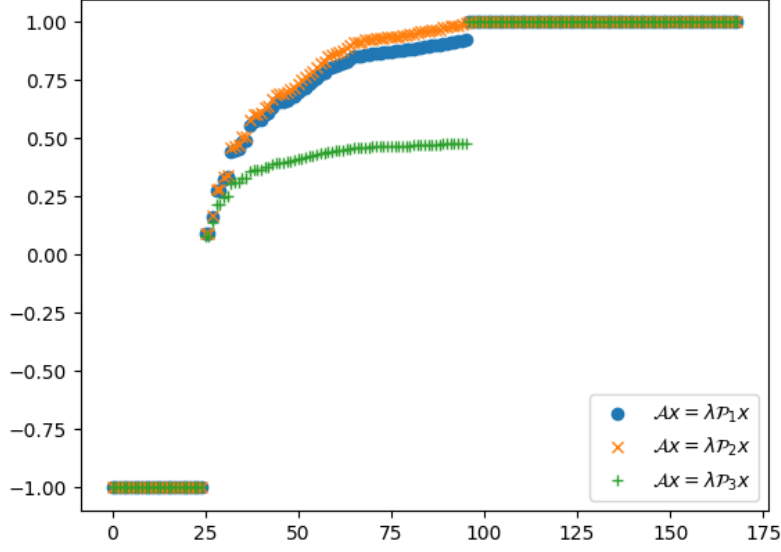
Figure 3: Eigenvalues of generalized eigenvalue problem with total 169 degrees of freedom, where $\mathcal{A}$ is the discretized Maxwell problem and $\mathcal{P}_1$, $\mathcal{P}_2$ and $\mathcal{P}_3$ are given in (6.11), (6.12), and (6.14), respectively.

DoF	Dual precond in (6.11)		Primal precond in (6.14)		Primal-ε precond in (6.12)		Natural norm precond in (6.10)	
	Time	Iteration	Time	Iteration	Time	Iteration	Time	Iteration
16,641	0.6	30	0.7	28	0.6	30	0.8	30
66,049	1.9	33	2.8	28	3.1	33	2.8	33
263,169	6.6	33	7.3	29	8.5	33	7.7	33
1,050,625	27.6	33	29.5	29	31.3	33	31.4	33
4,198,401	116.1	34	107.6	30	124.6	35	119.9	34
16,785,409	606.6	35	619.1	36	619.8	37	600.7	35

Table 5: Timing and iterations results for the mixed Maxwell equations with $\text{Re}_m = 100$ and $\varepsilon = $1e-6.

6.3. The Navier-Stokes equations. Finally, we consider the Navier-Stokes equations, given by

$$-\frac{1}{\text{Re}}\nabla^2\vec{u} + \vec{u}\cdot\nabla\vec{u} + \nabla p = \vec{f}$$
$$\nabla\cdot\vec{u} = 0.$$

The presence of the advection term $(\vec{u} \cdot \nabla \vec{u})$ provides both non-linearity and non-symmetry in the model, which are two key differences between this example and the previous ones. The nonlinear scheme we use is Newton's method with absolute and relative tolerances of 10^{-6} and 10^{-8}, respectively.

A standard mixed formulation of this model is: find $(\mathbf{u}, p) \in H_0^1(\Omega)^d \times L^2(\Omega)$ such that

$$
\begin{aligned}
\frac{1}{\text{Re}}(\nabla \vec{u}, \nabla \vec{v})_\Omega + (\vec{u} \cdot \nabla \vec{u}, \vec{v})_\Omega + (p, \nabla \cdot \vec{v})_\Omega &= (\vec{f}, \vec{v})_\Omega, \quad & \forall \vec{v} \in H_0^1(\Omega)^d, \\
(q, \nabla \cdot \vec{u})_\Omega &= 0, \quad & \forall q \in L^2(\Omega).
\end{aligned}
$$

Upon discretization and linearization, we obtain a saddle point system of the form (1.2) with A and B given by $\frac{1}{\text{Re}}G + O$, where G is the vector Laplacian, O is the advection matrix, and B is the divergence matrix.

Due to their non-symmetric nature, the discrete Navier Stokes equations do not fit into the framework for natural norm preconditioners, thus we cannot apply the theory in Sections 4 and 5. However, we can construct the element Schur complement preconditioner, and we include this method to show that such a preconditioner can be applied even if the natural norm preconditioner is not available.

For nonsymmetric systems GMRES [51] is often used, but in this case, as shown by Greenbaum, Pták and Strakoš [21], clustered eigenvalues are not necessarily enough to guarantee rapid convergence. However, we can still form local Schur complements and use these as a heuristic preconditioner. We compare this approach with the Pressure Convection Diffusion (PCD) preconditioner of Kay, Loghin and Wathen [30]. The implementation of PCD that we use in our tests is based the version bundled with Firedrake, which we modified slightly to give better performance.

For this example we will use the full $L^T D L$ decomposition:

$$
P^{-1} = \begin{bmatrix} I & -\hat{A}^{-1}B^T \\ 0 & I \end{bmatrix} \begin{bmatrix} \hat{A}^{-1} & 0 \\ 0 & \hat{S}^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -B\hat{A}^{-1} & I \end{bmatrix}.
$$

where $\hat{A}^{-1}$ and $\hat{S}^{-1}$ are approximations to the inverse of the advection diffusion operator and the Schur complement, respectively. In Table 6 we present the definitions of $\hat{A}$ and $\hat{S}$ for the four different preconditioners we will test. The action of the inverse of $\hat{A}$ and $\hat{S}$ is done via a multigrid V-cycle. For PCD, $\hat{S}^{-1} = -U^{-1}F_p Q^{-1}$, where again we use a multigrid V-cycle to approximate the inverses. We note that for the Primal$_{\text{alt}}$ preconditioner, the inverse of the approximate Schur complement is the sum of two matrix inverses. This preconditioner is based on the work of Benzi and Olshanskii [8].

Tables 7 and 8 show timing and iteration results for a two and three dimensional cavity driven flow (described in Elman, Silvester and Wathen [17, Example 8.1.2]) and flow over a two-dimensional backwards facing step (described in Elman, Silvester and Wathen [17, Example 8.1.3]), respectively. The backwards facing step is more complicated due to both the nonconvex domain and the natural outflow conditions on the rightmost boundary.

Figure 4 shows the iteration and timing results for both the PCD and Primal$_{\text{alt}}$ preconditioners in three dimensions whilst varying Reynolds number on a fixed mesh. Neither preconditioner exhibits Reynolds number independent iterations and/or timings, however, the PCD preconditioner fails to converge from Reynolds numbers higher than 380 whilst the Primal$_{\text{alt}}$

Preconditioner	Approx matrix
Dual	$\hat{A} = G + O$ $\hat{S} = -\mathrm{Re}N^T B_e(G_e + O_e + \epsilon T_e)^{-1}B_e^T N$
Primal	$\hat{A} = G + O + \frac{1}{\mathrm{Re}}LB_e^T Q_e^{-1}B_e L^T$ $\hat{S} = -\mathrm{Re}Q$
Primal$_\mathrm{alt}$	$\hat{A} = G + O + \frac{1}{\mathrm{Re}}LB_e^T Q_e^{-1}B_e L^T$ $\hat{S}^{-1} = -\frac{1}{\mathrm{Re}}\left((N^T B_e(G_e + O_e + \epsilon T_e)^{-1}B_e^T N)^{-1} + Q^{-1}\right)$
PCD	$\hat{A} = G + O$ $\hat{S} = -QF_p^{-1}U$

Table 6: $\hat{A}$ and $\hat{S}$ for the four different preconditioners.

	Dual precond		Primal precond		Primal$_\mathrm{alt}$ precond		PCD precond	
DoF	Time	Iteration	Time	Iteration	Time	Iteration	Time	Iteration
9,539	1.7	4/63.2	1.3	4/48.0	1.1	4/36.2	1.1	4/31.8
37,507	8.8	4/65.0	7.8	4/53.2	6.2	4/38.5	6.3	4/34.2
148,739	33.8	4/68.2	29.5	4/55.8	23.6	4/40.5	20.2	4/34.8
592,387	129.4	4/73.8	112.0	4/58.0	90.5	4/42.2	69.4	4/36.0
2,364,419	525.0	4/77.5	427.7	4/58.0	346.8	4/43.5	270.9	4/36.8
9,447,427	1635.6	3/79.0	1614.8	3/62.0	1303.4	3/45.3	994.2	3/40.0

(a) Two dimensional results

	Dual precond		Primal precond		Primal$_\mathrm{alt}$ precond		PCD precond	
DoF	Time	Iteration	Time	Iteration	Time	Iteration	Time	Iteration
402	-	-/-	0.5	5/55.2	9.4	5/46.6	0.5	5/48.6
2,312	1.7	4/80.8	1.4	4/61.5	1.4	4/50.2	1.4	4/59.2
15,468	18.5	4/63.8	21.4	4/59.5	17.3	4/46.5	18.6	4/68.0
112,724	170.3	4/71.5	217.1	4/61.0	174.3	4/47.2	159.9	4/68.0
859,812	1791.5	4/82.5	2106.0	4/64.5	1748.2	4/52.0	1434.2	4/64.8

(b) Three dimensional results

Table 7: Iteration and timing results for the two and three dimensional Navier-Stokes *leaky* cavity driven flow problem with Re = 100 and ε =1e-6.

preconditioner continues to converge to the solution for Reynolds number higher than 400. The PCD preconditioner iterations are significantly higher than the slightly more expensive Primal$_\mathrm{alt}$ preconditioner for all Reynolds numbers. The PCD preconditioner timing results are lower than the Primal$_\mathrm{alt}$ preconditioner until a Reynolds number of 320 but with the larger iteration count the timings for the PCD preconditioner grow faster than the Primal$_\mathrm{alt}$ precon-ditioner, and for the larger Reynolds numbers the PCD preconditioner takes longer to converge
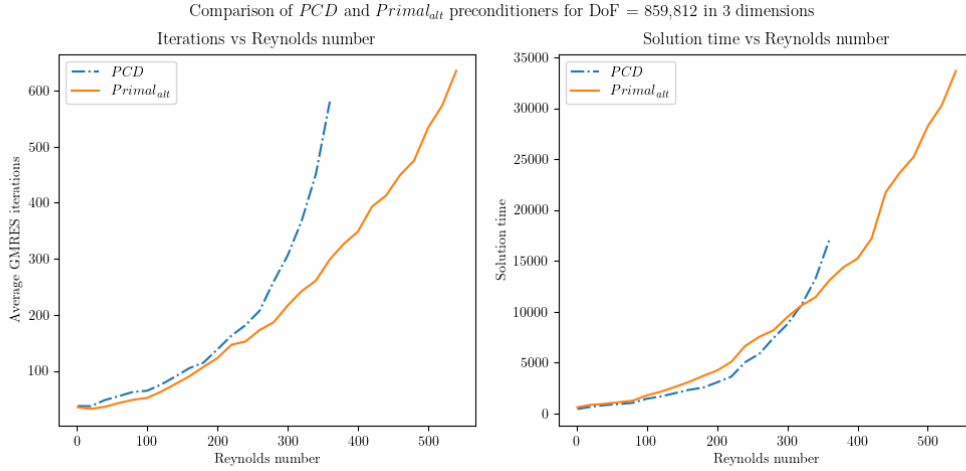
to the solution.

Figure 4: Comparison of the `PCD` and `Primal`_{alt} preconditioner whilst varying the Reynolds number for the three-dimensional cavity driven flow problem with DoF = 859,812.

Note, from the table we scale the `Primal` and `Primal`_{alt} preconditioners with $\mathrm{Re}Q_e$ rather than Q_e. Including the Reynolds number in the scaling is the natural choice, since $\mathrm{Re}Q_e$ is the appropriate matrix associated with norm of the dual space, and also the scaling of the Reynolds number arises in a similar fashion for the dual Schur complement. In fact, not including the Reynolds number scaling would cause the preconditioned iterations to degrade significantly and to not be mesh independent.

From Table 7, we see that for both the two and three dimensional results all the preconditioners exhibit near mesh independent iterations. It appears that all the preconditioned iterations increase slightly as the mesh is refined. For the two dimensional results, the `PCD` preconditioner performs best with the `Primal`_{alt} preconditioner differing by just a few iterations. This is to be expected since the `PCD` preconditioner is designed specifically for this problem, however it is interesting to note that it is possible to achieve similar iteration numbers based solely on generic element Schur complement preconditioners. The timings for the two dimensional results scale linearly with the problem size. For the three dimensional results, the `Primal`_{alt} preconditioner performs best in terms of iteration numbers, however, the `PCD` preconditioner performs best with respect to time. Similarly to the Stokes problem, the timing differences are likely due to the higher cost associated with the algebraic multigrid solver for the leading block.

For the backwards facing step, Table 8, the `Firedrake` implementation of the `PCD` preconditioner does not converge, and so we only report the element Schur complement results. While the iterations still grow with the mesh size, this growth appears to be stable. Again, the timings scale linearly with the problem size.

We remark that by carefully considering the boundary conditions, the `PCD` preconditioner can be modified to yield scalable iterations for outflow problems (see, for example, Elman,

| | Dual precond | | Primal precond | | Primal$_{alt}$ precond | |
DoF	Time	Iteration	Time	Iteration	Time	Iteration
24,498	7.1	3/79.3	5.1	3/54.0	12.6	3/48.7
35,885	10.8	3/82.7	8.0	3/55.0	7.2	3/49.7
91,950	31.3	3/97.3	22.2	3/63.7	18.3	3/50.7
133,092	39.1	3/87.3	30.9	3/63.3	25.6	3/51.7
271,286	93.1	3/103.3	65.5	3/65.7	53.9	3/53.3
599,892	237.2	3/117.3	152.0	3/69.7	128.7	3/58.7
2,455,263	975.0	3/119.0	646.7	3/74.7	578.4	3/65.7

Table 8: Results for preconditioning GMRES for the Navier-Stokes equations with Re = 10 for the backwards facing step problem.

Silvester and Wathen [17, Section 9.2.2] and Bootland [10, Chapter 5]), but these are not yet implemented in `Firedrake`. We speculate that similar considerations would also improve the performance of the element Schur complement approach considered here; this is beyond the scope of this paper.

7. Conclusion and outlook. In this work, we presented effective sparse Schur complement approximations for inf-sup stable mixed finite element discretizations of self-adjoint problems. Our new preconditioner is based on forming local, or element, Schur complements of an appropriate system which is elementally inf-sup stable and shares either a primal or dual space with the original system. By utilizing the existence of matrices associated with the natural inner products of the underlying function spaces, we show spectral equivalence between the Schur complement of the original system and the matrix formed by projecting the local Schur complements back to the global degrees of freedom.

From the numerical results, the element based Schur complement preconditioners perform similarly to the natural norm preconditioners for the self-adjoint multi-physics problems we have considered. While our theory suggests that we should not expect the element Schur complement preconditioner to be better, in terms of iteration count, than a natural norm preconditioner, its construction requires less knowledge of the underlying mathematical structure, and can be used as a *black-box* technique for implementing optimal block preconditioners. We also envisage this technique will more widely be applicable in cases where the natural norm preconditioner is not known. For example, in the nonsymmetric and nonlinear Navier-Stokes example, where there is no concept of a natural norm preconditioner, our new Schur complement approximation performs similarly to the well-known PCD preconditioner in our tests, while being somewhat more robust with respect to the Reynolds number in some cases. One possible area of future work would be to extend the theory to cover nonsymmetric systems. For example, one might be able to consider a field-of-values equivalence which could give rise to theory in the nonsymmetric cases, see Loghin and Wathen in the context of general saddle point systems [35], Ma, Hu, Hu, and Xu [36] in the context of magnetohydrodynamics preconditioners, or Benzi and Olshanskii [9] in the context of linearized Navier-Stokes equations.

We only consider single element contributions, but grouping multiple local elements to-

gether to form slightly larger local Schur complements may lead to a better approximation, requiring only marginally more storage.

We focus on problems from mixed finite elements, but the ideas should work for any partially separable functions which can be represented as a sum of element functions (see Daydé, L'Excellent and Gould [16]). This may allow us use similar ideas to build fast solvers for problems that do not come from PDEs, and where the natural norms are not obvious. The extension of element preconditioners to these problems is left for future work.

Code availability. Together with the written manuscript, we provide the code used to generate the results [65].

References.

[1] M. S. ALNÆS, J. BLECHTA, J. HAKE, A. JOHANSSON, B. KEHLET, A. LOGG, C. RICHARDSON, J. RING, M. E. ROGNES, AND G. N. WELLS, *The FEniCS Project Version 1.5*, Archive of Numerical Software, 3 (2015), https://doi.org/10.11588/ans.2015.100.20553.

[2] P. R. AMESTOY, I. S. DUFF, J. KOSTER, AND J.-Y. L'EXCELLENT, *A Fully Asynchronous Multifrontal Solver Using Distributed Dynamic Scheduling*, SIAM Journal on Matrix Analysis and Applications, 23 (2001), pp. 15–41, https://doi.org/10.1137/S0895479899358194.

[3] P. R. AMESTOY, A. GUERMOUCHE, J.-Y. L'EXCELLENT, AND S. PRALET, *Hybrid scheduling for the parallel solution of linear systems*, Parallel Computing, 32 (2006), pp. 136–156, https://doi.org/10.1016/j.parco.2005.07.004.

[4] O. AXELSSON, R. BLAHETA, AND M. NEYTCHEVA, *Preconditioning of boundary value problems using elementwise schur complements*, SIAM Journal on Matrix Analysis and Applications, 31 (2009), pp. 767–789, https://doi.org/10.1137/070679673.

[5] S. BALAY, S. ABHYANKAR, M. F. ADAMS, J. BROWN, P. BRUNE, K. BUSCHELMAN, L. DALCIN, A. DENER, V. EIJKHOUT, W. D. GROPP, D. KARPEYEV, D. KAUSHIK, M. G. KNEPLEY, D. A. MAY, L. C. MCINNES, R. T. MILLS, T. MUNSON, K. RUPP, P. SANAN, B. F. SMITH, S. ZAMPINI, H. ZHANG, AND H. ZHANG, *PETSc users manual*, Tech. Report ANL-95/11 - Revision 3.13, Argonne National Laboratory, 2020, https://www.mcs.anl.gov/petsc.

[6] S. BALAY, W. D. GROPP, L. C. MCINNES, AND B. F. SMITH, *Efficient management of parallelism in object oriented numerical software libraries*, in Modern Software Tools in Scientific Computing, E. Arge, A. M. Bruaset, and H. P. Langtangen, eds., Birkhäuser Press, 1997, pp. 163–202, https://doi.org/10.1007/978-1-4612-1986-6_8.

[7] M. BENZI, G. H. GOLUB, AND J. LIESEN, *Numerical solution of saddle point problems*, Acta Numer., 14 (2005), p. 1137, https://doi.org/10.1017/S0962492904000212.

[8] M. BENZI AND M. A. OLSHANSKII, *An augmented Lagrangian-based approach to the Oseen problem*, SIAM Journal on Scientific Computing, 28 (2006), pp. 2095–2113, https://doi.org/10.1137/050646421.

[9] M. BENZI AND M. A. OLSHANSKII, *Field-of-Values Convergence Analysis of Augmented Lagrangian Preconditioners for the Linearized Navier–Stokes Problem*, SIAM Journal on Numerical Analysis, 49 (2011), pp. 770–788, https://doi.org/10.1137/100806485.

[10] N. Bootland, *Scalable two-phase flow solvers*, PhD thesis, University of Oxford, 2018, https://ora.ox.ac.uk/objects/uuid:d315682a-4308-4b2a-8deb-0b2f7cb4bfc3.

[11] D. Braess, *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, Cambridge University Press, 3rd ed., 2007, https://doi.org/10.1017/CBO9780511618635.

[12] F. Brezzi, *On the existence, uniqueness and approximation of saddle-point problems arising from lagrangian multipliers*, R.A.I.R.O. Analyse Numérique, 8 (1974), pp. 129–151, http://www.numdam.org/item/M2AN_1974__8_2_129_0.

[13] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, Berlin, Heidelberg, 1991, https://doi.org/10.1007/978-1-4612-3172-1.

[14] L. C. Cowsar, J. Mandel, and M. F. Wheeler, *Balancing Domain Decomposition for Mixed Finite Elements*, Mathematics of Computation, 64 (1995), pp. 989–1015, https://doi.org/10.2307/2153480.

[15] L. D. Dalcin, R. R. Paz, P. A. Kler, and A. Cosimo, *Parallel distributed computing using Python*, Advances in Water Resources, 34 (2011), pp. 1124–1139, https://doi.org/10.1016/j.advwatres.2011.04.013.

[16] M. J. Daydé, J.-Y. L'Excellent, and N. I. M. Gould, *Element-by-Element Preconditioners for Large Partially Separable Optimization Problems*, SIAM J. Sci. Comput., 18 (1997), pp. 1767–1787, https://doi.org/10.1137/S1064827594274796.

[17] H. Elman, D. Silvester, and A. Wathen, *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, Numerical mathematics and scientific computation, Oxford University Press, 2014, https://doi.org/10.1093/acprof:oso/9780199678792.001.0001.

[18] R. D. Falgout and U. M. Yang, *hypre: A Library of High Performance Preconditioners*, in Computational Science — ICCS 2002, P. M. A. Sloot, A. G. Hoekstra, C. J. K. Tan, and J. J. Dongarra, eds., Berlin, Heidelberg, 2002, Springer Berlin Heidelberg, pp. 632–641, https://doi.org/10.1007/3-540-47789-6_66.

[19] P. E. Farrell, M. G. Knepley, F. Wechsung, and L. Mitchell, *PCPATCH: software for the topological construction of multigrid relaxation methods*, 2019, http://arxiv.org/abs/1912.08516, https://arxiv.org/abs/1912.08516.

[20] P. E. Farrell, L. Mitchell, and F. Wechsung, *An Augmented Lagrangian Preconditioner for the 3D Stationary Incompressible Navier–Stokes Equations at High Reynolds Number*, SIAM Journal on Scientific Computing, 41 (2019), pp. A3073–A3096, https://doi.org/10.1137/18m1219370.

[21] A. Greenbaum, V. Pták, and Z. Strakoš, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 465–469, https://doi.org/10.1137/S0895479894275030.

[22] C. Greif and D. Schötzau, *Preconditioners for saddle point linear systems with highly singular (1, 1) blocks*, ETNA, Special Volume on Saddle Point Problems, 22 (2006), pp. 114–121, http://etna.mcs.kent.edu/volumes/2001-2010/vol22/abstract.php?vol=22&pages=114-121.

[23] C. Greif and D. Schötzau, *Preconditioners for the discretized time-harmonic maxwell equations in mixed form*, Numerical Linear Algebra with Applications, 14 (2007), pp. 281–297, https://doi.org/10.1002/nla.515.

[24] C. Greif and M. Wathen, *Conjugate gradient for nonsingular saddle-point systems*

with a maximally rank-deficient leading block, Journal of Computational and Applied Mathematics, 358 (2019), pp. 1 – 11, https://doi.org/10.1016/j.cam.2019.02.016.

[25] I. Gustafsson and G. Lindskog, *A preconditioning technique based on element matrix factorizations*, Computer methods in applied mechanics and engineering, 55 (1986), pp. 201–220, https://doi.org/10.1016/0045-7825(86)90053-8.

[26] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Stand., 49 (1952), pp. 409–436, https://doi.org/10.6028/jres.049.044.

[27] R. Hiptmair and J. Xu, *Nodal Auxiliary Space Preconditioning in H(curl) and H(div) Spaces*, SIAM Journal on Numerical Analysis, 45 (2007), pp. 2483–2509, https://doi.org/10.1137/060660588.

[28] T. J. Hughes, I. Levit, and J. Winget, *An element-by-element solution algorithm for problems of structural and solid mechanics*, Computer Methods in Applied Mechanics and Engineering, 36 (1983), pp. 241–254, https://doi.org/10.1016/0045-7825(83)90115-9.

[29] I. C. F. Ipsen, *A note on preconditioning nonsymmetric matrices*, SIAM Journal on Scientific Computing, 23 (2001), pp. 1050–1051, https://doi.org/10.1137/S1064827599355153.

[30] D. Kay, D. Loghin, and A. Wathen, *A preconditioner for the steady-state navier–stokes equations*, SIAM Journal on Scientific Computing, 24 (2002), pp. 237–256, https://doi.org/10.1137/S106482759935808X.

[31] J. Kraus, *Algebraic multilevel preconditioning of finite element matrices using local schur complements*, Numerical linear algebra with applications, 13 (2006), pp. 49–70, https://doi.org/10.1002/nla.462.

[32] J. Kraus, *Additive Schur Complement Approximation and Application to Multilevel Preconditioning*, SIAM Journal on Scientific Computing, 34 (2012), pp. A2872–A2895, https://doi.org/10.1137/110845082.

[33] J. Kraus, M. Lymbery, and S. Margenov, *Auxiliary space multigrid method based on additive Schur complement approximation*, Numerical Linear Algebra with Applications, 22 (2015), pp. 965–986, https://doi.org/10.1002/nla.1959.

[34] A. Logg, K.-A. Mardal, G. N. Wells, et al., *Automated Solution of Differential Equations by the Finite Element Method*, Springer, 2012, https://doi.org/10.1007/978-3-642-23099-8.

[35] D. Loghin and A. J. Wathen, *Analysis of Preconditioners for Saddle-Point Problems*, SIAM Journal on Scientific Computing, 25 (2004), pp. 2029–2049, https://doi.org/10.1137/S1064827502418203.

[36] Y. Ma, K. Hu, X. Hu, and J. Xu, *Robust preconditioners for incompressible MHD models*, Journal of Computational Physics, 316 (2016), pp. 721 – 746, https://doi.org/10.1016/j.jcp.2016.04.019.

[37] K.-A. Mardal and R. Winther, *Preconditioning discretizations of systems of partial differential equations*, Numerical Linear Algebra with Applications, 18 (2011), pp. 1–40, https://doi.org/10.1002/nla.716.

[38] P. Monk, *Finite Element Methods for Maxwell's Equations*, Numerical Mathematics and Scientific Computation, Clarendon Press, 2003, https://doi.org/10.1093/acprof:oso/9780198508885.001.0001.

[39] M. F. Murphy, G. H. Golub, and A. J. Wathen, *A note on preconditioning for indefinite linear systems*, SIAM J. Sci. Comput., 21 (2000), pp. 1969–1972, https://doi.org/10.1137/S1064827599355153.

[40] J. Mlek and Z. Strako, *Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2014, https://doi.org/10.1137/1.9781611973846.

[41] M. Neytcheva, *On element-by-element schur complement approximations*, Linear Algebra and Its Applications, 434 (2011), pp. 2308–2324, https://doi.org/10.1016/j.laa.2010.03.031.

[42] M. Neytcheva and E. Bängtsson, *Preconditioning of nonsymmetric saddle point systems as arising in modelling of viscoelastic problems*, Electronic Transactions on Numerical Analysis, 29 (2008), pp. 193–211, http://eudml.org/doc/130827.

[43] M. Neytcheva, M. Do-Quang, and H. Xin, *Element-by-element schur complement approximations for general nonsymmetric matrices of two-by-two block form*, in International Conference on Large-Scale Scientific Computing, Springer, 2009, pp. 108–115, https://doi.org/10.1007/978-3-642-12535-5_11.

[44] B. Nour-Omid and B. Parlett, *Element preconditioning using splitting techniques*, SIAM journal on scientific and statistical computing, 6 (1985), pp. 761–770, https://doi.org/10.1137/0906051.

[45] M. Ortiz, P. M. Pinsky, and R. L. Taylor, *Unconditionally stable element-by-element algorithms for dynamic problems*, Computer Methods in Applied Mechanics and Engineering, 36 (1983), pp. 223–239, https://doi.org/10.1016/0045-7825(83)90114-7.

[46] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629, https://doi.org/10.1137/0712047, http://link.aip.org/link/?SNA/12/617/1.

[47] J. Pestana and A. J. Wathen, *Natural preconditioning and iterative methods for saddle point systems*, SIAM Review, 57 (2015), pp. 71–91, https://doi.org/10.1137/130934921.

[48] E. G. Phillips, J. N. Shadid, and E. C. Cyr, *Scalable Preconditioners for Structure Preserving Discretizations of Maxwell Equations in First Order Form*, SIAM Journal on Scientific Computing, 40 (2018), pp. B723–B742, https://doi.org/10.1137/17M1135827.

[49] C. E. Powell and D. Silvester, *Optimal preconditioning for Raviart–Thomas mixed formulation of second-order elliptic problems*, SIAM journal on matrix analysis and applications, 25 (2003), pp. 718–738, https://doi.org/10.1137/S0895479802404428.

[50] F. Rathgeber, D. A. Ham, L. Mitchell, M. Lange, F. Luporini, A. T. T. McRae, G.-T. Bercea, G. R. Markall, and P. H. J. Kelly, *Firedrake: automating the finite element method by composing abstractions*, ACM Trans. Math. Softw., 43 (2016), pp. 24:1–24:27, https://doi.org/10.1145/2998441.

[51] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869, https://doi.org/10.1137/0907058.

[52] J. Schöberl, *Multigrid methods for a parameter dependent problem in primal variables*, Numerische Mathematik, 84 (1999), pp. 97–119, https://doi.org/10.1007/s002110050465.

[53] D. Silvester and A. Wathen, *Fast iterative solution of stabilised Stokes systems Part*

II: Using general block preconditioners, SIAM J. Numer. Anal., 31 (1994), pp. 1352–1367, https://doi.org/10.1137/0731070.

[54] N. Spillane, V. Dolean, P. Hauret, F. Nataf, C. Pechstein, and R. Scheichl, *A robust two-level domain decomposition preconditioner for systems of PDEs*, Comptes Rendus Mathematique, 349 (2011), pp. 1255 – 1259, https://doi.org/10.1016/j.crma.2011.10.021.

[55] N. Spillane and D. Rixen, *Automatic spectral coarse spaces for robust finite element tearing and interconnecting and balanced domain decomposition algorithms*, International Journal for Numerical Methods in Engineering, 95 (2013), pp. 953–990, https://doi.org/10.1002/nme.4534.

[56] C. Taylor and P. Hood, *A numerical solution of the Navier-Stokes equations using the finite element technique*, Computers & Fluids, 1 (1973), pp. 73 – 100, https://doi.org/10.1016/0045-7930(73)90027-3.

[57] R. Temam, *Navier–Stokes Equations*, American Mathematical Soc., 1984, https://www.elsevier.com/books/navier-stokes-equations/temam/978-0-444-85307-3.

[58] A. Toselli and O. Widlund, *Domain Decomposition Methods - Algorithms and Theory*, Springer Science & Business Media, June 2006, https://doi.org/10.1007/b137868.

[59] S. Turek, *Efficient Solvers for Incompressible Flow Problems - An Algorithmic and Computational Approach*, in Lecture Notes in Computational Science and Engineering, 1999, https://doi.org/10.1007/978-3-642-58393-3.

[60] M. Van Gijzen, *An analysis of element-by-element preconditioners for nonsymmetric problems*, Computer methods in applied mechanics and engineering, 105 (1993), pp. 23–40, https://doi.org/10.1016/0045-7825(93)90114-D.

[61] S. Vanka, *Block-implicit multigrid solution of Navier-Stokes equations in primitive variables*, Journal of Computational Physics, 65 (1986), pp. 138–158, https://doi.org/10.1016/0021-9991(86)90008-2.

[62] A. Wathen, *An analysis of some element-by-element techniques*, Computer Methods in Applied Mechanics and Engineering, 74 (1989), pp. 271–287, https://doi.org/10.1016/0045-7825(89)90052-2.

[63] A. Wathen and D. Silvester, *Fast iterative solution of stabilised Stokes systems. part I: Using simple diagonal preconditioners*, SIAM J. Numer. Anal., 30 (1993), pp. 630–649, https://doi.org/10.1137/0730031.

[64] A. J. Wathen, *Preconditioning*, Acta Numerica, 24 (2015), pp. 329–376, https://doi.org/10.1017/S0962492915000021.

[65] M. Wathen and T. Rees, *ralna/elementschur: v1.1*, 2020, https://doi.org/10.5281/zenodo.3801732.

[66] W. Zulehner, *Nonstandard Norms and Robust Estimates for Saddle Point Problems*, SIAM Journal on Matrix Analysis and Applications, 32 (2011), pp. 536–560, https://doi.org/10.1137/100814767.