SADDLE POINT LEAST SQUARES PRECONDITIONING OF MIXED METHODS

CONSTANTIN BACUTA AND JACOB JACAVAGE

ABSTRACT. We present a simple way to discretize and precondition mixed variational formulations. Our theory connects with, and takes advantage of, the classical theory of symmetric saddle point problems and the theory of preconditioning symmetric positive definite operators. Efficient iterative processes for solving the discrete mixed formulations are proposed and choices for discrete spaces that are always compatible are provided. For the proposed discrete spaces and solvers, a basis is needed only for the test spaces and assembly of a global saddle point system is avoided. We prove sharp approximation properties for the discretization and iteration errors and also provide a sharp estimate for the convergence rate of the proposed algorithm in terms of the condition number of the elliptic preconditioner and the discrete $\inf - \sup$ and $\sup - \sup$ constants of the pair of discrete spaces.

1. INTRODUCTION

We provide a general approach in preconditioning mixed problems of the form: Given $f \in V^*$, find $p \in Q$ such that

$$b(v,p) = \langle f, v\rangle \quad \text{for all } v \in V, \tag{1.1}$$

where V and Q are Hilbert spaces and $b(\cdot,\cdot)$ is a continuous bilinear form on $V \times Q$ satisfying an $\inf - \sup$ condition. In [8, 15], a connection was made between problems of the form (1.1) and a natural saddle point formulation. More specifically, if $a(\cdot,\cdot)$ is the inner product on V, then p is the unique solution of (1.1) if and only if $(u = 0, p)$ is the unique solution to: Find $(u,p) \in V \times Q$ such that

$$\begin{array}{lllll} a(u,v) & + & b(v,p) & = \langle f, v\rangle & \text{for all } v \in V, \\ b(u,q) & & & = 0 & \text{for all } q \in Q, \end{array} \tag{1.2}$$

where appropriate assumptions on f and the form $b(\cdot,\cdot)$ hold, see Section 2.1. Thus, (1.2) is a saddle point reformulation of (1.1). It is clear that the solution component p of this reformulation is independent of the inner product (hence the norm) considered on V. This observation is essential for the discretization and preconditioning of (1.1).

2000 *Mathematics Subject Classification.* 74S05, 74B05, 65N22, 65N55.

Key words and phrases. least squares, saddle point systems, mixed methods, multilevel methods, conjugate gradient, preconditioning.

The work was supported by NSF, DMS-1522454.

For finite dimensional approximation spaces $V_h \subset V$ and $\mathcal{M}_h \subset Q$ satisfying a discrete $\inf - \sup$ condition, we consider the discrete problem of finding $(u_h, p_h) \in V_h \times \mathcal{M}_h$ such that

$$\begin{array}{lllll} a(u_h, v_h) & + & b(v_h, p_h) & = \langle f, v_h \rangle & \text{for all } v_h \in V_h, \\ b(u_h, q_h) & & & = 0 & \text{for all } q_h \in \mathcal{M}_h, \end{array} \tag{1.3}$$

which approximates the solution $(u = 0, p)$ of (1.2). The discrete variational formulation (1.3) is in fact a *saddle point least squares discretization* of (1.1), see Section 2.2. This saddle point discretization of (1.1) is also adapted by Demkowicz and Gopalakrishnan in [16, 17]. When solving the above problem, finding bases for the discrete trial space $\mathcal{M}_h$ and assembling a block stiffness matrix for (1.3) can be avoided by applying an Uzawa type algorithm. However, any attempt to solve (1.3) by an Uzawa iterative process requires the exact inversion of the operator A_h associated with the inner product $a(\cdot, \cdot)$ on V_h. To avoid exact inversion and to speed up the iterative solvers, we consider another form $\tilde{a}(\cdot, \cdot)$ on V_h, which leads to an equivalent norm on V_h, and introduce a preconditioned discrete saddle point problem: Find $(\tilde{u}_h, \tilde{p}_h) \in V_h \times \mathcal{M}_h$ such that

$$\begin{array}{lllll} \tilde{a}(\tilde{u}_h, v_h) & + & b(v_h, \tilde{p}_h) & = \langle f, v_h \rangle & \text{for all } v_h \in V_h, \\ b(\tilde{u}_h, q_h) & & & = 0 & \text{for all } q_h \in \mathcal{M}_h, \end{array} \tag{1.4}$$

where the action of the operator $\tilde{A}_h^{-1}$ associated with the inner product $\tilde{a}(\cdot, \cdot)$ on V_h is assumed to be fast and easy to implement.

The goals of this paper are: describe how well the component solution $\tilde{p}_h$ of (1.4) approximates the solution p of (1.1), describe possible choices for the discrete pairs $(V_h, \mathcal{M}_h)$, and propose an efficient iterative solver for (1.4) and estimate its convergence rate.

The paper is organized as follows. In Section 2, an abstract theory for saddle point least squares formulations is presented. Section 3 describes the general preconditioning theory and approximation results. In addition, convergence rates for the proposed iterative solver are estimated. Possible choices for discrete pairs of spaces are discussed in Section 4. Numerical results are presented in Section 5. Advantages of the proposed method are also discussed.

2. Abstract Saddle Point Least Squares Formulation for Mixed Methods

2.1. **Notation and the continuous problem.** Let V and Q be infinite dimensional Hilbert spaces and assume the inner products $a(\cdot, \cdot)$ and $(\cdot, \cdot)$ induce the norms $|\cdot|_V = |\cdot| = a(\cdot, \cdot)^{1/2}$ and $\|\cdot\|_Q = \|\cdot\| = (\cdot, \cdot)^{1/2}$. The duals of V and Q will be denoted by V^* and Q^*, respectively. The dual pairings on $V^* \times V$ and $Q^* \times Q$ will both be denoted by $\langle \cdot, \cdot \rangle$. With the inner products $a(\cdot, \cdot)$ and $(\cdot, \cdot)$, we associate the operators $\mathcal{A} : V \to V^*$ and $\mathcal{C} : Q \to Q^*$ defined by

$$\langle \mathcal{A}u, v \rangle = a(u, v) \quad \text{for all } u, v \in V,$$

and

$$\langle \mathcal{C}p, q\rangle = (p, q) \quad \text{for all} \;\; p, q \in Q.$$

Assume that $b(\cdot,\cdot)$ is a continuous bilinear form on $V \times Q$ satisfying the $\inf - \sup$ condition

$$\inf_{p\in Q} \sup_{v\in V} \frac{b(v,p)}{|v|\,\|p\|} = m > 0, \tag{2.1}$$

and is bounded, i.e.

$$\sup_{p\in Q} \sup_{v\in V} \frac{b(v,p)}{|v|\,\|p\|} = M < \infty. \tag{2.2}$$

With the form $b(\cdot,\cdot)$, we associate the linear operators $B : V \to Q^*$ and $B^* : Q \to V^*$ defined through the duality pairings

$$\langle Bv, q\rangle = b(v, q) = \langle B^*q, v\rangle \quad \text{for all } v \in V,\; q \in Q.$$

It is well known that if, in addition to the assumptions on $b(\cdot,\cdot)$, f satisfies the compatibility condition

$$\langle f, v\rangle = 0 \quad \text{for all } v \in V_0 := \{v \in V \mid b(v, q) = 0, \text{ for all } q \in Q\}, \tag{2.3}$$

then (1.1) has a unique solution p, see e.g. [1, 2]. Furthermore, $(u = 0, p)$ is the unique solution of (1.2).

Remark 2.1. *The saddle point problem* (1.2) *has a unique solution* (u, p) *regardless of the compatibility condition* (2.3)*. The operator form of problem* (1.1) *is equivalent to finding* $p \in Q$ *such that*

$$\mathcal{A}^{-1}B^*p = \mathcal{A}^{-1}f,$$

and solving for p *from* (1.2) *gives*

$$(\mathcal{C}^{-1}B)(\mathcal{A}^{-1}B^*)p = (\mathcal{C}^{-1}B)\mathcal{A}^{-1}f.$$

Since $\mathcal{C}^{-1}B$ *is the Hilbert transpose of* $\mathcal{A}^{-1}B^*$*, we have that the* p *component of the solution of* (1.2) *is the least squares solution of* (1.1)*.*

For the rest of this paper, we assume that the compatibility condition (2.3) holds. Consequently, problem (1.1) has a unique solution.

2.2. **Saddle point least squares discretization.** Let $V_h \subset V$ and $\mathcal{M}_h \subset Q$ be finite dimensional approximation spaces and A_h be the discrete version of the operator $\mathcal{A}$, i.e. A_h satisfies

$$\langle A_h u_h, v_h\rangle = a(u_h, v_h) \quad \text{for all } u_h, v_h \in V_h.$$

We define the discrete operators $B_h : V_h \to \mathcal{M}_h$ and $B_h^* : \mathcal{M}_h \to V_h^*$ by

$$(B_h v_h, q_h) = b(v_h, q_h) = \langle B_h^* q_h, v_h\rangle \quad \text{for all } v_h \in V_h, q_h \in \mathcal{M}_h.$$

Note that the operator B_h is defined using the inner product on $\mathcal{M}_h$ and not with the duality on $\mathcal{M}_h^* \times \mathcal{M}_h$. Thus, we can define the discrete Schur complement $S_h : \mathcal{M}_h \to \mathcal{M}_h$ as $S_h = B_h\, A_h^{-1} B_h^*$. We further assume the following discrete $\inf - \sup$ condition holds for the pair of spaces $(V_h, \mathcal{M}_h)$:

$$\inf_{p_h \in \mathcal{M}_h} \sup_{v_h \in V_h} \frac{b(v_h, p_h)}{|v_h|\, \|p_h\|} = m_h > 0. \tag{2.4}$$

It is well known that the spectrum of S_h satisfies $\sigma(S_h) \subset [m_h^2, M_h^2]$, where

$$M_h := \sup_{p_h \in \mathcal{M}_h} \sup_{v_h \in V_h} \frac{b(v_h, p_h)}{|v_h|\, \|p_h\|} \le M < \infty, \tag{2.5}$$

and that m_h^2, M_h^2 are (the extreme) eigenvalues of S_h. Define

$$V_{h,0} := \{v_h \in V_h \,|\, b(v_h, q_h) = 0 \quad \text{for all } q_h \in \mathcal{M}_h\},$$

to be the kernel of the discrete operator B_h. We define $f_h \in V_h^*$ to be the restriction of f to V_h, i.e. $\langle f_h, v_h\rangle := \langle f, v_h\rangle$ for all $v_h \in V_h$.

Remark 2.2. *In the case $V_{h,0} \subset V_0$, the compatibility condition* (2.3) *implies the discrete compatibility condition*

$$\langle f, v_h\rangle = 0 \quad \textit{for all } v_h \in V_{h,0}.$$

Hence, under assumption (2.4)*, the problem of finding $p_h \in \mathcal{M}_h$ such that*

$$b(v_h, p_h) = \langle f, v_h\rangle,\ v_h \in V_h,\ \textit{or } B_h^* p_h = f_h,\ \textit{or } A_h^{-1} B_h^* p_h = A_h^{-1} f_h, \tag{2.6}$$

has a unique solution. In general, (2.3) *may not hold on $V_{h,0}$ and problem* (2.6) *may not be well-posed. However, if the form $b(\cdot,\cdot)$ satisfies* (2.4) *then the problem of finding $(u_h, p_h) \in V_h \times \mathcal{M}_h$ satisfying* (1.3) *does have a unique solution. Solving for p_h from* (1.3)*, we obtain*

$$S_h\, p_h = B_h(A_h^{-1} B_h^*)\, p_h = B_h A_h^{-1} f_h. \tag{2.7}$$

Since the Hilbert transpose of B_h is $B_h^T = A_h^{-1} B_h^*$, we call the component p_h of the solution (u_h, p_h) of (1.3) *the saddle point least squares approximation of the solution p of the original mixed problem* (1.1)*.*

The following error estimate for $\|p - p_h\|$ was proved in [9].

Theorem 2.3. *Let $b : V \times Q \to \mathbb{R}$ satisfy* (2.1) *and* (2.2) *and assume that $f \in V^*$ is given and satisfies* (2.3)*. Assume that p is the solution of* (1.1) *and $V_h \subset V$, $\mathcal{M}_h \subset Q$ are chosen such that the discrete* $\inf - \sup$ *condition* (2.4) *holds. If (u_h, p_h) is the solution of* (1.3)*, then the following error estimate holds:*

$$\frac{1}{M} |u_h| \le \|p - p_h\| \le \frac{M}{m_h} \inf_{q_h \in \mathcal{M}_h} \|p - q_h\|. \tag{2.8}$$

2.3. **An Uzawa CG iterative solver.** Note that a global linear system may be difficult to assemble when solving (1.3) as bases for the trial spaces $\mathcal{M}_h$, which are chosen to satisfy (2.4), may be difficult to find. Nevertheless, we can solve (1.3) and avoid building a basis for $\mathcal{M}_h$ by using an Uzawa type algorithm, e.g. the Uzawa Conjugate Gradient (UCG) algorithm.

Algorithm 2.4. *(UCG) Algorithm*

Step 1: Choose any $p_0 \in \mathcal{M}_h$. **Compute** $u_1 \in V_h$, $q_1, d_1 \in \mathcal{M}_h$ *by*

$$\begin{aligned} a(u_1, v_h) &= \langle f, v_h \rangle - b(v_h, p_0) && \text{for all } v_h \in V_h, \\ (q_1, q_h) &= b(u_1, q_h) && \text{for all } q_h \in \mathcal{M}_h, \quad d_1 := q_1. \end{aligned}$$

Step 2: For $j = 1, 2, \dots,$ **compute** $h_j, \alpha_j, p_j, u_{j+1}, q_{j+1}, \beta_j, d_{j+1}$ *by*

$$\begin{aligned} &(\mathbf{UCG1}) && a(h_j, v_h) = -\, b(v_h, d_j) \qquad \text{for all } v_h \in V_h \\ &(\mathbf{UCG}\alpha) && \alpha_j = -\frac{(q_j, q_j)}{b(h_j, q_j)} \\ &(\mathbf{UCG2}) && p_j = p_{j-1} + \alpha_j \; d_j \\ &(\mathbf{UCG3}) && u_{j+1} = u_j + \alpha_j \; h_j \\ &(\mathbf{UCG4}) && (q_{j+1}, q_h) = b(u_{j+1}, q_h) \qquad \text{for all } q_h \in \mathcal{M}_h \\ &(\mathbf{UCG}\beta) && \beta_j = \frac{(q_{j+1}, q_{j+1})}{(q_j, q_j)} \\ &(\mathbf{UCG6}) && d_{j+1} = q_{j+1} + \beta_j d_j. \end{aligned}$$

Note that the only inversions needed in the algorithm involve the form $a(\cdot, \cdot)$ in **Step 1** and (**UCG1**). In operator form, these steps become

$$u_1 = A_h^{-1}(f_h - B_h^* p_0), \qquad \text{and} \qquad h_j = -A_h^{-1}(B_h^* d_j), \tag{2.9}$$

respectively. In practical implementations of Algorithm 2.4, we would like to replace the action of A_h^{-1} with the action of a suitable preconditioner. The properties of the new preconditioned algorithm are discussed in the next section.

Remark 2.5. *In particular, Algorithm 2.4 recovers the steps of the conjugate gradient algorithm for solving the problem* (2.7)*. Due to the assumption* (2.4)*, the Schur complement S_h is a symmetric positive definite operator. Consequently, the conjugate iterations p_j converge to the solution p_h of* (2.7)*, and the rate of convergence for the iteration error $\|p_j - p_h\|_{S_h}$ or $\|p_j - p_h\|$ depends on the condition number of S_h, which is $\kappa(S_h) = \frac{M_h^2}{m_h^2}$.*

The following *sharp error estimation* result was proved in [3].

Theorem 2.6. *If (u_h, p_h) is the discrete solution of* (1.3) *and (u_j, p_{j-1}) is the j^{th} iteration for Algorithm 2.4, then $(u_j, p_{j-1}) \to (u_h, p_h)$ and*

$$
\begin{aligned}
\frac{1}{M^2}\,\|q_j\| &\le \|p_{j-1}-p_h\| \le \frac{1}{m_h^2}\,\|q_j\|, \\
\frac{m_h}{M^2}\,\|q_j\| &\le |u_j-u_h| \le \frac{M}{m_h^2}\,\|q_j\|.
\end{aligned}
\tag{2.10}
$$

3. Preconditioning techniques

In this section, we develop a general preconditioning framework to approximate the solution of (1.1) based on (1.3) and elliptic preconditioning of the operator associated with the inner product on V_h. More precisely, we replace the original form $a(\cdot,\cdot)$ in (1.3) with a uniformly equivalent form $\tilde{a}(\cdot,\cdot)$ on V_h that leads to an implementably fast operator $\tilde{A}_h^{-1}$. We assume that $V_h \subset V$ and $\mathcal{M}_h \subset Q$ are finite dimensional approximation spaces satisfying (2.4) and (2.5).

3.1. **The preconditioned saddle point problem.** First, we introduce a general preconditioner operator $P_h : V_h^* \to V_h$ that is equivalent to A_h^{-1} in the following sense

$$
\langle g, P_h f\rangle = \langle f, P_h g\rangle \quad \text{for all } f, g \in V_h^*,
\tag{3.1}
$$

and

$$
m_1^2 |v_h|^2 \le a(P_h A_h v_h, v_h) \le m_2^2 |v_h|^2,
\tag{3.2}
$$

where m_1, m_2 are positive constants. We note that assumption (3.1) implies that the operator $P_h A_h$ is symmetric with respect to the $a(\cdot,\cdot)$ inner product. Indeed, from the definition of the operator A_h and (3.1), we have

$$
\begin{aligned}
a(P_h A_h u_h, v_h) = \langle A_h v_h, P_h A_h u_h\rangle &= \langle A_h u_h, P_h A_h v_h\rangle \\
&= a(u_h, P_h A_h v_h).
\end{aligned}
$$

Remark 3.1. *Assuming that m_1^2, m_2^2 are the smallest and the largest eigenvalues of $P_h A_h$, respectively, the inequality* (3.2) *gives us that the condition number of $P_h A_h$ satisfies*

$$
\kappa(P_h A_h) = \frac{m_2^2}{m_1^2}.
\tag{3.3}
$$

With the preconditioner $P_h : V_h^* \to V_h$, we define the form $\tilde{a} : V_h \times V_h \to \mathbb{R}$ by

$$
\tilde{a}(u_h, v_h) := a((P_h A_h)^{-1} u_h, v_h) \quad \text{for all } u_h, v_h \in V_h.
\tag{3.4}
$$

Proposition 3.2. *Under assumptions* (3.1) *and* (3.2)*, we have that $\tilde{a}(\cdot,\cdot)$ is symmetric and equivalent with $a(\cdot,\cdot)$ on V_h.*

Proof. For symmetry, we just use that $P_h A_h$ is symmetric w.r.t. the $a(\cdot,\cdot)$ inner product. For the equivalence, note that (3.2) and (3.4) imply

$$
\frac{1}{m_2^2} |v_h|^2 \le \tilde{a}(v_h, v_h) \le \frac{1}{m_1^2} |v_h|^2.
\tag{3.5}
$$

□

By Proposition 3.2, $\tilde{a}(\cdot,\cdot)$ defines an equivalent inner product on V_h. Let $|v_h|_P := \tilde{a}(v_h, v_h)^{1/2}$ be the norm induced by the inner product $\tilde{a}(\cdot,\cdot)$ and define the operator $\tilde{A}_h : V_h \to V_h^*$ by

$$\langle \tilde{A}_h u_h, v_h\rangle := \tilde{a}(u_h, v_h) \quad \text{for all } u_h, v_h \in V_h.$$

Note that for any $u_h, v_h \in V_h$

$$\begin{aligned}\langle \tilde{A}_h u_h, v_h\rangle = \tilde{a}(u_h, v_h) &= a((P_h A_h)^{-1} u_h, v_h)\\ &= \langle A_h (P_h A_h)^{-1} u_h, v_h\rangle,\end{aligned}$$

which implies $\tilde{A}_h = A_h(P_h A_h)^{-1} = P_h^{-1}$. Hence, we can view $\tilde{a}(\cdot,\cdot)$ as a preconditioned version of the form $a(\cdot,\cdot)$. The preconditioned discrete saddle point problem consists of finding $(\tilde{u}_h, \tilde{p}_h) \in V_h \times \mathcal{M}_h$ such that (1.4) holds. To simplify the notation, we will drop the $\tilde{}$ notation from $(\tilde{u}_h, \tilde{p}_h)$. Thus, for the remainder of this paper, the *preconditioned saddle point least squares* formulation is: Find $(u_h, p_h) \in V_h \times \mathcal{M}_h$ such that

$$\begin{array}{lllll} \tilde{a}(u_h, v_h) & + & b(v_h, p_h) & = \langle f, v_h\rangle & \text{for all } v_h \in V_h,\\ b(u_h, q_h) & & & = 0 & \text{for all } q_h \in \mathcal{M}_h. \end{array} \tag{3.6}$$

Using that $V_h \subset V$ and $\mathcal{M}_h \subset Q$ satisfy (2.4) and (2.5), we obtain

$$\tilde{m}_h := \inf_{p_h \in \mathcal{M}_h} \sup_{v_h \in V_h} \frac{b(v_h, p_h)}{|v_h|_P \, \|p_h\|} \geq m_1\, m_h > 0, \tag{3.7}$$

and

$$\tilde{M}_h := \sup_{p_h \in \mathcal{M}_h} \sup_{v_h \in V_h} \frac{b(v_h, p_h)}{|v_h|_P \, \|p_h\|} \leq m_2\, M_h \leq m_2\, M. \tag{3.8}$$

Hence, the *preconditioned saddle point least squares* formulation (3.6) has a unique solution.

The Schur complement associated with problem (3.6) is

$$\tilde{S}_h = B_h \tilde{A}_h^{-1} B_h^* = B_h P_h B_h^*.$$

Solving for p_h from (3.6), we obtain

$$\tilde{S}_h\, p_h = B_h (P_h B_h^*)\, p_h = B_h P_h f_h. \tag{3.9}$$

We call the component p_h of the solution (u_h, p_h) of (3.6) the *(preconditioned) saddle point least squares* approximation of the solution p of the original mixed prolem (1.1). To estimate $\|p - p_h\|$ in this case, we will prove the analog to Theorem 2.3 based on the Xu-Zikatanov argument, see [25].

Theorem 3.3. *Let* $b : V \times Q \to \mathbb{R}$ *satisfy* (2.1) *and* (2.2) *and assume that* $f \in V^*$ *is given and satisfies* (2.3). *Assume that* $V_h \subset V$, $\mathcal{M}_h \subset Q$ *are chosen such that the discrete* $\inf - \sup$ *condition* (2.4) *holds. If* p *is the* *solution of* (1.1) *and* (u_h, p_h) *is the solution of* (3.6)*, then the following error estimate holds:*

$$\frac{1}{M}\frac{1}{m_2^2}|u_h| \le \|p - p_h\| \le \frac{M}{m_h}\frac{m_2}{m_1} \inf_{q_h \in \mathcal{M}_h} \|p - q_h\|. \tag{3.10}$$

Proof. Define the operator $T_h : Q \to Q$ by $T_h p = p_h$. Note that T_h is linear and idempotent. To show the latter, consider the problem: Find $(u_h^*, p_h^*) \in V_h \times \mathcal{M}_h$ such that

$$\begin{array}{llll} \tilde{a}(u_h^*, v_h) & + \quad b(v_h, p_h^*) & = b(v_h, p_h) & \text{for all } v_h \in V_h, \\ b(u_h^*, q_h) & & = 0 & \text{for all } q_h \in \mathcal{M}_h. \end{array} \tag{3.11}$$

Since b satisfies (2.4), we have that (3.7) is satisfied as described above. Thus, problem (3.11) has a unique solution. Since $(u_h^*, p_h^*) = (0, p_h)$ solves the problem, we conclude $T_h p_h = p_h$ which gives us $T_h^2 = T_h$. From Kato [20] and Xu and Zikatanov [25], this implies

$$\|I - T_h\|_{\mathcal{L}(Q,Q)} = \|T_h\|_{\mathcal{L}(Q,Q)}.$$

Using the above equality, for an arbitrary $q_h \in \mathcal{M}_h$ we have

$$\|p - q_h\| = \|(I - T_h)p\| = \|(I - T_h)(p - q_h)\| \le \|T_h\| \, \|p - q_h\|. \tag{3.12}$$

We now estimate $\|T_h\|$. First, define $\tilde{V}_{h,0}^{\perp}$ to be the orthogonal complement of $V_{h,0}$ w.r.t. the $\tilde{a}(\cdot,\cdot)$ inner product. Note that from the first equation of (3.6) and the fact p solves (1.1) we have that

$$b(v_h, p_h) = b(v_h, p) - \tilde{a}(u_h, v_h). \tag{3.13}$$

Also, since b satisfies (2.2) we have that (3.8) holds. Hence, from (3.7), (3.8), and (3.13) we obtain

$$\begin{aligned} \|T_h p\| \le \frac{1}{m_h\, m_1} \sup_{v_h \in V_h} \frac{b(v_h, T_h p)}{|v_h|_P} &= \frac{1}{m_h\, m_1} \sup_{v_h \in \tilde{V}_{h,0}^{\perp}} \frac{b(v_h, p_h)}{|v_h|_P} \\ &= \frac{1}{m_h\, m_1} \sup_{v_h \in \tilde{V}_{h,0}^{\perp}} \frac{b(v_h, p) - \tilde{a}(u_h, v_h)}{|v_h|_P} \\ &\le \frac{M m_2}{m_h m_1} \|p\|. \end{aligned} \tag{3.14}$$

The right inequality now follows from (3.12) and (3.14). For the left inequality, note that

$$|u_h|_P = \sup_{v_h \in V_h} \frac{\tilde{a}(u_h, v_h)}{|v_h|_P} = \sup_{v_h \in V_h} \frac{b(v_h, p - p_h)}{|v_h|_P} \le M\, m_2 \|p - p_h\|,$$

and

$$|u_h| \le m_2 |u_h|_P.$$

□

3.2. **An iterative solver for the preconditioned variational formulation.** We use a modified version of Algorithm 2.4 to solve (3.6) by replacing the form $a(\cdot,\cdot)$ by $\tilde{a}(\cdot,\cdot)$ in **Step 1** and (**UCG1**). With this modification, we obtain the following (Uzawa) Preconditioned Conjugate Gradient (PCG) algorithm for mixed methods.

Algorithm 3.4. *(PCG) Algorithm for Mixed Methods*

Step 1: Choose any $p_0 \in \mathcal{M}_h$. **Compute** $u_1 \in V_h,\ q_1, d_1 \in \mathcal{M}_h$ *by*

$$\begin{aligned} u_1 &= P_h(f_h - B_h^* p_0) \\ q_1 &= B_h u_1, \quad d_1 := q_1. \end{aligned}$$

Step 2: For $j = 1, 2, \ldots,$ **compute** $h_j, \alpha_j, p_j, u_{j+1}, q_{j+1}, \beta_j, d_{j+1}$ *by*

$$\begin{aligned}
&(\mathbf{PCG1}) & h_j &= -\ P_h(B_h^* d_j) \\
&(\mathbf{PCG}\alpha) & \alpha_j &= -\ \frac{(q_j, q_j)}{b(h_j, q_j)} \\
&(\mathbf{PCG2}) & p_j &= p_{j-1} + \alpha_j\ d_j \\
&(\mathbf{PCG3}) & u_{j+1} &= u_j + \alpha_j\ h_j \\
&(\mathbf{PCG4}) & q_{j+1} &= B_h u_{j+1}, \\
&(\mathbf{PCG}\beta) & \beta_j &= \frac{(q_{j+1}, q_{j+1})}{(q_j, q_j)} \\
&(\mathbf{PCG6}) & d_{j+1} &= q_{j+1} + \beta_j d_j.
\end{aligned}$$

Note that only the actions of P_h, B_h, and B_h^* are needed in the above algorithm. For any preconditioner P_h and trial space $\mathcal{M}_h$ that is not defined via a global projection, these actions do not involve inversion processes, see Section 3.3 for the case P_h -an additive multilevel Schwarz preconditioner. Similar to the remark in Section 2.3, we have the following.

Remark 3.5. *Algorithm 3.4 recovers in particular the steps of the conjugate gradient algorithm for solving the problem* (3.9). *Due to the discrete* $\inf - \sup$ *assumption* (2.4) *and the assumptions* (3.1) *and* (3.2) *on the preconditioner* P_h*, the Schur complement* $\tilde{S}_h$ *is a symmetric positive definite operator. Consequently, the conjugate iterations* p_j *converge to the solution* p_h *of* (3.9)*, and the rate of convergence for* $\|p_j - p_h\|_{\tilde{S}_h}$ *or* $\|p_j - p_h\|$ *depends on the condition number of* $\tilde{S}_h$*, which is* $\kappa(\tilde{S}_h) = \frac{\tilde{M}_h^2}{\tilde{m}_h^2}$.

The following result is analogous to Theorem 2.6.

Theorem 3.6. *If* (u_h, p_h) *is the discrete solution of* (3.6) *and* (u_j, p_{j-1}) *is the* j^{th} *iteration for Algorithm 3.4, then* $(u_j, p_{j-1}) \to (u_h, p_h)$ *and*

$$\begin{aligned} \frac{1}{M^2}\frac{1}{m_2^2}\,\|q_j\| \le \|p_{j-1}-p_h\| \le \frac{1}{m_h^2}\frac{1}{m_1^2}\,\|q_j\|, \\ \frac{m_h}{M^2}\frac{m_1^2}{m_2^2}\,\|q_j\| \le |u_j-u_h| \le \frac{M}{m_h^2}\frac{m_2^2}{m_1^2}\,\|q_j\|. \end{aligned} \tag{3.15}$$

Proof. By induction over j, we have that

$$\tilde{a}(u_j, v_h) + b(v_h, p_{j-1}) = \langle f, v_h\rangle \quad \text{for all } v_h \in V_h.$$

Combining this with the first equation of (3.6) gives us

$$\tilde{a}(u_j - u_h, v_h) = b(v_h, p_h - p_{j-1}) \quad \text{for all } v_h \in V_h. \tag{3.16}$$

Note that $\sigma(\tilde{S}_h) \subset [\tilde{m}_h^2, \tilde{M}_h^2]$. Hence,

$$\tilde{m}_h\|q_h\| = (\tilde{S}_h q_h, q_h)^{1/2} \le \tilde{M}_h\|q_h\| \quad \text{for all } q_h \in \mathcal{M}_h. \tag{3.17}$$

By substituting $v_h = \tilde{A}_h^{-1}B_h^*(p_h - p_{j-1})$ into (3.16),

$$|u_j - u_h|_P^2 = (\tilde{S}_h(p_h - p_{j-1}), p_h - p_{j-1}) = \|p_h - p_{j-1}\|_{\tilde{S}_h}^2.$$

The above equality, (3.5), and (3.17) gives us that

$$m_1\tilde{m}_h\|p_h - p_{j-1}\| \le |u_j - u_h| \le m_2\tilde{M}_h\|p_h - p_{j-1}\|. \tag{3.18}$$

From (**PCG4**), the second equation of (3.6), and (3.16) we have that

$$q_j = B_h u_j = B_h(u_j - u_h) = \tilde{S}_h(p_h - p_{j-1}).$$

Thus,

$$\tilde{m}_h^2\|p_h - p_{j-1}\| \le \|\tilde{S}_h(p_h - p_{j-1})\| = \|q_j\| \le \tilde{M}_h^2\|p_h - p_{j-1}\|. \tag{3.19}$$

The inequalities (3.15) follow from (3.18), (3.19), and the fact that $\tilde{m}_h \ge m_h m_1$ and $\tilde{M}_h \le M m_2$. From Remark 3.5 and the standard estimate for the convergence rate of the conjugate gradient algorithm, [11, 18], we have that

$$\|p_h - p_j\|_{\tilde{S}_h} \le 2\left(\frac{\tilde{M}_h - \tilde{m}_h}{\tilde{M}_h + \tilde{m}_h}\right)^j \|p_h - p_0\|_{\tilde{S}_h}. \tag{3.20}$$

Hence, $p_j \to p_h$. From (3.15), we conclude that $u_j \to u_h$ as well. □

The following estimates are a direct consequence of (3.7), (3.8), (3.20), and the formula $\kappa(\tilde{S}_h) = \frac{\tilde{M}_h^2}{\tilde{m}_h^2}$.

Proposition 3.7. *The condition number of the Schur complement* $\tilde{S}_h = B_h P_h B_h^*$ *satisfies*

$$\kappa(\tilde{S}_h) \le \frac{M_h^2}{m_h^2}\frac{m_2^2}{m_1^2} = \kappa(S_h)\cdot\kappa(P_h A_h). \tag{3.21}$$

Consequently, the convergence rate ρ_h *for* $\|p_j - p_h\|_{\tilde{S}_h}$ *satisfies*

$$\rho_h \leq \frac{\frac{M_h}{m_h}\frac{m_2}{m_1} - 1}{\frac{M_h}{m_h}\frac{m_2}{m_1} + 1}.$$

Remark 3.8. *We can relate our preconditioned SPLS discretization method for solving the general mixed problem* (1.1) *with the Bramble-Pasciak least squares approach presented in* [12]*. In our notation, the Bramble-Pasciak least squares discretization can be formulated as: Find* $p_h \in \mathcal{M}_h$ *such that*

$$b(A_h^{-1}B_h^* q_h, p_h) = \langle f_h, A_h^{-1}B_h^* q_h\rangle = b(A_h^{-1} f_h, q_h) \quad \text{for all } q_h \in \mathcal{M}_h.$$

With a suitable preconditioner P_h *replacing* A_h^{-1}*, the problem becomes: Find* $p_h \in \mathcal{M}_h$ *such that*

$$b(P_h B_h^* q_h, p_h) = b(P_h f_h, q_h) \quad \text{for all } q_h \in \mathcal{M}_h. \tag{3.22}$$

We note that (3.22) *is equivalent to our Schur complement problem* (3.9)*. While we arrive at essentially the same normal equation for solving* (2.6)*, our saddle point approach is more direct and allows sharp error estimates for the error* $\|p - p_h\|$*. The two approaches are also essentially different in the way the trial spaces are chosen, see Section 4 for our choices of trial spaces. In* [12]*, to iteratively solve* (3.22) *bases for both the test and trial spaces are needed. In contrast, we solve the coupled preconditioned saddle point problem* (3.6) *using Algorithm 3.4 which avoids the need of a basis for the trial space.*

3.3. **An example of a preconditioner.** In order to illustrate the applicability of the theory presented thus far, we consider the case when P_h is given by the additive multilevel Schwarz or BPX preconditioner, see [13, 14, 26]. Assume that we have a nested sequence of approximation spaces $V_1 \subset V_2 \subset \cdots \subset V_J = V_h$ and let $\{\phi_1^k, \phi_2^k, \ldots, \phi_{n_k}^k\}$ be a basis for V_k. For $f_h \in V_h^*$, the action of P_h is given by

$$P_h f_h = \sum_{k=1}^{J}\sum_{i=1}^{n_k} \frac{\langle f_h, \phi_i^k\rangle}{a(\phi_i^k, \phi_i^k)}\phi_i^k. \tag{3.23}$$

It is known that for $V = H_0^1(\Omega)$ and a nested sequence $\{V_k\}$ of piecewise linear functions that, under standard mesh uniformity conditions, P_h is a preconditioner for A_h satisfying (3.1) and (3.2), see [13, 19, 23, 24, 26].

In this case, the first equation in **(Step 1)** of Algorithm 3.4 becomes

$$u_1 = P_h(f_h - B_h^* p_0) = \sum_{k=1}^{J}\sum_{i=1}^{n_k} \frac{\langle f_h, \phi_i^k\rangle - b(\phi_i^k, p_0)}{a(\phi_i^k, \phi_i^k)}\phi_i^k. \tag{3.24}$$

Furthermore, the iterates for h_j in **(PCG1)** are given by

$$h_j = -\sum_{k=1}^{J}\sum_{i=1}^{n_k} \frac{b(\phi_i^k, d_j)}{a(\phi_i^k, \phi_i^k)}\phi_i^k, \tag{3.25}$$

which implies that

$$b(h_j, q_j) = -\sum_{k=1}^{J}\sum_{i=1}^{n_k} \frac{b(\phi_i^k, d_j) b(\phi_i^k, q_j)}{a(\phi_i^k, \phi_i^k)}, \tag{3.26}$$

in **(PCGα)**. Thus, the implementation of Algorithm 3.4 does not involve matrix inversion. Certainly, any elliptic preconditioner, including the standard multigrid ones, can be used for P_h. We decided to show details of a general additive multilevel Schwarz (or BPX) preconditioner to emphasize the simplicity of implementation when dealing with mixed methods preconditioning. More details on implementing the matrix action of multilevel preconditioners (including BPX) can be found in [24].

3.4. **Computational complexity of the proposed PCG algorithm.** From **Step 1** or **Step 2** of Algorithm 3.4, we observe that at each step the number of operations depends on the complexity of P_h and the dimension of the test space V_h, say $n = n_h$. A preconditioner P_h is of optimal complexity if $O(n)$ operations are needed to compute its action, where n is the dimension of V_h. Preconditioners such as multigrid or the BPX preconditioner are of optimal complexity. For the BPX preconditioner defined in (3.23), this is because for each k (using a standard refinement strategy, e.g. in 2D split each triangle in four smaller triangles), we have $n_k = O(\alpha^k)$ for some $\alpha > 1$ depending on the dimension of the domain, where $n = n_J$ is the dimension of V_h. Thus, in this case the action of P_h needs

$$O\left(\sum_{k=1}^{J} n_k\right) = O\left(\sum_{k=1}^{J} \alpha^k\right) = O(\alpha^J) = O(n),$$

operations. Using that the action of B_h is the action of a differential operator (most often of first order) on a finite element function in V_h, we can conclude from formulas (3.24), (3.25), and (3.26) that the rest of the operations in **Step 1** or **Step 2** of Algorithm 3.4 sum up to at most $O(n)$ operations.

Regarding the global complexity and optimality of the algorithm, it is known that if the condition number of the symmetric positive definite operator $\tilde{S}_h$, defined in (3.9), is independent of h then the number of iterations of the PCG algorithm is bounded independent of h. Thus, if P_h is an optimal complexity preconditioner which is also a uniform preconditioner, i.e. the constants m_1, m_2 in (3.2) are independent of h, and the discrete $\inf - \sup$ constant m_h of (2.4) is independent of h, then Algorithm 3.4 is optimal. That is, to achieve a certain accuracy, it needs a number of operations that is proportional to the dimension of the space V_h.

4. Discrete spaces that satisfy an $\inf - \sup$ condition

In this section, we describe two pairs of discrete spaces, introduced in [9], which satisfy the discrete $\inf - \sup$ condition (2.4) in the general abstract framework of Section 2. In light of (3.21), we would like to provide families of spaces $\{(V_h, \mathcal{M}_h)\}$ such that $\kappa(S_h)$ is small. Let $V_h \subset V$ be a finite element test space and assume the action of $\mathcal{C}^{-1}$, where $\mathcal{C}$ was defined in Section 2, is easy to obtain at the continuous level.

4.1. **No projection trial space.** The first choice defines $\mathcal{M}_h \subset Q$ by

$$\mathcal{M}_h := \mathcal{C}^{-1} B V_h.$$

In this case, $V_{h,0} \subset V_0$ and a discrete $\inf - \sup$ condition holds. Indeed, for a generic $p_h = \mathcal{C}^{-1} B w_h \in \mathcal{M}_h$ where $w_h \in V_{h,0}^{\perp}$, we have

$$\begin{aligned} m_{h,0} := \inf_{p_h \in \mathcal{M}_h} \sup_{v_h \in V_h} \frac{b(v_h, p_h)}{|v_h| \, \|p_h\|} &= \inf_{w_h \in V_{h,0}^{\perp}} \sup_{v_h \in V_h} \frac{(\mathcal{C}^{-1} B v_h, \mathcal{C}^{-1} B w_h)}{|v_h| \, \|\mathcal{C}^{-1} B w_h\|} \\ \geq \inf_{w_h \in V_{h,0}^{\perp}} \frac{\|\mathcal{C}^{-1} B w_h\|^2}{|w_h| \, \|\mathcal{C}^{-1} B w_h\|} &= \inf_{w_h \in V_{h,0}^{\perp}} \frac{\|\mathcal{C}^{-1} B w_h\|}{|w_h|} > 0. \end{aligned} \tag{4.1}$$

Hence, by Remark 2.2 we have that (2.6) has a unique solution $p_h \in \mathcal{M}_h$ and $(u_h = 0, p_h)$ solves (1.3). In this case, p_h is an optimal approximation to the solution p of (1.1). Indeed, for any $v_h \in V_h$

$$\begin{aligned} 0 = b(v_h, p - p_h) &= \langle B v_h, p - p_h \rangle \\ &= (\mathcal{C}^{-1} B v_h, p - p_h). \end{aligned}$$

Thus, p_h is the orthogonal projection of p onto $\mathcal{M}_h$ which implies

$$\|p - p_h\| = \inf_{q_h \in \mathcal{M}_h} \|p - q_h\|. \tag{4.2}$$

While (4.2) gives optimal approximation error, to efficiently approximate p_h using Algorithm 2.4 or 3.4, it requires spaces $\{(V_h, \mathcal{M}_h)\}$ for which $\kappa(S_h) = \frac{M_h^2}{m_h^2}$ is small or independent of h.

4.2. **Projection type trial space.** The second choice defines $\mathcal{M}_h \subset Q$ by

$$\mathcal{M}_h := R_h \mathcal{C}^{-1} B V_h,$$

where $R_h : Q \to \tilde{M}_h$ is defined by

$$(R_h p, q_h)_h := (p, q_h) \quad \text{for all } q_h \in \tilde{M}_h, \tag{4.3}$$

and $\tilde{M}_h$ is a finite dimensional subspace of Q equipped with the inner product $(\cdot, \cdot)_h$.

Remark 4.1. *If the $(\cdot, \cdot)_h$ inner product coincides with the inner product on Q, then by definition R_h is the orthogonal projection onto $\tilde{M}_h$.*

In general, the inner product on $\tilde{M}_h$ could be different from the inner product on Q, but we assume that $(\cdot, \cdot)_h$ induces an equivalent norm independent of h. The following proposition provides a sufficient condition on R_h which implies well-posedness of problems (1.3) and (2.6) and relates the stability of the family of spaces $\{(V_h, R_h \mathcal{C}^{-1} B V_h)\}$ with the stability of the family of spaces $\{(V_h, \mathcal{C}^{-1} B V_h)\}$ defined in Section 4.1.

Proposition 4.2. *Assume that*

$$\|R_h q_h\|_h \geq \tilde{c}\|q_h\| \quad \text{for all } q_h \in \mathcal{C}^{-1}BV_h, \tag{4.4}$$

with a constant $\tilde{c}$ independent of h. Then $V_{h,0} \subset V_0$. Furthermore, the stability of the family $\{(V_h, \mathcal{C}^{-1}BV_h)\}$ implies the stability of the family $\{V_h, R_h\mathcal{C}^{-1}BV_h)\}$.

Proof. Let $v_h \in V_{h,0}$. Then, for any $p_h \in \mathcal{M}_h$,

$$0 = b(v_h, p_h) = (\mathcal{C}^{-1}Bv_h, p_h) = (R_h\mathcal{C}^{-1}Bv_h, p_h)_h.$$

Taking $p_h = R_h\mathcal{C}^{-1}Bv_h$ gives us $\|R_h\mathcal{C}^{-1}Bv_h\|_h = 0$ and the inclusion $V_{h,0} \subset V_0$ follows from (4.4). For the stability, note that for a generic function $p_h = R_h\mathcal{C}^{-1}Bw_h \in \mathcal{M}_h$, where $w_h \in V_{h,0}^{\perp}$, we have

$$\begin{aligned}
m_h = \inf_{p_h\in\mathcal{M}_h}\sup_{v_h\in V_h}\frac{b(v_h,p_h)}{|v_h|\ \|p_h\|_h} &= \inf_{w_h\in V_{h,0}^{\perp}}\sup_{v_h\in V_h}\frac{(\mathcal{C}^{-1}Bv_h, R_h\mathcal{C}^{-1}Bw_h)}{|v_h|\ \|R_h\mathcal{C}^{-1}Bw_h\|_h}\\
&= \inf_{w_h\in V_{h,0}^{\perp}}\sup_{v_h\in V_h}\frac{(R_h\mathcal{C}^{-1}Bv_h, R_h\mathcal{C}^{-1}Bw_h)_h}{|v_h|\ \|R_h\mathcal{C}^{-1}Bw_h\|_h}\\
&\geq \inf_{w_h\in V_{h,0}^{\perp}}\frac{\|R_h\mathcal{C}^{-1}Bw_h\|_h^2}{|w_h|\ \|R_h\mathcal{C}^{-1}Bw_h\|_h}\\
&\geq \tilde{c}\inf_{w_h\in V_{h,0}^{\perp}}\frac{\|\mathcal{C}^{-1}Bw_h\|}{|w_h|} = \tilde{c}\, m_{h,0},
\end{aligned}$$

where $m_{h,0}$ is defined in (4.1). □

In this case, we have that p_h is a quasi-optimal approximation of the solution p of (1.1) by Theorem 3.3.

Remark 4.3. *The benefit of using the projection type trial space is that it could lead to a better approximation of the continuous solution p. Indeed, for the case when preconditioning is not used, super-convergence of $\|p-p_h\|$ is observed, see* [6, 7, 9, 10]. *Using uniform preconditioners and Theorem 3.3, we expect the same order of super-convergence for $\|p-p_h\|$.*

Remark 4.4. *With this choice of trial space, Algorithms 2.4 and 3.4 need to be modified to account for the $(\cdot,\cdot)_h$ inner product on $\mathcal{M}_h \subset \tilde{M}_h$. This modification is nothing more than replacing the $(\cdot,\cdot)$ inner product with the $(\cdot,\cdot)_h$ inner product where it appears in the algorithms.*

5. Numerical Results

In this section, we compare the performance of Algorithm 3.4, referred to in this section as the *UPCG* algorithm, with the standard PCG algorithm. The problem considered is: Find $u \in H_0^1(\Omega)$ such that

$$-\nabla\cdot(A\nabla u) = f \text{ in } \Omega, \tag{5.1}$$

where $\Omega \subset \mathbb{R}^2$, $f \in L^2(\Omega)$ is given, and the matrix A is symmetric and satisfies

$$a_{min}|\xi|^2 \leq \langle A(x)\xi, \xi\rangle \leq a_{max}|\xi|^2 \quad \text{for all } x \in \Omega, \xi \in \mathbb{R}^2, \tag{5.2}$$

for positive constants $a_{min} \leq a_{max}$. From [10], we have that equation (5.1) fits the SPLS framework outlined in Section 2.1 with $V = H_0^1(\Omega)$, $Q = A\nabla V$, and $b : V \times Q \to \mathbb{R}$ and $f \in V^*$ defined as

$$b(v, q) := (q, \nabla v) \quad \text{for all } v \in V, q \in Q,$$

and

$$\langle f, v\rangle := (f, v) \quad \text{for all } v \in V.$$

On V, we can consider the inner products

$$a(u, v) := (\nabla u, \nabla v), \text{ or } a(u, v) := (A\nabla u, \nabla v) \quad \text{for all } u, v \in V,$$

and on Q we consider the inner product

$$(p, q)_Q = (A\nabla u, A\nabla v)_Q := (A\nabla u, A\nabla v)_{A^{-1}} = (A\nabla u, \nabla v).$$

We solve (3.6) using the UPCG algorithm with a suitable choice of preconditioner P_h. The test space $V_h \subset H_0^1(\Omega)$ is chosen to be the space of all *continuous* piecewise linear functions with respect to quasi uniform or locally quasi uniform meshes $\mathcal{T}_h$. The trial space $\mathcal{M}_h$ is chosen to be of the no projection type outlined in Section 4.1, i.e. of the form $\mathcal{M}_h = A\nabla V_h$. This choice of trial space is independent of the choice of preconditioner. Furthermore, the pair $(V_h, \mathcal{M}_h)$ is stable, i.e. $m_{h,0}$, defined in (4.1), satisfies $m_{h,0} \geq c > 0$ where c is a constant independent of h, see [10].

Two approaches are taken for the initial iterate p_0 in **Step 1** of the UPCG algorithm. The first takes $p_0 = 0$ for each level of refinement. The second uses a *cascadic* approach in which $p_0 = 0$ on the coarsest mesh, but for all successive refinements p_0 is chosen as the extension of the final iterate from the previous level [3, 5]. This approach is referred to as the *UPCG cascadic* algorithm for the remainder of this section. The stopping criteria is chosen to be the same across all examples.

5.1. **Application of multilevel preconditioners.** We consider (5.1) when $\Omega = (0, 1) \times (0, 1)$, $A = I$, and f is computed such that the exact solution is $u(x, y) = x(1 - x)y(1 - y)$. We apply the standard PCG, UPCG (Algorithm 3.4), and UPCG cascadic algorithms with the preconditioner P_h the standard BPX, see e.g. [13, 24]. Table 1 compares the performance of the UPCG algorithm, as well as the cascadic version, with the standard PCG algorithm.

From Table 1, we see the performance of the UPCG algorithm is comparable with standard PCG. In addition, there is a significant reduction in the number of iterations using the cascadic approach.

error = $\|\nabla u - \nabla u_h\|$

$h = 2^{-k}$	PCG			UPCG			UPCG cascadic		
k	error	rate	it	error	rate	it	error	rate	it
1	0.045	0.000	1	0.045	0.000	2	0.045	0.000	2
2	0.024	0.903	5	0.024	0.896	4	0.025	0.837	2
3	0.012	0.974	7	0.012	0.944	5	0.013	0.952	3
4	0.006	0.991	8	0.006	1.016	7	0.007	0.974	3
5	0.003	0.996	9	0.003	0.988	8	0.003	0.994	3
6	0.001	1.002	11	0.001	1.012	10	0.002	1.002	3

Table 1: Comparison on unit square example.

For the next example, we consider (5.1) when $\Omega = (0,1) \times (0,1)$ with interface $\Gamma := \Omega \cap \{(x,y) |\ x = 1/2\}$. We computed f such that for

$$A(x,y) = a(x,y)I_2, \text{ where } a(x,y) = \begin{cases} \beta & \text{if } x \geq \frac{1}{2}, \\ 1 & \text{if } x < \frac{1}{2}, \end{cases}$$

the exact solution is

$$u(x,y) = \begin{cases} \beta\ x(x-\frac{1}{2})y(y-1) & \text{if } x < \frac{1}{2}, \\ (x-\frac{1}{2})(x-1)y(1-y) & \text{if } x \geq \frac{1}{2}. \end{cases}$$

The preconditioner is taken to be the additive multilevel Schwarz preconditioner as described in Section 3.3. Table 2 compares the performance of the UPCG algorithm, as well as the cascadic version, with the standard PCG algorithm for $\beta = 10$. Table 3 compares the same algorithms for $\beta = 100$.

error = $\|A\nabla u - A\nabla u_h\|$

$h = 2^{-k}$	PCG			UPCG			UPCG cascadic		
k	error	rate	it	error	rate	it	error	rate	it
1	0.244	0.000	1	0.255	0.000	1	0.254	0.000	1
2	0.127	0.939	4	0.133	0.936	3	0.132	0.939	3
3	0.064	0.985	6	0.067	0.986	5	0.068	0.963	3
4	0.032	0.995	7	0.034	0.989	6	0.034	0.978	3
5	0.016	1.000	9	0.017	1.001	8	0.017	0.999	3
6	0.008	0.999	10	0.008	0.996	9	0.009	1.007	3

Table 2: Comparison on interface example with $\beta = 10$.

From Tables 2 and 3, we see a similar behavior in the performance of the UPCG and standard PCG algorithms as in the previous example.

The impact of using the projection type of trial space, presented in Section 4.2, will be tested in the near future. For the case of the interface problem (5.1) and no preconditioning, we can point out numerical results showing super-convergence in [6]. As mentioned in Remark 4.3, by using (uniform) error = $\|A\nabla u - A\nabla u_h\|$

$h = 2^{-k}$	PCG			UPCG			UPCG cascadic		
k	error	rate	it	error	rate	it	error	rate	it
1	2.427	0.000	1	2.439	0.000	2	2.439	0.000	2
2	1.265	0.940	4	1.271	0.940	5	1.271	0.940	4
3	0.639	0.985	6	0.642	0.985	7	0.642	0.985	5
4	0.320	0.994	7	0.322	0.996	9	0.322	0.996	6
5	0.160	1.000	9	0.161	0.999	11	0.161	0.999	5
6	0.080	0.999	10	0.081	0.999	12	0.081	0.999	6

Table 3: Comparison on interface example with $\beta = 100$.

preconditioning we expect the same order of super-convergence for $\|p - p_h\|$. Similar observations can be made for saddle point least squares discretization of div − curl systems, as presented in [9], and the Maxwell equations in [7].

6. Conclusion and future work

We presented a general preconditioning approach to mixed variational formulations of the form (1.1) that relies on the classical theory of symmetric saddle point problems and on the theory of preconditioning symmetric positive definite operators. First, a discrete saddle point variational formulation (1.3), that approximates the solution of the original mixed problem in a least squares sense, is considered. In this formulation, the inner product $a(\cdot,\cdot)$ is replaced by an equivalent bilinear form that give rise to efficient elliptic inversion or preconditioning.

In addition, an Uzawa preconditioned conjugate gradient algorithm for solving the new *symmetric* saddle point system was proposed that requires bases only for the space V_h and avoids costly inversion processes. Due to the saddle point interpretation of the preconditioned system, we were able to prove sharp approximability properties for the discretization and iteration errors and were able to provide practical estimates for the rate of convergence of the final preconditioned conjugate algorithm. Using a common test space, two choices of compatible discrete spaces were given. The numerical experiments demonstrate that the method is comparable in performance to the standard PCG algorithm for the no projection type of trial space.

We plan to combine this approach with multilevel and adaptive techniques. We are currently working on implementing 3D preconditioners for elliptic problems with various types of boundary conditions. We plan to extend this preconditioning technique to the case of non-conforming choices of trial spaces and further investigate the efficiency of the proposed preconditioned saddle point least squares discretization to first order systems of PDEs, including div − curl and Maxwell Equations, as well as more general second order elliptic problems. Non-conforming test spaces are also possible, [4, 22, 21], but the theory and its advantages need to be further investigated.

We would like to thank the three referees for all the help with improving the original version of the manuscript.

References

[1] A. Aziz and I. Babuška. Survey lectures on mathematical foundations of the finite element method. *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, A. Aziz, editor*, 1972.

[2] C. Bacuta. Schur complements on Hilbert spaces and saddle point systems. *J. Comput. Appl. Math.*, 225(2):581–593, 2009.

[3] C. Bacuta. Cascadic multilevel algorithms for symmetric saddle point systems. *Comput. Math. Appl.*, 67(10):1905–1913, 2014.

[4] C. Bacuta, J. H. Bramble, and J. Pasciak. New interpolation results and applications to finite element methods for elliptic boundary value problems. *East-West J. Numer. Math*, 9(3):179–198, 2001.

[5] C. Bacuta, J. H. Bramble, and J. Pasciak. Using finite element tools in proving shift theorems for elliptic boundary value problems. *Num. Lin. Alg. Appl.*, (1-2):33–64, 2003.

[6] C. Bacuta and J. Jacavage. A nonconforming saddle point least squares approach for elliptic interface problems. *arXiv:1808.10409 [math.NA]*.

[7] C. Bacuta, J. Jacavage, K. Qirko, and F.J. Sayas. Saddle point least squares iterative solvers for the time harmonic Maxwell equations. *Comput. Math. Appl.*, 74(11):2915–2928, 2017.

[8] C. Bacuta and P. Monk. Multilevel discretization of symmetric saddle point systems without the discrete LBB condition. *Appl. Numer. Math.*, 62(6):667–681, 2012.

[9] C. Bacuta and K. Qirko. A saddle point least squares approach to mixed methods. *Comput. Math. Appl.*, 70(12):2920–2932, 2015.

[10] C. Bacuta and K. Qirko. A saddle point least squares approach for primal mixed formulations of second order PDEs. *Comput. Math. Appl.*, 73(2):173–186, 2017.

[11] D. Braess. *Finite elements*. Cambridge University Press, Cambridge, third edition, 2007. Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker.

[12] J.H. Bramble and J.E. Pasciak. A new approximation technique for div-curl systems. *Math. Comp.*, 73:1739–1762, 2004.

[13] J.H. Bramble, J.E. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Math. Comp.*, 55(191):1–22, 1990.

[14] J.H. Bramble and X. Zhang. The analysis of multigrid methods. In *Handbook of numerical analysis, Vol. VII*, pages 173–415. North-Holland, Amsterdam, 2000.

[15] A. Cohen, W. Dahmen, and G. Welper. Adaptivity and variational stabilization for convection-diffusion equations. *ESAIM Math. Model. Numer. Anal.*, 46(5):1247–1273, 2012.

[16] L. Demkowicz and J. Gopalakrishnan. A class of discontinuous Petrov-Galerkin methods. Part I: the transport equation. *Comput. Methods Appl. Mech. Engrg.*, 199(23-24):1558–1572, 2010.

[17] L. Demkowicz and J. Gopalakrishnan. A primal DPG method without a first-order reformulation. *Comput. Math. Appl.*, 66(6):1058–1064, 2013.

[18] M.R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *J. Research Nat. Bur. Standards*, 49:409–436 (1953), 1952.

[19] M. Jung, S. Nicaise, and J. Tabka. Some multilevel methods on graded meshes. *Journal of Computational and Applied Mathematics*, 138(1):151 – 171, 2002.

[20] T. Kato. Estimation of iterated matrices, with application to the Von Neumann condition. *Numer. Math.*, 2:22–29, 1960.

[21] A. Veeser and P. Zanotti. Quasi-optimal nonconforming methods for symmetric elliptic problems. I—Abstract theory. *SIAM J. Numer. Anal.*, 56(3):1621–1642, 2018.
[22] A. Veeser and P. Zanotti. Quasi-optimal nonconforming methods for symmetric elliptic problems. III—Discontinuous Galerkin and other interior penalty methods. *SIAM J. Numer. Anal.*, 56(5):2871–2894, 2018.
[23] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34:581–613, 1992.
[24] J. Xu and J. Qin. Some remarks on a multigrid preconditioner. *SIAM Journal on Scientific Computing*, 15(1):172–184, 1994.
[25] J. Xu and L. Zikatanov. Some observations on Babuška and Brezzi theories. *Numer. Math.*, 94(1):195–202, 2003.
[26] X. Zhang. Multilevel schwarz methods. *Numerische Mathematik*, 63(1):521–539, Dec 1992.

University of Delaware, Department of Mathematics, 501 Ewing Hall 19716
E-mail address: bacuta@udel.edu

University of Delaware, Department of Mathematics, 501 Ewing Hall 19716
E-mail address: jjacav@udel.edu