A SADDLE POINT LEAST SQUARES APPROACH FOR PRIMAL MIXED FORMULATIONS OF SECOND ORDER PDES

CONSTANTIN BACUTA AND KLAJDI QIRKO

ABSTRACT. We present a Saddle Point Least Squares (SPLS) method for discretizing second order elliptic problems written as primal mixed variational formulations. A stability LBB condition and a data compatibility condition at the continuous level are automatically satisfied. The proposed discretization method follows a general SPLS approach and has the advantage that a discrete $\inf - \sup$ condition is automatically satisfied for standard choices of the test and trial spaces. For the proposed iterative processes a nodal basis for the trial space is not required. Efficient preconditioning techniques that involve inversion only on the test space can be considered. Stability and approximation properties for two choices of discrete spaces are investigated. Applications of the new approach include discretization of second order problems with highly oscillatory coefficient, interface problems, and higher order approximation of the flux for elliptic problems with smooth coefficients.

1. INTRODUCTION

The SPLS method for discretizing variational formulations with different types of test and trial spaces was introduced in [9]. The method is related with the Bramble-Pasciak least squares approach introduced in [15]. The SPLS method combines known theory and discretization techniques for approximating elliptic problems with theory and techniques for solving symmetric saddle point problems, [2, 5, 6, 13, 16, 17, 18, 31, 36, 40, 42, 44]. It provides a unified framework for discretizing variational formulations of PDEs formulated as first order differential equations or systems. Both the test and trial spaces for the SPLS discretization are conforming finite element spaces. The *test space* is *chosen first*, and the discrete *trial space* is built from the *test space, second*, using the action of the continuous differential operator (associated with the given problem) on the *test space*. For the proposed method, assembly of stiffness matrices for the trial spaces is avoided. A detailed review of the SPLS method is presented in the next section. In [9] the method is designed for discretization of *first order systems* such as $\operatorname{div} - \operatorname{curl}$ systems. The main goal of the paper is to show that

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

Key words and phrases. least squares, saddle point systems, mixed methods, multilevel methods, Uzawa type algorithms, conjugate gradient, cascadic algorithm, dual DPG.

The work was supported by NSF, DMS-1522454.

(SPLS) approach can be applied to efficiently solve primal mixed formulations for *second order elliptic problems.*

The paper is organized as follows. In Section 2, we introduce notation and review the SPLS approach as presented in [9]. In addition, we present stability and approximation properties for two choices of discrete spaces. In Section 3, we apply the general method to approximate solutions of second order elliptic problems with variable coefficients. In Section 4, we present numerical results obtained via SPLS discretization. After the conclusion of Section 5, we include Section 6 as an appendix, where the proposed Uzawa type iterative solvers are reviewed.

2. The general SPLS approach

The general problem that can be discretized using a SPLS is:
Find $p \in Q$ such that

$$b(v, p) = \langle f, v\rangle, \quad \text{for all } v \in V \quad \text{or} \quad B^*p = f, \tag{2.1}$$

where V and Q are infinite dimensional Hilbert spaces and $b(\cdot,\cdot)$ is a continuous bilinear form on $V \times Q$, that satisfies a standard $\inf - \sup$ condition and $f \in V^*$. We assume that on V and Q the inner products $a_0(\cdot,\cdot)$ and $(\cdot,\cdot)$ induce the norms $|\cdot|_V = |\cdot| = a_0(\cdot,\cdot)^{1/2}$ and $\|\cdot\|_Q = \|\cdot\| = (\cdot,\cdot)^{1/2}$, respectively. The dual pairings on $V^* \times V$ and $Q^* \times Q$ are denoted by $\langle\cdot,\cdot\rangle$. Here, V^* and Q^* denote the duals of V and Q, respectively. With the inner products $a_0(\cdot,\cdot)$ and $(\cdot,\cdot)$, we associate the operators $\mathcal{A} : V \to V^*$ and $\mathcal{C} : Q \to Q^*$ defined by

$$\langle \mathcal{A}u, v\rangle = a_0(u, v) \quad \text{for all } u, v \in V$$

and

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

The operators $\mathcal{A}^{-1} : V^* \to V$ and $\mathcal{C}^{-1} : Q^* \to Q$ are the Riesz-canonical isometries. We assume that $b(\cdot,\cdot)$ satisfies

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

and that the following $\inf - \sup$ condition holds,

$$\inf_{p\in Q} \sup_{v\in V} \frac{b(v, p)}{\|p\| \, |v|} = m > 0. \tag{2.3}$$

With the form b, we associate the linear operators $B : V \to Q^*$ and $B^* : Q \to V^*$ defined by

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

It is known that under the assumption (2.3), the operator $\mathcal{C}^{-1}B : V \to Q$ is onto, see e.g., [3]. We let V_0 be the kernel of B or $\mathcal{C}^{-1}B$, i.e.,

$$V_0 = Ker(B) = \{v \in V |\ Bv = 0\} = \{v \in V |\ \mathcal{C}^{-1}Bv = 0\}.$$

The existence and the uniqueness of (2.1) were first studied by Aziz and Babuška in [2]. It is well known that if a bounded form $b : V \times Q \to \mathbb{R}$ satisfies (2.3) and the data $f \in V^*$ satisfies the *compatibility condition*

$$\langle f, v\rangle = 0, \quad \text{for all } v \in V_0, \tag{2.4}$$

then, the problem (2.1) has a unique solution, see e.g. [2, 3].

With the problem (2.1) we associate the SPLS formulation:
Find $(u, p) \in (V, Q)$ such that

$$\begin{array}{llll} a_0(u, v) & + \quad b(v, p) & = \langle f, v\rangle & \quad \text{for all } v \in V, \\ b(u, q) & & = 0 & \quad \text{for all } q \in Q. \end{array} \tag{2.5}$$

The following statement summarizes the connection between the two variational formulations.

Proposition 2.1. *In the presence of the continuous* $\inf - \sup$ *condition* (2.3) *and the compatibility condition* (2.4)*, we have that* p *is the unique solution of* (2.1) *if and only if* $(u = 0, p)$ *is the unique solution of* (2.5)*.*

2.1. **SPLS discretization.** Due to Proposition 2.1, the *SPLS discretization* of (2.1) is defined as a standard saddle point discretization of (2.5). We let $V_h \subset V$ and $M_h \subset Q$ be finite dimensional approximation spaces and consider the restrictions of the forms $a_0(\cdot, \cdot)$ and $b(\cdot, \cdot)$ to the discrete spaces V_h and M_h. Assume that the following discrete $\inf - \sup$ condition holds for the pair (V_h, M_h).

$$\inf_{p_h \in M_h} \sup_{v_h \in V_h} \frac{b(v_h, p_h)}{\|p_h\| \, |v_h|} = m_h > 0. \tag{2.6}$$

We define $V_{h,0}$ to be the kernel of the discrete operator B_h, i.e.,

$$V_{h,0} := \{v_h \in V_h | \; b(v_h, q_h) = 0, \quad \text{for all } q_h \in M_h\}.$$

and let $V_{h,0}^{\perp}$ denote the orthogonal complement of $V_{h,0}$ with respect to $a_0(\cdot, \cdot)$ inner product on V_h.

Remark 2.2. *If* $V_{h,0} \subset V_0$*, then the compatibility condition* (2.4) *implies a discrete compatibility condition. Consequently, under the discrete stability assumption* (2.6)*, the problem: Find* $p_h \in M_h$ *such that*

$$b(v_h, p_h) = \langle f, v_h\rangle, \quad \textit{for all } v_h \in V_h, \tag{2.7}$$

has unique solution.

For a general choice of discrete spaces where compatibility condition (2.4) does not hold on $V_{h,0}$, the following standard saddle point discrete variational formulation: Find $(u_h, p_h) \in V_h \times M_h$ such that

$$\begin{array}{llll} a_0(u_h, v_h) & + \quad b(v_h, p_h) & = \langle f, v_h\rangle & \quad \text{for all } v_h \in V_h, \\ b(u_h, q_h) & & = 0 & \quad \text{for all } q_h \in M_h, \end{array} \tag{2.8}$$

does have a unique solution. We call the variational formulation (2.8) the *saddle point least squares discretrization* of (2.1). The reason we call it a *least squares method* is that the p_h part of the solution of (2.8) is in fact the solution of the normal equation associated with (2.7), see [9] for more details.

Remark 2.3. *We should mention that the SPLS formulation* (2.5) *and its discrete version* (2.8) *are connected with known minimum residual methods, as presented e.g., in the works of Dahmen, Huang, Schwab and Welper,* [26], *Demkowicz and Gopalakrishnan,* [27, 28], *and Bramble and Pasciak,* [15].

Using the extra compatibility condition (2.4), a sharp error estimate for $\|p - p_h\|$ was proved in [9] based on the Xu-Zikatanov argument, see [44].

Theorem 2.4. *Let $b : V \times Q \to \mathbb{R}$ satisfy* (2.3) *and* (2.2), *and assume that $f \in V^*$ is given and satisfies* (2.4). *Assume that p is the solution of* (2.1) *and $V_h \subset V$, $M_h \subset Q$ are chosen such that the discrete* $\inf - \sup$ *condition* (2.6) *holds. If (u_h, p_h) is the solution of* (2.8), *then the following error estimate holds:*

$$\frac{1}{M}|u_h| \leq \|p - p_h\| \leq \frac{M}{m_h} \inf_{q_h \in M_h} \|p - q_h\|. \tag{2.9}$$

The following observation leads to an efficient preconditioning approach for SPLS discretization.

Remark 2.5. *The considerations made regarding SPLS discretization in this section, make sense if the form $a_0(\cdot, \cdot)$, as an inner product on V_h, is replaced by another inner product which gives rise to an (independent of h) equivalent norm on V_h. The error estimate* (2.9) *remains valid with different estimating constants that factor in the norm equivalence constants. If A_h is the discrete operator associated with the form $a_0(\cdot, \cdot)$ on V_h and an Uzawa type algorithm is involved to solve* (2.8), *then the action of the discrete operator A_h^{-1} can be replaced by the action of any equivalent preconditioner.*

2.2. **Special iterative solvers.** One challenge we could face when solving the *least squares saddle point discretization* (2.8) is that one might not be able to find pairs $\{(V_h, M_h)\}$ that satisfy a (uniform or not) discrete $\inf - \sup$ condition. Even in the case when stable pairs are available, a global linear system corresponding to (2.8) might be difficult to assemble and solve. It is possible to solve (2.8) using inversion processes only on the test space V_h, without having explicit bases for the trial space M_h. The proposed idea in [9] is to use an iterative process such as the Uzawa (U), Uzawa Gradient (UG), or the Uzawa Conjugate Gradient (UCG), see the Appendix.

Each one of the described algorithms can be applied to approximate the solution (u_h, p_h) of (2.8). The following *sharp error estimation* result was proved by one of the authors in a slightly more general context in [4].

Theorem 2.6. *If (u_h, p_h) is the discrete solution of* (2.8), *and (u_{j+1}, p_j) is the j^{th} iteration for U, UG, or UCG, then $(u_{j+1}, p_j) \to (u_h, p_h)$ and*

$$\begin{aligned} \frac{1}{M^2}\,\|q_{j+1}\| \le \|p_j - p_h\| \quad &\le \frac{1}{m_h^2}\,\|q_{j+1}\|, \\ \frac{m_h}{M^2}\,\|q_{j+1}\| \le |u_{j+1} - u_h| &\le \frac{M}{m_h^2}\,\|q_{j+1}\|, \end{aligned} \tag{2.10}$$

where q_j is the orthogonal projection of $\mathcal{C}^{-1}Bu_j$ onto M_h.

Remark 2.7. *Theorem 2.6 entitles $\|q_{j+1}\|$ as a computable, robust, efficient, and uniform-modulo m_h estimator for the iteration error for all three algorithms.*

If the discretization error order is available, say $O(\|p-p_h\|)$, and an estimate for m_h is also available, the iteration error can match the discretization error by imposing the stopping criterion

$$\|q_{j+1}\| \le c_0\, m_h^2\, O(\|p-p_h\|), \tag{2.11}$$

where c_0 is a constant independent of h.

We note that if preconditioning is involved, i.e., the form $a_0(\cdot,\cdot)$ is replaced in (2.8) by $a_{prec}(\cdot,\cdot)$, Theorem 2.6 remains valid with a different constant m_h in (2.10) that depends on the quality of the preconditioner.

2.3. **The SPLS method.** In [9], we introduced the following five steps that define the *saddle point least squares discretization* method:

Step 1) Write the general problem (2.1) as a *saddle point least squares* formulation (2.5), using the natural inner product $a_0(\cdot,\cdot)$ on $V \times V$.

Step 2) Choose a standard *conforming approximation space* V_h for the *variational space* V.

Step 3) Construct a discrete *trial space* $M_h \subset Q$ using the operator B associated with the form $b(\cdot,\cdot)$. For example, take $M_h := \mathcal{C}^{-1}BV_h$, or $M_h := \tilde{\mathcal{Q}}_h\mathcal{C}^{-1}BV_h$, where $\mathcal{C}^{-1}$ is the Riesz representation operator for the space Q and $\tilde{\mathcal{Q}}_h$ is an *orthogonal projection* from Q to a subspace $\tilde{M}_h \subset Q$. The pair (V_h, M_h) will automatically satisfy a discrete $\inf-\sup$ condition.

Step 4) Write (2.8) - the discrete version of the SPLS formulation and, if available, replace $a_0(\cdot,\cdot)$ by an equivalent form $a_{prec}(\cdot,\cdot)$ on $V_h \times V_h$.

Step 5) Solve the new discrete SPLS problem using an Uzawa type iterative process that requires only the action of A_h^{-1} (or preconditioner), and the action of $\mathcal{C}^{-1}B$ or $\tilde{\mathcal{Q}}_h\mathcal{C}^{-1}B$ on functions in V_h.

2.4. **Special discrete spaces.** Let V_h be a *finite element subspace* of V. Assume that the action of $\mathcal{C}^{-1}$ at the continuous level is easy to obtain.

2.4.1. *No projection trial space.* The first choice for $M_h \subset Q$ is

$$M_h := \mathcal{C}^{-1}BV_h. \tag{2.12}$$

In this case, we have that $V_{h,0} \subset V_0$. *A discrete* $\inf-\sup$ *condition holds.* Indeed, using a generic function $p_h = \mathcal{C}^{-1}Bw_h \in M_h$, with $w_h \in V_{h,0}^{\perp}$, and the fact that $V_{h,0}^{\perp}$ is a finite dimensional space, we have

$$(*)\quad \begin{aligned} &\inf_{p_h\in M_h}\sup_{v_h\in V_h}\frac{b(v_h,p_h)}{\|p_h\|\;|v_h|} = \inf_{w_h\in V_{h,0}^{\perp}}\sup_{v_h\in V_h}\frac{(\mathcal{C}^{-1}Bv_h,\mathcal{C}^{-1}Bw_h)}{\|p_h\|\;|v_h|}\\ &\geq \inf_{w_h\in V_{h,0}^{\perp}}\frac{\|\mathcal{C}^{-1}Bw_h\|^2}{\|\mathcal{C}^{-1}Bw_h\|\;|v_h|} = \inf_{w_h\in V_{h,0}^{\perp}}\frac{\|\mathcal{C}^{-1}Bw_h\|}{|w_h|} := m_{h,0} > 0. \end{aligned}$$

Approximability: Using that $V_{h,0} \subset V_0$ and Remark 2.2, the variational formulation (2.7) is well posed, has unique solution $p_h \in M_h$, and by using Proposition 2.1 for the discrete pair (V_h, M_h), we have that $(u_h = 0, p_h)$ is the solution of (2.8). In this case, if p is the solution of (2.1) and p_h is the solution of (2.7) (or the SPLS solution of (2.8)), from (2.1) and (2.7) we obtain

$$0 = b(v_h, p - p_h) = (\mathcal{C}^{-1}Bv_h, p - p_h), \qquad \text{for all} \;\; v_h \in V_h.$$

Thus, we have that p_h is the orthogonal projection of p onto M_h, and consequently,

$$\|p - p_h\| = \inf_{q_h\in M_h}\|p - q_h\|. \tag{2.13}$$

Using that $\mathcal{C}^{-1}B$ is onto Q we can represent $p = \mathcal{C}^{-1}Bw$ for some $w \in V$, and write $q_h = \mathcal{C}^{-1}Bv_h$ for some $v_h \in V_h$. Thus, we have

$$\|p - p_h\| = \inf_{v_h\in V_h}\|\mathcal{C}^{-1}Bw - \mathcal{C}^{-1}Bv_h\| \leq M \inf_{v_h\in V_h}|w - v_h|. \tag{2.14}$$

The estimate (2.13) gives optimal approximability and is independent of the $\inf-\sup$ constant m_h, see also (2.9). The estimate (2.14) reveals that even though the trial space M_h is not chosen to be a standard approximation space, the approximability of the solution $p \in Q$ with discrete functions in M_h, reduces to approximability of the (best) representation $w \in V$ of p by discrete functions in the standard approximation test space V_h.

2.4.2. *Projection trial space.* Let $\tilde{M}_h$ be a finite dimensional subspace of Q that has good approximability properties. Typical examples of spaces $\tilde{M}_h$ are the spaces of piecewise polynomials. We equip $\tilde{M}_h$ with the restriction of the inner product on Q. If $\tilde{\mathcal{Q}}_h : Q \to \tilde{M}_h$ is the orthogonal projection onto $\tilde{M}_h$, we define the space M_h by

$$M_h := \tilde{\mathcal{Q}}_h\mathcal{C}^{-1}BV_h. \tag{2.15}$$

A discrete $\inf-\sup$ *condition always holds* for (V_h, M_h), see Section 5 of [9]. Regarding the stability of this type of discretization, we have the following:

Remark 2.8. *Assume that the following condition holds*

$$\|\tilde{\mathcal{Q}}_h q_h\| \geq \tilde{c}\,\|q_h\|, \qquad \textit{for all } q_h \in \mathcal{C}^{-1}BV_h, \tag{2.16}$$

with a constant $\tilde{c}$ independent of h. Then, $V_{h,0} \subset V_0$ and the variational formulation (2.7) *has a unique solution $p_h \in M_h$. Using Proposition 2.1 for the discrete pair (V_h, M_h), we have that $(u_h = 0, p_h)$ is the solution of* (2.8)*. The stability of the family $\{(V_h, \mathcal{C}^{-1}BV_h)\}$, which means that $m_{h,0}$ defined in $(*)$ satisfies $m_{h,0} > c_0 > 0$ for some positive constant c_0 independent of h, implies stability of the family $\{(V_h, \tilde{\mathcal{Q}}_h\mathcal{C}^{-1}BV_h)\}$.*

Next, we justify the last statement of Remark 2.8. Using a generic function $p_h = \tilde{\mathcal{Q}}_h\mathcal{C}^{-1}Bw_h \in M_h$, with $w_h \in V_{h,0}^{\perp}$ and (2.16), we have

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

where $m_{h,0}$ is defined in the *no projection trial space* subsection.

Approximability: Due to Theorem 2.4, in order to expect small discretization error $\|p - p_h\|$, besides stability, one needs to investigate the minimization problem $\inf_{q_h \in M_h} \|p - q_h\|$ in the special case when M_h is a proper subspace of $\tilde{M}_h$ and M_h might not be a standard approximation space for functions in Q. Using the surjectivity of $\mathcal{C}^{-1}B : V \to Q$, we can represent $p = \mathcal{C}^{-1}Bw$ for some $w \in V$, and write $q_h = \tilde{\mathcal{Q}}_h\mathcal{C}^{-1}Bv_h$ with $v_h \in V_h$. Next, we have

$$\begin{aligned} \|p - q_h\| &= \|\mathcal{C}^{-1}Bw - \tilde{\mathcal{Q}}_h\mathcal{C}^{-1}Bv_h\| \\ &\leq \|\mathcal{C}^{-1}Bw - \tilde{\mathcal{Q}}_h\mathcal{C}^{-1}Bw\| + \|\tilde{\mathcal{Q}}_h\mathcal{C}^{-1}Bw - \tilde{\mathcal{Q}}_h\mathcal{C}^{-1}Bv_h\| \\ &\leq \|\mathcal{C}^{-1}Bw - \tilde{\mathcal{Q}}_h\mathcal{C}^{-1}Bw\| + M|w - v_h|. \end{aligned} \tag{2.17}$$

Combining the above estimate with Theorem 2.4 for $p = \mathcal{C}^{-1}Bw$, we get

$$\|p - p_h\| \leq \frac{M}{m_h}\|\mathcal{C}^{-1}Bw - \tilde{\mathcal{Q}}_h\mathcal{C}^{-1}Bw\| + \frac{M^2}{m_h} \inf_{v_h \in V_h} |w - v_h|. \tag{2.18}$$

Consequently, in order to obtain good SPLS approximation for the solution p of (2.1), it would be enough to ask for some regularity of a solution w of $\mathcal{C}^{-1}Bw = p$, for approximation properties of the test space V_h, and for approximation properties of the projection $\tilde{\mathcal{Q}}_h : Q \to \tilde{M}_h$.

For our numerical experiments in Section 4, we observed higher order approximability than the estimate (2.18) suggests. While (2.9) could lead to a sharper estimate, due to the non standard choice of the space M_h, an order of approximation of $\|p - p_h\|$ obtained via the right part of (2.9) is more difficult to investigate.

3. SPLS for second order elliptic problems

In this section we apply the SPLS discretization method to discretize second order problems that can be formulated as:

Find $u \in H_0^1(\Omega)$ such that

$$-\operatorname{div}(A\nabla u) = f, \text{ in } \Omega, \tag{3.1}$$

where $\Omega \subset \mathbb{R}^n$, the data $f \in L^2(\Omega)$ is given, and the matrix A of (piecewise continuous) coefficients is known and satisfies

$$a_{min}|\xi|^2 \leq \langle A(x)\xi, \xi\rangle \leq a_{max}|\xi|^2, \quad \text{for all } x \in \Omega, \xi \in \mathbb{R}^n, \tag{3.2}$$

for some *positive* constants $a_{min} \leq a_{max}$. Here and in what follows, for vectors in $\mathbb{R}^k$, we denote the standard Euclidian inner product and norm by $\langle\cdot,\cdot\rangle$ and $|\cdot|$, respectively.

We also consider that (3.1) could model second order elliptic interface problems with a smooth interface $\Gamma \subset \Omega$. In this case the entries of the matrix A could be discontinuous across Γ, and we require that the normal component of the solution flux $n\cdot(A\nabla u)$ be continuous across Γ-with normal vector n.

In what follows, $(\cdot,\cdot)$ and $\|\cdot\|$ denote the standard L^2 inner product and norm for scalar or vector functions, and $|v| := \|\nabla v\|$ for any $v \in H_0^1(\Omega)$.

The primal mixed variational formulation that we consider is to:
Find $\sigma = A\nabla u$ with $u \in H_0^1(\Omega)$ such that

$$(\sigma, \nabla v) = (A\nabla u, \nabla v) = \langle f, v\rangle \quad \text{for all } v \in H_0^1(\Omega). \tag{3.3}$$

The formulation (3.3) is known as *primal mixed formulation* for second order elliptic problems. For the Laplace operator a *primal mixed formulation* is described in Chapter 5 of Braess' book, [14]. Other related works using *primal mixed formulation* can be found in [10, 11, 12, 33].

To fit the variational formulation (3.3) into the abstract formulation (2.1), we take $V := H_0^1(\Omega)$, $Q := A\nabla V$, define $b : V \times Q \to \mathbb{R}$ by

$$b(\sigma, v) := (\sigma, \nabla v) = (A\nabla u, \nabla v),$$

and let

$$\langle f, v\rangle := (f, v), \quad \text{for all } v \in V = H_0^1(\Omega).$$

We notice that for this setting, the representation $\sigma = \mathcal{C}^{-1}Bu = A\nabla u$ with $u \in H_0^1(\Omega)$ *is unique.*

On V we consider the standard inner product

$$a_0(u, v) := (\nabla u, \nabla v), \quad \text{for all } u, v \in V.$$

On $Q = A\nabla V$, we define the inner product

$$(\sigma, \tau)_Q = (A\nabla u, A\nabla v)_Q := (A\nabla u, A\nabla v)_{A^{-1}} := (A\nabla u, \nabla v). \tag{3.4}$$

One can immediately check that $B : V \to Q^*$ and $\mathcal{C}^{-1}B : V \to Q$ become

$$Bv = \nabla v, \text{ and } \mathcal{C}^{-1}Bv = A\nabla v, \quad \text{for all } v \in V.$$

We further note that

$$V_0 := Ker(B) = \{v \in V|\ Bv = 0\} = \{v \in H_0^1(\Omega)|\ \nabla v = 0\} = \{0\},$$

and the compatibility condition (2.4) is trivially satisfied. We also note that the continuity and the $\inf - \sup$ constants satisfy:

(3.5)
$$M = \sup_{\sigma = A\nabla u \in Q} \sup_{v \in V} \frac{b(v,\sigma)}{\|\sigma\| \, |v|} = \sup_{u \in V} \sup_{v \in V} \frac{(A\nabla u, \nabla v)}{(A\nabla u, \nabla u)^{1/2} \, |v|}$$
$$\leq \sup_{u \in V} \sup_{v \in V} \frac{\|A\nabla u\| \, \|\nabla v\|}{(A\nabla u, \nabla u)^{1/2} \, |v|} = \sup_{u \in V} \frac{\|A\nabla u\|}{(A\nabla u, \nabla u)^{1/2}} \leq \sqrt{a_{max}} < \infty,$$

and

$$\begin{aligned} m = & \inf_{\sigma = A\nabla u \in Q} \sup_{v \in V} \frac{b(v,\sigma)}{\|\sigma\| \, |v|} = \inf_{u \in V} \sup_{v \in V} \frac{(A\nabla u, \nabla v)}{(A\nabla u, \nabla u)^{1/2} \, |v|} \\ & \geq \inf_{u \in V} \frac{(A\nabla u, \nabla u)}{(A\nabla u, \nabla u)^{1/2} \, \|\nabla u\|} \geq \sqrt{a_{min}} > 0. \end{aligned} \tag{3.6}$$

Consequently, the mixed variational formulation (3.3) is well posed and suitable for a SPLS formulation and discretization.

3.1. **SPLS discretization for second order elliptic problems.** We take $V_h \subset V = H_0^1(\Omega)$ to be the space of *continuous* piecewise polynomials of degree m with respect to a regular mesh $\mathcal{T}_h$.

3.1.1. *The "no projection" trial space leads to stability and approximability.* The corresponding trial space $M_h \subset Q$ is

$$M_h := \mathcal{C}^{-1} B V_h = A\nabla V_h. \tag{3.7}$$

We do have *stability* in this case. More precisely, similar arguments used to establish (3.6) give

$$m_h := \inf_{\sigma_h = A\nabla u_h \in M_h} \sup_{v_h \in V_h} \frac{b(v_h, \sigma_h)}{\|\sigma_h\| \, |v_h|} \geq \sqrt{a_{min}} > 0.$$

The discrete mixed variational formulation in this case is to:
Find $\sigma_h = A\nabla u_h$, with $u_h \in V_h$ uniquely representing σ_h, such that

$$(\sigma_h, \nabla v_h) = (A\nabla u_h, \nabla v_h) = (f, v_h) \quad \text{for all } v_h \in V_h. \tag{3.8}$$

The SPLS discretization (2.8)-to be solved by an Uzawa type algorithm is:
Find $(w_h = 0, \sigma_h = A\nabla u_h) \in V_h \times M_h$ such that

$$\begin{array}{llll} (\nabla w_h, \nabla v_h) & + & (A\nabla u_h, \nabla v_h) & = (f, v_h) \quad \text{for all } v_h \in V_h, \\ A\nabla w_h & & & = 0. \end{array} \tag{3.9}$$

We note that the form $(\nabla \cdot, \nabla \cdot)$ in (3.9) can be replaced by $a_{prec}(\cdot,\cdot)$.

If $\sigma = A\nabla u$ is the solution of (3.3) and $\sigma_h = A\nabla u_h$ is the solution of (3.8) or the SPLS formulation (3.9), using also (3.5), the general estimate (2.13) gives,

$$\|\sigma - \sigma_h\|_Q = (A\nabla(u - u_h), \nabla(u - u_h))^{1/2} \leq \sqrt{a_{max}} \inf_{v_h \in V_h} \|\nabla(u - v_h)\|.$$

3.1.2. *The "projection" trial space.* Guided by the general theory of Section 2.4.2, we define $\tilde{M}_h$ as a finite dimensional subspace of $Q = A\nabla V \subset (L^2(\Omega))^n$, to be $\tilde{M}_h = A\, M_{h,0}$ were each component of $M_{h,0}$ consists of all *continuous* piecewise polynomials of degree m with respect to the mesh $\mathcal{T}_h$ (used to define V_h) with no restrictions on $\partial\Omega$. We equip $\tilde{M}_h$ with the restriction of the inner product on Q that is defined in (3.4). We let $\tilde{\mathcal{Q}}_h : Q \to \tilde{M}_h$ be the orthogonal projection onto $\tilde{M}_h$ and define the space M_h by

$$M_h := \tilde{\mathcal{Q}}_h \mathcal{C}^{-1} B V_h = \tilde{\mathcal{Q}}_h A \nabla V_h. \tag{3.10}$$

As discussed in Section 2.4.2, we do have a *discrete* $\inf - \sup$ *condition* for the pair (V_h, M_h), and approximability properties on M_h that depend on the approximability of the space V_h and the approximation quality of the orthogonal projection $\tilde{\mathcal{Q}}_h : A\nabla V_h \to \tilde{M}_h$. The discrete mixed variational formulation in this case is: Find $\sigma_h = \tilde{\mathcal{Q}}_h A \nabla u_h$ with $u_h \in V_h$ such that

$$(\sigma_h, \nabla v_h) = (\tilde{\mathcal{Q}}_h A \nabla u_h, \nabla v_h) = (f, v_h) \quad \text{for all } v_h \in V_h. \tag{3.11}$$

The SPLS discretization (2.8)-to be solved by an Uzawa type algorithm is: Find $(w_h, \sigma_h = \tilde{\mathcal{Q}}_h A \nabla u_h)$ such that

$$\begin{array}{llll} (\nabla w_h, \nabla v_h) & + & (\tilde{\mathcal{Q}}_h A \nabla u_h, \nabla v_h) & = (f, v_h) \quad \text{for all } v_h \in V_h, \\ \tilde{\mathcal{Q}}_h A \nabla w_h & & & = 0, \end{array} \tag{3.12}$$

where, again, the bilinear form $(\nabla\cdot, \nabla\cdot)$ can be replaced by $a_{prec}(\cdot,\cdot)$.

In order to discuss stability for families of pairs $\{(V_h, M_h)\}$, we will make further assumptions. We assume that Ω is a polygonal domain in $\mathbb{R}^2$ and the triangular mesh $\mathcal{T}_h$ is locally quasi-uniform. The result could be easily extended to polyhedral domains in $\mathbb{R}^3$, but to simplify the presentation, we will focus on the case $\Omega \subset \mathbb{R}^2$. We let $\{z_1, z_2, \cdots, z_N\}$ be the set of all nodes of $\mathcal{T}_h$ and assume that all triangles adjacent to z_j are of regular shape and their area is of order h_j^2. Then, the mesh size is $h := max\{h_1, h_2, \cdots h_N\}$. We further assume that $m = 1$, i.e. V_h consists of all *continuous* piecewise linear functions with respect to the mesh $\mathcal{T}_h$ that vanish on $\partial\Omega$, and each component of $M_{h,0}$, consists of all *continuous* piecewise linear functions with respect to the mesh $\mathcal{T}_h$ with no restrictions on the boundary. We let $\{\Phi_1, \Phi_2, \cdots \Phi_{2N}\}$ be the nodal basis for $M_{h,0}$ and assume for example that $\Phi_j = (\phi_j, 0)^t$ and $\Phi_{N+j} = (0, \phi_j)^t$, for $j = 1, 2, \cdots, N$, where $\{\phi_1, \phi_2, \cdots, \phi_N\}$ is the nodal basis for the space of scalar *continuous* piecewise linear functions with respect to the mesh $\mathcal{T}_h$. We note that $\{A\Phi_1, A\Phi_2, \cdots A\Phi_{2N}\}$ is a basis for $\tilde{M}_h$ and let M_A be the Gramm matrix of this basis with respect to the $(\cdot,\cdot)_Q$ inner product. We let D_1 be the diagonal matrix with diagonal entries $h_1^2, h_2^2, \cdots h_N^2$ and let D be the $2N \times 2N$ diagonal matrix, which is the 2×2 block diagonal matrix with $D1$ repeated as the diagonal blocks.

To prove (2.16) and conclude stability for $\{(V_h, M_h)\}$, we will prove first the following two lemmata.

Lemma 3.1. *Under the above assumptions we have*

$$\langle M_A\gamma, \gamma\rangle \le c\, a_{max} \langle D\gamma, \gamma\rangle, \quad \textit{for all } \gamma \in \mathbb{R}^{2N}. \tag{3.13}$$

Consequently,

$$\langle M_A^{-1}\gamma, \gamma\rangle \ge \frac{c}{a_{max}} \langle D^{-1}\gamma, \gamma\rangle, \quad \textit{for all } \gamma \in \mathbb{R}^{2N}. \tag{3.14}$$

Here, and in what follows, c is a generic constant that does not depend on h and can be different at different occurrences.

Proof. For any $\gamma \in \mathbb{R}^{2N}$, we define $q_h := \sum_{i=1}^{2N} \gamma_i \Phi_i$ and $q_h^A := Aq_h$. Next, we note that

$$\langle M_A\gamma, \gamma\rangle = (q_h^A, q_h^A)_Q = (Aq_h, q_h) \le a_{max}\|q_h\|^2 = a_{max} \sum_{\tau\in\mathcal{T}_h} \|q_h\|_\tau^2. \tag{3.15}$$

If $\tau = [z_{1_\tau}, z_{2_\tau}, z_{3_\tau}]$, then $q_h\big|_\tau = \begin{pmatrix} \sum_{j=1}^{3} \gamma_{j_\tau}\phi_{j_\tau} \\ \sum_{j=1}^{3} \gamma_{j_\tau+N}\ \phi_{j_\tau} \end{pmatrix}$, and we have:

$$\|q_h\|_\tau^2 \le c\,|\tau| \left(\sum_{j=1}^{3} \gamma_{\tau_j}^2 + \sum_{j=1}^{3} \gamma_{\tau_{j+N}}^2 \right). \tag{3.16}$$

Using (3.15), (3.16), and the fact that each coefficient γ_k can repeat at most three times, we get:

$$\langle M_A\gamma, \gamma\rangle \le c\, a_{\max} \left(\sum_{j=1}^{N} h_j^2\gamma_j^2 + h_j^2 \sum_{j=1}^{N} \gamma_{j+N}^2 \right) = c\, a_{max} \langle D\gamma, \gamma\rangle.$$

The estimate (3.14) follows from (3.13). ☐

Lemma 3.2. *Under the above assumptions, there exists a constant c independent of h such that*

$$\|\tilde{\mathcal{Q}}_h A\nabla v_h\|_Q \ge c\,\frac{a_{min}}{a_{max}}\ \|A\nabla v_h\|_Q \quad \textit{for all } v_h \in V_h. \tag{3.17}$$

Proof. For a fixed $A\nabla v_h$ with $v_h \in V_h$ we define the dual vector $G_h \in \mathbb{R}^{2N}$ by $(G_h)_i := (A\nabla v_h, A\Phi_i)_Q = (A\nabla v_h, \Phi_i)$, and let

$$\tilde{\mathcal{Q}}_h A\nabla v_h = \sum_{i=1}^{2N} \alpha_i A\Phi_i.$$

Thus, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{2N})^t$ is the solution of

$$M_A\,\alpha = G_h,$$

and, by using (3.14), we have

$$\|\tilde{\mathcal{Q}}_h A\nabla v_h\|_Q^2 = \sum_{i,j=1}^{2N} \alpha_i\,\alpha_j\,(A\Phi_i,\Phi_j) = \left\langle M_A^{-1} G_h, G_h\right\rangle$$
$$\geq \frac{1}{a_{max}}\left(D^{-1}G_h, G_h\right) \geq \frac{1}{a_{max}}\sum_{i=1}^{2N} h_i^{-2}\,(G_h, G_h)$$
$$= \frac{1}{a_{max}}\sum_{i=1}^{2N} h_i^{-2}\left[\left(a_{11}\frac{\partial v_h}{\partial x} + a_{12}\frac{\partial v_h}{\partial y}, \phi_i\right)^2 + \left(a_{21}\frac{\partial v_h}{\partial x} + a_{22}\frac{\partial v_h}{\partial y}, \phi_i\right)^2\right]$$
$$= \frac{1}{a_{max}}\sum_{i=1}^{2N} h_i^{-2}\sum_{\tau\subset supp(\phi_j)}\left|\begin{pmatrix} (a_{11},\phi_i)_\tau & (a_{12},\phi_i)_\tau \\ (a_{21},\phi_i)_\tau & (a_{22},\phi_i)_\tau\end{pmatrix}\begin{pmatrix}\frac{\partial v_h|_\tau}{\partial x} \\ \frac{\partial v_h|_\tau}{\partial y}\end{pmatrix}\right|^2$$
$$\overset{(*)}{\geq} \frac{a_{min}^2}{a_{max}}\sum_{i=1}^{2N}\sum_{\tau\subset supp(\phi_j)} h_i^2\|\nabla v_h\|_\tau^2 = c\frac{a_{min}^2}{a_{max}}\|\nabla v_h\|^2 \geq c\frac{a_{min}^2}{a_{max}^2}\|A\nabla v_h\|_Q^2.$$

Here, $\|\cdot\|_\tau$ denotes the L^2 - norm on τ. To justify the inequality $(*)$ above, we note that the lowest eigenvalue of the matrix $\begin{pmatrix} (a_{11},\phi_i)_\tau & (a_{12},\phi_i)_\tau \\ (a_{21},\phi_i)_\tau & (a_{22},\phi_i)_\tau\end{pmatrix}$ is bounded below by $c\,h_i^2 a_{min}^2$ with a constant c independent of τ and h. For the last estimate, we used that $\|A\nabla v_h\|_Q^2 = (A\nabla v_h, \nabla v_h) \leq a_{max}\|\nabla v_h\|^2$. □

As a direct consequence of Remark 2.8 and Lemma 3.2, we obatin:

Theorem 3.3. *Let $\Omega \subset \mathbb{R}^2$ be a polygonal domain and let $\{\mathcal{T}_h\}$ be a family of locally quasi-uniform triangular meshes for Ω. For each h let V_h be the space of continuous piecewise linear functions with respect to the mesh $\mathcal{T}_h$ that vanish on $\partial\Omega$, and let M_h be the corresponding projection space, as defined above in this section. Then, the family of spaces $\{(V_h, M_h)\}$ associated with the SPLS discretization* (3.12) *is stable.*

4. Numerical Examples

In this section, we apply the SPLS discretization method to second order PDEs on polygonal domains $\Omega \subset \mathbb{R}^2$. For all presented examples, we choose the test space $V_h \subset H_0^1(\Omega)$ to be the P_1 conforming space, i.e, the space of all *continuous* piecewise linear functions with respect to quasi uniform or locally quasi uniform meshes $\mathcal{T}_h$, and use the Uzawa conjugate gradient without preconditioning to obtain the numerical approximation of the flux $A\nabla u$.

The stopping criterion we imposed was based on (2.11) and the stability Theorem 3.3. The maximum possible order $O(\|A\nabla u - \sigma_h\|)$ was taken to be h^1 for the *no projection* case ($\sigma_h = A\nabla u_h$), and h^2 for the *projection* case ($\sigma_h = \tilde{\mathcal{Q}}_h A\nabla u_h$). For the non-uniform refinement cases, the parameter h

is defined as $h := N_{dof}^{-1/2}$, where N_{dof} is the number of degrees of freedom associated with the discrete test space V_h.

4.1. **A highly oscillatory coefficient example.** We consider the problem (3.1) on $\Omega = [-1, 1]^2$. We define f such that the exact solution is given by:

$$u(x, y) = \frac{\sin(\pi x)\sin(\pi y)}{\pi^2} + \frac{1}{n}\frac{\sin(n\pi x)\sin(n\pi y)}{\pi^2},$$

where the matrix A is given by:

$$A(x, y) = \left[\begin{array}{cc} 2 + \cos(n\pi y) & 0 \\ 0 & 2 + \cos(n\pi x) \end{array}\right].$$

We use SPLS discretization with both *no projection* and *projection* cases for M_h. For the numerical results shown on Table 1 with the *no projection* trial space we notice optimal rate of convergence (for n not too large). Here, $A\nabla u_c$ (the computed $A\nabla u$) is defined to be the iteration $A\nabla u_j$ with j representing the maximum number of iterations performed on level k due to the imposed stopping criterion. For the *projection* case on Table 2, we notice higher order than h^1 approximation of $A\nabla u$, using a small number of iterations.

$h = 2^{-k}$ k	$P_1 - \mathcal{C}^{-1}BP_1$, error = $\|A\nabla u - A\nabla u_c\|$								
	n = 4			n = 16			n = 32		
	error	*rate*	it	*error*	*rate*	it	*error*	*rate*	it
5	0.06314	0.95194	5	0.20942	0.24834	4	0.24469	0.33069	3
5	0.03184	0.98768	6	0.11937	0.81104	5	0.20870	0.22952	5
7	0.01596	0.99690	7	0.06186	0.94838	6	0.11915	0.80871	6
8	0.00798	0.99922	8	0.03121	0.98684	7	0.06177	0.94787	7

Table 1: Highly Oscillatory coefficients with no projection

$h = 2^{-k}$ k	$P_1 - \tilde{\mathcal{Q}}_h\mathcal{C}^{-1}B(P_1)$, error = $\|A\nabla u - \tilde{\mathcal{Q}}_h A\nabla u_c\|$								
	n = 4			n = 16			n = 32		
	error	*rate*	it	*error*	*rate*	it	*error*	*rate*	it
6	0.00166	1.91	6	0.0235	2.95	8	0.1834	12.82	15
7	0.00047	1.83	7	0.0051	2.21	11	0.0231	2.99	11
8	0.00014	1.74	8	0.0012	2.04	13	0.0049	2.22	17
9	0.00004	1.87	11	0.0003	1.99	16	0.0012	2.05	24

Table 2: Highly Oscillatory coefficients with orthogonal projection

4.2. **Interface problem examples.** Many approaches have been designed to efficiently approximate interface problems, [19, 23, 29, 30, 34, 37, 38, 39]. Next, we would like to illustrate with two examples that the SPLS discretization leads to simple and efficient discretization, too. For both examples, we implemented the *no projection* choice for M_h.

For the first example, we solve (3.1) with $\Omega = [0,1] \times [0,1]$ and a family of quasi uniform meshes $\{\mathcal{T}_h\}$ obtained by a standard uniform refinement strategy, starting with a uniform coarse mesh. We defined f such that for

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

with the interface $\Gamma := \Omega \cap \{(x,y)|\ x = 1/2\}$, the exact solution is

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

In Table 3, we compute the *error* $:= \|A\,\nabla u - A\,\nabla u_c\|$.

$h = 2^{-k}$	c = 4			c = 64			c = 1024		
k	*error*	*rate*	it	*error*	*rate*	it	*error*	*rate*	it
2	0.056	0.94	3	0.815	0.94	3	12.96	0.94	3
3	0.028	0.98	3	0.412	0.98	3	6.55	0.98	3
4	0.014	0.996	3	0.207	0.996	3	3.28	0.996	3
5	0.007	0.999	3	0.103	0.999	3	1.64	0.999	3
6	0.004	0.999	3	0.051	0.999	3	0.82	0.999	3

Table 3: Interface problem with jump discontinuity c across $x = \frac{1}{2}$

The experiments show an optimal convergence rate $O(h)$ in only three iterations, independent of the jump discontinuity c.

As a second example of interface problem, we consider Kellog's example [32], a standard benchmark problem with intersecting interfaces. For this example, we solve again the interface problem (3.1), on $\Omega = (-1,1)\times(-1,1)$ with the interface $\Gamma := \Omega \cap \{(x,y)|\ xy = 0\}$. The exact solution given in polar coordinates is:

$$u(r,\theta) = r^\beta(1-r^2)\mu(\theta)$$

with:

$$\mu(\theta) = \begin{cases} \cos(\beta(\frac{\pi}{2}-\sigma))\cos(\beta(\theta-\frac{\pi}{2}+\rho)) & \text{if } \theta \in \left[0, \frac{\pi}{2}\right] \\ \cos(\beta\rho)\cos(\beta(\theta-\pi+\sigma)) & \text{if } \theta \in \left[\frac{\pi}{2}, \pi\right] \\ \cos(\beta\sigma)\cos(\beta(\theta-\pi-\rho)) & \text{if } \theta \in \left[\pi, \frac{3\pi}{2}\right] \\ \cos(\beta(\frac{\pi}{2}-\rho))\cos(\beta(\theta-\frac{3\pi}{2}-\sigma)) & \text{if } \theta \in \left[\frac{3\pi}{2}, 2\pi\right]. \end{cases}$$

For

$$A(x,y) = a(x,y)I_2, \text{ where } a(x,y) = \begin{cases} R & \text{if } (x,y) \in [0,1]^2 \cup [-1,0]^2 \\ 1 & \text{if } (x,y) \in \Omega \backslash ([0,1]^2 \cup [-1,0]^2), \end{cases}$$

and the scalars $R,\ \beta,\ \rho,\ \sigma$ satisfying nonlinear relations designed by Kellogg, [32]. For numerical experiments, we test with:

$$\beta = 0.1,\ R \cong 161.4476387975881, \rho = \pi/4, \sigma \approx -14.92256510455152.$$

Special adaptive methods to deal with this (very) singular problem have been designed to obtain optimal approximation with piecewise linear functions, see e.g. [20, 22, 25]. Our SPLS approach produces good approximations of the solution for the simple refinement we consider. The family of locally quasi uniform meshes $\{\mathcal{T}_h\}$ is obtained by a graded refinement strategy, which depends on a refining parameter κ, [7, 8]. We refine by dividing all the edges that contain the singular point (the origin) under a fixed ratio κ such that the segment containing the singular point is κ - times the other segment. Numerical results using graded meshes with $\kappa = 0.022$ are shown in Table 4 below. Here $h = N^{-1/2}$, where N is the number of degrees of freedom for V_h on level k.

level k	$u - u_c$				$\|A\nabla u - A\nabla u_c\|$		
	L^2-err	$rate$	H^1-err	$rate$	L^2	$rate$	it
5	0.00181	0.99	0.16	0.39	0.25	0.49	17
6	0.00089	1.03	0.12	0.45	0.17	0.51	19
7	0.00043	1.05	0.08	0.49	0.12	0.53	24
8	0.00020	1.08	0.06	0.52	0.08	0.54	27
9	0.00009	1.09	0.04	0.53	0.06	0.55	31

Table 4: Interface problem with singularity at (0, 0)

The number of iterations needed is larger here because, by using the stopping criterion given by (2.11), to match the maximum possible order of the discretization error, we imposed $\|q_j\| \leq c_0 h\, m_h^2$. From Remark 2.8 and Lemma 3.2, we have $m_h \approx 1/R$, which leads to a very small iteration error.

We note that an optimal order of approximation for the flux was not achieved. This is because we used a simple way to refine (graded meshes) when changing levels. This technique takes care of the singularity caused by r^β, and can find optimal approximation for ∇u, which in this case is close to $h^{1/2}$. In other words, by using graded refinement, the best we can expect for the *discretization error* is close to order $h^{1/2}$. An adaptive method that can be combined with the SPLS discretization approach to obtain optimal approximation for the flux $A\nabla u$, for this example and similar interface problems, remains to be investigated.

4.3. **Higher order flux recovery example.** We solved (3.1) on $\Omega = [-1,1]^2 \setminus ((0,1)\times(0,-1))$ with $A = I$ and the data f computed such that the exact solution in polar coordinates is given by:

$$u = r^a \sin(a\theta)(1-r^2), \qquad \text{where } a = \frac{2}{3}.$$

The results obtained using graded meshes with $\kappa = 0.1$ are in Table 5.

Level	$u-u_c$					$\|\nabla u - Q_h\nabla u_c\|$		
	L^2-err	$rate$	H^1-err	$rate$	it	L^2	$rate$	it
3	0.01818	1.61	0.201	0.76	3	0.07368	1.81	3
4	0.00521	1.80	0.108	0.89	4	0.02622	1.49	4
5	0.00137	1.92	0.056	0.96	5	0.00864	1.60	5
6	0.00035	1.96	0.028	0.98	6	0.00287	1.59	6
7	0.00009	1.98	0.014	0.99	6	0.00098	1.55	6
8	0.00002	1.98	0.007	0.99	6	0.00033	1.55	8

Table 5: Flux recovery on L-shaped domain

We obtained a convergence order higher than $h^{\frac{3}{2}}$ for the flux, (where $h = N_{dof}^{-1/2}$). The presence of higher order *gradient recovery* for this singular solution example can be justified by Theorem 3.3 and the approximation estimate (2.9). Indeed, in this case $\tilde{M}_h$ is the space of all continuous piecewise linear vector functions as subspace of $(L^2(\Omega))^2$ and $\tilde{\mathcal{Q}}_h$ becomes the orthogonal projection $Q_h : (L^2(\Omega))^2 \to \tilde{M}_h$. Then, $M_h = Q_h \nabla V_h$. From Theorem 3.3 and (2.9), the discrete solution $\sigma_h = Q_h(\nabla u_h)$ of (3.12) satisfies

$$\|\nabla u - Q_h(\nabla u_h)\| \leq c \inf_{v_h \in V_h} \|\nabla u - Q_h(\nabla v_h)\|, \tag{4.1}$$

with a constant c independent of h. Due to known results for gradient recovery, see e.g., [1, 21, 41, 45], we have that if w_h is the Galerkin projection of u on V_h then, under mesh assumptions, we have that $\|\nabla u - Q_h \nabla w_h\|$ is of higher order than $\|\nabla u - \nabla w_h\|$. By using adapted or graded meshes, optimal order of approximation can be achieved for $\|\nabla u - \nabla w_h\|$. Thus, higher order than $h = N_{dof}^{-1/2}$ of the right hand side of (4.1) is expected. For various types of meshes, a proof for the exact order of the error $\|\nabla u - Q_h(\nabla u_h)\|$ with $Q_h(\nabla u_h)$ the SPLS discrete solution, remains to be investigated and it might be a difficult problem. Nevertheless, the important fact to notice here is that, due to (4.1), the SPLS discrete solution $Q_h(\nabla u_h)$ is a nearly optimal approximation of ∇u with functions in $M_h = Q_h \nabla V_h$, and the SPLS iteration process approximates $Q_h(\nabla u_h)$ well in few iterations. We plan to investigate optimal gradient approximation rates for the space $M_h = Q_h \nabla V_h$ and relate our findings with the works of [24, 35] in a future paper.

5. Conclusions

We presented applications of the *saddle point least squares* discretization method to second order PDEs. The proposed iterative solver can be computationally expensive if compared with some direct methods for solving the same problems. Nevertheless, we expect that the method will be efficient for problems with high degrees of freedom where inverting discrete Laplacian operator is slow and fast preconditioners can be used instead. In addition, specially in the case when a *projection trial space* is used, we observe higher order of approximation of the flux if we compare our approach with standard finite element approximation techniques that use linear elements.

We mention that the SPLS discretization method is a minimum residual approach, related with the work of Dahmen, Huang, Schwab and Welper, [26], Demkowicz and Gopalakrishnan, [27, 28], and Bramble and Pasciak, [15]. The essential novelty of the proposed discretization is in the way the discrete trial spaces are chosen and the iterative process that is involved in solving the system. The main goal was to show that the SPLS approach, which is easy to implement and suitable for standard preconditioning techniques, can be applied to solve efficiently a large class of second order problems with variable coefficients, including highly oscillatory coefficients and interface problems. We plan to apply the approach to solve other problems such as linear elasticity or Maxwell equations formulated as first order systems.

The proposed SPLS approach is different from the DPG formulations as presented in [27, 28], where a trial space is chosen first, and a close to optimal test space that provides stability of the pairs is chosen second. Due to the role the test and the trial space play and the order of choosing the discrete *test* and *trial* spaces in these two methods, the *SPLS discretization* can be viewed as *dual to the DPG* method.

6. Appendix

6.1. **Uzawa type iterative algorithms for SPP.** We review the Uzawa type iterative methods that are suitable for solving the SPLS discrete problem (2.8) in the special case when bases for the test space M_h are not available. Following [4, 9], we have that the standard U and UG algorithms can be rewritten such that they differ only by the way the parameter α is chosen. For the Uzawa algorithm, we have to choose a fixed number $\alpha = \alpha_0$ in the interval $\left(0, \frac{2}{M}\right)$. For the UG algorithm, the parameter α is chosen to impose the orthogonality of consecutive residuals associated with the second equation in (2.8). The first step for Uzawa is identical with the first step of UG. We combine the two algorithms in:

Algorithm 6.1. *(U-UG) Algorithms*

Step 1: Set $u_0 = 0 \in V_h$, $p_0 \in M_h$, **compute** $u_1 \in V_h$, $q_1 \in M_h$ *by*

$$\begin{aligned} a_0(u_1, v) &= \langle f_h, v\rangle - b(v, p_0), \qquad \text{for all } v \in V_h \\ (q_1, q) &= b(u_1, q), \qquad \text{for all } q \in M_h. \end{aligned}$$

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

$$
\begin{aligned}
&(\mathbf{U-UG1}) && a_0(h_j, v) = -\, b(v, q_j),\ v \in V_h\\
&(\mathbf{U}\alpha) && \alpha_j = \alpha_0 \textit{ for the Uzawa algorithm or}\\
&(\mathbf{UG}\alpha) && \alpha_j = -\frac{(q_j, q_j)}{b(h_j, q_j)} \textit{ for the UG algorithm}\\
&(\mathbf{U-UG2}) && p_j = p_{j-1} + \alpha_j\ q_j\\
&(\mathbf{U-UG3}) && u_{j+1} = u_j + \alpha_j\ h_j\\
&(\mathbf{U-UG4}) && (q_{j+1}, q) = b(u_{j+1}, q), \qquad \textit{for all } q \in M_h.
\end{aligned}
$$

Here, f_h is the restriction of f (as functional) to V_h. To obtain the UCG algorithm, the UG algorithm is modified as in [14, 43] as follows: First, we define $d_1 := q_1$ in **Step 1**, and then modify **Step 2** by replacing $b(\cdot, q_j)$ with $b(\cdot, d_j)$, where $\{d_j\}$ is a sequence of conjugate directions:

Algorithm 6.2. *(UCG) Algorithm*

Step 1: Set $u_0 = 0 \in V_h,\ p_0 \in M_h$. **Compute** $u_1 \in V_h,\ q_1, d_1 \in M_h$ *by*

$$
\begin{aligned}
a_0(u_1, v) &= \langle f_h, v\rangle - b(v, p_0),\ v \in V_h\\
(q_1, q) &= b(u_1, q), \qquad \textit{for all } q \in M_h, \quad d_1 := q_1.
\end{aligned}
$$

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*

$$
\begin{aligned}
&(\mathbf{UCG1}) && a_0(h_j, v) = -\, b(v, d_j),\ v \in V_h\\
&(\mathbf{UCG}\alpha) && \alpha_j = -\frac{(q_j, q_j)}{b(h_j, q_j)}\\
&(\mathbf{UCG2}) && p_j = p_{j-1} + \alpha_j\ d_j\\
&(\mathbf{UCG3}) && u_{j+1} = u_j + \alpha_j\ h_j\\
&(\mathbf{UCG4}) && (q_{j+1}, q) = b(u_{j+1}, q), \qquad \textit{for all } q \in M_h\\
&(\mathbf{UCG}\beta) && \beta_j = \frac{(q_{j+1}, q_{j+1})}{(q_j, q_j)}\\
&(\mathbf{UCG6}) && d_{j+1} = q_{j+1} + \beta_j d_j
\end{aligned}
$$

We note that for the *no projection trial space* case the q_{j+1} from **Step 1**, **(U-UG4)**, and **(UCG4)** can be computed by

$$q_{j+1} = \mathcal{C}^{-1} B u_{j+1},$$

and for the *projection trial space* case we have

$$q_{j+1} = \tilde{\mathcal{Q}}_h(\mathcal{C}^{-1} B u_{j+1}).$$

In both cases, M_h-bases for computing q_j are not needed, and at each iteration step *only one inversion* involving the form $a_0(\cdot, \cdot)$ is required.

References

[1] M. Ainsworth and J.T. Oden. *A Posteriori Error Estimation in Finite Element Analysis*. Wiley-Interscience, New York, 2000.

[2] 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.

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

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

[5] C. Bacuta, J. H. Bramble, and J. Xu. Regularity estimates for elliptic boundary value problems in Besov spaces. *Math. Comp.*, 72:1577 –1595, 2003.

[6] C. Bacuta, J. H. Bramble, and J. Xu. Regularity estimates for elliptic boundary value problems with smooth data on polygonal domains. *J. Numer. Math.*, 11(2):75–94, 2003.

[7] C. Bacuta, V. Nistor, and L. Zikatanov. Improving the rate of convergence of 'high order finite elements' on polygons and domains with cups. *Numerische Mathematik*, 100(2):165 –184, 2005.

[8] C. Bacuta, V. Nistor, and L. Zikatanov. Improving the rate of convergence of high-order finite elements on polyhedra. I. A priori estimates. *Numer. Funct. Anal. Optim.*, 26(6):613–639, 2005.

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

[10] E. Bank and J. Xu. Asymptotically exact a posteriori error estimators. I. Grids with superconvergence. *SIAM J. Numer. Anal.*, 41(6):2294–2312 (electronic), 2003.

[11] R. E. Bank and J. Xu. Asymptotically exact a posteriori error estimators. II. General unstructured grids. *SIAM J. Numer. Anal.*, 41(6):2313–2332 (electronic), 2003.

[12] G. R. Barrenechea, G. N. Gatica, and J.-M. Thomas. Primal mixed formulations for the coupling of FEM and BEM. I. Linear problems. *Numer. Funct. Anal. Optim.*, 19(1-2):7–32, 1998.

[13] M. Benzi, G. Golub, and J. Liesen. Numerical solutions of saddle point problems. *Acta Numerica*, pages 1–137, 2005.

[14] D. Braess. *Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, Cambridge, 1997.

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

[16] S. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, New York, 1994.

[17] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, 8(R-2):129–151, 1974.

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

[19] E. Burman, J. Guzman, M. A. Sanchez, and M. Sarkis. Robust flux error estimation of kitschs method for high contrast interface problems. *arXiv 1602.00603v1*, 2016.

[20] Zh. Cai, C. He, and Sh. Zhang. Finite element methods for interface problems: Robust residual a posteriori error estimates for discontinuous approximations. *arXiv 1510.06481v2*, 2016.

[21] Zh. Cai and Sh. Zhang. Flux recovery and a posteriori error estimators: conforming elements for scalar elliptic equations. *SIAM J. Numer. Anal.*, 48(2):578–602, 2010.

[22] Zh. Cai and Sh. Zhang. Recovery-based error estimators for interface problems: mixed and nonconforming finite elements. *SIAM J. Numer. Anal.*, 48(1):30–52, 2010.

[23] L. Chen, M. Holst, J. Xu, and Y. Zhu. Local multilevel preconditioners for elliptic equations with jump coefficients on bisection grids. *Comput. Vis. Sci.*, 15(5):271–289, 2012.

[24] L. Chen and H. Li. Superconvergence of gradient recovery schemes on graded meshes for corner singularities. *J. Comput. Math.*, 28(1):11–31, 2010.

[25] L. Chen and Ch. Zhang. A coarsening algorithm on adaptive grids by newest vertex bisection and its applications. *J. Comput. Math.*, 28(6):767–789, 2010.

[26] W. Dahmen, C. Huang, C. Schwab, and G. Welper. Adaptive Petrov-Galerkin methods for first order transport equations. *SIAM J. Numer. Anal.*, 50(5):2420–2445, 2012.

[27] 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.

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

[29] M. Dryja, M. V. Sarkis, and O. B. Widlund. Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. *Numer. Math.*, 72(3):313–348, 1996.

[30] R. Ewing, O. Iliev, and R. Lazarov. A modified finite volume approximation of second-order elliptic equations with discontinuous coefficients. *SIAM J. Sci. Comput.*, 23(4):1335–1351, 2001.

[31] V. Girault and P.A. Raviart. *Finite Element Methods for Navier-Stokes Equations*, volume 15. Springer-Verlag, Berlin, 1986.

[32] R. B. Kellogg. On the Poisson equation with intersecting interfaces. *Applicable Anal.*, 4:101–129, 1974/75. Collection of articles dedicated to Nikolai Ivanovich Muskhelishvili.

[33] D. Kim and E.-J. Park. Primal mixed finite-element approximation of elliptic equations with gradient nonlinearities. *Comput. Math. Appl.*, 51(5):793–804, 2006.

[34] A. Knyazev and O. Widlund. Lavrentiev regularization + Ritz approximation = uniform finite element error estimates for differential equations with rough coefficients. *Math. Comp.*, 72(241):17–40 (electronic), 2003.

[35] H. Li, A. Mazzucato, and V. Nistor. Analysis of the finite element method for transmission/mixed boundary value problems on general polygonal domains. *Electron. Trans. Numer. Anal.*, 37:41–69, 2010.

[36] P. Monk. *Finite element methods for Maxwell's equations*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2003.

[37] L. Mu, J. Wang, G. Wei, X. Ye, and Sh. Zhao. Weak Galerkin methods for second order elliptic interface problems. *J. Comput. Phys.*, 250:106–125, 2013.

[38] L. Mu, J. Wang, G. Wei, X. Ye, and Sh. Zhao. Weak Galerkin methods for second order elliptic interface problems. *J. Comput. Phys.*, 250:106–125, 2013.

[39] S. Nicaise and A.-M. Sändig. Transmission problems for the Laplace and elasticity operators: regularity and boundary integral formulation. *Math. Models Methods Appl. Sci.*, 9(6):855–898, 1999.

[40] A. Quarteroni and A. Valli. *Domain Decomposition Methods for Partial Differential Equations*. Clarendon Press, Oxford, 1999.

[41] R.E.Bank and J. Xu. Asymptotically exact a posteriori error estimators, part II: General unstructured grids. Technical Report AM-245, Department of Mathematics, The Pennsylvania State University, 2001. submitted to SIAM J. Numer. Anal.

[42] F.J. Sayas. Infimum-supremum. *Bol. Soc. Esp. Mat. Apl. S$\vec{e}$MA*, pages 19–40, 2007.

[43] R. Verfürth. A combined conjugate gradient-multigrid algorithm for the numerical solution of the Stokes problem. *IMA J. Numer. Anal.*, 4(4):441–455, 1984.

[44] J. Xu and L. Zikatanov. Some observations on Babuška and Brezzi theories. *Numer. Math.*, 94(1):195–202, 2003.

[45] Zh. Zhang and A. Naga. A new finite element gradient recovery method: superconvergence property. *SIAM J. Sci. Comput.*, 26(4):1192–1213 (electronic), 2005.

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: kqirko@udel.edu