Title

Multigrid Methods for Constrained Minimization Problems and Application to Saddle Point Problems

Permalink

https://escholarship.org/uc/item/0468g396

Author

Chen, Long

Publication Date

2016-01-15

Copyright Information

Peer reviewed

MULTIGRID METHODS FOR CONSTRAINED MINIMIZATION PROBLEMS AND APPLICATION TO SADDLE POINT PROBLEMS

LONG CHEN

ABSTRACT. The first order condition of the constrained minimization problem leads to a saddle point problem. A multigrid method using a multiplicative Schwarz smoother for saddle point problems can thus be interpreted as a successive subspace optimization method based on a multilevel decomposition of the constraint space. Convergence theory is developed for successive subspace optimization methods based on two assumptions on the space decomposition: stable decomposition and strengthened Cauchy-Schwarz inequality, and successfully applied to the saddle point systems arising from mixed finite element methods for Poisson and Stokes equations. Uniform convergence is obtained without the full regularity assumption of the underlying partial differential equations. As a byproduct, a V-cycle multigrid method for non-conforming finite elements is developed and proved to be uniform convergent with even one smoothing step.

1. INTRODUCTION

Given a quadratic energy $E(v)$ defined on a Hilbert space $\mathcal{V}$, we consider the constrained minimization problem:

$$\min_{v \in \mathcal{K}} E(v), \tag{1}$$

where $\mathcal{K} \subset \mathcal{V}$ is the null space of a linear and bounded operator B defined on $\mathcal{V}$. By introducing the Lagrange multiplier for the constraint, we can find the minimizer of (1) by solving a saddle point system. In this paper, we shall design and analyze multigrid methods for the constrained minimization problem (1) and apply them to the saddle point systems arisings from mixed finite element discretization of Poisson, Darcy, and Stokes equations.

We shall adapt the constraint decomposition methods developed by Tai for nonlinear variational inequalities [55] to the constrained minimization problem. Let $\mathcal{K} = \sum_{i=1}^{N} \mathcal{K}_i$ be a space decomposition. Our method consists of solving a local constrained minimization problem in each subspace $\mathcal{K}_i$ which is equivalent to solving a small saddle point problem. Thus our relaxation can be interpreted as a multiplicative overlapping Schwarz method which is known as Vanka smoother [58] in the context of computational fluid dynamics. With a proper multilevel decomposition, our method becomes the classical V-cycle multigrid method.

Assuming that the decomposition $\mathcal{K} = \sum_{i=1}^{N} \mathcal{K}_i$ satisfies two assumptions: energy stable decomposition (SD) and strengthened Cauchy-Schwarz inequality (SCS), we are

1

able to prove the convergence of our method

$$E(u^{k+1}) - E(u) \leq \left(1 - \frac{1}{1 + C_A C_S}\right) \left[E(u^k) - E(u)\right],$$

where u^k is the k-th iteration, and C_A and C_S are positive constants in (SD) and (SCS). We also extend the analysis to the case where the local constrained minimization problem is not solved exactly but one gradient iteration is applied.

It is known that numerically multiplicative Schwarz smoother leads to an efficient multigrid methods for saddle point problems [52, 53], however, theoretical analysis for the convergence is only available for less efficient additive versions [52, 53]. Our new framework can fill this gap. Furthermore, the optimal choice of the relaxation parameter used in the inexact solvers of local problems can be derived from the minimization point of view.

We then apply our method to the saddle point systems arising from mixed finite element methods of Poisson, Darcy, and Stokes equations. By verifying assumptions (SD) and (SCS) for multilevel decompositions of H(div) element spaces, we will prove the uniform convergence of a V-cycle multigrid method for mixed finite element methods for the Poisson and Darcy equations. Our smoother is related to the overlapping Schwarz method developed for H(div) problems in [32, 59, 42, 41, 3]. But our analysis from the energy minimization point of view is more transparent. We note that a similar stable multilevel decomposition for the Raviart-Thomas space has been proposed in [59] in two dimensions and in [34, 3] in three dimensions. Our decomposition for three dimensional case is new and does not require the duality argument and thus relax the full regularity assumption needed in [34, 3].

We use the equivalence between Crouzeix-Raviart (CR) non-conforming methods and mixed methods to develop a V-cycle multigrid method for non-conforming methods of Poisson equation and prove its uniform convergence. Existing convergence proofs of multigrid methods for non-conforming methods [11, 48, 15, 16, 49] cannot cover V-cycles with few smoothing steps while our new framework can. The two ingredients of our new multigrid method for non-conforming methods are: the overlapping Schwarz smoothers, and inter-grid transfer operators through the nested flux spaces.

For discrete Stokes equations, we apply our theory to divergence free and nested finite element spaces, e.g., Scott-Voligious elements [54]. Again traditional multigrid convergence proofs for Stokes equations requires the full regularity assumption [60, 12, 14, 8, 68, 47]. Using the framework developed in this paper, we can obtain multigrid convergence without the full regularity assumption. Very recently, Brenner, Li, and Sung [17] have developed new multigrid methods for Stokes equations and have proved the uniform convergence without the full regularity assumption. The convergence result of [17] is, however, restricted to W-cycle multigrid methods with sufficient many smoothing steps. Here we consider V-cycle multigrid with only one smoothing. Furthermore, smoothers developed in [17] are less efficient than Vanka-type smoothers considered here; see numerical examples in [52, 17]. On the other hand, the framework developed in [17] can be applied to any stable mixed finite element discretization of Stokes equation and in [18] such convergence theory is also extended to the Darcy systems, while the current theory can be only applied to the case when the constrained subspaces are nested. For non-nested constrained subspaces, an additional projector is needed and an analysis for W-cycle multigrid without the full regularity assumption can be found in [23]. For popular finite element pairs of Stokes equations, a fast multigrid method using least square distributive Gauss-Sedel smoother has been developed in [62] for Stokes equations and generalize to Oseen problem in [25].

Although most of the abstract theory, either based on the Xu-Zikatanov identity [67] or following the Tai-Xu approach [56], has been developed in certain form in the literature, the application to multigrid methods for solving saddle point systems are new and lead to several new contribution of the multigrid theory for saddle point systems: a convergence proof of V-cycle with even one smoothing step, a convergence proof without full regularity assumption, and a convergence proof for the multiplicative Schwarz smoother. Stable decomposition of several finite element spaces established in this paper also have their own interest.

The rest of this paper is structured as follows. In Section 2, we introduce the algorithm. In Section 3, we give a convergence proof using the X-Z identity and in Section 4, we present an alternative proof based on the constraint subspace optimization method. We extend the convergence proof to the inexact local solver in Section 5. In Section 6, 7, and 8, we apply our method to mixed finite element methods for the Poisson and Darcy equations, non-conforming finite element methods for the Poisson equation, and mixed finite element methods for the Stokes equations, respectively. In the last section, we give conclusion and outlook for future work.

2. Algorithm

Let $\mathcal{H}$ be a Hilbert space equipped with inner product $(\cdot, \cdot)$ and $\mathcal{V} \subset \mathcal{H}$ be a closed subspace and thus $\mathcal{V}$ is also a Hilbert space. Suppose $A : \mathcal{V} \to \mathcal{V}$ is a symmetric and positive definite (SPD) operator with respect to $(\cdot, \cdot)$, which introduces a new inner product $(u, v)_A := (Au, v) = (u, Av)$ on $\mathcal{V}$. The norm associated to $(\cdot, \cdot)$ or $(\cdot, \cdot)_A$ will be denoted by $\| \cdot \|$ or $\| \cdot \|_A$, respectively. Let $\mathcal{P}$ be another Hilbert space and let $B : \mathcal{V} \to \mathcal{P}$ be a linear operator. With a slight abuse of notation, we still denote the inner product of $\mathcal{P}$ by $(\cdot, \cdot)$. In most problems of consideration, the inner product $(\cdot, \cdot)$ for $\mathcal{H}$ is the vector L^2-inner product while for $\mathcal{P}$ it is the scalar L^2-inner product. The transpose $B^T : \mathcal{P} \to \mathcal{V}$ is the adjoint of B in the $(\cdot, \cdot)$ inner product, i.e., $(Bv, q) = (v, B^T q)$ for all $v \in \mathcal{V}, q \in \mathcal{P}$.

For an $f \in \mathcal{H}$, we define the Dirichlet-type energy:

$$(2) \qquad E(v) = \frac{1}{2}\|v\|_A^2 - (f, v), \quad \text{for } v \in \mathcal{V}.$$

In this paper we always identify a functional in the dual space $\mathcal{H}'$ as an element in $\mathcal{H}$ through the Riesz map induced by $(\cdot, \cdot)$. Denote by $\mathcal{K} = \ker(B)$ the subspace satisfying the constraint $Bv = 0$, i.e., the null space of B. We are interested in the following constrained minimization problem:

$$(3) \qquad \min_{v \in \mathcal{K}} E(v).$$

Since the energy is quadratic and convex, there exists a unique solution to (3) and the minimizer u of (3) is characterized as the solution of the equation: Find $u \in \mathcal{K}$ such that

$$(4) \qquad (Au, v) = (f, v) \quad \text{for all } v \in \mathcal{K}.$$

We introduce the operator $A_\mathcal{K} : \mathcal{K} \to \mathcal{K}$ as $(A_\mathcal{K} u, v) = (Au, v)$ for all $u, v \in \mathcal{K}$ and the operator $Q_\mathcal{K} : \mathcal{H} \to \mathcal{K}$ as the $(\cdot, \cdot)$-projection, i.e., for a given function $f \in \mathcal{H}$, $Q_\mathcal{K} f \in \mathcal{K}$ satisfies $(Q_\mathcal{K} f, v) = (f, v)$ for all $v \in \mathcal{K}$. Then the operator form of (4) is: Find $u \in \mathcal{K}$ such that

$$(5) \qquad A_\mathcal{K} u = Q_\mathcal{K} f \quad \text{in } \mathcal{K}.$$

As it might be difficult to find bases for the subspace $\mathcal{K}$, instead of solving the symmetric positive definite formulation (5), we shall consider an equivalent saddle point formulation.

Let us introduce the Lagrange multiplier $p \in \mathcal{P}$, equation (4) can be rewritten as the following saddle point system: Find $u \in \mathcal{V}, p \in \mathcal{P}$ such that

$$
\begin{aligned}
(Au, v) + (p, Bv) &= (f, v) && \text{for all } v \in \mathcal{V}, \\
(Bu, q) &= 0 && \text{for all } q \in \mathcal{P},
\end{aligned}
$$

which will be written in the operator form

$$
(6) \qquad \begin{pmatrix} A & B^T \\ B & O \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}.
$$

Let $\| \cdot \|_V$ and $\| \cdot \|_P$ be two appropriate norms for space $\mathcal{V}$ and $\mathcal{P}$, respectively. It is well known that (6) is well posed if and only if the following so-called Brezzi conditions [19] hold:

(1) Continuity of operators A and B: there exist constants $c_a, c_b > 0$ such that

$$
(Au, v) \leq c_a \|u\|_V \|v\|_V, \quad (Bv, q) \leq c_b \|v\|_V \|q\|_P, \quad \text{for all } u, v \in \mathcal{V}, q \in \mathcal{P}.
$$

(2) Coercivity of A in the kernel space. There exists a constant $\alpha > 0$ such that

$$
(Au, u) \geq \alpha \|u\|_V^2 \quad \text{for all } u \in \ker(B).
$$

(3) Inf-sup condition of B. There exists a constant $\beta > 0$ such that

$$
\inf_{p \in \mathcal{P}, p \neq 0} \sup_{v \in \mathcal{V}, \tau \neq 0} \frac{(Bv, p)}{\|v\|_V \|p\|_P} \geq \beta.
$$

Choices of norms $\| \cdot \|_V$ and $\| \cdot \|_P$ are not unique [69] and $\| \cdot \|_V = \| \cdot \|_A$ may not be always a good choice since B may not be continuous in $\| \cdot \|_A$ norm, c.f. the mixed formulation of Poisson equation in Section 6. Throughout this paper, we will assume the well-posedness of (6) and focus on its efficient solvers.

Problems (5) and (6) are equivalent theoretically but will lead to different algorithms. In practice, the saddle point formulation will be easier to solve when bases of $\mathcal{K}$ are not available or expensive to form.

We shall develop and analyze multigrid methods for solving the saddle point system (6) based on subspace correction methods [65] and its adaptation to optimization problems [56, 55]. Let

$$
\mathcal{V} = \mathcal{V}_1 + \mathcal{V}_2 + \cdots + \mathcal{V}_N, \ \mathcal{V}_i \subset \mathcal{V}, i = 1, \ldots, N,
$$

be a space decomposition of $\mathcal{V}$ satisfying the condition

$$
\mathcal{K} = \mathcal{K}_1 + \mathcal{K}_2 + \cdots + \mathcal{K}_N, \ \mathcal{K}_i = \mathcal{V}_i \cap \ker(B), i = 1, \ldots, N.
$$

For $k \geq 0$ and a given approximated solution $u^k \in \mathcal{K}$, one step of the Successive Subspace Optimization (SSO) method [56] is as follows:

Algorithm: $u^{k+1} = \mathrm{SSO}(u^k)$
$v_0 = u^k$;
for $i = 1 : N$ **do**
    $e_i = \operatorname{argmin}_{w_i \in \mathcal{K}_i} E(v_{i-1} + w_i)$;
    $v_i = v_{i-1} + e_i$;
end
$u^{k+1} = v_N$;

Algorithm 1: Successive Subspace Optimization Method.

If we write the Euler equation of the local minimization problem, it reads as

$$(7) \qquad (Ae_i, \phi_i) = (f - Av_{i-1}, \phi_i) \quad \text{for all } \phi_i \in \mathcal{K}_i.$$

Namely e_i is the solution of the residual equation restrict to $\mathcal{K}_i$. We can thus treat SSO as the subspace correction method for solving (4) using the space decomposition $\mathcal{K} = \sum_{i=1}^{N} \mathcal{K}_i$. We can analyze the convergence from this point of view.

Using the fact $Au = f$ in $\mathcal{K}'$ and $v_i = v_{i-1} + e_i$, equation (7) is also equivalent to the A-orthogonality

$$(8) \qquad (u - v_i, \phi_i)_A = 0 \quad \text{for all } \phi_i \in \mathcal{K}_i,$$

which can be also written as

$$(9) \qquad (E'(v_i), \phi_i) = 0 \quad \text{for all } \phi_i \in \mathcal{K}_i.$$

Let $\mathcal{P}_i = \mathcal{P} \cap B(\mathcal{V}_i)$. Define $A_i : \mathcal{V}_i \to \mathcal{V}_i$ as for $u_i \in \mathcal{V}_i$, $Au_i \in \mathcal{V}_i$ such that $(A_i u_i, v_i) = (Au_i, v_i)$ for all $v_i \in \mathcal{V}_i$, and $B_i : \mathcal{V}_i \to \mathcal{P}_i$ as for $u_i \in \mathcal{V}_i$, $Bu_i \in \mathcal{P}_i$ such that $(B_i u_i, q_i) = (Bu_i, q_i)$ for all $q_i \in \mathcal{P}_i$. Let $Q_i : \mathcal{H} \to \mathcal{V}_i$ be the projection in $(\cdot, \cdot)$ inner product. The constrained minimization problem in the constraint subspace $\mathcal{K}_i$ will be solved by solving a small saddle point system in $\mathcal{V}_i$:

$$(10) \qquad \begin{pmatrix} A_i & B_i^T \\ B_i & O \end{pmatrix} \begin{pmatrix} e_i \\ p_i \end{pmatrix} = \begin{pmatrix} Q_i(f - Av_{i-1}) \\ 0 \end{pmatrix}.$$

A typical multilevel decomposition is given as follows. First we construct a macro-decomposition $\mathcal{V} = \sum_{k=1}^{J} \mathcal{V}_k$ with nested subspaces $\mathcal{V}_1 \subset \mathcal{V}_2 \subset \ldots \subset \mathcal{V}_J = \mathcal{V}$. Usually they are based on a sequence of successively refined meshes. For each subspace $\mathcal{V}_k, k = 1, \ldots, J$, we introduce a micro-decomposition $\mathcal{V}_k = \sum_{i=1}^{N_k} \mathcal{V}_{k,i}$ and set $\mathcal{K}_{k,i} = \mathcal{V}_{k,i} \cap \ker(B)$. Note that the assumption $\mathcal{K} = \sum_{k=1}^{J} \sum_{i=1}^{N_k} \mathcal{K}_{k,i}$ requires a careful choice of the micro-decomposition of $\mathcal{V}_k$. Roughly speaking, each subspace $\mathcal{V}_{k,i}$ should be big enough to contain a basis function of $\mathcal{K}$ and each basis function of $\mathcal{K}$ should be contained in at least one $\mathcal{V}_{k,i}$. Similar decomposition is required to design robust multigrid methods for nearly singular system [38].

Remark 2.1. Solving local saddle problems in $\mathcal{V}_{k,i}$ sequentially in the k-th level can be interpret as a multiplicative Schwarz smoother which is better known as the Vanka smoother [58] for Navier-Stokes equations. $\qquad \square$

Due to the nestedness of the macro-decomposition, restriction and prolongation operators are needed only for two consecutive levels. In summary, SSO based on this multilevel decomposition leads to a V-cycle multigrid method for the saddle point problem (6) with a multiplicative Schwarz smoother.

Thanks to the assumption $\mathcal{K}_{k,i} \subset \mathcal{K}$, if $u^k \in \mathcal{K}$, then $u^{k+1} = \text{SSO}(u^k)$ is still in $\mathcal{K}$. Namely the iteration remains in the constrained subspace. Uzawa method [57], another popular iterative method for solving the saddle point problem, will not preserve the constraint and thus is not considered here.

We shall use either the unconstrained SPD formulation (4) and (7) or constrained saddle point formulation (6) and (10). They are equivalent forms for the convergence analysis but different algorithmically.

We end this section with a discussion of the non-homogenous constraint, i.e., the saddle point problem

$$(11) \qquad \begin{pmatrix} A & B^T \\ B & O \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}.$$

To change to the form (6), we can first find a $u_* \in \mathcal{V}$ satisfying $Bu_* = g$ and let $u = u_* + \delta u$. Then the equation for δu is in the form (6).

There are several ways to find such u_*. One way is to solve

$$(12) \qquad \begin{pmatrix} I & B^T \\ B & O \end{pmatrix} \begin{pmatrix} u_* \\ p_* \end{pmatrix} = \begin{pmatrix} 0 \\ g \end{pmatrix},$$

which is supposed to be easier than solving (11). For Stokes equations, solving (12) essentially requires a Poisson solver for pressure for which fast solvers are available. For Darcy equations, A is a weighted mass matrix with possibly highly oscillatory coefficients, while (12) is again just a Poisson operator.

When the space $\mathcal{P}$ consists of discontinuous elements, which is the case of most applications considered in this paper, we can find such u_* by one V-cycle with post-smoothing only; see Section 6 for details.

3. CONVERGENCE ANALYSIS BASED ON THE XZ IDENTITY

In this section, we provide a convergence analysis using the SPD formulation (4) and (7). The analysis is based on the XZ identity [67] for the multiplicative iterative methods and can be found in [66].

Denoted by P_i the A-orthogonal projection onto $\mathcal{K}_i$ for $i = 1, \ldots, N$. Then the error operator of SSO can be written as $(I - P_N)(I - P_{N-1}) \cdots (I - P_1)$, i.e., $u - u^{k+1} = \prod_{i=1}^{N}(I - P_i)(u - u^k)$, where $u^{k+1} = SSO(u^k)$. The following XZ identity was established in [67]

$$(13) \qquad \left\| \prod_{i=1}^{N}(I - P_i) \right\|_A^2 = 1 - \frac{1}{1 + c_0},$$

where

$$c_0 = \sup_{\|v\|_A = 1} \inf_{\sum_{i=1}^{J} v_i = v, v_i \in \mathcal{K}_i} \sum_{i=1}^{N} \left\| P_i \sum_{j=i+1}^{J} v_j \right\|_A^2.$$

For an elementary proof of (13), we refer to Chen [22].

In order to estimate the constant c_0, we propose two important properties of the space decomposition.

Stable decomposition (SD): for every $v \in \mathcal{K}$, there exists $v_i \in \mathcal{K}_i, i = 1, \ldots, N$ such that

$$v = \sum_{i=1}^{N} v_i, \quad \text{and} \quad \sum_{i=1}^{N} \|v_i\|_A^2 \le C_A \|v\|_A^2.$$

Strengthened Cauchy Schwarz inequality (SCS): for any $u_i \in \mathcal{K}_i$ and $v_j \in \mathcal{K}_j$

$$\sum_{i=1}^{N} \sum_{j=i+1}^{N} (u_i, v_j)_A \le C_S^{1/2} \left(\sum_{i=1}^{N} \|u_i\|_A^2 \right)^{1/2} \left(\sum_{j=1}^{N} \|v_j\|_A^2 \right)^{1/2}.$$

With assumptions (SD) and (SCS), we shall provide a upper bound of c_0 and thus obtain a convergence proof of SSO method for solving the saddle point problem (6).

Theorem 3.1. *Assume that the space decomposition $\mathcal{K} = \sum_{i=1}^{N} \mathcal{K}_i$ satisfy assumptions (SD) and (SCS). For SSO method, we have*

$$\left\| \prod_{i=1}^{N}(I - P_i) \right\|_A^2 \le 1 - \frac{1}{1 + C_A C_S}.$$

Proof. We apply (SCS) with $u_i = P_i \sum_{j=i+1}^{N} v_j$ to obtain

$$\sum_{i=1}^{N} \|u_i\|_A^2 = \sum_{i=1}^{N} (u_i, P_i \sum_{j=i+1}^{N} v_j)_A = \sum_{i=1}^{N} \sum_{j=i+1}^{N} (u_i, v_j)_A$$

$$\leq C_S^{1/2} \left(\sum_{i=1}^{N} \|u_i\|_A^2 \right)^{1/2} \left(\sum_{i=1}^{N} \|v_i\|_A^2 \right)^{1/2},$$

which leads to the inequality

(14)
$$\sum_{i=1}^{N} \|u_i\|_A^2 \leq C_S \sum_{i=1}^{N} \|v_i\|_A^2.$$

Consequently, we choose $v = \sum_{i=1}^{N} v_i$ as a stable decomposition satisfying (SD) to get

$$\sum_{i=1}^{N} \left\| P_i \sum_{j=i+1}^{N} v_j \right\|_A^2 = \sum_{i=1}^{N} \|u_i\|_A^2 \leq C_S \sum_{i=1}^{N} \|v_i\|_A^2 \leq C_S C_A \|v\|_A^2,$$

which implies $c_0 \leq C_S C_A$. The desired result then follows from the X-Z identity (13). $\quad\square$

The assumption (SCS) is relatively easy to verify. The key is to construct a stable decomposition of the constraint space $\mathcal{K}$.

4. CONVERGENCE ANALYSIS BASED ON CONSTRAINED OPTIMIZATION

In this section we provide an alternative proof using the constraint optimization approach established by Tai [55]. It also provides a better approach to extend the convergence proof to inexact and/or nonlinear local solvers.

We will always denote by u the global minimizer of (3). Given an initial guess $u^0 \in \mathcal{K}$, let u^k be the kth iteration in SSO algorithm for $k = 1, 2, \cdots$. We aim to prove a linear reduction of the energy difference

(15)
$$E(u^{k+1}) - E(u) \leq \rho \left[E(u^k) - E(u) \right],$$

with a contraction factor $\rho \in (0, 1)$. Ideally ρ is independent of the size of the problem. The proof is developed in [56, 55] for a nonlinear and convex energy but simplified here for the quadratic energy.

We first explore the relation between the energy and the A-norm of the error.

Lemma 4.1. *For any $w, v \in \mathcal{V}$, we have*

(16)
$$E(w) - E(v) = \frac{1}{2} \|w - v\|_A^2 + (E'(v), w - v).$$

Consequently for the minimizer $u \in \mathcal{K}$ and any $w \in \mathcal{K}$,

(17)
$$E(w) - E(u) = \frac{1}{2} \|w - u\|_A^2.$$

Proof. Verification of (16) and (17) is straightforward. $\quad\square$

Based on the identity (17), the target inequality (15) becomes a more familiar one

(18)
$$\|u^{k+1} - u\|_A \leq \rho^{1/2} \|u^k - u\|_A.$$

Let $d_k = E(u^k) - E(u)$ and $\delta_k = E(u^k) - E(u^{k+1})$. The quantity d_k is the distance of the current energy to the lowest one, δ_k is the amount of the energy decreased in one iteration, and they are connected by the identity $\delta_k = d_k - d_{k+1}$. By Lemma 4.1, we

have $d_k = \frac{1}{2}\|u^k - u\|_A^2$ but in general $\delta_k \neq \frac{1}{2}\|u^k - u^{k+1}\|_A^2$ since u^{k+1} may not be the minimizer. For each $v_i, i = 1, \ldots, N$, in SSO, we do have

$$E(v_{i-1}) - E(v_i) = \frac{1}{2}\|v_{i-1} - v_i\|_A^2,$$

since v_i is the local minimizer and $v_{i-1} - v_i = -e_i \in \mathcal{K}_i$; see also the orthogonality (9). Borrowing the terminology of the convergence theory of adaptive finite element methods [46], we shall present our proof based on the following two inequalities.

Discrete Lower Bound. There exists a positive constant C_L such that for $k = 0, 1, 2, \ldots$

$$\delta_k \geq C_L \sum_{i=1}^{N} \|e_i\|_A^2.$$

Upper Bound. There exists a positive constant C_U such that for $k = 0, 1, 2, \ldots$

$$d_{k+1} \leq C_U \sum_{i=1}^{N} \|e_i\|_A^2.$$

Theorem 4.2. *Assume that the discrete lower bound and upper bound hold with constants C_L and C_U respectively. We then have*

$$d_{k+1} \leq \frac{c_0}{1 + c_0} d_k,$$

where $c_0 = C_U/C_L$.

Proof. The proof is straightforward by assumptions and rearrangement of the following inequality

$$d_{k+1} \leq C_U \sum_{i=1}^{N} \|e_i\|_A^2 \leq C_U/C_L \delta_k = c_0(d_k - d_{k+1}).$$

$\square$

Verifying the lower bound is relatively easy since E is convex. Indeed we have the following identity which characterizes exactly the amount of energy decreased in one step of SSO. Again in the sequel, $u^{k+1} = \text{SSO}(u^k)$ and e_i is the ith correction in $\mathcal{K}_i$, for $i = 1, \ldots, N$.

Theorem 4.3.

$$E(u^k) - E(u^{k+1}) = \frac{1}{2} \sum_{i=1}^{N} \|e_i\|_A^2.$$

Proof. By the identity (16) and the orthogonality (9), we have, for $i = 1, \ldots, N$,

$$E(v_{i-1}) - E(v_i) = \frac{1}{2}\|v_{i-1} - v_i\|_A^2 = \frac{1}{2}\|e_i\|_A^2,$$

and consequently

$$E(u^k) - E(u^{k+1}) = \sum_{i=1}^{N} [E(v_{i-1}) - E(v_i)] = \frac{1}{2} \sum_{i=1}^{N} \|e_i\|_A^2.$$

$\square$

Proving the upper bound is more delicate. We first present a lemma which can be verified directly by definition and Lemma 4.1.

Lemma 4.4.

$$(19) \qquad (E'(u^{k+1}) - E'(u), u^{k+1} - u) = \|u^{k+1} - u\|_A^2 = 2\left[E(u^{k+1}) - E(u)\right].$$

We then give a multilevel decomposition of the left-hand side of (19).

Lemma 4.5. *For any decomposition* $u^{k+1} - u = \sum_{i=1}^{N} w_i, w_i \in \mathcal{K}_i, i = 1, 2, \ldots, N,$

$$(E'(u^{k+1}) - E'(u), u^{k+1} - u) = \sum_{i=1}^{N} \sum_{j>i}^{N} (e_j, w_i)_A.$$

Proof.

$$(E'(u^{k+1}) - E'(u), u^{k+1} - u) = (E'(u^{k+1}), u^{k+1} - u)$$

$$= \sum_{i=1}^{N} (E'(u^{k+1}) - E'(v_i), w_i) = \sum_{i=1}^{N} \sum_{j>i}^{N} (E'(v_j) - E'(v_{j-1}), w_i) = \sum_{i=1}^{N} \sum_{j>i}^{N} (e_j, w_i)_A.$$

In the first step, we use the fact $E'(u) = 0$ in $\mathcal{K}'$ since u is the minimizer and $u^{k+1} - u \in \mathcal{K}$. In the second step we use $E'(v_i) = 0$ in $\mathcal{K}'_i$ since v_i is the minimizer in $\mathcal{K}_i$ and $w_i \in \mathcal{K}_i$; see also (8). $\qquad \square$

Lemma 4.6. *Assume that the space decomposition satisfies assumptions (SD) and (SCS). Then we have the upper bound*

$$E(u^{k+1}) - E(u) \le \frac{1}{2} C_S C_A \sum_{i=1}^{N} \|e_i\|_A^2.$$

Proof. We shall chose a stable decomposition for $u^{k+1} - u = \sum_{i=1}^{N} w_i, w_i \in \mathcal{K}_i, i = 1, 2, \ldots, N$. By Lemma 4.5 and (SCS), we have

$$(E'(u^{k+1}) - E'(u), u^{k+1} - u) = \sum_{i=1}^{N} \sum_{j>i}^{N} (e_j, w_i)_A$$

$$\le C_S^{1/2} \left(\sum_{j=1}^{N} \|e_j\|_A^2 \right)^{1/2} \left(\sum_{i=1}^{N} \|w_i\|_A^2 \right)^{1/2}$$

$$\le (C_S C_A)^{1/2} \left(\sum_{i=1}^{N} \|e_i\|_A^2 \right)^{1/2} \|u^{k+1} - u\|_A.$$

Substituting the identity (see Lemma 4.4)

$$\|u^{k+1} - u\|_A^2 = (E'(u^{k+1}) - E'(u), u^{k+1} - u)$$

into the above inequality and canceling one $\|u^{k+1} - u\|_A$, we can obtain

$$\|u^{k+1} - u\|_A^2 \le C_S C_A \sum_{i=1}^{N} \|e_i\|_A^2.$$

Using the identity $E(u^{k+1}) - E(u) = \|u^{k+1} - u\|_A^2 / 2$, we obtain the desired result. $\qquad \square$

We summarize our convergence result into the following theorem.

Theorem 4.7. *Assume that the space decomposition $\mathcal{K} = \sum_{i=1}^{N} \mathcal{K}_i$ satisfies assumptions (SD) and (SCS). Then*

$$E(u^{k+1}) - E(u) \leq \left(1 - \frac{1}{1 + C_A C_S}\right) \left[E(u^k) - E(u)\right].$$

Remark 4.8. The estimate is consistent with the one obtained by the XZ identity which indicates that our energy estimate is sharp.

5. CONVERGENCE ANALYSIS WITH INEXACT LOCAL SOLVERS

In the algorithm SSO, we assume that the local problem is solved exactly which may be costly when the dimension of the local space is large. In this section, we consider inexact solvers using one gradient iteration and establish the corresponding convergence proof. Note that XZ identity cannot be applied to the nonlinear solvers considered here.

Recall that the local constrained minimization problem is: let $r_i = Q_i(f - Av_{i-1})$, find $e_i^* \in \mathcal{K}_i$ such that

$$(20) \qquad \begin{pmatrix} A_i & B_i^T \\ B_i & O \end{pmatrix} \begin{pmatrix} e_i^* \\ p_i^* \end{pmatrix} = \begin{pmatrix} r_i \\ 0 \end{pmatrix}.$$

Here we use e_i^*, p_i^* to denote the solution obtained by the exact solver. In the inexact solver proposed below, the constraint is still satisfied but operator A_i is replaced by a simpler one D_i, e.g., the diagonal of A_i. In general, let D_i be an SPD operator on $\mathcal{V}_i$, we first solve the local problem

$$(21) \qquad \begin{pmatrix} D_i & B_i^T \\ B_i & O \end{pmatrix} \begin{pmatrix} s_i \\ p_i \end{pmatrix} = \begin{pmatrix} r_i \\ 0 \end{pmatrix}.$$

Then we apply the line search along the direction s_i to find an optimal scaling:

$$(22) \qquad \min_{\alpha \in \mathbb{R}} E(v_{i-1} - \alpha s_i),$$

whose solution is

$$(23) \qquad \alpha = \frac{(r_i, s_i)}{(As_i, s_i)}.$$

We update

$$v_i = v_{i-1} - \alpha s_i.$$

This is one step of a preconditioned gradient method and D_i is a preconditioner of A_i.

In this section, we will always denote by e_i^* the solution of (20) and $e_i = \alpha s_i$ with s_i being the solution of (21) and α giving by (23). With such choice of α, we still have the first order condition

$$(24) \qquad (E'(v_i), e_i) = 0.$$

Remark 5.1. In the original Vanka smoother for Navier-Stokes equation, $D_i = \omega \operatorname{diag}(A_i)$ with a suitable parameter $\omega \in (0.5, 0.8)$ [58] and no line search is applied, i.e., $\alpha = 1$. $\square$

Using the first order condition (24), we still have the following identity.

Lemma 5.2.

$$E(u^k) - E(u^{k+1}) = \frac{1}{2} \sum_{i=1}^{N} \|e_i\|_A^2.$$

Again the upper bound is more delicate. We first adapt the analysis in [7] to establish the following inequalities. Recall that for an SPD operator M, $\kappa(M) = \lambda_{\max}(M)/\lambda_{\min}(M)$ is the condition number of M.

Lemma 5.3. *For the inexact local solver described above, we have*

$$(25) \qquad \|e_i^* - e_i\|_A \le \epsilon\|e_i^*\|_A, \quad \text{with } \epsilon = \frac{\kappa(D_i^{-1}A_i) - 1}{\kappa(D_i^{-1}A_i) + 1} \in (0, 1).$$

Consequently by the triangle inequality

$$(26) \qquad \|e_i^* - e_i\|_A \le \frac{\epsilon}{1 - \epsilon}\|e_i\|_A = \frac{1}{2}(\kappa(D_i^{-1}A_i) - 1)\|e_i\|_A.$$

Proof. To simplify the notation, we suppress the subscript i in the proof. Let $\tilde{e} = \omega s$ where s is determined by (21) and $\omega \in \mathbb{R}$ is a parameter. Then following [8, 7] we have the error equation

$$(27) \qquad e^* - \tilde{e} = P_D(I - \omega D^{-1}A)e^* = (I - \omega D^{-1}A_{\mathcal{K}})e^*,$$

where $P_D = I - D^{-1}B^T(BD^{-1}B^T)^{-1}B$ is the projection to $\mathcal{K}$ in the $(\cdot, \cdot)_D := (D\cdot, \cdot)$ inner product, and $A_{\mathcal{K}} = DP_D D^{-1}AP_D$.

Note that P_D is symmetric in $(\cdot, \cdot)_D$. We can then verify $A_{\mathcal{K}}$ is symmetric and semi-positive definite and $(\cdot, \cdot)_{A_{\mathcal{K}}} = (\cdot, \cdot)_A$ restricted to $\mathcal{K}$. Since the operator $D^{-1}A_{\mathcal{K}}$ is symmetric w.r.t. $(\cdot, \cdot)_{A_{\mathcal{K}}}$ and $e^*, \tilde{e} \in \mathcal{K}$, we have

$$\|e^* - \tilde{e}\|_A = \|e^* - \tilde{e}\|_{A_{\mathcal{K}}} \le \|I - \omega D^{-1}A_{\mathcal{K}}\|_{A_{\mathcal{K}}}\|e^*\|_A.$$

By subtracting a fixed energy $E(v_i^*)$ from $E(v_{i-1} - \alpha s_i)$, it is easy to see the line search (22) is equivalent to $\min_{\alpha \in \mathbb{R}} \|e_i^* - \alpha s_i\|_A$. Therefore

$$\|e^* - e\|_A \le \|e^* - \tilde{e}\|_A \le \|I - \omega D^{-1}A_{\mathcal{K}}\|_{A_{\mathcal{K}}}\|e^*\|_A.$$

Consequently

$$\|e^* - e\|_A \le \inf_{\omega \in \mathbb{R}}\|I - \omega D^{-1}A_{\mathcal{K}}\|_{A_{\mathcal{K}}}\|e^*\|_A = \frac{\kappa(D^{-1}A_{\mathcal{K}}) - 1}{\kappa(D^{-1}A_{\mathcal{K}}) + 1}\|e^*\|_A.$$

The condition number $\kappa(D^{-1}A_{\mathcal{K}})$, which is not easy to estimate since $A_{\mathcal{K}}$ is not formed explicitly, can be bounded by

$$\kappa(D^{-1}A_{\mathcal{K}}) = \kappa(P_D D^{-1}AP_D) \le \kappa(D^{-1}A).$$

$\square$

Lemma 5.4. *Assume that the space decomposition satisfies assumptions (SD) and (SCS). For SSO with the local in-exact solver described in this section, we have*

$$E(u^{k+1}) - E(u) \le \frac{1}{2}C_A\left[C_S^{1/2} + \frac{1}{2}\left(\max_{1 \le i \le N}\kappa(D_i^{-1}A_i) - 1\right)\right]^2 \sum_{i=1}^{N}\|e_i\|_A^2.$$

Proof. As before, we chose a stable decomposition for $u^{k+1} - u = \sum_{i=1}^{N} w_i, w_i \in \mathcal{K}_i, i = 1, 2, \ldots, N$ and split as

$$
\begin{aligned}
(E'(u^{k+1}) - E'(u), u^{k+1} - u) &= (E'(u^{k+1}), u^{k+1} - u) \\
&= \sum_{i=1}^{N} (E'(u^{k+1}) - E'(v_i), w_i) + (E'(v_i) - E'(v_i^*), w_i) \\
&= \sum_{i=1}^{N} \left[\sum_{j>i}^{N} (E'(v_j) - E'(v_{j-1}), w_i) + (v_i - v_i^*, w_i)_A \right].
\end{aligned}
$$

The first term can be bounded as before, c.f. Lemma 4.6

$$
\sum_{i=1}^{N} \sum_{j>i}^{N} (E'(v_j) - E'(v_{j-1}), w_i) \leq (C_S C_A)^{1/2} \left(\sum_{i=1}^{N} \|e_i\|_A^2 \right)^{1/2} \|u^{k+1} - u\|_A.
$$

For the second term, using (26), we have

$$
\begin{aligned}
\sum_{i=1}^{N} (v_i - v_i^*, w_i)_A &\leq \frac{\epsilon}{1-\epsilon} \sum_{i=1}^{N} \|e_i\|_A \|w_i\|_A \\
&\leq \frac{\epsilon}{1-\epsilon} \left(\sum_{i=1}^{N} \|e_i\|_A^2 \right)^{1/2} \left(\sum_{i=1}^{N} \|w_i\|_A^2 \right)^{1/2} \\
&\leq \frac{\epsilon}{1-\epsilon} C_A^{1/2} \left(\sum_{i=1}^{N} \|e_i\|_A^2 \right)^{1/2} \|u - u^{k+1}\|_A.
\end{aligned}
$$

Combining these two estimates, we then get the desired result. $\qquad\square$

Theorem 5.5. *Assume that the space decomposition satisfies assumptions (SD) and (SCS). For SSO with the local in-exact solver described in this section, we have*

$$
E(u^{k+1}) - E(u) \leq \rho \left[E(u^k) - E(u) \right],
$$

with contraction rate

$$
\rho = 1 - \frac{1}{1 + C_A \left[C_S^{1/2} + (\max_{1 \leq i \leq N} \kappa(D_i^{-1} A_i) - 1)/2 \right]^2}.
$$

We end this section with several remarks.

Remark 5.6. To be an efficient local solver, D_i is usually a diagonal matrix which may not be a good preconditioner for elliptic operators. The rate will deteriorate as ϵ becomes close to one, i.e., $\kappa(D_i^{-1} A_i) \gg 1$. On the other hand, for elliptic operators in $\mathbb{R}^n$, and for $D_i = \operatorname{diag}(A_i)$, we have estimate $\kappa(D_i^{-1} A_i) \lesssim \dim(\mathcal{V}_i)^{2/n}$ [5]. We can thus apply the estimate to a decomposition such that each local problem is of size $\mathcal{O}(1)$. $\square$

Remark 5.7. The solver considered here is one step of the preconditioned gradient method. The same analysis is applicable to a more efficient Preconditioned Conjugate-Gradient (PCG) solver with more than one iteration. The first order condition (24) still holds for the PCG iterations. $\square$

Remark 5.8. The local gradient method is a nonlinear iterative method since the parameter α depends on the iteration. To prove the energy contraction for the linear constraint smoother, i.e., with a fixed parameter α, we need to estimate the spectrum of the operator $P_D D^{-1} A P_D$ which is not easy since the projection P_D is in a L^2-type inner product not the A-inner product. Technically, the first order condition (24) may not hold for a fixed parameter α. $\square$

6. APPLICATION TO MIXED METHODS FOR POISSON AND DARCY EQUATION

In this section, we consider mixed finite element methods for solving the Poisson equation and Darcy equation in two and three dimensions. Let Ω be a polygon or polyhedron domain and triangulated into a quasi-uniform mesh $\mathcal{T}_h$ with mesh size h. Assume that $\mathcal{T}_h$ is obtained by uniform refinements from an initial mesh $\mathcal{T}_1$ of Ω, i.e., there exists a sequence of meshes $\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_J = \mathcal{T}_h$. The triangulation $\mathcal{T}_1$ is a shape regular triangulation of Ω and $\mathcal{T}_{k+1}$ is obtained by dividing each element in $\mathcal{T}_k$ into four congruent small elements (two dimensions) or eight small elements (three dimensions). The mesh size $\mathcal{T}_k$ will be denoted by h_k. By the construction $h_k/h_{k+1} = 2$.

6.1. **Problem setting.** We consider the Poisson equation with Neumann boundary

$$-\Delta p = f \text{ in } \Omega, \quad \partial_n p = 0 \text{ on } \partial\Omega$$

where n is the outwards normal vector of $\partial\Omega$. Let $u = \nabla p$. We obtain the mixed formulation of Poisson equation: find $u \in H_0(\text{div}; \Omega) := \{v \in (L^2(\Omega))^2, \text{div } v \in L^2(\Omega), v \cdot n|_{\partial\Omega} = 0\}$, where $v \cdot n$ should be understood in the trace sense, and $p \in L_0^2(\Omega) := \{q \in L^2(\Omega), \int_\Omega q \, dx = 0\}$ such that

$$(u, v) - (\text{div } v, p) = 0, \quad \forall v \in H_0(\text{div}; \Omega),$$
$$-(\text{div } u, q) = (f, q), \quad \forall q \in L_0^2(\Omega).$$

Choose finite element spaces $\mathcal{V} \subset H_0(\text{div}; \Omega)$ and $\mathcal{P} \subset L_0^2(\Omega)$ so that the following sequence is exact

$$(28) \qquad\qquad \mathcal{S} \xrightarrow{\text{curl}} \mathcal{V} \xrightarrow{\text{div}} \mathcal{P} \to 0,$$

where $\mathcal{S}$ is another appropriate finite element space. Choices of $\mathcal{S}, \mathcal{V}$, and $\mathcal{P}$ will be made clear in the context. Subscript k will be used when spaces are associated with triangulation $\mathcal{T}_k$ and when $k = J$ the subscript will be suppressed.

The saddle point problem can be written as follows: Given $f \in \mathcal{P}$, find $w \in \mathcal{V}, p \in \mathcal{P}$ such that

$$(29) \qquad\qquad \begin{pmatrix} M & B^T \\ B & O \end{pmatrix} \begin{pmatrix} w \\ p \end{pmatrix} = \begin{pmatrix} 0 \\ f \end{pmatrix},$$

where M is the mass matrix of $\mathcal{V}$ and B is the discretization of $-\text{div}$ operator. For this problem, $A = M$ and the A-norm is just the standard L^2-norm.

Our method and analysis can be readily adapted to the second order elliptic equation with variable coefficients K i.e., Darcy equation, for which the constitutive equation becomes $(K^{-1}u, v) - (\text{div } v, p) = 0$. The A-norm is a weighted L^2-norm and the exact sequence (28) still holds. The constant C_A, however, could depend on the condition number of K; see Remark 6.3.

To apply our framework, we should first find a $u_* \in \mathcal{V}$ satisfying $Bu_* = f$. Set $w = u_* + u$, the system (29) can be changed to the form of (6):

$$(30) \qquad \begin{pmatrix} M & B^T \\ B & O \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} -Mu_* \\ 0 \end{pmatrix}.$$

As discussed in Section 2, we can find such u_* by solving $BB^T u_* = Bf$. We now discuss a more efficient way utilizing the hierarchical structure of meshes. We start from a solution u_*^1 of (29) on the coarsest mesh $\mathcal{T}_1$ which can be found by direct solvers. For $k = 1, \ldots, J-1$, when u_*^k on $\mathcal{T}_k$ with property $Bu_*^k = f$ holds element-wise on $\mathcal{T}_k$ is found, for each element $T \in \mathcal{T}_k$, we solve (29) in $\mathcal{V}_{k+1}$ restricted to T and with boundary condition $u \cdot n|_{\partial T} = u_*^k \cdot n|_{\partial T}$. That is we use $\mathcal{T}_k$ to get a domain decomposition of $\mathcal{T}_{k+1}$ and u_*^k as the boundary condition to decompose a global problem into local problems on elements. The local problem is well defined since the compatible condition is enforced by $Bu_*^k = f$ on T and the solution of the local problem will give u_*^{k+1} with the property $Bu_*^{k+1} = f$ for each element in $\mathcal{T}_{k+1}$. The whole procedure is just one V-cycle with post-smoothing only and using a non-overlapping Schwarz method as a smoother. The computational cost is thus negligible.

Thanks to the exact sequence (28), we have a clear characterization of $\ker(B) = \mathrm{curl}\,(\mathcal{S})$ which will be helpful to construct a stable multilevel decomposition of $\mathcal{K}$. Based on the hierarchy of the meshes, we have a macro-decomposition of $\mathcal{S} = \sum_{k=1}^{J} \mathcal{S}_k$. For each space $\mathcal{S}_k$, we decompose into one dimensional subspaces $\Phi_{k,j}$ spanned by one basis function, i.e., $\mathcal{S}_k = \sum_{j=1}^{N_k} \Phi_{k,j}$ with $N_k = \dim \mathcal{S}_k$. Let $\Omega_{k,i}$ be the support of $\Phi_{k,i}$. We chose $\mathcal{V}_{k,i} = H_0(\mathrm{div}; \Omega_{k,i}) \cap \mathcal{V}_k$ for $i = 1, \ldots, N_k$. Then we have the decomposition

$$(31) \qquad \mathcal{V} = \sum_{k=1}^{J} \sum_{i=1}^{N_k} \mathcal{V}_{k,i},$$

and

$$\mathcal{K} = \sum_{k=1}^{J} \sum_{i=1}^{N_k} \mathcal{K}_{k,i} \text{ with } \mathcal{K}_{k,i} = \mathrm{curl}\,\Phi_{k,i}.$$

We shall apply SSO based on the space decomposition $\mathcal{K} = \sum_{k=1}^{J} \sum_{i=1}^{N_k} \mathcal{K}_{k,i}$ and prove its uniform convergence. As an example, for the lowest order RT element [50], the smoother is an overlapping multiplicative Schwarz smoother requiring solving a small saddle point system at in the patch of each vertex in two dimensions and in the patch of each edge in three dimensions. Since the construction of a stable decomposition in two and three dimensions is different, we split the discussion into two subsections.

Remark 6.1. One can use a basis of $\mathcal{S}$ to reduce the saddle point system into a SPD one and develop multigrid methods or domain decomposition methods for the SPD formulation; see e.g. [32, 34, 20]. □

6.2. **Two dimensions.** In two dimensions, the space $\mathcal{S} \subset H_0^1(\Omega)$ is a Lagrange element space based on the mesh $\mathcal{T}_h$. To be specific, we will consider the important case when $\mathcal{S}$ is the simplest linear finite element space, $\mathcal{V}$ is the lowest order Raviart-Thomas element space [50], and $\mathcal{P}$ is the piecewise constant space. Extension to high order elements is straightforward. The subspace $\mathcal{V}_{k,i}$ is spanned by basis vectors of edges connecting to the ith vertex in triangulation $\mathcal{T}_k$.

We first verify the stable decomposition for the macro-decomposition $\mathcal{K} = \sum_{k=1}^{J} \mathcal{K}_k$. We denoted by Q_k the L^2 projection $Q_k : \mathcal{S}_J \to \mathcal{S}_k$ for $k = 1, \ldots, J$ and set $Q_0 = 0$. Note that due to the nestedness $Q_l Q_k = Q_l$ for $l \leq k$.

Lemma 6.2. *For every $v \in \mathcal{K}$, there exists $v_k \in \mathcal{K}_i, k = 1, \ldots, J$ such that $v = \sum_{k=1}^{J} v_k$ and $\sum_{k=1}^{J} \|v_k\|^2 \lesssim \|v\|^2$.*

Proof. In two dimensions, we have the relation $(\operatorname{curl}\phi, \operatorname{curl}\psi) = (\nabla\phi, \nabla\psi)$. Therefore the stable decomposition (SD) comes from that for the Lagrange elements. More specifically, since $u \in \mathcal{K}$, there exists a unique $\phi \in \mathcal{S}$ such that $u = \operatorname{curl}\phi$. We then chose the H^1-stable decomposition of ϕ as $\phi = \sum_{k=1}^{J}(Q_k - Q_{k-1})\phi$ and let $u_k = \operatorname{curl}(Q_k - Q_{k-1})\phi$. The stable decomposition for the decomposition $u = \sum_k u_k$ in M-norm is equivalent to that of $\phi = \sum_{k=1}^{J}(Q_k - Q_{k-1})\phi$ in H^1-norm which is well known; see e.g. [65]. $\qquad\square$

Remark 6.3. For Darcy equation with variable coefficients K, the A-norm of v is changed to a weighted H^1 norm of ϕ for $v = \operatorname{curl}\phi$. If assuming K is piecewise constant on the coarsest mesh, we can find a multilevel decomposition using hierarchical basis such that the inequality $\sum_{k=1}^{J} \|v_k\|^2 \leq C|\log h|\|v\|^2$ holds with a penalty factor $|\log h|$ but with a constant C independent of the variation of K; see [4]. $\qquad\square$

We then verify the micro-decomposition is stable.

Lemma 6.4. *Let $\phi_k = (Q_k - Q_{k-1})\phi = \sum_{i=1}^{N_k} \phi_{k,i}$ be the nodal basis decomposition and let $u_{k,i} = \operatorname{curl}\phi_{k,i}$. Then the decomposition $u_k = \sum_{i=1}^{N_k} u_{k,i}$ is stable in L^2-norm.*

Proof. We apply the inverse inequality and the stability of the nodal basis decomposition in L^2-norm to get

$$\sum_{i=1}^{N_k} \|u_{k,i}\|^2 = \sum_{i=1}^{N_k} \|\operatorname{curl}\phi_{k,i}\|^2 \lesssim h_k^{-2} \sum_{i=1}^{N_k} \|\phi_{k,i}\|^2 \lesssim h_k^{-2}\|\phi_k\|^2.$$

We write the term $\phi_k = (Q_k - Q_{k-1})\phi = (I - Q_{k-1})(Q_k - Q_{k-1})\phi$ and bound it as

$$\|\phi_k\| \lesssim h_k\|\operatorname{curl}(Q_k - Q_{k-1})\phi\| = h_k\|u_k\|.$$

The desired inequality then follows. $\qquad\square$

Combination of Lemma 6.2 and 6.4 leads to a stable multilevel decomposition.

Theorem 6.5. *For every $v \in \mathcal{K}$, there exists $v_{k,i} \in \mathcal{K}_{k,i}, k = 1, \ldots, J, i = 1, \ldots, N_k$ such that $v = \sum_{k=1}^{J} \sum_{i=1}^{N_k} v_{k,i}$ and $\sum_{k=1}^{J} \sum_{i=1}^{N_k} \|v_{k,i}\|^2 \lesssim \|v\|^2$.*

To verify assumption (SCS), we first present the following inequality and refer to [65] for a proof.

Lemma 6.6. *For any $\phi_k \in \mathcal{S}_k, \phi_l \in \mathcal{S}_l, l \geq k$, we have*

$$(\operatorname{curl}\phi_k, \operatorname{curl}\phi_l) \lesssim \left(\frac{1}{2}\right)^{l-k} \|\operatorname{curl}\phi_k\|h_l^{-1}\|\phi_l\|.$$

We use the lexicographical order of the double index, i.e., $(l, j) > (k, i)$ if $l > k$ or $l = k, j > i$.

Theorem 6.7. *For any $u_{k,i} \in \mathcal{K}_{k,i}$ and $v_{l,j} \in \mathcal{K}_{l,j}$, we have*

$$\sum_{k=1}^{J}\sum_{i=1}^{N_k}\sum_{(l,j)>(k,i)} (u_{k,i}, v_{l,j}) \lesssim \left(\sum_{k=1}^{J}\sum_{i=1}^{N_k}\|u_{k,i}\|^2\right)^{1/2} \left(\sum_{l=1}^{J}\sum_{j=1}^{N_l}\|v_{l,j}\|^2\right)^{1/2}.$$

Proof. We can write $u_{k,i} = \operatorname{curl}\phi_{k,i}$ and $v_{l,j} = \operatorname{curl}\psi_{l,j}$ for some $\phi_{k,i} \in \mathcal{S}_k, \psi_{l,j} \in \mathcal{S}_l$. We split the summation $\sum_{(l,j)>(k,i)}$ into two parts $\sum_{l>k}\sum_{j=1}^{N_l}$ and $\sum_{l=k,j>i}^{N_k}$. For the first part, we apply Lemma 6.6 and note that $h_l^{-1}\|\psi_{l,j}\| \approx \|\operatorname{curl}\psi_{l,j}\|$ to get

$$\begin{aligned}
\sum_{k=1}^{J}\sum_{i=1}^{N_k}\sum_{l>k}\sum_{j=1}^{N_l}(u_{k,i}, v_{l,j}) &= \sum_{k=1}^{J}\sum_{i=1}^{N_k}\sum_{l>k}\sum_{j=1}^{N_l}(\operatorname{curl}\phi_{k,i}, \operatorname{curl}\psi_{l,j}) \\
&\leq \sum_{k=1}^{J}\sum_{i=1}^{N_k}\sum_{l>k}\sum_{j=1}^{N_l}\left(\frac{1}{2}\right)^{l-k}\|\operatorname{curl}\phi_{k,i}\|\|\operatorname{curl}\psi_{l,j}\| \\
&\lesssim \left(\sum_{k=1}^{J}\sum_{i=1}^{N_k}\|u_{k,i}\|^2\right)^{1/2}\left(\sum_{l=1}^{J}\sum_{j=1}^{N_l}\|v_{l,j}\|^2\right)^{1/2}.
\end{aligned}$$

For the second part, we use the finite overlapping property of finite element spaces. Namely, in the kth level, the index set $n_k(i) = \{j \in \{1,\ldots,N_k\}, \Omega_{k,i}\cap\Omega_{k,j} \neq \emptyset\}$ is finite. Then

$$\begin{aligned}
\sum_{k=1}^{J}\sum_{i=1}^{N_k}\sum_{j>i}^{N_k}(u_{k,i}, v_{k,j}) &= \sum_{k=1}^{J}\sum_{i=1}^{N_k}\sum_{j\in n_k(i)} (u_{k,i}, v_{k,j}) \\
&\lesssim \left(\sum_{k=1}^{J}\sum_{i=1}^{N_k}\|u_{k,i}\|^2\right)^{1/2}\left(\sum_{l=1}^{J}\sum_{j=1}^{N_l}\|v_{l,j}\|^2\right)^{1/2}.
\end{aligned}$$

$\square$

6.3. **Three dimensions.** We consider the same problem in three dimensions which is much more difficult than the two dimensional case. The reason is that the previous space $\mathcal{S}$ is an edge element space and a stable multilevel decomposition for $\mathcal{S}$ is non-trivial.

We again consider the lowest order case. Now $\mathcal{S}$ is the lowest order Nédélec edge element space [44, 45] of $H_0(\operatorname{curl},\Omega) := \{v \in (L^2(\Omega))^3, \operatorname{curl} v \in (L^2(\Omega))^3, v \times n|_{\partial\Omega} = 0\}$, $\mathcal{V}$ is the lowest order Raviart-Thomas element space of $H_0(\operatorname{div},\Omega)$, and $\mathcal{P} \subset L_0^2(\Omega)$ is the piecewise constant space. Furthermore let $\mathcal{U} \subset H_0^1(\Omega)$ be the linear finite element space. We have the following exact sequence [33, 1]

$$0 \hookrightarrow \mathcal{U} \xrightarrow{\operatorname{grad}} \mathcal{S} \xrightarrow{\operatorname{curl}} \mathcal{V} \xrightarrow{\operatorname{div}} \mathcal{P} \to 0.$$

To verify (SD), we need the following discrete regular decomposition for edge elements [35]. In the sequel, the operator $\Pi_h^{\operatorname{curl}}$ is the canonical interpolation to $\mathcal{S}$: for a smooth enough function w, $\Pi_h^{\operatorname{curl}} w \in S$ satisfying $\int_E \Pi_h^{\operatorname{curl}} w \cdot t\,\mathrm{d}s = \int_E w \cdot t\,\mathrm{d}s$ for all edges E of $\mathcal{T}_h$ where t is a tangential vector of E. Similarly $\Pi_k^{\operatorname{curl}}$ is the canonical interpolation to $\mathcal{S}_k$ on mesh $\mathcal{T}_k$ for $k = 1,\ldots,J$.

Lemma 6.8 (Discrete Regular Decomposition [35]). *For every $\phi \in \mathcal{S}$, there exist $\tilde{\phi} \in \mathcal{S}, w \in \mathcal{U}^3$, and $\psi \in \mathcal{S} \cap \ker(\mathrm{curl})$ such that*

$$\phi = \tilde{\phi} + \Pi_h^{\mathrm{curl}} w + \psi, \quad \text{and} \tag{32}$$

$$\|h^{-1}\tilde{\phi}\| + \|w\|_1 \lesssim \|\mathrm{curl}\,\phi\|. \tag{33}$$

In the decomposition (32), $\psi \in \ker(\mathrm{curl})$ and there is no need to control the norm of ψ. The component $\tilde{\phi}$ is of high frequency and the component $w \in \mathcal{U}^3$ for which a stable multilevel decomposition for the linear finite element can be applied. The following decomposition can be found in [66].

Lemma 6.9. *For every $\phi \in \mathcal{S}$, there exist $\tilde{\phi} \in \mathcal{S}$, $w_k \in \mathcal{U}_k^3$, and $\psi \in \mathcal{S} \cap \ker(\mathrm{curl})$ such that*

$$\phi = \tilde{\phi} + \sum_{k=1}^{J} \Pi_k^{\mathrm{curl}} w_k + \psi, \quad \text{and} \tag{34}$$

$$\|h^{-1}\tilde{\phi}\|^2 + \sum_{k=1}^{J} h_k^{-1}\|w_k\|^2 \lesssim \|\mathrm{curl}\,\phi\|^2. \tag{35}$$

Theorem 6.10. *For every $v \in \mathcal{V} \cap \ker(\mathrm{div})$, there exists a decomposition $v = \sum_{k=1}^{J} \sum_{i=1}^{N_k} v_{k,i}$ such that*

$$\sum_{k=1}^{J} \sum_{i=1}^{N_k} \|v_{k,i}\|^2 \lesssim \|v\|^2. \tag{36}$$

Proof. For $v \in \mathcal{V} \cap \ker(\mathrm{div})$, there exists $\phi \in \mathcal{S}$ such that $v = \mathrm{curl}\,\phi$. We then apply Lemma 6.9 to obtain a decomposition of ϕ in the form of (34). We can write the first two terms in (34) into multilevel basis decomposition, i.e.,

$$\tilde{\phi} + \sum_{k=1}^{J} \Pi_k^{\mathrm{curl}} w_k = \sum_{k=1}^{J} \sum_{i=1}^{N_k} \phi_{k,i}. \tag{37}$$

Decomposition of v is obtained by choosing $v_{k,i} = \mathrm{curl}\,\phi_{k,i}$. The stability (36) is from the inverse inequality, the stability of bases decomposition of edge element spaces in L^2-norm, and the stability of the decomposition (35):

$$\sum_{k=1}^{J} \sum_{i=1}^{N_k} \|v_{k,i}\|^2 \lesssim \sum_{k=1}^{J} h_k^{-2} \sum_{i=1}^{N_k} \|\phi_{k,i}\|^2 \lesssim \|h^{-1}\tilde{\phi}\|^2 + \sum_{k=1}^{J} h_k^{-1}\|w_k\|^2 \lesssim \|\mathrm{curl}\,\phi\|^2.$$

$\square$

The (SCS) can be proved similarly as in the two dimensional case.

6.4. **Numerical examples.** In this subsection we present two numerical examples to support our theory. We perform the numerical experiments using the iFEM package [21].

We consider four examples on the Darcy equations

$$K^{-1}u + \nabla p = 0, \quad \mathrm{div}\,u = f \quad \text{in } \Omega$$

with given flux boundary condition $u \cdot n = g$ on $\partial\Omega$. We chose $\Omega = (0,1)^2$. Since we focus on the performance of solvers, we only specify the tensor K used in these examples.

- Example 1. K is the identity 2×2 matrix, i.e., $Id_{2\times2}$ and the grid is uniform.

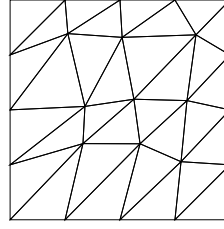- Example 2. The grid is still uniform but the tensor is non-diagonal

$$K = \begin{pmatrix} 1 + 4(x^2 + y^2) & 3xy \\ 3xy & 1 + 11(x^2 + y^2) \end{pmatrix}.$$

  This is the Example 5.2 considered in [51]. The spectrum of K is in $[1, 25]$ and thus contains certain anisotropy.
- Example 3. The tensor $K = a(x)Id_{2\times 2}$ with piecewise constant $a(x)$ on the initial 4×4 uniform partition of Ω. The scalar function $a = 10^{-p}$ where p is a random integer such that $0 \le p \le 5$.
- Example 4. The same tensor in Example 3, except the initial grid is distorted; see Fig. 6.4 (b). The interior grid points are randomly perturbed by up to 40% of the mesh size $h = 1/4$. Example 3 and 4 are two dimensional version of the example used in [64].
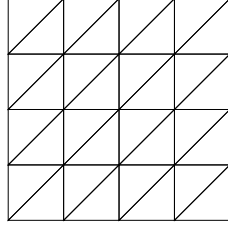
(a) The uniform mesh with $h = 1/4$ (b) A distorted mesh

FIGURE 1. The initial mesh of Example 1 - 3 is the uniform mesh in (a) and the initial mesh of Example 4 is a distorted mesh in (b).

We discretize the Darcy equations using the lowest order RT element and apply the SSO method with the decomposition (31). We implement SSO in a V-cycle formulation and perform only one pre-smoothing and one post-smoothing. The smoother is an overlapping multiplicative Schwarz smoother requiring solving a small saddle point system in the patch of each vertex. The local problem is solved exactly as the dimension of the local problem is small and admit a very efficient direct solver described below. Let n_i denote the number of edges connected to the i-th vertex patch. If we orientated these interior edges with normal direction counterclockwise, then locally the divergence free basis is represented by the constant vector $(1, \ldots, 1)_{n_i \times 1}^T$. Note that the local mass matrix M is tridiagonal, given a residual vector, the local problem can be solved in $4n_i$ addition and one division (no multiplication required as the divergence basis corresponds to a constant vector). Let N be the number of interior vertices. The total cost of the local solver is thus $4\sum_{i=1}^{N} n_i$. In average $n_i \approx 6$ and thus the cost is around $24N$. On the other hand, the size of the saddle point system is the number of interior edges plus the number of triangles, which is around $5N$, and the number of non zeros of this matrix is around $21N$. A matrix-vector product thus requires $21N$ multiplication which is way costly than the $24N$ addition needed in the local solvers. Similar calculation holds for the 3D local problem with different constant. We conclude that the dominated cost of the smoother will be the evaluation of the residual and one step of the smoother requires just one matrix-vector product.

We stop the iteration when an approximated relative error in the energy norm is less than or equal to 10^{-8}. Let r be the current residual of the iterate u and Br is the correction obtained by one V-cycle. Then we use the error formulae $\sqrt{(Br, r)/(u, f)}$ which is better

than using the relative residual error as $B \approx A^{-1}$. We report iteration steps of V-cycle required for the four examples. We did not include the CPU time since it depends on the implementation and testing environment: the programming language, optimization of codes, and the hardware (memory and cache), etc. The operation count we did before indicates that our method can be implemented very efficiently.

TABLE 1. Iteration steps of V-cycle multigrid for the saddle point system with 1 pre-smoothing and 1 post-smoothing step. Stopping criterion is the approximated relative error is less than 10^{-8}.

h	size	Ex 1	Ex 2	Ex 3	Ex 4
1/8	336	10	13	7	20
1/16	1,312	11	15	10	19
1/32	5,184	11	16	13	25
1/64	20,608	11	16	13	25

Based on the numerical results in Table 1, we conclude that our multigrid method is convergent uniformly to the mesh size and pretty robust to the variation of the tensor K and the distortion of meshes.

A popular Uzawa type preconditioned conjugate gradient (PCG) method for solving the Schur complement $BM^{-1}B^T$ equation requires the evaluation of M^{-1} (the so-called inner iteration) for each PCG iteration (the so-called outer iteration) and an effective preconditioner for the Schur complement. As noticed in [10, 51], the inner iteration of computing M^{-1} should be very accurate and thus the overall inner-outer iteration process is costly. And in [64], it is shown that preconditioners for $BM^{-1}B^T$ should be tuned to the variation of the tensor and the distortion of the mesh. Better preconditioned iterative methods have been developed in [10, 51].

7. APPLICATION TO NON-CONFORMING METHODS

In this section, we use the equivalence between non-conforming methods and mixed methods to develop a V-cycle multigrid method for non-conforming methods and prove its uniform convergence. The two ingredients of our new multigrid method for non-conforming methods are: the overlapping Schwarz smoothers, and inter-grid transfer operators through the nested flux spaces.

Again we consider the Poisson equation with Neumann boundary condition $-\Delta p = f$ in Ω with $\partial_n p|_{\partial\Omega} = 0$. Based on a triangulation $\mathcal{T}_h$ of Ω, the Crouzeix-Raviart (CR) non-conforming finite element space [31] is defined as follows

$$\Lambda_h = \{\lambda|_T \in \mathcal{P}_1(T), \forall T \in \mathcal{T}_h, \int_E \lambda \, ds \text{ is continuous for all sides } E \text{ of } \mathcal{T}_h\}.$$

The space Λ_h is not a subspace of $H^1(\Omega)$ due to the loss of continuity across the sides of elements. An elementwise gradient operator ∇_h is defined as

$$(\nabla_h \lambda)|_T := \nabla(\lambda|_T) \quad \forall T \in \mathcal{T}_h,$$

and the bilinear form is defined as

$$(\nabla_h \lambda, \nabla_h \mu) := \sum_{T \in \mathcal{T}_h} \int_T \nabla_h \lambda \cdot \nabla_h \mu \, dx \quad \text{for all } \lambda, \mu \in \Lambda_h.$$

The CR non-conforming finite element discretization is as follows: Given an $f \in L^2(\Omega)$, find $\lambda \in \Lambda_h \cap L_0^2(\Omega)$ such that

$$(38) \qquad (\nabla_h \lambda, \nabla_h \mu) = (f, \mu), \quad \text{for all } \mu \in \Lambda_h.$$

Let $u \in \mathcal{V}$ be the mixed finite element approximation of the flux using the lowest order RT element. It is well known that [40], for every $T \in \mathcal{T}_h$,

$$(39) \qquad u|_T = -\nabla_h \lambda|_T + \frac{1}{d} f_T (x - x_T), \quad \forall x \in T,$$

where x_T is the barycenter of the triangle T and f_T is the average of f over T. Throughout this section we shall always consider a piecewise constant function f. We always denote by u the solution to (29) and by λ the solution to (38).

We note that such equivalence has been used to design multigrid methods for mixed methods with the help of non-conforming methods [13, 26]. We are exploiting this equivalence in the other way around.

Based on a sequence of hierarchy meshes, we will have a sequence of spaces $\Lambda_1, \Lambda_2, \ldots$, $\Lambda_J = \Lambda_h$. We shall develop a V-cycle multigrid method for solving the equation (38) on the finest level. The notorious difficulty is the non-nestedness of hierarchies of non-conforming finite element spaces. Robust inter-grid operators (restriction and prolongation operators) should be designed carefully [11, 9, 48, 27, 15, 36, 37]. The existing convergence proof of multigrid methods for non-conforming methods [11, 15, 48, 16, 49] do not cover V-cycles with few smoothing steps but for multigrid cycles with sufficiently many smoothing steps. We shall design a V-cycle multigrid method for CR element and prove its convergence even for only one smoothing step.

Essentially our method is just a different interpretation of the SSO method applied to the mixed finite element discretization. Therefore during the iteration, we always keep two quantities (u^k, λ^k) which is the kth iteration of (u, λ).

The smoother in the finest level is an overlapping Schwarz smoother with Neumann boundary condition. It consists of solving a local problem $-\Delta \lambda_i^k = f$ in Ω_i with Neumann boundary condition $\partial_n \lambda_i^k|_{\partial\Omega_i} = u_{i-1}^k \cdot n$, where Ω_i is the patch of the ith vertex. Here we loop over the vertex $i = 1, \ldots, N$ of $\mathcal{T}_h$ and use subscript i to denote the iteration at the ith vertex. We set $u_0^k = u^k$ and $u^{k+1} = u_N^k$. Once λ_i^k is computed, it will be used to update the flux u_i^k by the relation (39) since the relation holds for the local problem as well. To begin with, we need to compute flux u_* on the finest level such that the local Neumann problem is well defined, i.e, the source f is compatible with the prescribed boundary flux. Such flux u_* can be found by a V-cycle multigrid iteration similar to the procedure for non-homogenous constraint discussed before.

In the implementation level, the matrix of local problems can be obtained by extracting sub-matrices of the global one. The right-hand side is the corresponding components of f plus the contribution from the boundary condition. It is the degree of freedom $\int_E u \cdot n$ that enters the computation which can be calculated by the formulae

$$(40) \qquad \int_E u_i^k \cdot n_E \, ds = \frac{|T|}{d+1} f + \int_E \nabla_h \lambda_i^k \cdot n_e \, ds.$$

The relation (40) can be used to eliminate the flux and get a direct updated formulation $\lambda_{i-1}^k \to \lambda_i^k$ without recording the flux approximation u_i^k. Algebraically it can be realized by matrix multiplication of λ_i^k. Conceptually it is better to record the flux explicitly.

We then discuss the prolongation from the coarse grid to the fine grid. Since now only two levels are involved, we will follow the convention to use subscript $(\cdot)_H$ for quantities in the coarse grid and $(\cdot)_h$ for that in the fine grid. In the coarse grid, we will solve a residual

equation to be considered in a moment. Suppose we have obtained a correction of the flux e_H, we prolongate e_H in the RT space in the coarse grid to that in the fine grid and denoted by $I_H^h e_H$. Note that although spaces of CR non-conforming elements are non-nested, the RT spaces for flux are, and I_H^h is just the natural inclusion. The correction is applied to the flux $u_h \leftarrow u_h + I_H^h e_H$. With the updated flux, we have different boundary conditions for the local problems (the source is always f in the finest level) and the smoother in the finest level can be applied again.

We then discuss in detail the residual equation to be solved in the coarse grid. We first describe the restriction. Denoted by u_h the current approximation of flux in the fine grid. The residual equation of the corresponding mixed method in the fine grid is

$$(41) \qquad \begin{pmatrix} M & B^T \\ B & O \end{pmatrix} \begin{pmatrix} e_h \\ p_h \end{pmatrix} = \begin{pmatrix} -Mu_h \\ 0 \end{pmatrix}.$$

So the restriction operator will apply to the residual $-Mu_h$, i.e., $r_H = -(I_H^h)^T M u_h$. On the coarse grid, we will still solve local problems on vertex patches. We start with the zero initial guess of flux e_H, i.e., $e_{H,0} = 0$, and solve local problems to update $e_{H,i}$ patchwise for $i = 1, \dots, N_H$. The updated flux $e_{H,i}$ in patch $\Omega_{H,i}$ will provide a boundary condition for the next patch $\Omega_{H,i+1}$. To use the non-conforming formulation, we need to figure out the source data for each local problem. This can be done as follows. Let r_H^i be the restriction of r_H to the ith patch $\Omega_{H,i}$, and let M_H^i be the corresponding mass matrix. Then the source for the local problem on $\Omega_{H,i}$ will be given by $\delta f_{H,i} = -\operatorname{div} M_{H,i}^{-1} r_H^i$ which is piecewise constant on $\Omega_{H,i}$. The inverse $M_{H,i}^{-1}$ can be computed efficiently since $M_{H,i}$ is tri-diagonal. Now we can solve the non-conforming discretization of the problem $-\Delta_H \lambda_{H,i} = \delta f_H$ with Neumann boundary condition $\partial_n \lambda_{H,i}|_{\partial\Omega_{H,i}} = e_{H,i-1} \cdot n$ and use $\lambda_{H,i}$ to update the flux correction $e_{H,i}$. Again such procedure can be implemented as one matrix multiplication which leads to a non-trivial restriction matrix.

As usual, a V-cycle multigrid method is obtained by applying the above two-level method recursively to the coarse grid problem.

Convergence of this multigrid algorithm is striaghtforward since it is just a different way to compute the same solution of the mixed formulation for each local problem. The quantity $2(E(u^k) - E(u)) = \|u - u^k\|^2 = \|\nabla_h \lambda - \nabla_h \lambda^k\|^2$ since the relation (39) always holds during the iteration.

The same algorithm and convergence proof can be applied to other non-conforming methods, e.g., hybridized discontinuous Galerkin (HDG) methods [2, 29, 30] and weak Galerkin (WG) method [61, 43], which are equivalent to the mixed methods. The only difference is the relation of λ and the flux u. For example, for WG, we can simply use the following formulae to update the flux: $u = \nabla_w \lambda$, where ∇_w is the weak gradient operator.

We thus have obtained a V-cycle multigrid method for non-conforming finite elements and have proved the uniform convergent with even one smoothing step. Such results are very rare in literature and a recent work on a multigrid method for HDG methods with only one smoothing step can be found in [28].

8. APPLICATION TO STOKES EQUATIONS

In this section, we apply our approach to designing a multigrid method for a discrete Stokes system in two dimensions and prove its uniform convergence.

Let Ω be a polygon and triangulated into a quasi-uniform mesh $\mathcal{T}_h$ with mesh size h. Again we assume that there exists a sequence of meshes $\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_J = \mathcal{T}_h$. The triangulation $\mathcal{T}_1$ is a shape regular triangulation of Ω and $\mathcal{T}_{k+1}$ is obtained by dividing

each triangle in $\mathcal{T}_k$ into four congruent small triangles. We further assume triangulations contain no singular vertex defined in [54].

Consider Stokes equations

$$-\Delta u + \nabla p = f, \quad \operatorname{div} u = 0$$

with Dirichlet boundary condition $u|_{\partial\Omega} = 0$. The homogenous boundary condition is not essential. As discussed before, the non-homogenous boundary condition will lead to a non-homogenous constraint and can be eliminated by one V-cycle or by a fast Poisson solver.

We shall use exact divergence free elements and assume $\mathcal{K}_i = \ker(\operatorname{div}) \cap \mathcal{V}_i, i = 1, \ldots, J$ are nested, i.e.,

$$\mathcal{K}_1 \subset \mathcal{K}_2 \subset \ldots \subset \mathcal{K}_J = \mathcal{K}.$$

Examples of such Stokes elements include Scott-Vogelius elements [54] for which the assumption that all triangulations contain non-singular vertex is needed. We chose $\mathcal{V} \subset (H_0^1(\Omega))^2$ and $\mathcal{P} \subset L_0^2(\Omega)$ as Scott-Vogelius elements [54]. For this problem, the A-norm is the H^1 semi-norm $|\cdot|_1 = \|\nabla(\cdot)\|$ which is a norm on H_0^1.

Let $\mathcal{U}$ be the C^1 finite element space on $\mathcal{T}$ such that $\operatorname{curl}\mathcal{U} = \mathcal{K}$. Namely we have the so-called Stokes complex:

$$(42) \qquad\qquad \mathcal{U} \xrightarrow{\operatorname{curl}} \mathcal{V} \xrightarrow{\operatorname{div}} \mathcal{P} \to 0.$$

Similar exact sequence exists in each level $k = 1, 2, \ldots, J$.

We further decompose each $\mathcal{K}_k$ into subspaces associated to vertices. For a vertex $x_{k,j} \in \mathcal{T}_k$, we denote by $\Omega_{k,j}$ the patch of $x_{k,j}$, i.e., union of all triangles containing $x_{k,j}$. Let $\mathcal{U}_{k,j} = C_0^1(\Omega_{k,j}) \cap \mathcal{U}_k$ and $\mathcal{V}_{k,j} = (H_0^1(\Omega_{k,j}))^2 \cap \mathcal{V}_k$ be the subspaces spanned by all basis functions with support in $\Omega_{k,j}$ and set $\mathcal{K}_{i,j} = \mathcal{V}_{i,j} \cap \ker(\operatorname{div})$. By the construction of Scott-Vogeligus element, $\mathcal{K}_{k,j} = \operatorname{curl}\mathcal{U}_{k,j}$. The final decomposition is

$$\mathcal{V} = \sum_{k=1}^{J} \sum_{j=1}^{N_k} \mathcal{V}_{k,j}, \text{ and } \mathcal{K} = \sum_{k=1}^{J} \sum_{j=1}^{N_k} \mathcal{K}_{k,j}.$$

In the correction form, the smoother applied to this decomposition is equivalent to solving a local Stokes problem with force $f - Av_i$ in the subdomain surrounding of a vertex with zero Dirichlet boundary condition on $\partial\Omega_{k,j}$. In the update form, it is solving the original Stokes problem with force f but the boundary condition on $\partial\Omega_{k,j}$ is given by the current approximation of the velocity. Now the local problem is of considerable size (around a 200×200 saddle point system) and the inexact solver using diagonal matrix of A can reduce it to a SPD problem of smaller size (around 60×60).

Remark 8.1. We are aware that more effective block preconditioners for the Stokes equations are available [6, 39]. Multigrid method based on solving local problems is of more theoretic value since multigrid convergence theory for Stokes equations with the partial regularity assumption and/or for V-cycle method with few smoothing steps is rare. $\qquad\square$

We define $Q_k : \mathcal{U} \to \mathcal{U}_k$ the L^2 projection, for $k = 1, 2, \ldots, J$. For $v \in \mathcal{K}$, since $\operatorname{div} v = 0$, we can find a unique $\phi \in \mathcal{U}$ such that $v = \operatorname{curl}\phi$. We then define $\Pi_k v = \operatorname{curl} Q_k \phi$. It is easy to show that $\Pi_l \Pi_k = \Pi_l$ for $l \leq k$ due to the nestedness of spaces.

We document the stability and error estimate of Π_k in the following lemmas.

Lemma 8.2. *The operator* Π_k *is stable in* L^2 *norm and*

$$(43) \qquad\qquad \|v - \Pi_k v\| \lesssim h_k |v|_1, \quad \text{for all } v \in \mathcal{K}.$$

Proof. It is well known that Q_k is stable in both L^2-norm and curl -norm on quasi-uniform meshes. Consequently $\Pi_k = \text{curl}\, Q_k$ is stable in L^2-norm. It is obvious that $\Pi_k v = v$ for all $v \in \mathcal{K}_k$. Therefore $v - \Pi_k v = (I - \Pi_k)(v - v_k)$ for any $v_k \in \mathcal{K}_k$ and consequently

$$\|v - \Pi_k v\| \lesssim \inf_{v_k \in \mathcal{K}_k} \|v - v_k\| = \inf_{\phi_k \in \mathcal{U}_k} \|\text{curl}\,\phi - \text{curl}\,\phi_k\| \lesssim h_k |\phi|_2 = h_k |v|_1.$$

$\square$

Lemma 8.3. *The operator Π_k is stable in H^σ-norm for $\sigma \in [0, 1/2)$, i.e.,*

$$(44) \qquad \|\Pi_k v\|_\sigma \lesssim \|v\|_\sigma, \quad \text{for all } v \in \mathcal{K}.$$

Proof. For $\sigma = 0$, i.e., the stability of Π_k in L^2-norm has been proved in Lemma 8.2. For $v \in \mathcal{K}$, we define $\bar{v}$ as the piecewise constant approximation of v defined by $\int_T \bar{v} = \int_T v$ for all $T \in \mathcal{T}_k$. Obviously $\|v - \bar{v}\| \lesssim h_k |v|_1$.

We prove the stability of Π_k in H^1-norm as follows:

$$|\Pi_k v|_1 = |\Pi_k v - \bar{v}|_1 \lesssim h_k^{-1} \|\Pi_k v - \bar{v}\| \lesssim h_k^{-1} (\|\Pi_k v - v\| + \|v - \bar{v}\|) \lesssim |v|_1.$$

In the last step, we have used the approximation property of Π_k; c.f. Lemma 8.2.

By the interpolation of divergence free spaces, c.f. Proposition 3.7 in [63], we obtain the desired inequality (44).

$\square$

Following Xu [65], we can obtain the following stable decomposition. For completeness, we include a proof here.

Theorem 8.4. *The decomposition $v = \sum_{k=1}^J (\Pi_k - \Pi_{k-1})v$ is stable in A-norm, i.e.,*

$$(45) \qquad \sum_{k=1}^J |(\Pi_k - \Pi_{k-1})v|_1^2 \lesssim |v|_1^2, \quad \text{for all } v \in \mathcal{K}.$$

Proof. Let $P_i : \mathcal{K} \to \mathcal{K}_i$ be the projection in A-inner product. Then by the duality argument, c.f. Theorem 6.9 in [23], we have the following error estimate, for some $\alpha \in (1/2, 1]$,

$$(46) \qquad \|v - P_i v\|_{1-\alpha} \lesssim h_i^\alpha \|v\|_1, \text{ for all } v \in H_0^1(\Omega).$$

Let $\tilde{\Pi}_k = \Pi_k - \Pi_{k-1}$, and $v_i = (P_i - P_{i-1})v$ for $i = 1, 2, \cdots, J$ with notation $P_0 = 0$. Using Cauchy-Swarchz inequality, it holds

$$\sum_{k=1}^J \|\nabla(\tilde{\Pi}_k v)\|^2 = \sum_{k=1}^J \sum_{i,j=k}^J \int_\Omega \nabla(\tilde{\Pi}_k v_i) \cdot \nabla(\tilde{\Pi}_k v_j)\, dx$$

$$= \sum_{i,j=1}^J \sum_{k=1}^{i \wedge j} \int_\Omega \nabla(\tilde{\Pi}_k v_i) \cdot \nabla(\tilde{\Pi}_k v_j)\, dx$$

$$\leq \sum_{i,j=1}^J \sum_{k=1}^{i \wedge j} \|\nabla(\tilde{\Pi}_k v_i)\| \|\nabla(\tilde{\Pi}_k v_j)\|,$$

where $i \wedge j = \min\{i, j\}$. According to the inverse inequality, the stability of Π_k, c.f., (44), and the error estimate of P_i c.f. (46), we have

$$\|\nabla(\tilde{\Pi}_k v_i)\| \lesssim h_k^{-\alpha} \|\tilde{\Pi}_k v_i\|_{1-\alpha} \lesssim h_k^{-\alpha} \|v_i\|_{1-\alpha} \lesssim h_k^{-\alpha} h_i^\alpha |v_i|_1.$$

Combining last two inequalities, we get from the strengthened Cauchy-Swarchz inequality

$$\sum_{k=1}^{J} \|\nabla (\Pi_k - \Pi_{k-1})v\|^2 \lesssim \sum_{i,j=1}^{J} \sum_{k=1}^{i \wedge j} h_k^{-2\alpha} h_j^\alpha h_i^\alpha |v_i|_1 |v_j|_1 \lesssim \sum_{i,j=1}^{J} h_{i \wedge j}^{-2\alpha} h_j^\alpha h_i^\alpha |v_i|_1 |v_j|_1$$

$$\lesssim \sum_{i,j=1}^{J} \left(\frac{1}{2}\right)^{\alpha|i-j|} |v_i|_1 |v_j|_1 \lesssim \sum_{i=1}^{J} |v_i|_1^2 = |v|_1^2.$$

$\square$

We continue to show the micro-decomposition of the slice $(\Pi_k - \Pi_{k-1})v$ is stable in the energy norm.

Lemma 8.5. *For $v_k = (\Pi_k - \Pi_{k-1})v \in \mathcal{K}_k$, there exists a decomposition $v_k = \sum_{j=1}^{N_k} v_{k,j}$ with $v_{k,j} \in \mathcal{K}_{k,j}$ such that*

$$\sum_{j=1}^{N_j} |v_{k,j}|_1^2 \lesssim |v_k|_1^2.$$

Proof. Recall that $\phi_k = Q_k \phi$ and $v = \text{curl}\,\phi$. Let $\phi_k - \phi_{k-1} = \sum_{j=1}^{N_k} \psi_{k,j}$ be a decomposition such that $\text{supp}\,\psi_{k,j} \in \Omega_{k,j}$. Such decomposition can be obtained by partition the basis decomposition. For example, for a basis function associated to an edge, it can be split as half and half to the patch of each vertex of this edge. We then set $v_{k,j} = \text{curl}\,\psi_{k,j}$ and obtain the decomposition $v_k = \sum_{j=1}^{N_k} v_{k,j}$. Then

$$\sum_{j=1}^{N_k} |v_{k,j}|_1^2 \leq \sum_{j=1}^{N_k} |\psi_{k,j}|_2^2 \lesssim \sum_{j=1}^{N_k} h_k^{-4} \|\psi_{k,j}\|^2 \lesssim h_k^{-4} \|\phi_k - \phi_{k-1}\|^2$$

$$= h_k^{-4} \|(I - Q_{k-1})(\phi_k - \phi_{k-1})\|^2 \lesssim h_k^{-2} \|\text{curl}\,(\phi_k - \phi_{k-1})\|^2 = h_k^{-2} \|v_k\|^2.$$

We write $v_k = (\Pi_k - \Pi_{k-1})v = (I - \Pi_{k-1})(\Pi_k - \Pi_{k-1})v = (I - \Pi_{k-1})v_k$ and use the L^2-norm estimate of Π_k to conclude

$$h_k^{-2} \|v_k\|^2 \lesssim |v_k|_1^2.$$

Then the desired inequality follows. $\square$

Combination of Theorem 8.4 and Lemma 8.5 leads to the stability of the decomposition $\mathcal{K} = \sum_{k=1}^{J} \sum_{j=1}^{N_j} \mathcal{K}_{k,j}$ in A-norm.

Theorem 8.6. *For every $v \in \mathcal{K}$, there exists a decomposition $v = \sum_{k=1}^{J} \sum_{j=1}^{N_k} v_{k,j}$ with $v_{k,j} \in \mathcal{K}_{k,j}$ such that*

$$\sum_{k=1}^{J} \sum_{j=1}^{N_j} |v_{k,j}|_1^2 \lesssim |v|_1^2.$$

The assumption (SCS) is just that for multilevel H^1 finite element spaces $\mathcal{V}_1 \subset \mathcal{V}_2 \subset \ldots \subset \mathcal{V}_J$ and can be proved similarly as before. Note that since $\mathcal{V}_{k,j} \subset (H_0^1(\Omega_{k,j}))^2$, for functions $v_{k,j}$ in $\mathcal{V}_{k,j}$, the norm equivalence $h_k^{-2} \|v_{k,j}\| \eqsim |v_{k,j}|_1$ holds with an $\mathcal{O}(1)$ constant.

9. Conclusion and Future Work

In this paper we have developed a multigrid method for saddle point systems based on a multilevel subspace decomposition of the constraint space $\mathcal{K}$. We have proved the convergence of such method based on the stable decomposition and strengthened Cauchy Schwarz inequality. For some mixed finite element discretizations of Poisson, Darcy, and Stokes equations, we have verified SD and SCS assumptions and consequently obtained a multigrid method for the resulting saddle point systems. In a forthcoming work [24], we shall examine a plate bending problem which is a fourth order elliptic equation. The key is to find a underlying exact sequence.

References

[1] D. Arnold, R. Falk, and R. Winther. Finite element exterior calculus, homological techniques, and applications. *Acta numerica*, 15:1–155, 2006.

[2] D. N. Arnold and F. Brezzi. Mixed and nonconforming finite element methods: Implementation, postporcessing and error estimates. *RAIRO Model Math. Anal. Numer.*, 19:7–32, 1985.

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

[4] R. E. Bank, T. F. Dupont, and H. Yserentant. The Hierarchical Basis Multigrid Method. *Numer. Math.*, 458:427–458, 1988.

[5] R. E. Bank and L. R. Scott. On the conditioning of finite element equations with highly refined meshes. *SIAM J. Numer. Anal.*, 26:1383–1394, 1989.

[6] M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numer.*, 14:1–137, may 2005.

[7] D. Braess and W. Dahmen. A cascadic multigrid algorithm for the Stokes equations. *Numer. Math.*, 82:179–191, 1999.

[8] D. Braess and R. Sarazin. An efficient smoother for the Stokes equation. *Appl. Numer. Math.*, 23(1):3–19, feb 1997.

[9] D. Braess and R. Verfürth. Multigrid Methods for Nonconforming Finite Element Methods. *SIAM J. Numer. Anal.*, 27:979–986, 1990.

[10] J. H. Bramble and J. E. Pasciak. A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems. *Math. Comp.*, 50(181):1–17, 1988.

[11] S. C. Brenner. An optimal order multigrid for P1 nonconforming finite elements. *Math. Comp.*, 52:1–15, 1989.

[12] S. C. Brenner. A nonconforming multigrid method for the stationary Stokes equations. *Math. Comp.*, 55:411–437, 1990.

[13] S. C. Brenner. A Multigrid algorithm for the lowest-order Raviart-Thomas mixed triangular finite element method. *SIAM J. Numer. Anal.*, 29:647–678, 1992.

[14] S. C. Brenner. Multigrid methods for parameter dependent problems. *RAIRO-M2AN Modelisation Math et Analyse Numerique*, 30(3):265–297, 1996.

[15] S. C. Brenner. Convergence of nonconforming multigrid methods without full elliptic regularity. *Math. Comp.*, 68(225):25–53, 1999.

[16] S. C. Brenner. Convergence of nonconforming V-cycle and F-cycle multigrid algorithms for second order elliptic boundary value problems. *Math. Comp.*, 73(247):1041–1066, 2003.

[17] S. C. Brenner, H. Li, and L.-Y. Sung. Multigrid methods for saddle point problems: Stokes and Lamé systems. *Numerische Mathematik*, (2):1–24, jan 2014.

[18] S. C. Brenner, D.-S. Oh, and L.-Y. Sung. Multigrid Methods for Saddle Point Problems: Darcy Systems. *arXiv*, pages 1–30, 2015.

[19] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*. Springer-Verlag, 1991.

[20] Z. Cai, R. R. Parashkevov, T. F. Russell, J. D. Wilson, and X. Ye. Domain Decomposition for a Mixed Finite Element Method in Three Dimensions. *SIAM Journal on Numerical Analysis*, 41(1):181–194, jan 2003.

[21] L. Chen. iFEM: An Integrated Finite Element Methods Package in MATLAB. *Technical Report, University of California at Irvine*, 2009.

[22] L. Chen. Deriving the X-Z Identity from Auxiliary Space Method. In Y. Huang, R. Kornhuber, O. Widlund, and J. Xu, editors, *Domain Decomposition Methods in Science and Engineering XIX*, pages 309–316. Springer Berlin Heidelberg, 2010.

[23] L. Chen. Multigrid methods for saddle point systems using constrained smoothers. *Computers & Mathematics with Applications*, (2):1–13, 2015.

[24] L. Chen, J. Hu, and X. Huang. Multigrid Methods for Hellan-Herrmann-Johnson Mixed Method of Kirchhoff Plate Bending Problems. *ArXiv*, pages 1–20, 2015.

[25] L. Chen, X. Hu, M. Wang, and J. Xu. A Multigrid Solver based on Distributive Smoother and Defect Correction for Oseen Problems. *Submitted*, 2013.

[26] Z. Chen. Equivalence between and multigrid algorithms for nonconforming and mixed methods for second order elliptic problems. *East-West Journal of Numerical Mathematics*, 4:1–33, 1996.

[27] Z. Chen and P. Oswald. Multigrid and multilevel methods for nonconforming rotated Q1 elements. *Math. Comp.*, 67(222):667–693, 1998.

[28] B. Cockburn, O. Dubois, J. Gopalakrishnan, and S. Tan. Multigrid for an HDG Method. *IMA J. Numer. Anal.*, pages 1–40, 2013.

[29] B. Cockburn and J. Gopalakrishnan. A Characterization of Hybridized Mixed Methods for Second Order Elliptic Problems. *SIAM J. Numer. Anal.*, 42(1):283–301, 2004.

[30] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous galerkin, mixed, and continuous galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.*, 47(2):1319–1365, 2009.

[31] P. M. Crouzeix and P. A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations I. *R.A.I.R.O*, 76:3–33, 1973.

[32] R. E. Ewing and J. Wang. Analysis of the schwarz algorithm for mixed finite elements methods. *M2AN*, 26:739–756, 1992.

[33] R. Hiptmair. Finite elements in computational electromagnetism. *Acta Numer.*, 11:237–339, 2002.

[34] R. Hiptmair and R. Hoppe. Multilevel methods for mixed finite elements in three dimensions. *Numerische Mathematik*, 82:253–279, 1999.

[35] R. Hiptmair and J. Xu. Nodal Auxiliary Space Preconditioning in H(curl) and H(div) Spaces. *SIAM J. Numer. Anal.*, 45(6):2483–2509, 2007.

[36] K. S. Kang and S. Y. Lee. New intergrid transfer operator in multigrid method for {P1}-nonconforming finite element method. *Appl. Math. Comp.*, 100(2-3):139–149, 1999.

[37] M. Koster, a. Ouazzi, F. Schieweck, S. Turek, P. Zajac, and M. Köster. New robust nonconforming finite elements of higher order. *Applied Numerical Mathematics*, 62(3):166–184, mar 2012.

[38] Y. J. Lee, J. Wu, J. Xu, and L. Zikatanov. Robust subspace correction methods for nearly singular systems. *Mathematical Models and Methods in Applied Sciences*, 17(11):1937–1963, 2006.

[39] K. Mardal and R. Winther. Preconditioning discretizations of systems of partial differential equations. *Numerical Linear Algebra with Applications*, 18:1–40, 2011.

[40] L. Marini. An Inexpensive Method for the Evaluation of the Solution of the Lowest Order Raviart–Thomas Mixed Method. *SIAM J. Numer. Anal.*, 22(3):493–496, 1985.

[41] T. Mathew. Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems, part II: convergence theory. *Numerische Mathematik*, 492:469–492, 1993.

[42] T. P. Mathew. Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems, part I: Algorithms and numerical results. *Numerische Mathematik*, 65(1):445–468, dec 1993.

[43] L. Mu, J. Wang, and X. Ye. Weak Galerkin Finite Element Methods on Polytopal Meshes. *arXiv preprint arXiv:1204.3655*, 72204:22, apr 2012.

[44] J. C. Nédélec. Mixed finite elements in R^3. *Numer. Math.*, 35:315–341, 1980.

[45] J. C. Nédélec. A new family of mixed finite elements in R^3. *Numer. Math.*, 50:57–81, 1986.

[46] R. H. Nochetto, K. G. Siebert, and A. Veeser. Theory of adaptive finite element methods: an introduction. In R. A. DeVore and A. Kunoth, editors, *Multiscale, Nonlinear and Adaptive Approximation*. Springer, 2009.

[47] M. A. Olshanskii. Multigrid analysis for the time dependent Stokes problem. *Mathematics of Computation*, 5718:1–23, 2011.

[48] P. Oswald. Intergrid transfer operators and multilevel preconditioners for nonconforming discretizations. *Appl. Numer. Math.*, 23(1):139–158, 1997.

[49] P. Oswald. Optimality of multilevel preconditioning for nonconforming P1 finite elements. *Numer. Math.*, 111(2):267–291, sep 2008.

[50] P. A. Raviart and J. Thomas. A mixed finite element method fo 2-nd order elliptic problems. In I. Galligani and E. Magenes, editors, *Mathematical aspects of the Finite Elements Method*, Lectures Notes in Math. 606, pages 292–315. Springer, Berlin, 1977.

[51] T. Rusten and R. Winther. A Preconditioned Iterative Method for Saddlepoint Problems. *SIAM Journal on Matrix Analysis and Applications*, 13(3):887–904, jul 1992.

[52] J. Schöberl. Multigrid Methods for a Parameter Dependent Problem in Primal Variables. *Numer. Math.*, 84:1–19, 1999.

[53] J. Schöberl and W. Zulehner. On Schwarz-type Smoothers for Saddle Point Problems. *Numerische Mathematik*, 95:377–399, 2003.

[54] L. R. Scott and M. Vogelius. Norm estimates for a maximal right inverse of the divergence operator in spaces of piecewise polynomials. *Mathematical Modelling And Numerical Analysis*, 19(1):111–143, 1985.

[55] X.-C. Tai. Rate of convergence for some constraint decomposition methods for nonlinear variational inequalities. *Numer. Math.*, 93(4):755–786, 2003.

[56] X.-C. Tai and J. Xu. Global and uniform convergence of subspace correction methods for some convex optimization problems. *Mathematics of Computation*, 71(237):105–125, may 2001.

[57] H. Uzawa. Iterative methods for concave programming. *Studies in linear and nonlinear programming*, 6, 1958.

[58] S. P. Vanka. Block-implicit multigrid solution of Navier-Stokes equations in primitive variables. *Journal of Computational Physics*, 65(1):138–158, 1986.

[59] P. S. Vassilevski and J. Wang. Multilevel iterative methods for mixed finite element discretizations of elliptic problems. *Numer. Math.*, 63(1):503–520, 1992.

[60] R. Verfürth. A multilevel algorithm for mixed problems. *SIAM J. Numer. Anal.*, 21(2):264–271, 1984.

[61] J. Wang and X. Ye. A weak Galerkin finite element method for second-order elliptic problems. *Journal of Computational and Applied Mathematics*, 241:103–115, 2013.

[62] M. Wang and L. Chen. Multigrid Methods for the Stokes Equations using Distributive Gauss-Seidel Relaxations based on the Least Squares Commutator. *Journal of Scientific Computing*, 56(2):409–431, feb 2013.

[63] H. Wendland. Divergence-Free Kernel Methods for Approximating the Stokes Problem. *SIAM J. Numer. Anal.*, 47(4):3158–3179, 2009.

[64] J. D. Wilson, R. L. Naff, and T. F. Russell. Multigrid preconditioned conjugate-gradient solver for mixed finite-element method. *Computational Geosciences*, 14(2):289–299, aug 2009.

[65] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Rev.*, 34:581–613, 1992.

[66] J. Xu, L. Chen, and R. H. Nochetto. Optimal Multilevel Methods for H(grad), H(curl), and H(div) Systems on Adaptive and Unstructured Grids. In R. A. DeVore and A. Kunoth, editors, *Multiscale, Nonlinear and Adaptive Approximation*. Springer, 2009.

[67] J. Xu and L. Zikatanov. The Method of Alternating Projections and the Method of Subspace Corrections in Hilbert Space. *J. Amer. Math. Soc.*, 15:573–597, 2002.

[68] W. Zulehner. A class of smoothers for saddle point problems. *Computing*, 65(3):227–246, 2000.

[69] W. Zulehner. Nonstandard Norms and Robust Estimates for Saddle Point Problems. *SIAM Journal on Matrix Analysis and Applications*, 32(2):536–560, apr 2011.

(L. Chen) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT IRVINE, IRVINE, CA 92697, USA

E-mail address: chenlong@math.uci.edu