A two-level algorithm for the weak Galerkin discretization of diffusion problems

Binjie Li† Xiaoping Xie ‡
School of Mathematics, Sichuan University, Chengdu 610064, China

Abstract

This paper analyzes a two-level algorithm for the weak Galerkin (WG) finite element methods based on local Raviart-Thomas (RT) and Brezzi-Douglas-Marini (BDM) mixed elements for two- and three-dimensional diffusion problems with Dirichlet condition. We first show the condition numbers of the stiffness matrices arising from the WG methods are of $O(h^{-2})$. We use an extended version of the Xu-Zikatanov (XZ) identity to derive the convergence of the algorithm without any regularity assumption. Finally we provide some numerical results.

Keywords. diffusion problem, weak Galerkin finite element, condition number, two-level algorithm, X-Z identity

1 Introduction

Let $\Omega \subset R^d$ ($d = 2, 3$) be a bounded polyhedral domain. Consider the following diffusion problem:

\begin{align}
-\text{div}(a \nabla u) &= f \quad \text{in } \Omega, \\
u &= 0 \quad \text{on } \partial \Omega,
\end{align}

(1.1)

where $a \in [L^\infty(\Omega)]^{d \times d}$ is a given symmetric positive-definite permeability tensor, $f \in L^2(\Omega)$.

*This work was supported by National Natural Science Foundation of China (11171239), Major Research Plan of National Natural Science Foundation of China (91430105) and Open Fund of Key Laboratory of Mountain Hazards and Earth Surface Processes, CAS.
†Email: libinjiefem@yahoo.com
‡Corresponding author. Email: xpxie@scu.edu.cn
The weak Galerkin (WG) finite element method was first introduced and analyzed by Wang and Ye [32] for general second order elliptic problems and later developed by their research group in [37, 38, 34, 39, 36, 33, 35]. It is designed by using a weakly defined gradient operator over functions with discontinuity. The method, based on local Raviart-Thomas (RT) elements [40] or Brezzi-Douglas-Marini (BDM) elements [18], allows the use of totally discontinuous piecewise polynomials in the finite element procedure, as is common in discontinuous Galerkin methods [3] and hybridized discontinuous Galerkin methods [22]. As shown in [37, 38, 34, 36], the WG method also enjoys an easy-to-implement formulation that inherits the physical property of mass conservation locally on each element. We note that when $a$ in (1.1) is a piecewise-constant matrix, the WG method, by introducing the discrete weak gradient as an independent variable, is equivalent to some hybridized version of the corresponding mixed RT or BDM method [2, 18] (cf. Remark 2.1).

As one knows, multigrid methods are among the most efficient methods for solving linear algebraic systems arising from the discretization of partial differential equations. By now, the research of the multigrid methods for second order elliptic problems has reached a mature stage in some sense (see [5, 6, 7, 8, 9, 10, 11, 12, 29, 41, 42, 43, 44, 45, 46] and the references therein). Especially, Xu, Chen, and Nochetto [46] presented an overview of the multigrid methods in an elegant fashion. For the model problem (1.1), Brenner [14] developed an optimal order multigrid method for the lowest-order Raviart-Thomas mixed triangular finite element. The algorithm and the convergence analysis are based on the equivalence between Raviart-Thomas mixed methods and certain nonconforming methods. In [28] Gopalakrishnan and Tan analyzed the convergence of a variable V-cycle multigrid algorithm for the hybridized mixed method for Poisson problems. Following the same idea, Cockburn et al. [23] analyzed the convergence of a non-nested multigrid V-cycle algorithm, with a single smoothing step per level, for one type of HDG method. One may refer to [13, 14, 15, 17, 21, 25, 27, 30, 31] for multigrid algorithms for nonconforming and DG methods.

This paper is to analyze a two-level algorithm for the WG methods. We show the condition numbers of the WG systems are of $O(h^{-2})$. We follow the basic ideas of [45, 46, 19] to establish an extended version of the Xu-Zikatanov (XZ) identity [45], and then derive the convergence of the algorithm without any regularity assumption.
The rest of this paper is organized as follows. Section 2 introduces the WG methods. Section 3 analyzes the conditioning of the WG systems. Section 4 describes the two-level algorithm, and analyzes its convergence. Section 5 provides some numerical experiments to verify our theoretical results.

2 Weak Galerkin finite element method

2.1 Preliminaries and Notations

Throughout this paper, we shall use the standard definitions of Sobolev spaces and their norms([1]), namely, for an arbitrary open set, \( D \), of \( \mathbb{R}^d \) and any nonnegative integer \( s \),

\[
H^s(D) := \{ v \in L^2(D) : \partial^\alpha v \in L^2(D), \forall |\alpha| \leq s \},
\]

\[
\|v\|_{s,D} := \left( \sum_{0 \leq j \leq s} |v|_{j,D}^2 \right)^{\frac{1}{2}}, \quad |v|_{j,D} := \left( \sum_{|\alpha|=j} \int_D |\partial^{\alpha} v|^2 \right)^{\frac{1}{2}}.
\]

We use \( \langle \cdot, \cdot \rangle_D \) and \( \langle \cdot, \cdot \rangle_{\partial D} \) to denote the standard \( L^2 \) inner products on \( L^2(D) \) and \( L^2(\partial D) \), respectively, and use \( \|\cdot\|_D \) and \( \|\cdot\|_{\partial D} \) to denote the norms induced by \( \langle \cdot, \cdot \rangle_D \) and \( \langle \cdot, \cdot \rangle_{\partial D} \), respectively. In particular, \( \|\cdot\| \) abbreviates \( \|\cdot\|_\Omega \).

Let \( T_h \) be a regular triangulation of \( \Omega \). For any \( T \in T_h \), we denote by \( h_T \) the diameter of \( T \) and set \( h := \max_{T \in T_h} h_T \). We denote by \( F_h \) the set of all faces of \( T_h \).

We introduce some mesh-dependent inner products and mesh-dependent norms as follows. We define \( \langle \cdot, \cdot \rangle_h : L^2(F_h) \times L^2(F_h) \to \mathbb{R} \) by

\[
\langle \lambda, \mu \rangle_h := \sum_{T \in T_h} h_T \int_{\partial T} \lambda \mu, \quad \forall \lambda, \mu \in L^2(F_h),
\]

and \( \langle \cdot, \cdot \rangle_h : [L^2(\Omega) \times L^2(F_h)] \times [L^2(\Omega) \times L^2(F_h)] \to \mathbb{R} \) by

\[
((u, \lambda), (v, \mu))_h := (u, \lambda)_\Omega + \langle \lambda, \mu \rangle_h, \quad \forall (u, \lambda), (v, \mu) \in L^2(\Omega) \times L^2(F_h).
\]

With a little abuse of notations, we use \( \|\cdot\|_h \) to denote the norms induced by the inner products \( \langle \cdot, \cdot \rangle_h \) and \( \langle \cdot, \cdot \rangle_h \), i.e.,

\[
\|\mu\|_h := \langle \mu, \mu \rangle_h^{\frac{1}{2}}, \quad \forall \mu \in L^2(F_h),
\]

\[
\|(v, \mu)\|_h := \left( \langle (v, \mu), (v, \mu) \rangle_h \right)^{\frac{1}{2}} = \left( \|v\|^2 + \|\mu\|^2 \right)^{\frac{1}{2}}, \quad \forall (v, \mu) \in L^2(\Omega) \times L^2(F_h).
\]

We also need the following elementwise norm and seminorms: for any \( \mu \in L^2(F_h) \),

\[
\|\mu\|_{h,\partial T} := h_T^{\frac{1}{2}} \|\mu\|_{\partial T}.
\]
\[ |\mu|_{h,\partial T}^2 := h^{-1}_T ||\mu - m_T(\mu)||_{\partial T}^2 \quad \text{with} \quad m_T(\mu) := \frac{1}{d+1} \sum_{F \in \mathcal{F}_T} \frac{1}{|F|} \int_F \mu, \]

and

\[ |\mu|_h := \left( \sum_{T \in \mathcal{T}_h} |\mu|_{h,\partial T}^2 \right)^{\frac{1}{2}}, \quad (2.5) \]

where \( \mathcal{F}_T \) denotes the set of all faces of \( T \), and \( |F| \) denotes the \((d-1)\)-dimensional Hausdorff measure of \( F \).

Throughout this paper, \( x \lesssim y \) (\( x \gtrsim y \)) means \( x \leq Cy \) (\( x \geq Cy \)), where \( C \) denotes a positive constant that is independent of the mesh size \( h \). The notation \( x \sim y \) abbreviates \( x \lesssim y \lesssim x \).

### 2.2 Weak Galerkin formulations

We first introduce two spaces:

\[ V_h := \{ v_h \in L^2(\Omega) : v_h|_T \in V(T), \forall T \in \mathcal{T}_h \}, \]
\[ M_h^0 := \{ \mu_h \in L^2(\mathcal{F}_h) : \mu_h|_F \in M(F), \forall F \in \mathcal{F}_h, \mu_h|_{\partial \Omega} = 0 \}, \]

where \( V(T) \) and \( M(F) \) denote two local finite dimensional spaces.

For \( T \in \mathcal{T}_h \), let \( \mathbf{W}(T) \) be a local finite dimensional subspace of \( [L^2(T)]^d \). Then, following [32], we introduce the discrete weak gradient \( \nabla_w : L^2(T) \times L^2(\partial T) \rightarrow \mathbf{W}(T) \) defined by

\[ \nabla_w(v, \mu) = \nabla^i_w v + \nabla^b_w \mu, \forall (v, \mu) \in L^2(T) \times L^2(\partial T), \quad (2.6) \]

where \( \nabla^i_w v, \nabla^b_w \mu \in \mathbf{W}(T) \) satisfy, for any \( q \in \mathbf{W}(T) \),

\[ \langle \nabla^i_w v, q \rangle_T = -(v, \text{div} q)_T, \quad (2.7) \]
\[ \langle \nabla^b_w \mu, q \rangle_T = \langle \mu, q \cdot n \rangle_{\partial T}. \quad (2.8) \]

The WG method for problem (1.1) reads as follows (32): Seek \( (u_h, \lambda_h) \in V_h \times M_h^0 \) such that

\[ a_h((u_h, \lambda_h), (v_h, \mu_h))_\Omega = (f, v_h)_\Omega, \forall (v_h, \mu_h) \in V_h \times M_h^0, \quad (2.9) \]

where

\[ a_h((u_h, \lambda_h), (v_h, \mu_h)) := (a \nabla_w(u_h, \lambda_h), \nabla_w(v_h, \mu_h))_\Omega. \]

For any set \( D \), we denote by \( P_j(D) \) the set of polynomials of degree \( \leq j \) on \( D \). This paper considers two type of WG methods [32] which are based on local RT and BDM mixed elements, respectively:
Type 1. \( V(T) = P_k(T), \ M(F) = P_k(F), \ W(T) = [P_k(T)]^d + P_k(T)x. \)

Type 2. \( V(T) = P_{k-1}(T), \ M(F) = P_k(F), \ W(T) = [P_k(T)]^d \) \((k \geq 1).\)

**Remark 2.1.** When \( a \) is a piecewise constant matrix, we can show that the two type of WG methods are equivalent to the hybridized version of the corresponding mixed RT and BDM method \([2, 15]\) respectively. In fact, by introducing the vector \( p_h := a \nabla_w (u_h, \lambda_h) \) and the space \( W_h := \{ q_h \in [L^2(\Omega)]^d : q_h|T \in W(T) \} \), it’s straightforward that the WG scheme \([2, 9]\) is equivalent to the following problem: Seek \((p_h, u_h, \lambda_h) \in W_h \times V_h \times M_h^0, \) such that

\[
(a^{-1}p_h, q_h)_\Omega + \sum_{T \in T_h} (u_h, \text{div}q_h)_T - \sum_{T \in T_h} \langle \lambda_h, q_h \cdot n \rangle_{\partial T} = 0,
\]

\[
- \sum_{T \in T_h} (v_h, \text{div}p_h)_T = (f, v_h)_\Omega,
\]

\[
\sum_{T \in T_h} \langle p_h \cdot n, \mu_h \rangle_{\partial T} = 0
\]

hold for all \((q_h, v_h, \mu_h) \in W_h \times V_h \times M_h^0.\) This scheme is no other than the hybridized version of the RT mixed element method (cf. (1.18) in \([2]\)) or the BDM mixed method (cf. (1.13) in \([18]\)).

In the following we give an operator form and a matrix form of the WG discretization \([2, 15]\). Let \( \{ \phi_i : i = 1, 2, \ldots, M \} \subset V_h \) and \( \{ \eta_i : i = 1, 2, \ldots, N \} \subset M_h^0 \) be nodal bases for \( V_h \) and \( M_h^0 \), respectively. Denote by \( \tilde{u}_h, \tilde{v}_h \in \mathbb{R}^N \) the vectors of coefficients of \( u_h, v_h \) in the \( \{ \phi_i \} \)-basis, and by \( \tilde{\lambda}_h, \tilde{\mu}_h \in \mathbb{R}^N \) the vectors of coefficients of \( \lambda_h, \mu_h \) in the \( \{ \eta_i \} \)-basis, respectively.

Define the operators \( C_h : V_h \rightarrow V_h, \ B_h : V_h \rightarrow M_h^0, \ B_h^T : M_h^0 \rightarrow V_h, \ D_h : M_h^0 \rightarrow M_h^0, \) and the matrices \( B_h \in \mathbb{R}^{N \times M}, \ C_h \in \mathbb{R}^{N \times M}, \ D_h \in \mathbb{R}^{N \times N} \) respectively by

\[
(C_h u_h, v_h)_\Omega := (a \nabla^i_w u_h, \nabla^i_w v_h)_\Omega =: \tilde{u}_h^T C_h \tilde{v}_h,
\]

\[
(B_h u_h, \lambda_h)_\Omega := (a \nabla^i_w u_h, \nabla^b_w \lambda_h)_\Omega =: (u_h, B_h^T \lambda_h)_\Omega =: \tilde{u}_h^T B_h^T \tilde{\lambda}_h, \quad \forall u_h, \lambda_h \in M_h^0,
\]

\[
(D_h \lambda_h, \mu_h)_\Omega := (a \nabla^b_w \lambda_h, \nabla^b_w \mu_h)_\Omega =: \tilde{\lambda}_h^T D_h \tilde{\mu}_h, \quad \forall \lambda_h, \mu_h \in M_h^0.
\]

Let \( A_h : V_h \times M_h^0 \rightarrow V_h \times M_h^0 \) and \( A_h \in \mathbb{R}^{(M+N) \times (M+N)} \) be defined by

\[
(A_h(u_h, \lambda_h), (v_h, \mu_h))_h := a_h((u_h, \lambda_h), (v_h, \mu_h)) =: (\tilde{u}_h^T \tilde{\lambda}_h^T) A_h \begin{pmatrix} \tilde{v}_h \\ \tilde{\mu}_h \end{pmatrix} \quad (2.10)
\]

for any \((u_h, \lambda_h), (v_h, \mu_h) \in V_h \times M_h^0.\) Then we have

\[
A_h = \begin{pmatrix} C_h & B_h^T \\ B_h & D_h \end{pmatrix}, \quad A_h = \begin{pmatrix} C_h & B_h^T \\ B_h & D_h \end{pmatrix}, \quad (2.11)
\]
and the WG discretization (2.9) is equivalent to the following system: Seek \((u_h, \lambda_h) \in V_h \times M_h^0\) such that
\[
A_h(u_h, \lambda_h) = b_h
\] (2.12)
with \(b_h := (f_h, 0)\) and \(f_h \in V_h\) denoting the standard \(L^2\)-orthogonal projection of \(f\) onto \(V_h\).

### 3 Conditioning of WG methods

In what follows we assume \(T_h\) to be a quasi-uniform triangulation. We recall that \(\|\cdot\|_h, \|\cdot\|_T, \|\cdot\|_{h, \partial T}\) and \(\|\cdot\|_{h, \partial T}\) are defined in Subsection 2.1.

We first present a basic estimate as follows.

**Lemma 3.1.** For any \(\mu_h \in M_h^0\), it holds
\[
\|\mu_h\|_h \lesssim |\mu_h|_h. \tag{3.1}
\]

**Proof.** See Appendix A.

For any simplex \(T\), define
\[
M(\partial T) := \{\mu \in L^2(\partial T) : \mu|_F \in M(F), \text{ for each face } F \text{ of } T\}.
\]

The following lemma gives some basic estimates of weak gradients.

**Lemma 3.2.** For any \(T \in T_h\) and \((v, \mu) \in V(T) \times M(\partial T)\), it holds
\[
\begin{align*}
\|\nabla w^e v\|_T &\sim h^{-1} \|v\|_T, \\
\|\nabla^b w \mu\|_T &\sim h^{-1} \|\mu\|_{h, \partial T}, \\
\|\nabla w(v, \mu)\|_T &\sim h^{-1} \|v - m_T(\mu)\|_T + |\mu|_{h, \partial T}.
\end{align*} \tag{3.2}
\]

**Proof.** See Appendix B.

In view of Lemmas 3.1–3.2 we have the following conclusion.

**Theorem 3.1.** For any \((v_h, \mu_h) \in V_h \times M_h^0\), it holds
\[
\|(v_h, \mu_h)\|_h^2 \lesssim a_h((v_h, \mu_h), (v_h, \mu_h)) \lesssim h^{-2} \|(v_h, \mu_h)\|_h^2. \tag{3.3}
\]
Proof. From (3.2c) it follows
\[
\| \nabla_w (v_h, \mu_h) \|^2 \sim \sum_{T \in T_h} h_T^{-2} \| v_h - m_T(\mu_h) \|_T^2 + |\mu_h|^2_h. \tag{3.4}
\]
Since
\[
|m_T(\mu_h)| \lesssim \frac{1}{d+1} \sum_{F \in F_T} |1/F| \int_F \mu_h | \lesssim h_T^{-d+1} \| \mu_h \|_{\partial T},
\]
we have
\[
\| m_T(\mu_h) \|_T \lesssim h_T^{\frac{1}{2}} \| \mu_h \|_{\partial T} \lesssim \| \mu_h \|_{h, \partial T}, \tag{3.5}
\]
which, together with Lemmas 3.1-3.2, implies
\[
\| (v_h, \mu_h) \|_h^2 \lesssim \sum_{T \in T_h} \left\{ \| v_h - m_T(\mu_h) \|_T^2 + \| m_T(\mu_h) \|_T^2 \right\} 
\lesssim \sum_{T \in T_h} \| v_h - m_T(\mu_h) \|_T^2 + \| \mu_h \|_h^2 
\lesssim \sum_{T \in T_h} \| v_h - m_T(\mu_h) \|_T^2 + |\mu_h|^2_h. \tag{3.6}
\]
A combination of (3.1), (3.4) and (3.6) yields
\[
\| (v_h, \mu_h) \|_h^2 = \| v_h \|^2 + \| \mu_h \|_h^2 \lesssim a_h((v_h, \mu_h), (v_h, \mu_h)). \tag{3.7}
\]
On the other hand, it holds
\[
a_h((v_h, \mu_h), (v_h, \mu_h)) \lesssim \| \nabla_w v_h \|^2 + \| \nabla_w \mu_h \|^2 
\lesssim h^{-2} \| v_h \|^2 + h^{-2} \| \mu_h \|_h^2 \text{ by (3.2a) and (3.2b)} 
\lesssim h^{-2} \| (v_h, \mu_h) \|_h^2. \tag{3.8}
\]
The estimates (3.7)-(3.8) lead to the desired result (3.3). \(\square\)

**Theorem 3.2.** It holds
\[
\sup_{(v_h, \mu_h) \in V_h \times M_0^h} \frac{a_h((v_h, \mu_h), (v_h, \mu_h))}{\| (v_h, \mu_h) \|_h^2} \gtrsim h^{-2}. \tag{3.9}
\]
In addition,
\[
\inf_{(v_h, \mu_h) \in V_h \times M_0^h} \frac{a_h((v_h, \mu_h), (v_h, \mu_h))}{\| (v_h, \mu_h) \|_h^2} \lesssim 1 \tag{3.10}
\]
holds if \(h\) is sufficiently small.
Proof. Given $v_h \in V_h$, from Lemma 3.2 it follows

\[ a_h((v_h,0),(v_h,0)) \sim h^{-2} \|v_h\|^2, \]  

which implies (3.9).

Let $s$ be the smallest eigenvalue of problem (1.1) with $f = su$ and let $u_0 \in H^1_0(\Omega)$ be the corresponding eigenvector function. Then it holds

\[ \|\nabla u_0\|^2 \sim s \|u_0\|^2. \]  

(3.12)

In the analysis below, we shall denote by $C$ a positive constant that is independent of the mesh size $h$ and may take a different value at its each occurrence.

We define $(v_h,\mu_h) \in V_h \times M^0_h$ by

\[ v_h|_T = m_T(u_0), \quad \forall T \in \mathcal{T}_h, \]

\[ \mu_h|_F = \frac{1}{|F|} \int_F u_0, \quad \forall F \in \mathcal{F}_h. \]

By the definition of $m_T(\cdot)$ it is easy to see

\[ m_T(\mu_h) = \frac{1}{d+1} \sum_{F \in \mathcal{F}_T} \frac{1}{|F|} \int_F \mu_h = m_T(u_0). \]  

(3.13)

Standard scaling arguments yield

\[ \|u_0 - m_T(u_0)\|_T \lesssim h_T|u_0|_{1,T}, \]

\[ \|\mu_h - m_T(\mu_h)\|_{\partial T} \lesssim h_T^2|u_0|_{1,T}. \]  

(3.14)  

(3.15)

Thus, in view of (3.14) and (3.12) we have

\[ \|v_h\|^2 = \sum_{T \in \mathcal{T}_h} \|m_T(u_0)\|^2_T \geq \sum_{T \in \mathcal{T}_h} \left\{ \frac{1}{2} \|u_0\|^2_T - \|u_0 - m_T(u_0)\|^2_T \right\} \]

\[ \gtrsim \sum_{T \in \mathcal{T}_h} \left\{ \|u_0\|^2_T - Ch^2_T|u_0|^2_{1,T} \right\} \]

\[ \gtrsim (1 - sCh^2) \|u_0\|^2, \]

which, together with (3.15) and (3.13), further implies

\[ \|\mu_h\|^2 \geq \sum_{T \in \mathcal{T}_h} h_T \left( \frac{1}{2} \|m_T(\mu_h)\|^2_{\partial T} - \|\mu_h - m_T(\mu_h)\|^2_{\partial T} \right) \]

\[ \gtrsim \sum_{T \in \mathcal{T}_h} h_T \|m_T(\mu_h)\|^2_{\partial T} - Ch^2|u_0|^2_{1,\Omega} \]

\[ \gtrsim \sum_{T \in \mathcal{T}_h} \|m_T(u_0)\|^2_T - Ch^2|u_0|^2_{1,\Omega} \]

\[ \gtrsim (1 - sCh^2) \|u_0\|^2. \]  

(3.16)  

(3.17)
On the other hand, from the definition (2.5) and the estimate (3.15) it follows
\[ |\mu_h|_h \lesssim |u_0|_{1,\Omega}. \] (3.18)
Therefore, it holds
\[
\frac{a_h((v_h, \mu_h), (v_h, \mu_h))}{\|(v_h, \mu_h)\|_h^2} \sim \frac{\|\nabla w(v_h, \mu_h)\|^2}{\|(v_h, \mu_h)\|_h^2} \\
\lesssim \frac{|\mu_h|^2}{\|v_h\|^2 + \|\mu_h\|^2} \quad \text{(by (3.2c))} \\
\lesssim \frac{|\mu_h|^2}{(1 - sC h^2) \|u_0\|^2} \quad \text{(by (3.16) and (3.17))} \\
\lesssim \frac{s \|u_0\|^2}{(1 - sC h^2) \|u_0\|^2} \quad \text{(by (3.18))} \\
\lesssim \frac{s}{1 - sC h^2},
\]
which indicates the inequality (3.10) immediately.

In light of Theorems 3.1-3.2, it’s straightforward to derive the following theorem.

**Theorem 3.3.** Let \( A_h \) be the operator defined by (2.10), then it holds
\[ \kappa(A_h) \lesssim h^{-2}, \] (3.19)
where \( \kappa(A_h) := \frac{\lambda_{\text{max}}(A_h)}{\lambda_{\text{min}}(A_h)} \), with \( \lambda_{\text{max}}(A_h), \lambda_{\text{min}}(A_h) \) denoting the largest and smallest eigenvalues of \( A_h \) respectively. Further more, it holds
\[ \kappa(A_h) = O(h^{-2}) \] (3.20)
if \( h \) is sufficiently small.

**Remark 3.1.** Let \( A_h \) be the stiffness matrix of \( a_h(\cdot, \cdot) \) defined by (2.10), then we easily have \( \kappa(A_h) \sim \kappa(A_h) = O(h^{-2}) \).

## 4 Two-level algorithm

In this section, we analyze a two-level algorithm for the discrete system (2.12). For the sake of clarity, our description is in operator form.
4.1 Algorithm definition

Set
\[ \tilde{V}_h := \{ \tilde{v}_h \in H^1_0(\Omega) : \tilde{v}_h|_T \in P_1(T), \forall T \in T_h \}. \tag{4.1} \]

We first define the prolongation operator \( I_h : \tilde{V}_h \rightarrow V_h \times M^0_h \) as follows: for any \( \tilde{v}_h \in \tilde{V}_h \),
\[ I_h \tilde{v}_h := (I_h^1 \tilde{v}_h, I_h^2 \tilde{v}_h) \in V_h \times M^0_h \]
satisfies
\[ \begin{align*}
    \int_T I_h^1 \tilde{v}_h v &= \int_T \tilde{v}_h v, \quad \forall v \in V(T), \quad \forall T \in T_h, \\
    \int_F I_h^2 \tilde{v}_h \mu &= \int_F \tilde{v}_h \mu, \quad \forall \mu \in M(F), \quad \forall F \in F_h.
\end{align*} \]

Then define the adjoint operator, \( I_h^t \), of \( I_h \) by
\[ (I_h^t(u_h, \mu_h), \tilde{v}_h)_\Omega := ((u_h, \mu_h), I_h \tilde{v}_h)_h, \quad \forall (u_h, \mu_h) \in V_h \times M^0_h, \forall \tilde{v}_h \in \tilde{V}_h. \]

Define \( \bar{A}_h : \tilde{V}_h \rightarrow \tilde{V}_h \) by
\[ (\bar{A}_h \tilde{v}_h, \tilde{v}_h)_\Omega := (a \nabla \tilde{v}_h, \nabla \tilde{v}_h)_\Omega, \quad \forall \tilde{v}_h \in \tilde{V}_h. \tag{4.2} \]

**Remark 4.1.** By the definition of \( I_h \), it’s trivial to verify that \( \nabla_a I_h \tilde{v}_h = \nabla \tilde{v}_h, \quad \forall \tilde{v}_h \in \tilde{V}_h. \)
Thus we have the following important relationship:
\[ \bar{A}_h = I_h^t A_h I_h. \tag{4.3} \]

Let \( \bar{R}_h : \tilde{V}_h \rightarrow \tilde{V}_h \) be a good approximation of \( \bar{A}_h^{-1} \) and define \( \bar{R}_h^t \) by
\[ (\bar{R}_h^t \tilde{u}_h, \tilde{v}_h)_\Omega := (\tilde{u}_h, \bar{R}_h \tilde{v}_h)_\Omega, \quad \forall \tilde{u}_h, \tilde{v}_h \in \tilde{V}_h. \]

Let \( \bar{R}_h : V_h \times M^0_h \rightarrow V_h \times M^0_h \) be a good approximation of \( \bar{A}_h^{-1} \) and let \( \bar{R}_h^t : V_h \times M^0_h \rightarrow V_h \times M^0_h \) be defined by
\[ (\bar{R}_h^t(u_h, \lambda_h), (v_h, \mu_h))_h := ((u_h, \lambda_h), \bar{R}_h(v_h, \mu_h))_h, \quad \forall (u_h, \lambda_h), (v_h, \mu_h) \in V_h \times M^0_h. \]

Using the above operators, we define an ingredient operator \( B_h : V_h \times M^0_h \rightarrow V_h \times M^0_h \) as follows:

**Algorithm 1.** For any \( b_h \in V_h \times M^0_h \), define \( B_h b_h = (v^1_h, \mu^1_h) \) by
1. Smooth: \( (v^1_h, \mu^1_h) := \bar{R}_h b_h \),
2. Correct: \( (v^2_h, \mu^2_h) := (v^1_h, \mu^1_h) + I_h \bar{R}_h I_h^t (b_h - A_h (v^1_h, \mu^1_h)) \),
3. Correct: \( (v^3_h, \mu^3_h) := (v^2_h, \mu^2_h) + I_h \bar{R}_h^t I_h^t (b_h - A_h (v^2_h, \mu^2_h)) \),
4. Smooth: \( (v^4_h, \mu^4_h) := (v^3_h, \mu^3_h) + \bar{R}_h^t (b_h - A_h (v^3_h, \mu^3_h)) \).

10
We are now in a position to present the two-level algorithm for the system (2.12).

Algorithm 2. Set \((u_0^h, \lambda_0^h) = (0, 0)\),

\[ \text{for } j = 1, 2, \ldots \text{ till convergence} \]
\[ (u_j^h, \lambda_j^h) := (u_{j-1}^h, \lambda_{j-1}^h) + B_h(b_h - A_h(u_{j-1}^h, \lambda_{j-1}^h)). \]

\[ \text{end} \]

4.2 Convergence analysis

At first, we introduce some abstract notations. Let \(X\) be a finite dimensional Hilbert space with inner product \((\cdot, \cdot)\) and its induced norm \(\|\cdot\|\). For any linear SPD operator \(A : X \to X\), the notation \((\cdot, \cdot)_A := (A\cdot, \cdot)\) defines an inner product on \(X\) and we denote by \(\|\cdot\|_A\) the norm induced by \((\cdot, \cdot)_A\). Let \(B : X \to X\) be a linear operator with

\[ \|B\|_A := \sup_{0 \neq x \in X} \frac{\|Bx\|_A}{\|x\|_A}. \]

From the definition of \(B_h\) in Algorithm 1, we easily obtain the following lemma.

Lemma 4.1. It holds

\[ I - B_hA_h = (I - R_hI_hA_h)(I - I_h\tilde{R}_hI_hA_h)(I - I_h\tilde{R}_hI_hA_h)(I - R_hA_h). \quad (4.4) \]

It’s trivial to verify that \(I - B_hA_h\) is symmetric semi-positive definite with respect to the inner product \((\cdot, \cdot)_A\), and thus it follows \(\lambda_{\text{max}}(B_hA_h) \leq 1\) and

\[ \|I - B_hA_h\|_A = 1 - \lambda_{\text{min}}(B_hA_h). \quad (4.5) \]

Now we introduce the symmetrizations of \(R_h\) and \(\tilde{R}_h\), i.e.

\[ \overline{R}_h := R_h^t + R_h - R_hA_hR_h, \quad (4.6) \]
\[ \overline{\tilde{R}}_h := \tilde{R}_h^t + \tilde{R}_h - \tilde{R}_hA_h\tilde{R}_h, \quad (4.7) \]

and make the following assumption.

Assumption I. The operators \(R_h\) and \(\tilde{R}_h\) are such that

\[ \|I - \overline{R}_hA_h\|_A < 1, \quad (4.8) \]
\[ \|I - \overline{\tilde{R}}_hA_h\|_A < 1. \quad (4.9) \]
Remark 4.2. It follows from (4.8) that $R_h$ is SPD with respect to the inner product $(\cdot, \cdot)_h$. Then it follows from

$$R_h = R_h^t(R_h^{-t} + R_h^{-1} - A_h)R_h$$

that $R_h^{-t} + R_h^{-1} - A_h$ is SPD with respect to the inner product $(\cdot, \cdot)_h$. Similarly, $\overline{R_h}$ and $\overline{R_h}^{-t} + \overline{R_h}^{-1} - \overline{A_h}$ are both SPD with respect to the inner product $(\cdot, \cdot)_\Omega$.

Following the basic idea of the X-Z identity ([15], [20], [19]), we have the following ingredient theorem.

Theorem 4.1. Under Assumption II $B_h$ is a SPD operator with respect to the inner product $(\cdot, \cdot)_h$, and, for any $(u_h, \lambda_h) \in V_h \times M^0_h$, it holds

$$(B_h^{-1}(u_h, \lambda_h), (u_h, \lambda_h))_h = \inf_{(v_h, \mu_h), \tilde{v}_h, \mu_h} \|(v_h, \mu_h) + R_h^t \mu_h I_h \tilde{v}_h\|_{R_h^{-1}}^2 + \|\tilde{v}_h\|_{R_h}^2.$$

Further more, it holds the following extended X-Z identity:

$$\|I - B_h A_h\|_{\mathcal{A}_h} = 1 - \frac{1}{K},$$

where

$$K = \sup_{\|u_h, \lambda_h\|_{\mathcal{A}_h} = 1} \inf_{(v_h, \mu_h), \tilde{v}_h, \mu_h} \|(v_h, \mu_h) + R_h^t \mu_h I_h \tilde{v}_h\|_{R_h^{-1}}^2 + \|\tilde{v}_h\|_{R_h}^2.$$  (4.12)

Proof. The desired results follow from a trivial modification of the proof of the X-Z identity in [19]. For completeness we sketch the proof of this theorem. We note that $\tilde{V}_h \not\subset V_h \times M^0_h$ means the corresponding spaces here are nonnested.

Denote $X_h := (V_h \times M^0_h) \times \tilde{V}_h$ and define the inner product $[\cdot, \cdot]$ on $X_h$ by

$$[(a, b), (c, d)] := (a, c)_h + (b, d)_\Omega, \quad \forall (a, b), (c, d) \in X_h.$$  

Introduce the operator $\Pi_h : X_h \to V_h \times M^0_h$ and its adjoint operator $\Pi_h^t : V_h \times M^0_h \to X_h$ with

$$\Pi_h := \begin{pmatrix} I & 0 \\ 0 & I_h \end{pmatrix}, \quad \text{i.e. } \Pi_h(a, b) = a + I_h b \text{ for any } (a, b) \in X_h,$$

$$\Pi_h^t := \begin{pmatrix} I & I_h \end{pmatrix}, \quad \text{i.e. } \Pi_h^t a = \begin{pmatrix} a \\ I_h a \end{pmatrix} \text{ for any } a \in V_h \times M^0_h.$$  

Obviously, we have $(\Pi_h \bar{a}, b)_h = [\bar{a}, \Pi_h^t b], \forall \bar{a} \in X_h, \forall b \in V_h \times M^0_h.$
Now define
\[
\begin{align*}
\widetilde{A}_h &:= \begin{pmatrix} A_h & A_h I_h \\ I_h A_h & I_h A_h I_h \end{pmatrix}, \\
\widetilde{B}_h &:= \begin{pmatrix} \widetilde{R}_h^{-1} & 0 \\ I_h A_h & \widetilde{R}_h^{t} \end{pmatrix}^{-1},
\end{align*}
\]
and denote by \(\widetilde{D}_h\) the diagonal of \(\widetilde{A}_h\).

For any \(b_h \in V_h \times M_0^h\), set
\[
\begin{align*}
w_1 &:= \widetilde{B}_h \Pi_t b_h, \\
w_2 &:= w_1 + \widetilde{B}_h (\Pi_t b_h - \widetilde{A}_h w_1).
\end{align*}
\]
Then it holds \(\Pi_h w_2 = \Pi_h \widetilde{B}_h \Pi_t b_h\), where
\[
\begin{align*}
\widetilde{B}_h := \widetilde{B}_h + \widetilde{B}_h - \widetilde{A}_h \widetilde{D}_h \widetilde{B}_h.
\end{align*}
\]
It’s easy to verify that \(\Pi_h w_2 = B_h b_h\), which yields
\[
B_h = \Pi_h \widetilde{B}_h \Pi_t^t. \tag{4.13}
\]
Denoting \(\widetilde{R}_h := \text{diag}(R_h, \widetilde{R}_h)\), we have
\[
\begin{align*}
\widetilde{B}_h &= \widetilde{B}_h (\widetilde{B}_h + \widetilde{R}_h - \widetilde{A}_h) \widetilde{B}_h \\
&= \widetilde{B}_h (\widetilde{R}_h + \widetilde{R}_h - \widetilde{D}_h) \widetilde{B}_h \\
&= \widetilde{B}_h \widetilde{R}_h \widetilde{R}_h \widetilde{B}_h,
\end{align*}
\]
where \(\widetilde{R}_h = \widetilde{R}_h + \widetilde{R}_h - \widetilde{D}_h \widetilde{R}_h\). By \(4.3\) we also have \(\widetilde{R}_h = \text{diag}(\widetilde{R}_h, \widetilde{R}_h)\). From Remark \(4.2\), it follows that \(\widetilde{R}_h\) is SPD with respect to \([\cdot, \cdot]\). Thus \(\widetilde{B}_h\) is SPD with respect to \([\cdot, \cdot]\). Then from Theorem 1 in \(19\) and \(4.13\) it follows
\[
(B_h^{-1}(u_h, \lambda_h), (u_h, \lambda_h))_h = \inf_{\Pi_h w_h = (u_h, \lambda_h)} [\widetilde{B}_h^{-1}]_{w_h, w_h}. \tag{4.14}
\]
In view of
\[
\begin{align*}
\widetilde{B}_h^{-1} &= \widetilde{B}_h \widetilde{R}_h \widetilde{R}_h \widetilde{B}_h = \begin{pmatrix} I & 0 \\ I_h A_h \widetilde{R}_h & I \end{pmatrix} \begin{pmatrix} \widetilde{R}_h^{-1} & 0 \\ 0 & I \end{pmatrix}.
\end{align*}
\]
the identity \(4.10\) follows immediately from \(4.14\). The extended X-Z identity \(4.11\) is just a trivial conclusion from \(4.10\).
Lemma 4.2. For any \( \lambda_h \in M_h^0 \), \( P_h \lambda_h \) satisfies
\[
\begin{cases}
P_h \lambda_h(x) &= \sum_{T \in \omega_x} \frac{\sum_{T \in \omega_x} m_T(\lambda_h)}{\sum_{T \in \omega_x} 1}, \quad \text{for each interior vertex } x \text{ of } T_h, \\
\quad P_h \lambda_h(x) &= 0, \quad \text{for each vertex } x \in \partial \Omega,
\end{cases}
\]
where the set \( \omega_x := \{ T \in T_h : x \text{ is a vertex of } T \} \).

As for the operator \( P_h \), we have the following important estimates.

**Lemma 4.2.** For any \( (u_h, \lambda_h) \in V_h \times M_h^0 \), it holds
\[
\begin{align*}
\| (I - P_h^T P_h) \lambda_h \|_{\Omega} &\lesssim h \| (u_h, \lambda_h) \|_{A_h}, \quad (4.15) \\
\| u_h - P_h^T P_h \lambda_h \| &\lesssim h \| (u_h, \lambda_h) \|_{A_h}, \quad (4.16)
\end{align*}
\]
which further indicate
\[
\| (u_h, \lambda_h) - I_h P_h \lambda_h \| \lesssim h \| (u_h, \lambda_h) \|_{A_h}. \quad (4.17)
\]

**Proof.** We denote by \( \omega_T \) the set \( \{ T' \in T_h : T' \text{ and } T \text{ share a vertex} \} \) and by \( N(T) \) the set of all vertices of \( T \). Since
\[
\begin{align*}
&h_T \left\| P_h^T P_h \lambda_h - m_T(\lambda_h) \right\|_{\partial T}^2 \\
\leq & \ h_T \left\| P_h \lambda_h - m_T(\lambda_h) \right\|_{\partial T}^2 \\
\lesssim & \ h_T^d \sum_{x \in N(T)} |P_h \lambda_h(x) - m_T(\lambda_h)|^2 \\
\lesssim & \ h_T^d \sum_{x \in N(T)} \sum_{T_1, T_2 \in \omega_x} |m_{T_1}(\lambda_h) - m_{T_2}(\lambda_h)|^2 \\
\lesssim & \ h_T^2 \sum_{T' \in \omega_T} |\lambda_h|_{h, \partial T'}^2,
\end{align*}
\]
we have
\[
\begin{align*}
&h_T \left\| (I - P_h^T P_h) \lambda_h \right\|_{\partial T}^2 \lesssim h_T \| \lambda_h - m_T(\lambda_h) \|_{\partial T}^2 + h_T \left\| P_h^T P_h \lambda_h - m_T(\lambda_h) \right\|_{\partial T}^2 \\
\lesssim & \ h_T^2 \sum_{T' \in \omega_T} |\lambda_h|_{h, \partial T'}^2.
\end{align*}
\]
Then the estimate \( (4.15) \) follows immediately from \( (4.18) \).

On the other hand, since
\[
\left\| P_h^T P_h \lambda_h - m_T(\lambda_h) \right\|_{T}^2 \leq \left\| P_h \lambda_h - m_T(\lambda_h) \right\|_{T}^2
\]
\[
\lesssim \ h_T \left\| P_h \lambda_h - m_T(\lambda_h) \right\|_{\partial T}^2
\]
\[
\lesssim \ h_T^2 \sum_{T' \in \omega_T} |\lambda_h|_{h, \partial T'}^2, \quad \text{(by \( (4.16) \))}
\]

14
it holds
\[ \|u_h - I_h^T P_h \lambda_h\|^2_T \lesssim \|u_h - m_T(\lambda_h)\|^2_T + \|I_h^T P_h \lambda_h - m_T(\lambda_h)\|^2_T \]
\[ \lesssim \|u_h - m_T(\lambda_h)\|^2_T + \sum_{T' \in \omega_T} h_{T'}^2 |\lambda_h|_{h,\partial T'}^2. \]

Then the estimate (4.16) also follows immediately from (3.2c).

Finally, the result (4.17) is a trivial conclusion from (4.15) and (4.16).

**Lemma 4.3.** For any \((u_h, \lambda_h) \in V_h \times M_h^0\), it holds
\[ \|I_h P_h \lambda_h\|_{\mathcal{A}_h} \lesssim \|(u_h, \lambda_h)\|_{\mathcal{A}_h}. \quad (4.19) \]

**Proof.** The relation \(\|I_h P_h \lambda_h\|_{\mathcal{A}_h} = \|P_h \lambda_h\|_{\tilde{\mathcal{A}}_h}\) follows from (4.3). It suffices to prove the inequality of (4.19). Since
\[ |P_h \lambda_h|_{h,1,T}^2 = |P_h \lambda_h - m_T(\lambda_h)|_{h,1,T}^2 \]
\[ \lesssim h_{1,T}^{-2} \|P_h \lambda_h - m_T(\lambda_h)\|^2_T \quad \text{(by inverse estimate)} \]
\[ \lesssim h_{1,T}^{-1} \|P_h \lambda_h - m_T(\lambda_h)\|_{\partial T}^2 \]
\[ \lesssim \sum_{T' \in \omega_T} h_{T'}^2 |\lambda_h|_{h,\partial T'}^2, \quad \text{(by (4.18))} \]
we have
\[ \|P_h \lambda_h\|_{\tilde{\mathcal{A}}_h} \sim |P_h \lambda_h|_{h,1,\Omega} \lesssim |\lambda_h|_h, \]
which, together with (3.2c), implies the desired conclusion.

**Assumption II.** The smoother \(\mathcal{R}_h : V_h \times M_h^0 \to V_h \times M_h^0\) is SPD with respect to \((\cdot, \cdot)_h\) and satisfies
\[ \sigma(\mathcal{R}_h \mathcal{A}_h) \subset (0, 1], \quad (4.20) \]
where \(\sigma(\mathcal{R}_h \mathcal{A}_h)\) denotes the set of all eigenvalues of \(\mathcal{R}_h \mathcal{A}_h\). What’s more, for any \((u_h, \lambda_h) \in V_h \times M_h^0\), it holds
\[ \|(u_h, \lambda_h)\|^2_{R_h^{-1}} \leq C_R \lambda_{\max}(\mathcal{A}_h) \|(u_h, \lambda_h)\|^2_h, \quad (4.21) \]
where \(R_h\) is the symmetrization of \(\mathcal{R}_h\), and \(C_R\) denotes a positive constant.

**Remark 4.3.** If we take \(\mathcal{R}_h = \frac{1}{\lambda_{\max}(\mathcal{A}_h)} I\), then it holds \(\mathcal{R}_h^{-1} = \frac{\lambda_{\max}(\mathcal{A}_h)}{2\lambda_{\max}(\mathcal{A}_h)}(2\lambda_{\max}(\mathcal{A}_h)I - \mathcal{A}_h)^{-1}. In this case it is obvious that \(C_R = 1\). If we take \(\mathcal{R}_h\) to be the symmetric Gauss-Seidel smoother, then \(C_R\) is a bounded positive constant independent of the mesh size \(h\).
Remark 4.4. Suppose Assumption III holds, then the relation \( I - \overline{R}_h A_h = (I - R_h A_h)^2 \) leads to \( \sigma(I - \overline{R}_h A_h) \subset [0, 1) \) and it follows \( \| I - \overline{R}_h A_h \|_{A_h} < 1 \).

Lemma 4.4. Under Assumption III, for any \((u_h, \lambda_h) \in V_h \times M^0_h\), it holds

\[
\| R_h A_h (u_h, \lambda_h) \|_{\overline{R}_h^{-1}} \leq \| (u_h, \lambda_h) \|_{A_h}. \tag{4.22}
\]

Proof. Denoting \( S_h := R_h A_h \) and thanks to

\[
\overline{R}_h = 2R_h - R_h A_h R_h = (2S_h - S_h^2) A_h^{-1},
\]

we have

\[
\| R_h A_h (u_h, \lambda_h) \|_{\overline{R}_h^{-1}}^2 = (\overline{R}_h^{-1} R_h A_h (u_h, \lambda_h), R_h A_h (u_h, \lambda_h))_h
\]

\[
= (A_h (2S_h - S_h^2)^{-1} S_h (u_h, \lambda_h), R_h A_h (u_h, \lambda_h))_h \tag{4.23}
\]

\[
= (S_h (2S_h - S_h^2)^{-1} S_h (u_h, \lambda_h), (u_h, \lambda_h))_{A_h},
\]

which, together with the fact that \( S_h \) is SPD with respect to \((\cdot, \cdot)_{A_h}\) and the inequality

\[
t(2t - t^2)^{-1} t \leq 1, t \in (0, 1],
\]

yields the desired estimate (4.22).

Finally, we state the following convergence theorem.

Theorem 4.2. Under Assumptions III, it holds

\[
\| (I - B_h A_h) \|_{A_h} \leq 1 - \frac{1}{K}, \tag{4.24}
\]

where

\[
K \lesssim \left( 1 + C_R + \frac{1}{1 - \| I - \overline{R}_h A_h \|_{A_h}} \right). \tag{4.25}
\]

Proof. For any \((u_h, \lambda_h) \in V_h \times M^0_h\), set

\[
\overline{v}_h := P_h \lambda_h, \quad (v_h, \mu_h) := (u_h, \lambda_h) - I_h \overline{v}_h,
\]

we then obtain

\[
\| (v_h, \mu_h) + R_h A_h I_h \overline{v}_h \|_{\overline{R}_h^{-1}}^2 \leq \| (v_h, \mu_h) \|_{\overline{R}_h^{-1}}^2 + \| R_h A_h I_h \overline{v}_h \|_{\overline{R}_h^{-1}}^2
\]

\[
\lesssim \| (v_h, \mu_h) \|_{\overline{R}_h^{-1}}^2 + \| I_h \overline{v}_h \|_{A_h}^2 \tag{by (4.22)}
\]

\[
\lesssim \| (v_h, \mu_h) \|_{\overline{R}_h^{-1}}^2 + \| (u_h, \lambda_h) \|_{A_h}^2 \tag{by (4.19)}
\]

\[
\lesssim C_R \lambda_{\text{max}} (A_h) \| (v_h, \mu_h) \|_{A_h}^2 + \| (u_h, \lambda_h) \|_{A_h}^2 \quad \text{(by Assumption III)}
\]

\[
\lesssim (1 + C_R) \| (u_h, \lambda_h) \|_{A_h}^2,
\]
where, in the last inequality, we have used the estimate (4.17) and the fact $\lambda_{\text{max}}(A_h) \sim h^{-2}$ derived from Theorems 3.1-3.2. Similar to (4.5), we have

$$\| I - R_h \tilde{A}_h \|_{\tilde{A}_h} = 1 - \lambda_{\text{min}}(R_h \tilde{A}_h),$$

and it follows

$$\frac{1}{\lambda_{\text{min}}(R_h \tilde{A}_h)} \| \tilde{v}_h \|^2_{\tilde{A}_h} = 1 - \frac{1}{\| I - R_h \tilde{A}_h \|_{\tilde{A}_h}} \| \tilde{v}_h \|^2_{\tilde{A}_h} \leq \frac{1}{1 - \| I - R_h \tilde{A}_h \|_{\tilde{A}_h}} \| (u_h, \lambda_h) \|^2_{A_h}.$$ (by (4.19))

Therefore, we have

$$\| (v_h, \mu_h) + R_h A_h I_h \tilde{v}_h \|^2_{R_h^{-1}} + \| \tilde{v}_h \|^2_{R_h^{-1}} \leq (1 + C_R + \frac{1}{1 - \| I - R_h \tilde{A}_h \|_{\tilde{A}_h}}) \| (u_h, \lambda_h) \|^2_{A_h},$$

which implies

$$\sup_{(u_h, \lambda_h) = 1} \inf_{(v_h, \mu_h) = (u_h, \lambda_h)} \| (v_h, \mu_h) + R_h A_h I_h \tilde{v}_h \|^2_{R_h^{-1}} + \| \tilde{v}_h \|^2_{R_h^{-1}} \leq (1 + C_R + \frac{1}{1 - \| I - R_h \tilde{A}_h \|_{\tilde{A}_h}}).$$

As a result, the desired estimate (4.24) follows from the extended X-Z identity (4.11) in Theorem 4.1.

**Remark 4.5.** In our analysis, we do not use any regularity assumption of the model problem (1.1). Thus our theory applies to the case that (1.1) doesn’t have full elliptic regularity. However, if $R_h$ is constructed by standard multigrid methods, as shown in [8]-[10], the lack of full regularity will affect the convergence rate $\| I - R_h \tilde{A}_h \|_{\tilde{A}_h}$.

## 5 Numerical experiments

This section reports some numerical results in two space dimensions to verify our theoretical results. For the model problem (1.1), we set $a \in \mathbb{R}^{2 \times 2}$ to be the identity matrix, $\Omega = (0, 1) \times (0, 1)$ and we shall use the **Type 2** WG method $(k = 1)$. When given a coarse triangulation $\mathcal{T}_0$, we produce a sequence of uniformly refined triangulations...
\{T_i : i = 0, 1, \ldots, 5\}$ (cf. Figure 1 for $T_0$ and $T_1$) by a simple procedure: $T_{j+1}$ is obtained by connecting the midpoints of all edges of $T_j$ for $j = 0, 1, 2, 3, 4$. 

In our first experiment, we compute the smallest eigenvalue $\lambda_{\text{min}}(A_h)$, the largest eigenvalue $\lambda_{\text{max}}(A_h)$ and the condition number $\kappa(A_h)$ of the stiffness matrix $A_h$ on each triangulation $T_i$ and list them in Table 1. The results imply $\kappa(A_h) \sim \kappa(A_h) = O(h^{-2})$, which is conformable to Theorem 3.3.

|      | $T_0$ | $T_1$ | $T_2$ | $T_3$ | $T_4$ | $T_5$ |
|------|-------|-------|-------|-------|-------|-------|
| $\lambda_{\text{min}}(A_h)$ | 0.55  | 0.28  | 0.075 | 0.019 | 0.0048| 0.0012|
| $\lambda_{\text{max}}(A_h)$ | 27.63 | 33.31 | 33.46 | 33.47 | 33.47 | 33.47 |
| $\kappa(A_h)$         | 50.1  | 121.1 | 444.8 | 1746.7| 6954.6| 27792 |

In our second experiment, for each triangulation $T_j$, we set $T_h = T_j$ and take $R_h$ to be the $m$-times symmetric Gauss-Seidel iteration with $\tilde{R}_h = \tilde{A}_h^{-1}$. We are to solve the problem $A_h x = b$, where $b$ is a zero vector. In order to verify the convergence, in Algorithm 2 we take $x_0 = (1, 1, \ldots, 1)^t$ as the initial value, rather than the zero vector. We stop the two-level algorithm when the initial error, i.e. $\sqrt{x_0^t A_h x_0}$, is reduced by a factor of $10^{-8}$. The corresponding results listed in Table 2 show that the two-level algorithm is efficient.

Our third experiment is a modification of the second one. In this experiment, the operator $\tilde{R}_h$ is constructed by using the standard V-cycle multigrid method based on the nested triangulations $T_0, T_1, \ldots, T_j$, rather than by simply setting $\tilde{R}_h = \tilde{A}_h^{-1}$. Here we set all smoothers encountered to be the $m$-times symmetric Gauss-Seidel iterations. This

Figure 1: $T_0$ (left) and $T_1$ (right)
is a practical multi-level algorithm. The numerical results listed in Tables 3-4 show that the multi-level algorithm is efficient.

Table 2: Numbers of iterations for two-level algorithm

| m   | T_0 | T_1 | T_2 | T_3 | T_4 | T_5 |
|-----|-----|-----|-----|-----|-----|-----|
| m = 1 | 13  | 23  | 28  | 31  | 31  | 31  |
| m = 2 | 8   | 12  | 14  | 17  | 17  | 17  |
| m = 3 | 7   | 9   | 10  | 12  | 13  | 13  |
| m = 4 | 6   | 8   | 9   | 9   | 10  | 10  |
| m = 10 | 4   | 6   | 6   | 7   | 7   | 7   |

Table 3: Number of iterations for multi-level algorithm

| m   | T_1 | T_2 | T_3 | T_4 | T_5 |
|-----|-----|-----|-----|-----|-----|
| m = 1 | 22  | 28  | 31  | 31  | 31  |
| m = 2 | 12  | 14  | 17  | 17  | 17  |
| m = 3 | 9   | 10  | 12  | 13  | 13  |

Table 4: Average error reduction rates for multi-level algorithm

| m   | T_1 | T_2 | T_3 | T_4 | T_5 |
|-----|-----|-----|-----|-----|-----|
| m = 1 | 0.53| 0.56| 0.56| 0.56| 0.56|
| m = 2 | 0.31| 0.36| 0.36| 0.36| 0.36|
| m = 3 | 0.19| 0.25| 0.26| 0.26| 0.27|

A Appendix: Proof of Lemma 3.1

For any simplex $T$ with vertexes $a_1, a_2, \ldots, a_{d+1}$, let $\lambda_i$ be the barycentric coordinate function associated with the vertex $a_i$ for $i = 1, 2, \ldots, d + 1$. We first introduce

$$\Lambda(T) := Q_1(T) + Q_2(T) + \ldots + Q_{d+1}(T),$$

19
where

$$Q_i(T) = \left( \prod_{j \neq i} \lambda_j \right) \text{span} \left\{ \prod_j \lambda_j^{\alpha_j} : \sum_j \alpha_j = k, \alpha_i = 0 \right\}, i = 1, 2, \ldots, d + 1.$$ 

Then we define the operator $S : L^2(\partial T) \rightarrow \Lambda(T)$ as follows: For any $\mu \in L^2(\partial T)$, $S\mu$ satisfies

$$\int_F S\mu q = \int_F \mu q, \ \forall q \in P_k(F),$$

for each face $F$ of $T$.

Finally, we define $\mathcal{R} : L^2(\partial T) \rightarrow P_1(T) + \Lambda(T)$ by

$$\mathcal{R}\mu := \Pi^{CR} \mu + S(\mu - \Pi^{CR} \mu),$$

where $\Pi^{CR} \mu \in P_1(T)$ satisfies

$$\int_F \Pi^{CR} \mu := \int_F \mu,$$

for each face $F$ of $T$.

By recalling $M(\partial T) := \{ \mu \in L^2(\partial T) : \mu|_F \in M(F), \ \text{for each face } F \text{ of } T \}$ and using standard scaling arguments, it is easy to derive the following lemma.

**Lemma A.1.** For any $\mu \in M(\partial T)$, it holds

$$\|\mu\|_{h,\partial T} \sim \|\mathcal{R}\mu\|_T, \quad (A.1)$$

$$|\mu|_{h,\partial T} \sim |\mathcal{R}\mu|_{1,T}. \quad (A.2)$$

For any $\mu_h \in M^0_\partial$, it is obvious that $\mathcal{R}\mu_h$ satisfies the 0-th order weak continuity, i.e., $\mathcal{R}\mu_h$ is continuous at the gravity point of each interior face of $T_h$. In addition, it holds $\mathcal{R}\mu_h|_{\partial T} = 0$. Therefore, from discrete Poincaré-Friedrichs inequalities ([16]) we have

$$\|\mathcal{R}\mu_h\| \lesssim \left( \sum_{T \in T_h} |\mathcal{R}\mu_h|_{1,T}^2 \right)^{1/2}. \quad (A.1)$$

Then it follows

$$\|\mu_h\|_h^2 = \sum_{T \in T_h} \|\mu_h\|_{h,\partial T}^2 \sim \sum_{T \in T_h} \|\mathcal{R}\mu_h\|_T^2 \quad \text{(by (A.1))}$$

$$\lesssim \sum_{T \in T_h} |\mathcal{R}\mu_h|_{1,T}^2 \quad \lesssim |\mu_h|_{h}^2, \quad \text{(by (A.2))}$$
Appendix: Proof of Lemma [3.2]

Denote by \( \hat{T} \) the referential unit simplex. For any simplex \( T \), there exists an invertible affine map \( F : \hat{T} \to T \) with \( F(\hat{x}) = A\hat{x} + b \) for \( \hat{x} \in \hat{T} \), \( A \in \mathbb{R}^{d \times d} \) a nonsingular matrix and \( b \in \mathbb{R}^d \). For any \( p \in [L^2(T)]^s (s = 1, 2, 3) \) and \( \mu \in L^2(\partial T) \), we understand \( \hat{p} \) and \( \hat{\mu} \) by

\[
\hat{p}(\hat{x}) = p(x), \quad (B.1)
\]

\[
\hat{\mu}(\hat{x}) = \mu(x), \quad (B.2)
\]

where \( x = F(\hat{x}) \) for \( \hat{x} \in \hat{T} \).

We state two well-known results as follows [21]:

\[
\|A\| \sim h_T, \quad (B.3)
\]

\[
\|A^{-1}\| \sim h_T^{-1}, \quad (B.4)
\]

where the matrix norm \( \|\cdot\| : \mathbb{R}^{d \times d} \to \mathbb{R} \) is defined by

\[
\|A\| = \max_{0 \neq x \in \mathbb{R}^d} \frac{\|Ax\|}{\|x\|}, \quad \forall A \in \mathbb{R}^{d \times d}. \quad (B.5)
\]

Based on the above two results, it’s straightforward to obtain

\[
\|A^{-T}x\| \sim h_T^{-1} \|x\|, \quad \forall x \in \mathbb{R}^d. \quad (B.6)
\]

Using the same techniques as in the proof the properties of the famous Piola transformation (\[4\]), we easily obtain the lemma below.

**Lemma B.1.** For any \((v, \mu) \in L^2(T) \times L^2(\partial T)\), it holds

\[
\hat{\nabla}^b_w \hat{\mu} = A^T \nabla^b_w \mu, \quad (B.7)
\]

\[
\hat{\nabla}^i_w \hat{v} = A^T \nabla^i_w v, \quad (B.8)
\]

\[
\hat{\nabla}_w(\hat{v}, \hat{\mu}) = A^T \nabla_w(v, \mu). \quad (B.9)
\]

**Lemma B.2.** For any simplex \( T \), there exist two positive constants \( c_T \) and \( C_T \), which only depend on \( T \) and \( k \), such that

\[
c_T \|\mu\|_{\partial T} \leq \|\nabla^b_w \mu\|_T \leq C_T \|\mu\|_{\partial T}, \quad \forall \mu \in M(\partial T). \quad (B.10)
\]

**Proof.** Assuming \( \nabla^b_w \mu = 0 \), by the definition of \( \nabla^b_w \), i.e. (2.8) we have

\[
\langle \mu, q \cdot n \rangle_{\partial T} = 0, \quad \forall q \in W(T),
\]
which implies \( \mu = 0 \). This means the semi-norm \( \| \nabla_w^b \|_T \) is a norm on \( M(\partial T) \). Since different norms on a finite dimensional space are equivalent, this lemma follows immediately.

\[ \text{Theorem B.1.} \quad \text{For any simplex } T, \text{ it holds} \]
\[ \| \nabla_w^b \mu \|_T \sim h_T^{-1} \| \mu \|_{h, \partial T}, \forall \mu \in M(\partial T). \quad (B.11) \]

\[ \text{Proof.} \quad \text{In view of } T = A\hat{T} + b, \text{ we have} \]
\[ \| \nabla_w^b \mu \|_T \sim \hat{h}_T^{\frac{k}{2}} \| \nabla_w^b \mu \|_{\hat{T}} \]
\[ \sim \hat{h}_T^{\frac{k}{2}} \| A^{-T} \nabla_w^b \hat{\mu} \|_{\hat{T}} \quad (\text{by Lemma B.1}) \]
\[ \sim \hat{h}_T^{\frac{k}{2}-1} \| \nabla_w^b \hat{\mu} \|_{\hat{T}} \quad (\text{by B.6}) \]
\[ \sim \hat{h}_T^{\frac{k}{2}-1} \| \hat{\mu} \|_{\partial \hat{T}} \quad (\text{by Lemma B.2}) \]
\[ \sim \hat{h}_T^{\frac{k}{2}-1} \| \mu \|_{h, \partial T} \]
\[ \sim h_T^{-1} \| \mu \|_{h, \partial T}. \]

\[ \square \]

Similarly, we can easily prove the following theorem.

\[ \text{Theorem B.2.} \quad \text{For any simplex } T, \text{ it holds} \]
\[ \| \nabla_w^i v \|_T \sim h_T^{-1} \| v \|_T, \forall v \in V(T). \quad (B.12) \]

\[ \text{Lemma B.3.} \quad \text{For any simplex } T, \text{ there exist two positive constants } c_T \text{ and } C_T \text{ that only depend on } T \text{ and } k, \text{ such that} \]
\[ c_T(\| v \|_T + \| \mu \|_{\partial T}) \leq \| \nabla_w(v, \mu) \|_T \leq C_T(\| v \| + \| \mu \|_{\partial T}), \forall (v, \mu) \in \Sigma(T), \quad (B.13) \]

where \( \Sigma(T) := \{(v, \mu) \in V(T) \times M(\partial T) : m_T(\mu) = 0\} \).

\[ \text{Proof.} \quad \text{It’s easy to know} \]
\[ (v, \mu) \mapsto \| v \|_T + \| \mu \|_{\partial T}, \forall (v, \mu) \in \Sigma(T) \]
defines a norm on \( \Sigma(T) \).

Next we show
\[ (v, \mu) \mapsto \| \nabla_w(v, \mu) \|_T, \forall (v, \mu) \in \Sigma(T) \]
also defines a norm on $\Sigma(T)$. In fact, if $\| \nabla_w (v, \mu) \|^T = 0$, then by the definition of $\nabla_w$, i.e. (2.6) we have

$$ (\nabla v, q)_T + (\mu - v, q \cdot n)_{\partial T} = 0, \forall q \in W(T). $$ \hspace{1cm} (B.14)

This relation, together with the properties of the BDM elements ([18]) and the RT elements [40], shows $v = \mu = \text{constant}$. Thus the relation $m_T(\mu) = 0$ leads to $(v, \mu) = 0$.

Finally, the desired conclusion follows from the equivalence of the above two norms. 

**Lemma B.4.** For any simplex $T$, it holds

$$ \| \nabla_w (v, \mu) \|^T \sim h_T^{-1} \| v \|^T + h_T^{-\frac{1}{2}} \| \mu \|_{\partial T}, \forall (v, \mu) \in \Sigma(T). $$ \hspace{1cm} (B.15)

**Proof.** In light of $T = A\hat{T} + b$ and $m_T(\mu) = m_{\hat{T}}(\hat{\mu})$ for all $\mu \in L^2(\partial T)$, we obtain

$$ \| \nabla_w (v, \mu) \|^T \sim h_T^\frac{1}{2} \| \hat{\nabla}_w (v, \mu) \|_{\hat{T}} $$

$$ \sim h_T^\frac{1}{2} \| A^{-T} \hat{\nabla}_w (\hat{v}, \hat{\mu}) \|_{\hat{T}} \quad \text{(by Lemma B.1)} $$

$$ \sim h_T^{\frac{1}{2}} - 1 \| \hat{\nabla}_w (\hat{v}, \hat{\mu}) \|_{\hat{T}} \quad \text{(by (2.6))} $$

$$ \sim h_T^{\frac{1}{2}} - 1 (\| \hat{v} \|_{\hat{T}} + \| \hat{\mu} \|_{\partial \hat{T}}) \quad \text{(by Lemma B.3)} $$

$$ \sim h_T^{-1} \| v \|^T + h_T^{-\frac{1}{2}} \| \mu \|_{\partial T}. $$

**Theorem B.3.** For any simplex $T$, it holds

$$ \| \nabla_w (v, \mu) \|^T \sim h_T^{-1} \| v - m_T(\mu) \|^T + |\mu|_{k, \partial T}, \forall (v, \mu) \in V(T) \times M(\partial T). $$ \hspace{1cm} (B.16)

**Proof.** By (2.6) we have

$$ (\nabla_w (v, \mu), q)_T = -(v, \text{div} q) + (\mu, q \cdot n)_{\partial T} $$

$$ = -(v - m_T(\mu), \text{div} q) + (\mu - m_T(\mu), q \cdot n)_{\partial T} $$

$$ = (\nabla_w (v - m_T(\mu), \mu - m_T(\mu)), q)_T, \forall q \in W(T), $$

which implies

$$ \nabla_w (v, \mu) = \nabla_w (v - m_T(\mu), \mu - m_T(\mu)). $$

Thus it follows

$$ \| \nabla_w (v, \mu) \|^T = \| \nabla_w (v - m_T(\mu), \mu - m_T(\mu)) \|^T $$

$$ \sim h_T^{-1} \| v - m_T(\mu) \|^T + h_T^{-\frac{1}{2}} \| \mu - m_T(\mu) \|_{\partial T} \quad \text{(by Lemma B.4)} $$

$$ \sim h_T^{-1} \| v - m_T(\mu) \|^T + |\mu|_{k, \partial T}. $$
A combination of Theorems B.1-B.3 proves Lemma 3.2.

References

[1] R. A. ADAMS, J. J. F. FOURNIER, Sobolev Spaces, Academic Press, 2nd ed., 2003.

[2] D. N. ARNOLD, F. BREZZI, Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates, RAIRO Modél. Math. Anal. Numér., 19 (1985), 7-32.

[3] D. N. ARNOLD, F. BREZZI, B. COCKBURN, L. D. MARINI, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM J. Numer. Anal., 39 (2002), 1749-1779.

[4] D. BOFFI, F. BREZZI, L. DEMKOWICZ, R. DURN, R. FALK, M. FORTIN, Mixed finite elements, compatibility conditions, and applications, Lecture Notes in Mathematics 939. Springer-Verlag, Berlin, Germany (2008), 12-14.

[5] R. E. BANK, T. DUPONT, An optimal order process for solving finite element equations, Math. Comp., 36 (1981), 35-51.

[6] R. E. BANK, C. C. DOUGLAS, Sharp estimates for multigrid rates of convergence with general smoothing and acceleration, SIAM J. Numer. Anal., 22 (1985), 617-633.

[7] D. BRAESS, W. HACKBUSCH, A new convergence proof for the multigrid method including the V-cycle, SIAM J. Numer. Anal., 20 (1983), 967-975.

[8] J. H. BRAMBLE, J. E. PASCIAK, New convergence estimates for multigrid algorithms, Math. Comp., 49 (1987), 311-329.

[9] J. H. BRAMBLE, J. E. PASCIAK, J. XU, The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms, Math. Comp., 56 (1991), 1-34.

[10] J. H. BRAMBLE, J. E. PASCIAK, J. WANG, J. XU, Convergence estimates for multigrid algorithms without regularity assumptions, Math. Comp., 57 (1991), 23-45.

[11] J. H. BRAMBLE, J. E. PASCIAK, New estimates for multilevel algorithms including the V-cycle, Math. Comp., 60 (1993), 447-471.

[12] A. BRANDT, Multi-level adaptive solutions to boundary-value problems, Math. Comp., 31 (1977), 333-390.

[13] S. C. BRENNER, An optimal-order multigrid method for P1 nonconforming finite elements, Math. Comp. 52 (1989), 1-16.
[14] S. C. BRENNER, A multigrid algorithm for the lowest-order Raviart-Thomas mixed triangular finite element method, SIAM J. Numer. Anal., 29 (1992), 647-678.

[15] S. C. BRENNER, Convergence of nonconforming multigrid methods without full elliptic regularity. Math. Comp., 68 (1999), 25-53.

[16] S. C. BRENNER, Poincaré-Friedrichs inequalities for piecewise $H^1$ functions, SIAM J. Numer. Anal., 41 (2003), 306-324.

[17] S. C. BRENNER, Convergence of nonconforming V-cycle and F-cycle multigrid algorithms for second order elliptic boundary value problems, Math. Comp., 73 (2004), 1041-1066 (electronic).

[18] F. BREZZI, J. DOUGLAS, JR., L. D. MARINI, Two families of mixed finite elements for second order elliptic problems, Numer. Math., 47 (1985), 217-235.

[19] L. CHEN, Deriving the X-Z identity from auxiliary space method, Domain Decomposition Methods in Science and Engineering XIX, Lecture Notes in Computational Science and Engineering, 78 (2011), 309-316.

[20] D. CHO, J. XU, L. ZIKATANOV, New estimates for the rate of convergence of the method of subspace corrections, Numer. Math. Theor. Meth. Appl., 1 (2008), 44-56.

[21] P. CIARLET, The finite element method for elliptic problems, North-Holland, Amsterdam, 1978.

[22] B. COCKBURN, J. GOPALAKRISHNAN, R. LAZAROV, Unified hybridization of discontinuous Galerkin, mixed, and conforming Galerkin methods for second order elliptic problems, SIAM J. Numer. Anal., 47 (2009), 1319-1365.

[23] B. COCKBURN, O. DUBOIS, J. GOPALAKRISHNAN, S. TAN, Multigrid for an HDG method, IMA Journal of Numerical Analysis, 34 (2014), 1386-1425.

[24] V. A. DOBREV, R. D. LAZAROV, P. S. VASSILEVSKI, L. T. ZIKATANOV, Two-level preconditioning of discontinuous Galerkin approximations of second-order elliptic equations. Numer. Linear Algebra Appl., 13 (2006), 753-770.

[25] H. Y. DUAN, S. Q. GAO, R. TAN, S. ZHANG, A generalized BPX multigrid framework covering nonnested V-cycle methods. Math. Comp., 76 (2007), 137-152.

[26] J. GOPALAKRISHNAN, A Schwarz preconditioner for a hybridized mixed method, Computational Methods in Applied Mathematics, 3 (2003), 116-134.
[27] J. GOPALAKRISHNAN, G. KANSCHAT, A multilevel discontinuous Galerkin method, Numer. Math., 95 (2003), 527-550.

[28] J. GOPALAKRISHNAN, S. TAN, A convergent multigrid cycle for the hybridized mixed method, Numer. Linear Algebra Appl., 16 (2009), 689-714.

[29] W. HACKBUSCH, Multi-grid methods and applications, Springer series in computational mathematics, vol. 4, Spring-Verlag, Berlin, New York, 1985.

[30] J. K. KRAUS, S. K. TOMAR, A multilevel method for discontinuous Galerkin approximation of three-dimensional anisotropic elliptic problems. Numer. Linear Algebra Appl., 15 (2008), 417-438.

[31] J. K. KRAUS, S. K. TOMAR, Multilevel preconditioning of two-dimensional elliptic problems discretized by a class of discontinuous Galerkin methods, SIAM J. Sci. Comput., 30 (2008), 684-706.

[32] J. WANG, X. YE, A weak Galerkin finite element method for second-order elliptic problems, J. Comp. and Appl. Math, 241 (2013), 103-115.

[33] J. WANG, X. YE, A weak Galerkin finite element method for the Stokes equations, arXiv:1302.2707v1 [math.NA].

[34] L. MU, J. WANG, Y. WANG, X. YE, A computational study of the weak Galerkin method for second-order elliptic equations, arXiv:1111.0618v1, 2011, Numerical Algorithms, 2012, DOI:10.1007/s11075-012-9651-1.

[35] L. MU, J. WANG, Y. WANG, X. YE, A weak Galerkin mixed finite element method for biharmonic equations, Springer Proceedings in Mathematics & Statistics, 45 (2013), 247-277.

[36] L. MU, J. WANG, G. WEI, X. YE, S. ZHAO, Weak Galerkin methods for second order elliptic interface problems, arXiv:1201.6438v2, 2012, Journal of Computational Physics, doi:10.1016/j.jcp.2013.04.042, to appear.

[37] L. MU, J. WANG, X. YE, A weak Galerkin finite element methods with polynomial reduction, arXiv:1304.6481, submitted to SIAM J on Scientific Computing.

[38] L. MU, J. WANG, X. YE, Weak Galerkin finite element methods on polytopal meshes, arXiv:1204.3655v2, submitted to International J of Numerical Analysis and Modeling.

[39] L. MU, J. WANG, X. YE, S. ZHAO, A numerical study on the weak Galerkin method for the Helmholtz equation with large wave numbers, arXiv:1111.0671v1, 2011.
[40] P. RAVIART, J. THOMAS, A mixed finite element method for second order elliptic problems, Mathematical Aspects of the Finite Element Method, I. Galligani, E. Magenes, eds., Lectures Notes in Math. 606, Springer-Verlag, New York, 1977.

[41] Y. WU, L. CHEN, X. XIE, J. XU, Convergence analysis of V-Cycle multigrid methods for anisotropic elliptic equations, IMA Journal of Numerical Analysis, 32 (2012), 1329-1347.

[42] J. XU, Iterative methods by space decomposition and subspace correction. SIAM Rev., 34 (1992), 581-613.

[43] J. XU, The auxiliary space method and optimal multigrid preconditioning techniques for unstructured grids. Computing, 56 (1996), 215-235.

[44] J. XU, An introduction to multigrid convergence theory. Iterative Methods in Scientific Computing (R. Chan, T. Chan & G. Golub eds). Springer, 1997.

[45] J. XU, L. ZIKATANOV, The method of alternating projections and the method of subspace corrections in Hilbert space, J. Am. Math. Soc., 15 (2002), 573-597.

[46] J. XU, L. CHEN, R. H. NOCHETTO, Optimal multilevel methods for H(grad), H(curl), and H(div) systems on graded and unstructured grids, Multiscale, Nonlinear and Adaptive Approximation, 2009, 599-659.