BPX preconditioner for nonstandard finite element methods for diffusion problems *

Binjie Li†, Xiaoping Xie ‡

School of Mathematics, Sichuan University, Chengdu 610064, China

Abstract

This paper proposes and analyzes an optimal preconditioner for a general linear symmetric positive definite (SPD) system by following the basic idea of the well-known BPX framework. The SPD system arises from a large number of nonstandard finite element methods for diffusion problems, including the well-known hybridized Raviart-Thomas (RT) and Brezzi-Douglas-Marini (BDM) mixed element methods, the hybridized discontinuous Galerkin (HDG) method, the Weak Galerkin (WG) method, and the nonconforming Crouzeix-Raviart (CR) element method. We prove that the presented preconditioner is optimal, in the sense that the condition number of the preconditioned system is independent of the mesh size. Numerical experiments are provided to confirm the theoretical results.

Keywords. BPX preconditioner, RT element, BDM element, HDG method, WG method, nonconforming CR element

1 Introduction

This paper is to design an efficient preconditioner for a large class of nonstandard finite element methods for solving the diffusion model

$$\begin{cases} -\operatorname{div}(\boldsymbol{A}\boldsymbol{\nabla} u) = f & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega, \end{cases} \tag{1.1}$$

where $\Omega \subset \mathbb{R}^d$ ($d = 2, 3$) is a bounded polyhedral domain, the diffusion tensor $\boldsymbol{A} : \Omega \to \mathbb{R}^{d\times d}$ is a matrix function that is assumed to be symmetric and uniformly positive definite, $f \in L^2(\Omega)$ and $g \in H^{\frac{1}{2}}(\partial\Omega)$.

*This work was supported by National Natural Science Foundation of China (11171239) and Major Research Plan of National Natural Science Foundation of China (91430105).

†Email: libinjiefem@yahoo.com

‡Corresponding author. Email: xpxie@scu.edu.cn

Let $\mathcal{T}_h$ be a triangulation of Ω, and $\mathcal{F}_h$ be the set of all faces of $\mathcal{T}_h$. We introduce a finite dimensional space

$$\mathbb{M}_{h,k} := \{\mu_h \in L^2(\cup_{F\in\mathcal{F}_h} F) : \mu_h|_F \in P_k(F) \text{ for all } F \in \mathcal{F}_h \text{ and } \mu_h|_{\partial\Omega} = 0\}, \tag{1.2}$$

with $P_k(F)$ denoting the set of polynomials of degree $\leqslant k$ on F. Consider the following general symmetric and positive definite (SPD) system for equation (1.1): Seek $\lambda_h \in \mathbb{M}_{h,k}$ such that

$$d_h(\lambda_h, \mu_h) = b(\mu_h) \text{ for all } \mu_h \in \mathbb{M}_{h,k}. \tag{1.3}$$

Here $d_h(\cdot,\cdot) : \mathbb{M}_{h,k} \times \mathbb{M}_{h,k} \to \mathbb{R}$ is an inner-product on $\mathbb{M}_{h,k}$ and $b_h(\cdot) : \mathbb{M}_{h,k} \to \mathbb{R}$ is a linear functional on $\mathbb{M}_{h,k}$.

The first class of nonstandard finite element methods that fall into the framework (1.3) are hybrid or hybridized finite element methods ([5, 35, 38, 39, 3, 14, 17, 18, 19, 4, 21, 20, 31]). Due to the relaxation of the constraint of continuity at the inter-element boundaries by introducing some Lagrange multipliers, the corresponding hybrid method allows for piecewise-independent approximation to the potential or flux solution. Thus, after local elimination of unknowns defined in the interior of elements, the method leads to a SPD discrete system of the form (1.3), where the unknowns are only the globally coupled degrees of freedom describing the Lagrange multiplier. In [3, 14], the Raviart-Thomas (RT) [37] and Brezzi-Douglas-Marini (BDM) mixed methods were shown to have equivalent hybridized versions. A new characterization of the approximate solution of hybridized mixed methods was developed and applied in [17] to obtain an explicit formula for the entries of the matrix equation for the Lagrange multiplier unknowns. An overview of some new hybridization techniques was presented in [18]. In [21] a unifying framework for hybridization of finite element methods was developed. Error estimates of some hybridized discontinuous Galerkin (HDG) methods were derived in [19, 20, 31].

The weak Galerkin (WG) method [42, 34, 33] is the second class of nonstandard approach that applies to the framework (1.3). The WG method is designed by using a weakly defined gradient operator over functions with discontinuity, and allows the use of totally discontinuous functions in the finite element procedure. The concept of weak gradients provides a systematic framework for dealing with discontinuous functions defined on elements and their boundaries in a near classical sense [42]. Similar to the hybrid methods, the WG scheme can be reduced to the form (1.3) after local elimination of unknowns defined in the interior of elements. We note that when $\boldsymbol{A}$ in (1.1) is a piecewise-constant matrix, the WG method is, by introducing the discrete weak gradient as an independent variable, equivalent to the hybridized version of the RT or BDM mixed methods. For the discretization of the diffusion model (1.1) on simplicial 2D or 3D meshes, we refer to [30] for a multigrid WG algorithm, and to [16] for an auxiliary space multigrid preconditioner for the WG method as well as a reduced system of the weak Galerkin method involving only the degrees of freedom on edges/faces.

Besides, some noncnforming methods, e.g. the nonconforming Crouzeix-Raviart element method [23], can also lead to a SPD discrete system of the form (1.3). To this end, one needs to introduce a special projection of the flux solution to the element boundaries as the trace approximation. We refer to [12, 6, 13, 1, 29, 36, 28, 40, 48] for multigrid algorithms or preconditioning for the CR or CR-related nonconforming finite element methods. In particular, in [13], an optimal-order multigrid method was proposed and analyzed for the lowest-order Raviart-Thomas mixed element based on the equivalence between Raviart-Thomas mixed methods and certain nonconforming methods.

As far as we know, the first preconditioner for the system (1.3) was developed in [25], where a Schwarz preconditioner was designed for the hybridized RT and BDM mixed element methods. In [26] a convergent V-cycle multigrid method was proposed for the hybridized mixed methods for Poisson problems with full elliptic regularity. By following the idea of [26], a non-nested multigrid V-cycle algorithm, with a single smoothing step per level, was analyzed in [22] for the system (1.3) arising from one type of HDG method, where only a weak elliptic regularity is required. In [32], a general framework for designing fast solvers for the system (1.3) was presented without any regularity assumption.

It is well known that the BPX multigrid framework, developed by Bramble, Pasciak and Xu [10], is widely used in the analysis of multigrid and domain decomposition methods. We refer to [7, 8, 9, 11, 24, 27, 41, 46, 44, 45, 47] for the development and applications of the BPX framework. In [43] an abstract framework of auxiliary space method was proposed and an optimal multigrid technique was developed for general unstructured grids. Especially, in [44] an overview of multilevel methods, such as V-cycle multigrid and BPX preconditioner, was given for solving various partial differential equations on quasi-uniform meshes, and the methods were extended to graded meshes and completely unstructured grids.

In this paper, we shall follow the basic ideas of ([10], [43], [44]) to construct a BPX preconditioner for the system (1.3), which is, due to the definition of the discrete space $\mathbb{M}_{h,k}$, corresponding to nonnested multilevel finite element spaces. We will show the proposed preconditioner is optimal.

We arrange the rest of the paper as follows. Section 2 introduces some notation and preliminaries. Section 3 introduces and analyzes a general auxiliary space preconditioner. Section 4 constructs the BPX preconditioner and derives the condition number estimation of the preconditioned system. Section 5 shows some applications of the proposed preconditioner. Finally, Section 6 provides some numerical results.

2 Notations and preliminaries

Throughout this paper, we use the standard definitions of Sobolev spaces and their norms and semi-norms (cf. [2]), namely for an arbitrary open set $D \subset \mathbb{R}^d$ and any nonnegative integer s,

$$\begin{aligned} H^s(D) &:= \{v \in L^2(D) : \partial^\alpha v \in L^2(D), \forall |\alpha| \leqslant s\}, \\ \|v\|_{s,D} &:= (\textstyle\sum_{|\alpha|\leqslant s} \int_D |\partial^\alpha v|^2)^{\frac{1}{2}}, \quad |v|_{s,D} := (\textstyle\sum_{|\alpha|=s} \int_D |\partial^\alpha v|^2)^{\frac{1}{2}}. \end{aligned}$$

We denote respectively by $(\cdot,\cdot)_D$ and $\langle\cdot,\cdot\rangle_{\partial D}$ the L^2 inner products on $L^2(D)$ and $L^2(\partial D)$, and respectively by $\|\cdot\|_D$ and $\|\cdot\|_{\partial D}$ the L^2-norms on $L^2(D)$ and $L^2(\partial D)$. In particular, $(\cdot,\cdot)$ and $\|\cdot\|$ abbreviate $(\cdot,\cdot)_\Omega$ and $\|\cdot\|_\Omega$, respectively.

Let $\mathcal{T}_h$ be a conforming shape-regular triangulation of the polyhedral domain Ω. For any $T \in \mathcal{T}_h$, h_T denotes the diameter of T, and we set $h := \max_{T\in\mathcal{T}_h} h_T$. We define the mesh-dependent inner product $\langle\cdot,\cdot\rangle_h : \mathbb{M}_{h,k}\times\mathbb{M}_{h,k} \to \mathbb{R}$ and the norm $\|\cdot\|_h : \mathbb{M}_{h,k} \to \mathbb{R}$ as follows: for any $\lambda_h, \mu_h \in \mathbb{M}_{h,k}$,

$$\langle\lambda_h,\mu_h\rangle_h := \sum_{T\in\mathcal{T}_h} h_T \int_{\partial T} \lambda_h\mu_h, \quad \|\mu_h\|_h := \langle\mu_h,\mu_h\rangle_h^{1/2}. \tag{2.1}$$

We also need the following notation: for any $\mu \in L^2(\partial T)$,

$$\begin{aligned} \|\mu\|_{h,\partial T} &:= h_T^{\frac{1}{2}} \|\mu\|_{\partial T}, \\ |\mu|_{h,\partial T} &:= h_T^{-\frac{1}{2}} \|\mu - m_T(\mu)\|_{\partial T} \quad \text{with} \quad m_T(\mu) := \tfrac{1}{d+1} \sum_{F\in\mathcal{F}_T} \tfrac{1}{|F|}\int_F \mu, \\ |\mu|_h &:= (\textstyle\sum_{T\in\mathcal{T}_h} |\mu|_{h,\partial T}^2)^{\frac{1}{2}}, \end{aligned}$$

where $\mathcal{F}_T := \{F : F \subset \partial T \text{ is a face of } T\}$ and $|F|$ denotes the $(d-1)$-dimensional Hausdorff measure of F.

In the context, we use $x \lesssim y$ to denote $x \leqslant cy$, where c is a positive constant independent of h which may be different at its each occurrence. The notation $x \sim y$ abbreviates $x \lesssim y \lesssim x$. For the bilinear form $d_h(\cdot,\cdot)$ in the system (1.3), we shall make the following abstract assumption.

Assumption 2.1. *For any $\mu_h \in \mathbb{M}_{h,k}$, it holds*

$$d_h(\mu_h,\mu_h) \sim |\mu_h|_h^2. \tag{2.2}$$

Remark 2.1. *This assumption is valid for many nonstandard finite element methods, as will be shown in Section 5. We note that the Schwarz preconditioner constructed in [25] can also be extended to the system* (1.3) *under* **Assumption 2.1**.

Based on **Assumption 2.1**, we are ready to present an estimate that describes the conditioning of the system (1.3).

Theorem 2.1. *Suppose $\mathcal{T}_h$ to be quasi-uniform. Under* **Assumption 2.1**, *it holds*

$$\|\mu_h\|_h^2 \lesssim d_h(\mu_h,\mu_h) \lesssim h^{-2}\|\mu_h\|_h^2, \ \forall\mu_h \in \mathbb{M}_{h,k}. \tag{2.3}$$

Proof. By Lemma 3.1 of [30], we have

$$\|\mu_h\|_h^2 \lesssim |\mu_h|_h^2. \tag{2.4}$$

Then the desired conclusion follows from **Assumption 2.1** and the fact $|\mu_h|_h^2 \lesssim h^{-2}\,\|\mu_h\|_h^2$. $\square$

Remark 2.2. *In Theorem 2.3 of [25], a similar result was derived in the two-dimensional case. But the proof there could not be extended to three-dimensional case directly.*

We define the operator $D_h : \mathbb{M}_{h,k} \to \mathbb{M}_{h,k}$ by

$$\langle D_h\lambda_h, \mu_h\rangle_h := d_h(\lambda_h, \mu_h) \quad \text{ for all } \lambda_h, \mu_h \in \mathbb{M}_{h,k}. \tag{2.5}$$

Obviously, D_h is an SPD operator and, from Theorem 2.1, it follows the condition number estimate

$$\kappa(D_h) \lesssim h^{-2}, \tag{2.6}$$

where $\kappa(D_h) := \frac{\lambda_{\max}(D_h)}{\lambda_{\min}(D_h)}$ and $\lambda_{\max}(D_h), \lambda_{\min}(D_h)$ denote the maximum and minimum eigenvalues of D_h, respectively. In fact, with a slight modification of the proof of Theorem 2.1 of [30], we can show that $\kappa(D_h) \sim h^{-2}$ holds under the condition that h is sufficiently small.

3 Auxiliary space preconditioning

In this section, we shall follow the basic idea of [43] to introduce a general auxiliary space preconditioner for D_h. It should be stressed that we only require the triangulation $\mathcal{T}_h$ to be conforming and shape regular.

Let V be a finite dimensional Hilbert space endowed with inner product $(\cdot,\cdot)$ and its induced norm $\|\cdot\|$. Let $S : V \to V$ be SPD with respect to $(\cdot,\cdot)$. We use $(\cdot,\cdot)_S$ to denote the inner product $(S\cdot,\cdot)$, and use $\|\cdot\|_S$ to denote the norm induced by $(\cdot,\cdot)_S$.

We choose the H^1-conforming piecewise linear element space as the so-called auxiliary space V_h^c, namely

$$V_h^c := \{v_h \in H_0^1(\Omega) : v_h|_T \in P_1(T) \ \text{ for all } T \in \mathcal{T}_h\}. \tag{3.1}$$

Then we introduce two different prolongation operators that map V_h^c into $\mathbb{M}_{h,k}$ as follows:

- $\Pi_h^1 : V_h^c \to \mathbb{M}_{h,k}$ is defined by

$$\Pi_h^1 v_h|_F := \frac{1}{|F|}\int_F v_h \quad \text{ for all } F \in \mathcal{F}_h \text{ and } v_h \in V_h^c. \tag{3.2}$$

- $\Pi_h^2 : V_h^c \to \mathbb{M}_{h,k}$ is defined by

$$\int_F \Pi_h^2 v_h q := \int_F v_h q \quad \text{ for all } F \in \mathcal{F}_h, v_h \in V_h^c \text{ and } q \in P_k(F). \tag{3.3}$$

Obviously, Π_h^1 coincides with Π_h^2 in the case that $k=0$. For the sake of convenience, in the rest of this paper, unless otherwise specified, we shall use Π_h to denote both Π_h^1 and Π_h^2 at the same time. Define the adjoint operator, $\Pi_h^t : \mathbb{M}_{h,k} \to V_h^c$, of Π_h, by

$$(\Pi_h^t \mu_h, v_h) := \langle \mu_h, \Pi_h v_h \rangle_h \quad \text{for all } \mu_h \in \mathbb{M}_{h,k} \text{ and } v_h \in V_h^c. \tag{3.4}$$

Before defining the auxiliary space preconditioner, we introduce two linear operators, S_h and $\widetilde{B_h}$, in the following two assumptions.

Assumption 3.1. *Let $S_h : M_{h,k} \to M_{h,k}$ be SPD with respect to $\langle \cdot, \cdot \rangle_h$ and satisfy the following estimates: for all $\mu_h \in M_{h,k}$,*

$$\langle S_h \mu_h, \mu_h \rangle_h \lesssim \langle D_h^{-1} \mu_h, \mu_h \rangle_h, \tag{3.5}$$

$$\|\mu_h\|_{S_h^{-1}}^2 \lesssim \textstyle\sum_{T \in \mathcal{T}_h} h_T^{-2} \|\mu_h\|_{h,\partial T}^2 . \tag{3.6}$$

Assumption 3.2. *Let $\widetilde{B_h} : V_h^c \to V_h^c$ be SPD with respect to $(\cdot, \cdot)$ and satisfy the estimate*

$$(\widetilde{B_h}^{-1} v_h, v_h) \sim |v_h|_{1,\Omega}^2 \quad \textit{for all } v_h \in V_h^c.$$

Then we define the general auxiliary space preconditioner $B_h^G : M_{h,k} \to M_{h,k}$ by

$$B_h^G := S_h + \Pi_h \widetilde{B_h} \Pi_h^t. \tag{3.7}$$

Remark 3.1. *We note that the Jacobi iteration and the symmetric Gauss-Seidel iteration satisfy* **Assumption 3.1** *if $\mathcal{T}_h$ is conforming and shape regular, while the Richardson iteration does if $\mathcal{T}_h$ is quasi-uniform.*

Remark 3.2. *The preconditioner B_h^G was also analyzed recently in [16] for two types of WG methods, where $\mathcal{T}_h$ is assumed to be quasi-uniform. In our analysis below we only require $\mathcal{T}_h$ to be conforming and shape regular. We refer to Theorem 2.1 of [43] for a more general result for auxiliary space preconditioning under quasi-uniform meshes.*

For the auxiliary space preconditioner B_h^G, we have the following main result.

Theorem 3.1. *Under* **Assumptions 2.1, 3.1** *and* **3.2**, *it holds*

$$\kappa(B_h^G D_h) \lesssim 1, \tag{3.8}$$

where $\kappa(B_h^G D_h) := \frac{\lambda_{max}(B_h^G D_h)}{\lambda_{min}(B_h^G D_h)}$, and $\lambda_{max}(B_h^G D_h)$ and $\lambda_{min}(B_h^G D_h)$ denote the maximum and minimum eigenvalues of $B_h^G D_h$, respectively.

Remark 3.3. *Since we only require $\mathcal{T}_h$ to be conforming and shape regular, Theorem 3.1 is not a trivial application of Theorem 2.1 in [43].*

Before proving Theorem 3.1, we introduce a key ingredient operator $P_h : M_{h,k} \to V_h^c$ as follows: For each node $\boldsymbol{a}$ of $\mathcal{T}_h$,

$$P_h\mu_h(\boldsymbol{a}) := \begin{cases} \frac{\sum_{T\in\omega_{\boldsymbol{a}}} m_T(\mu_h)}{\sum_{T\in\omega_{\boldsymbol{a}}} 1} & \text{if } \boldsymbol{a} \text{ is an interior node,} \\ 0 & \text{if } \boldsymbol{a} \in \partial\Omega, \end{cases} \tag{3.9}$$

where $\omega_{\boldsymbol{a}}$ denotes the set of simplexes that share the vertex $\boldsymbol{a}$.

Lemma 3.1. *For any $\mu_h \in \mathbb{M}_{h,k}$, it holds*

$$|P_h\mu_h|_{1,\Omega} \lesssim |\mu_h|_h, \tag{3.10}$$

$$\sum_{T\in\mathcal{T}_h} h_T^{-2} \|(I - \Pi_h P_h)\lambda_h\|_{h,\partial T}^2 \lesssim |\mu_h|_h^2. \tag{3.11}$$

Proof. For any $T \in \mathcal{T}_h$, we use $\mathcal{N}(T)$, ω_T to denote the set of all vertexes of T and the set $\{T' \in \mathcal{T}_h : T' \in \omega_{\boldsymbol{a}} \text{ for some } \boldsymbol{a} \in \mathcal{N}(T)\}$, respectively. For $\boldsymbol{a} \in \mathcal{N}(T)$, if $\boldsymbol{a} \in \Omega$, then we have

$$\begin{aligned}
&h_T^{d-2}|m_T(\mu_h) - (P_h\mu_h)(\boldsymbol{a})|^2 \\
\lesssim& h_T^{d-2} \sum_{\substack{T_1,T_2\in\omega_{\boldsymbol{a}} \\ T_1,T_2 \text{ share a same face}}} |m_{T_1}(\mu_h) - m_{T_2}(\mu_h)|^2 \\
\lesssim& h_T^{-1} \sum_{\substack{T_1,T_2\in\omega_{\boldsymbol{a}} \\ T_1,T_2 \text{ share a same face}}} \|m_{T_1}(\mu_h) - m_{T_2}(\mu_h)\|_{\partial T_1\cap\partial T_2}^2 \\
\lesssim& h_T^{-1} \sum_{\substack{T_1,T_2\in\omega_{\boldsymbol{a}} \\ T_1,T_2 \text{ share a same face}}} \left(\|\mu_h - m_{T_1}(\mu_h)\|_{\partial T_1\cap\partial T_2}^2 + \|\mu_h - m_{T_2}(\mu_h)\|_{\partial T_1\cap\partial T_2}^2 \right) \\
\lesssim& h_T^{-1} \sum_{T'\in\omega_{\boldsymbol{a}}} \|\mu_h - m_{T'}(\mu_h)\|_{\partial T'}^2 \\
\lesssim& \sum_{T'\in\omega_{\boldsymbol{a}}} |\mu_h|_{h,\partial T'}^2.
\end{aligned}$$

If $\boldsymbol{a} \in \partial\Omega$, suppose that $F \subset \partial\Omega$ is a face of T such that $\boldsymbol{a} \in \partial F$. Since $\mu_h|_{\partial\Omega} = 0$, we have

$$\begin{aligned}
h_T^{d-2}|m_T(\mu_h) - (P_h\mu_h)(\boldsymbol{a})|^2 = h_T^{d-2}|m_T(\mu_h)|^2 &\sim h_T^{-1} \|m_T(\mu_h)\|_F^2 \\
&\lesssim h_T^{-1} \|\mu_h - m_T(\mu_h)\|_F^2 \\
&\lesssim |\mu_h|_{h,\partial T}^2.
\end{aligned}$$

In light of the above two estimates, we immediately get

$$h_T^{d-2} \sum_{\boldsymbol{a}\in\mathcal{N}(T)} |m_T(\mu_h) - (P_h\mu_h)(\boldsymbol{a})|^2 \lesssim \sum_{T'\in\omega_T} |\mu_h|_{h,\partial T'}^2. \tag{3.12}$$

Since $m_T(\mu_h)$ is a constant on T, it follows

$$\begin{aligned}
|P_h\mu_h|_{1,T}^2 &= |m_T(\mu_h) - P_h\mu_h|_{1,T}^2 \\
&\lesssim h_T^{-2} \|m_T(\mu_h) - P_h\mu_h\|_T^2 && \text{(by inverse estimate)} \\
&\lesssim h_T^{d-2} \textstyle\sum_{\boldsymbol{a}\in\mathcal{N}(T)} |m_T(\mu_h) - (P_h\mu_h)(\boldsymbol{a})|^2 \\
&\lesssim \textstyle\sum_{T'\in\omega_T} |\mu_h|_{h,\partial T'}^2, && \text{(by (3.12))}
\end{aligned}$$

which implies

$$|P_h\mu_h|_{1,\Omega}^2 = \sum_{T\in\mathcal{T}_h} |P_h\mu_h|_{1,T}^2 \lesssim |\mu_h|_h^2,$$

i.e., the estimate (3.10) holds.

We recall that $\mathcal{F}_T$ is the set of all faces of T. For any $F \in \mathcal{F}_T$, we use $\mathcal{N}(F)$ to denote the set of all vertexes of F. Since

$$\begin{aligned}
\|m_T(\mu_h) - \Pi_h P_h\mu_h\|_{\partial T}^2 &= \sum_{F\in\mathcal{F}_T} \|m_T(\mu_h) - \Pi_h P_h\mu_h\|_F^2 \\
&\lesssim h_T^{d-1} \sum_{F\in\mathcal{F}_T} \sum_{\boldsymbol{a}\in\mathcal{N}(F)} |m_T(\mu_h) - P_h\mu_h(\boldsymbol{a})|^2 \quad (\text{by } (3.2) \text{ and } (3.3)) \\
&\lesssim h_T^{d-1} \sum_{\boldsymbol{a}\in\mathcal{N}(T)} |m_T(\mu_h) - P_h\mu_h(\boldsymbol{a})|^2 \\
&\lesssim h_T \sum_{T'\in\omega_T} |\mu_h|_{h,T'}^2, \quad (\text{by } (3.12))
\end{aligned}$$

we get

$$\begin{aligned}
\|\mu_h - \Pi_h P_h\mu_h\|_{\partial T}^2 &\lesssim h_T|\mu_h|_{h,\partial T}^2 + \|m_T(\mu_h) - \Pi_h P_h\mu_h\|_{\partial T}^2 \\
&\lesssim \sum_{T'\in\omega_T} h_{T'}|\mu_h|_{h,\partial T'}^2.
\end{aligned}$$

Therefore,

$$\|(I - \Pi_h P_h)\mu_h\|_h^2 = \sum_{T\in\mathcal{T}_h} h_T \|(I - \Pi_h P_h)\mu_h\|_{\partial T}^2 \lesssim h^2|\mu_h|_h^2,$$

i.e. (3.11) holds. □

Now, we are ready to prove Theorem 3.1.

Proof of Theorem 3.1. For any $T \in \mathcal{T}_h$, standard scaling arguments yield

$$|\Pi_h v|_{h,\partial T} \sim |v|_{1,T} \quad \text{for all } v \in P_1(T). \tag{3.13}$$

Define $\widetilde{D_h} := \Pi_h^t D_h \Pi_h$. Then, for any $v_h \in V_h^c$, we have

$$\begin{aligned}
(\widetilde{D_h}v_h, v_h) &= \langle \Pi_h v_h, \Pi_h v_h\rangle_{D_h} \\
&\sim \textstyle\sum_{T\in\mathcal{T}_h} |\Pi_h v_h|_{h,\partial T}^2 && (\text{by } \textbf{Assumption 2.1}) \\
&\sim \textstyle\sum_{T\in\mathcal{T}_h} |v_h|_{1,T}^2 && (\text{by } (3.13)) \\
&\sim (\widetilde{B_h}^{-1} v_h, v_h), && (\text{by } \textbf{Assumption 3.2})
\end{aligned}$$

i.e.,

$$(\widetilde{D_h}v_h, v_h) \sim (\widetilde{B_h}^{-1} v_h, v_h) \quad \text{for all } v_h \in V_h^c. \tag{3.14}$$

By the definition of B_h^G, it holds, for any $\mu_h \in M_{h,k}$,

$$\begin{aligned}
\langle B_h^G D_h\mu_h, \mu_h\rangle_{D_h} &= \langle S_h D_h\mu_h, \mu_h\rangle_{D_h} + (\widetilde{B_h}\Pi_h^t D_h\mu_h, \Pi_h^t D_h\mu_h) \\
&\lesssim \|\mu_h\|_{D_h}^2 + (\widetilde{B_h}\Pi_h^t D_h\mu_h, \Pi_h^t D_h\mu_h) \qquad (\text{by } \textbf{Assumption 3.1}) \\
&\lesssim \|\mu_h\|_{D_h}^2 + (\widetilde{D_h}^{-1}\Pi_h^t D_h\mu_h, \Pi_h^t D_h\mu_h) \quad (\text{by } (3.14)) \\
&\lesssim \|\mu_h\|_{D_h}^2 + \left\|\widetilde{D_h}^{-1}\Pi_h^t D_h\mu_h\right\|_{\widetilde{D_h}}^2,
\end{aligned}$$

which, together with

$$\begin{aligned}\left\|\widetilde{D_h}^{-1}\Pi_h^t D_h\mu_h\right\|_{\widetilde{D_h}} &= \sup_{v_h\in V_h^c}\frac{(\widetilde{D_h}^{-1}\Pi_h^t D_h\mu_h, v_h)_{\widetilde{D_h}}}{\|v_h\|_{\widetilde{D_h}}}\\ &= \sup_{v_h\in V_h^c}\frac{(\mu_h, \Pi_h v_h)_{D_h}}{\|v_h\|_{\widetilde{D_h}}}\\ &\leqslant \sup_{v_h\in V_h^c}\frac{\|\mu_h\|_{D_h}\|\Pi_h v_h\|_{D_h}}{\|v_h\|_{\widetilde{D_h}}}\\ &= \|\mu_h\|_{D_h},\end{aligned}$$

yields

$$\langle B_h^G D_h\mu_h, \mu_h\rangle_{D_h} \lesssim \|\mu_h\|_{D_h}^2 \quad \text{for all } \mu_h \in M_{h,k}. \tag{3.15}$$

Thus it follows

$$\lambda_{max}(B_h^G D_h) \lesssim 1. \tag{3.16}$$

On the other hand, by Theorem 1 of [15], we have, for any $\lambda_h \in M_{h,k}$,

$$\begin{aligned}&\langle (B_h^G)^{-1}\lambda_h, \lambda_h\rangle_h\\ &= \inf_{\mu_h+\Pi_h v_h=\lambda_h}\langle S_h^{-1}\mu_h, \mu_h\rangle_h + (\widetilde{B_h}^{-1}v_h, v_h)\\ &\leqslant \|(I-\Pi_h P_h)\lambda_h\|_{S_h^{-1}}^2 + \|P_h\lambda_h\|_{\widetilde{B_h}^{-1}}^2\\ &\lesssim \sum_{T\in T_h} h_T^{-2}\|(I-\Pi_h P_h)\lambda_h\|_{h,\partial T}^2 + |P_h\lambda_h|_{1,\Omega}^2 \quad (\text{by } \textbf{Assumptions 3.1-3.2})\\ &\lesssim \|\lambda_h\|_{D_h}^2, \qquad (\text{by Lemma 3.1})\end{aligned}$$

which implies

$$\lambda_{min}(B_h^G D_h) \gtrsim 1. \tag{3.17}$$

As a result, the desired estimate (3.8) follows immediately from (3.16) and (3.17). This finishes the proof.

4 BPX preconditioner

4.1 Preconditioner construction

Suppose we are given a coarse quasi-uniform triangulation $\mathcal{T}_0$. Then we obtain a nested sequence of triangulations $\{\mathcal{T}_j : 0 \leqslant j \leqslant J\}$ through a successive refinement process, i.e., $\mathcal{T}_j$ is the uniform refinement of $\mathcal{T}_{j-1}$ for $j = 1, 2, \ldots, J$. We use h_j to denote the mesh size of $\mathcal{T}_j$, i.e., the maximum diameter of the simplexes in $\mathcal{T}_j$. For each triangulation $\mathcal{T}_j$, we define V_j^c by

$$V_j^c := \{v \in H_0^1(\Omega) : v|_T \in P_1(T) \text{ for all } T \in \mathcal{T}_j\}, \tag{4.1}$$

and let $\{\phi_{j,i} : i = 1, 2, \cdots, N_j\}$ be the standard nodal basis of V_j^c, where N_j is the dimension of V_j^c. We set $\{\eta_i : i = 1, 2, \ldots, M\}$ to be the standard nodal basis of $\mathbb{M}_{h,k}$. Set $h = h_J$, $\mathcal{T}_h = \mathcal{T}_J$ and $V_h^c = V_J^c$.

With the operators Π_h (defined by (3.2) or (3.3)), Π_h^t (defined by (3.4)), the nodal basis, $\{\phi_{j,i} : i = 0, 1, \ldots, N_j\}$, of V_j^c, and the nodal basis, $\{\eta_i : i = 1, 2, \ldots, M\}$, of $\mathbb{M}_{h,k}$, we define the BPX preconditioner (in operator form) for the operator D_h given in (2.5) as follows:

$$B_h\mu_h = h^{2-d}\sum_{i=1}^{M}\langle \mu_h, \eta_i\rangle_h \eta_i + \sum_{(j,i)\in\Lambda} h_j^{2-d}(\Pi_h^t\mu_h, \phi_{j,i})\Pi_h\phi_{j,i} \quad \text{for all } \mu_h \in \mathbb{M}_{h,k}, \tag{4.2}$$

where $\Lambda := \{(j, i) : 0 \leqslant j \leqslant J, 1 \leqslant i \leqslant N_j\}$. It is trivial to verify that B_h is SPD with respect to $\langle\cdot,\cdot\rangle_h$.

Remark 4.1. *We shall prove in the next subsection that both Π_h^1 and Π_h^2 lead to optimal preconditioners in the case $k \geqslant 1$, although numerical results in Section 6 show that Π_h^2 is much more efficient. We note that Π_h^2 was also used in [26], [22] and [32] to construct multilevel methods for HDG methods.*

4.2 Conditioning of B_hD_h

In this subsection, we shall use the framework of auxiliary space preconditioning introduced in Section 3 to analyze the BPX conditioning of B_hD_h. For the sake of convenience, in this subsection we assume

$$\mu_i \in \text{span}\{\eta_i\} \quad \text{for all } i = 1, \cdots, M.$$

We define two SPD operators $S_h : M_{h,k} \to M_{h,k}$ and $\widetilde{B}_h : V_h^c \to V_h^c$, respectively as follows:

$$S_h\mu_h := h^{2-d}\sum_{i=1}^{M}\langle \mu_h, \eta_i\rangle_h \eta_i \quad \text{for all } \mu_h \in M_{h,k}, \tag{4.3}$$

$$\widetilde{B}_h v_h = \sum_{(j,i)\in\Lambda} h_j^{2-d}(v_h, \phi_{j,i})\phi_{j,i} \quad \text{for all } v_h \in V_h^c. \tag{4.4}$$

Apparently we have

$$B_h = S_h + \Pi_h \widetilde{B}_h \Pi_h^t. \tag{4.5}$$

Thus, according to Theorem 3.1, to show $\kappa(B_hD_h) \lesssim 1$ it suffices to prove that S_h and $\widetilde{B}_h$ satisfy **Assumption 3.1** and **Assumption 3.2**, respectively.

Lemma 4.1. *The operator S_h defined by (4.3) satisfies* **Assumption 3.1**.

Proof. For any $\mu_h \in M_{h,k}$, by using the same technique as in the proof of Lemma 2.4 in [46], it is easy to verify

$$\langle S_h^{-1}\mu_h, \mu_h\rangle_h = \inf_{\sum_i \mu_i = \mu_h} \sum_i \frac{h^{d-2}}{\|\eta_i\|_h^2}\|\mu_i\|_h^2 . \tag{4.6}$$

By a standard scaling argument, it holds $\|\eta_i\|_h \sim h^{\frac{d}{2}}$. Then from (4.6) it follows

$$\langle S_h^{-1}\mu_h, \mu_h\rangle_h \sim h^{-2} \inf_{\sum_i \mu_i = \mu_h} \sum_i \|\mu_i\|_h^2 . \tag{4.7}$$

Further more, since $\mathcal{T}_h$ is shape regular, a standard scaling argument also yields $\sum_i \|\mu_i\|_h^2 \sim \|\sum_i \mu_i\|_h^2$. Thus it holds

$$\langle S_h^{-1}\mu_h, \mu_h\rangle_h \sim h^{-2} \|\mu_h\|_h^2 , \tag{4.8}$$

and the desired estimate (3.6) follows from (4.8) immediately by the quasi-uniform assumption of $\mathcal{T}_h$.

The thing left is to prove (3.5). In fact, from

$$\begin{aligned}\langle D_h\mu_h, \mu_h\rangle_h &\sim |\mu_h|_h^2 \qquad (\text{by } \textbf{Assumption 2.1}) \\ &\lesssim h^{-2} \|\mu_h\|_h^2 \\ &\lesssim \langle S_h^{-1}\mu_h, \mu_h\rangle_h, \qquad (\text{by } (4.8))\end{aligned}$$

which implies immediately $\langle S_h\mu_h, \mu_h\rangle_h \lesssim \langle D_h^{-1}\mu_h, \mu_h\rangle_h$. Hence, (3.5) follows immediately. □

The following lemma follows from Theorems 3.1-3.2 of [44].

Lemma 4.2. *The operator* $\widetilde{B_h}$ *defined by* (4.4) *satisfies* **Assumption 3.2**.

Finally, thanks to Theorem 3.1, we obtain immediately the following main conclusion for the BPX preconditioning.

Theorem 4.1. *Under* **Assumption 2.1**, *it holds*

$$\kappa(B_h D_h) \lesssim 1, \tag{4.9}$$

where D_h *and* B_h *are defined by* (2.5) *and* (4.2), *respectively.*

4.3 Implementation

We recall that $\{\eta_i : 1 \leqslant i \leqslant M\}$ is the standard nodal basis of $\mathbb{M}_{h,k}$ and $\{\phi_{j,i} : i = 0, 1, \ldots, N_j\}$ is the standard nodal basis of V_j^c for $j = 0, 1, \cdots, J$. For each $\mu_h \in \mathbb{M}_{h,k}$, we use $\widetilde{\mu}_h \in \mathbb{R}^M$ to denote the vector of coefficients of μ_h with respect to the basis $\{\eta_1, \eta_2, \ldots, \eta_M\}$. Let $\mathcal{D}_h \in \mathbb{R}^{M\times M}$ be the stiffness matrix with respective to the operator D_h defined in (2.5) with

$$\widetilde{\lambda}_h^T \mathcal{D}_h \widetilde{\mu}_h := \langle D_h\mu_h, \eta_h\rangle_h \qquad \text{for all } \lambda_h, \mu_h \in \mathbb{M}_{h,k}.$$

Then it follows from Theorem 2.1, or the estimate (2.6), that

$$\kappa(\mathcal{D}_h) \lesssim h^{-2}.$$

By the definition, (3.2), of Π_h, there exists a matrix $\mathcal{I}_j \in \mathcal{R}^{M\times N_j}$ for $j = 0, 1, \cdots, J$, such that

$$\Pi_h(\phi_{j,1}, \phi_{j,2}, \dots, \phi_{j,N_j}) = (\eta_1, \eta_2, \dots, \eta_M)\mathcal{I}_j. \tag{4.10}$$

We set $\mathcal{I}_h \in \mathbb{R}^{M\times M}$ to be the identity matrix. From the definition, (4.2), of B_h, it follows, for any $\mu_h \in \mathbb{M}_{h,k}$,

$$\begin{aligned} B_h D_h \mu_h &= h^{2-d}\sum_{i=1}^{M}\langle D_h\mu_h, \eta_i\rangle_h \eta_i + \sum_{(j,i)\in\Lambda} h_j^{2-d}(\Pi_h^t D_h\mu_h, \phi_{j,i})\Pi_h\phi_{j,i} \\ &= h^{2-d}\sum_{i=1}^{M}\langle D_h\mu_h, \eta_i\rangle_h \eta_i + \sum_{(j,i)\in\Lambda} h_j^{2-d}\langle D_h\mu_h, \Pi_h\phi_{j,i}\rangle_h \Pi_h\phi_{j,i}. \end{aligned}$$

Thus, in view of (4.10), we have

$$\widetilde{B_h D_h \mu_h} = \mathcal{B}_h\mathcal{D}_h\widetilde{\mu}_h, \ \forall \mu_h \in \mathbb{M}_{h,k}, \tag{4.11}$$

where $\mathcal{B}_h$, a preconditioner for $\mathcal{D}_h$, is given by

$$\mathcal{B}_h = h^{2-d}\mathcal{I}_h + \sum_{k=0}^{J} h_j^{2-d}\mathcal{I}_j\mathcal{I}_j^t. \tag{4.12}$$

From Theorem 4.1 it follows

$$\kappa(\mathcal{B}_h\mathcal{D}_h) \lesssim 1. \tag{4.13}$$

This means that the matrix $\mathcal{B}_h$ is an optimal preconditioner for the stiffness matrix $\mathcal{D}_h$.

5 Applications

We begin by introducing some notation. For any $T \in \mathcal{T}_h$, let $V(T)$ and $\boldsymbol{W}(T)$ be two local finite dimensional spaces. Define

$$\begin{aligned} V_h &:= \{v \in L^2(\Omega) : v_h|_T \in V(T) \text{ for all } T \in \mathcal{T}_h\}, \\ \boldsymbol{W}_h &:= \{\boldsymbol{\tau} \in [L^2(\Omega)]^d : \boldsymbol{\tau}_h|_T \in \boldsymbol{W}(T) \text{ for all } T \in \mathcal{T}_h\}. \end{aligned}$$

Then we introduce another local space as follows:

$$M(\partial T) := \left\{\mu \in L^2(\partial T) : \mu|_F \in P_k(F) \text{ for all } F \in \mathcal{F}_T\right\}.$$

We recall

$$\mathbb{M}_{h,k} := \{\mu_h \in L^2(\mathcal{F}_h) : \mu_h|_F \in P_k(F) \ \text{ for all } F \in \mathcal{F}_h \text{ and } \mu_h|_{\partial\Omega} = 0\}. \tag{5.1}$$

For the sake of clarity, in what follows we assume $g = 0$ for the model problem (1.1).

5.1 Hybridized discontinuous Galerkin method

The general framework of HDG method for the problem (1.1) reads as follows ([21]): Seek $(u_h, \lambda_h, \boldsymbol{\sigma}_h) \in V_h \times \mathbb{M}_{h,k} \times \boldsymbol{W}_h$ such that

$$
\begin{aligned}
(\boldsymbol{C}\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + (u_h, \mathrm{div}_h \boldsymbol{\tau}_h) - \sum_{T\in\mathcal{T}_h} \langle \lambda_h, \boldsymbol{\tau}_h \cdot \boldsymbol{n}\rangle_{\partial T} &= 0, && (5.2a)\\
-(v_h, \mathrm{div}_h \boldsymbol{\sigma}_h) + \sum_{T\in\mathcal{T}_h} \langle \alpha_T (P_T^{\partial} u_h - \lambda_h), v_h\rangle_{\partial T} &= (f, v_h), && (5.2b)\\
\sum_{T\in\mathcal{T}_h} \langle \boldsymbol{\sigma}_h \cdot \boldsymbol{n} - \alpha_T (P_T^{\partial} u_h - \lambda_h), \mu_h\rangle_{\partial T} &= 0 && (5.2c)
\end{aligned}
$$

hold for all $(v_h, \mu_h, \boldsymbol{\tau}_h) \in V_h \times \mathbb{M}_{h,k} \times \boldsymbol{W}_h$, where $\boldsymbol{C} = \boldsymbol{A}^{-1}$, div_h is the broken div operator with respective to the triangulation $\mathcal{T}_h$, $\boldsymbol{n}$ denotes the unit outward normal of T, $P_T^{\partial} : H^1(T) \to M(\partial T)$ denotes the standard L^2-orthogonal projection operator, and α_T denotes a nonnegative penalty function defined on ∂T.

For any $T \in \mathcal{T}_h$, we introduce two local problems as follows.

Local problem 1: For any given $\lambda \in L^2(\partial T)$, seek $(u_\lambda, \boldsymbol{\sigma}_\lambda) \in V(T) \times \boldsymbol{W}(T)$ such that

$$
\begin{aligned}
(\boldsymbol{C}\boldsymbol{\sigma}_\lambda, \boldsymbol{\tau})_T + (u_\lambda, \mathrm{div}\boldsymbol{\tau})_T &= \langle \lambda, \boldsymbol{\tau} \cdot \boldsymbol{n}\rangle_{\partial T}, && (5.3a)\\
-(v, \mathrm{div}\boldsymbol{\sigma}_\lambda)_T + \langle \alpha_T P_T^{\partial} u_\lambda, v\rangle_{\partial T} &= \langle \alpha_T \lambda, v\rangle_{\partial T}, && (5.3b)
\end{aligned}
$$

hold for all $(v, \boldsymbol{\tau}) \in V(T) \times \boldsymbol{W}(T)$.

Local problem 2: For any given $f \in L^2(T)$, seek $(u_f, \boldsymbol{\sigma}_f) \in V(T) \times \boldsymbol{W}(T)$ such that

$$
\begin{aligned}
(\boldsymbol{C}\boldsymbol{\sigma}_f, \boldsymbol{\tau})_T + (u_f, \mathrm{div}\boldsymbol{\tau})_T &= 0, && (5.4a)\\
-(v, \mathrm{div}\boldsymbol{\sigma}_f)_T + \langle \alpha_T P_T^{\partial} u_f, v\rangle_{\partial T} &= (f, v)_T, && (5.4b)
\end{aligned}
$$

hold for all $(v, \boldsymbol{\tau}) \in V(T) \times \boldsymbol{W}(T)$.

Theorem 5.1. *[21] Suppose $(u_h, \lambda_h, \boldsymbol{\sigma}_h) \in V_h \times \mathbb{M}_{h,k} \times \boldsymbol{W}_h$ to be the solution to the system (5.2), and suppose, for any $T \in \mathcal{T}_h$, $(u_{\lambda_h}, \boldsymbol{\sigma}_{\lambda_h})|_T \in V(T) \times \boldsymbol{W}(T)$ and $(u_f, \boldsymbol{\sigma}_f)|_T \in V(T) \times \boldsymbol{W}(T)$ to be the solutions to the local problems (5.3) (by replacing λ with λ_h) and (5.4), respectively. Then it holds*

$$
\begin{aligned}
\boldsymbol{\sigma}_h &= \boldsymbol{\sigma}_{\lambda_h} + \boldsymbol{\sigma}_f, && (5.5)\\
u_h &= u_{\lambda_h} + u_f, && (5.6)
\end{aligned}
$$

and $\lambda_h \in \mathbb{M}_{h,k}$ is the solution to the system (1.3), i.e.

$$
d_h(\lambda_h, \mu_h) = b_h(\mu_h) \quad \text{for all } \mu_h \in \mathbb{M}_{h,k},
$$

where

$$
\begin{aligned}
d_h(\lambda_h, \mu_h) &:= (\boldsymbol{C}\boldsymbol{\sigma}_{\lambda_h}, \boldsymbol{\sigma}_{\mu_h}) + \sum_{T\in\mathcal{T}_h} \langle \alpha_T (P_T^{\partial} u_{\lambda_h} - \lambda_h), P_T^{\partial} u_{\mu_h} - \mu_h\rangle_{\partial T}, && (5.7)\\
b_h(\mu_h) &:= (f, u_{\mu_h}), && (5.8)
\end{aligned}
$$

and, for any $T \in \mathcal{T}_h$, $(u_{\mu_h}, \boldsymbol{\sigma}_{\mu_h})|_T \in V(T) \times \boldsymbol{W}(T)$ denotes the solution to the local problem (5.3) *by replacing λ with μ_h.*

We list four types of HDG methods as follows.

Type 1. $V(T) = P_k(T)$, $\boldsymbol{W}(T) = [P_k(T)]^d + P_k(T)\boldsymbol{x}$ and $\alpha_T = 0$.

Type 2. $V(T) = P_{k-1}(T)$ $(k \geqslant 1)$, $\boldsymbol{W}(T) = [P_k(T)]^d$ and $\alpha_T = 0$.

Type 3. $V(T) = P_k(T)$, $\boldsymbol{W}(T) = [P_k(T)]^d$ and $\alpha_T = O(1)$.

Type 4. $V(T) = P_{k+1}(T)$, $\boldsymbol{W}(T) = [P_k(T)]^d$ and $\alpha_T = O(h_T^{-1})$.

Type 1 HDG method turns out to be the well-known hybridized RT mixed element method ([3]), and **Type 2** HDG method turns out to be the well-known hybridized BDM mixed element method ([14]). For both **Types 1-2** HDG methods, it was shown in [25] that **Assumption 2.1** holds.

Type 3 HDG method was proposed in [21] and analyzed in [20]. In [22] it was shown that Assumption 2.1 holds for this method.

Type 4 HDG method was proposed and analyzed in [31]. The proof of **Assumption 2.1** can also be found there. For completeness we sketch the proof here.

Lemma 5.1. *Let $T \in \mathcal{T}_h$. For any given $\lambda \in M(\partial T)$, it holds*

$$(\boldsymbol{C}^{-1}\boldsymbol{\sigma}_\lambda, \boldsymbol{\sigma}_\lambda)_T + \langle \alpha_T(P_T^\partial u_\lambda - \lambda), P_T^\partial u_\lambda - \lambda\rangle_{\partial T} \sim |\lambda|_{h,\partial T}^2. \tag{5.9}$$

Proof. We first show

$$h_T^{-\frac{1}{2}} \left\|\lambda - \bar{\lambda}\right\|_{\partial T} \lesssim \|\boldsymbol{\sigma}_\lambda\|_T + \left\|\alpha_T^{\frac{1}{2}}(P_T^\partial u_\lambda - \lambda)\right\|_{\partial T}, \tag{5.10}$$

where $\bar{\lambda} = \frac{1}{|\partial T|}\int_{\partial T} \lambda$. In fact, from (5.3a) it follows

$$(\boldsymbol{\nabla} u_\lambda, \boldsymbol{\tau})_T = (\boldsymbol{C}\boldsymbol{\sigma}_\lambda, \boldsymbol{\tau})_T + \langle P_T^\partial u_\lambda - \lambda, \boldsymbol{\tau}\cdot\boldsymbol{n}\rangle_{\partial T} \quad \text{for all } \boldsymbol{\tau} \in \boldsymbol{W}(T). \tag{5.11}$$

Taking $\boldsymbol{\tau} = \boldsymbol{\nabla} u_\lambda$ in (5.11), we immediately get

$$|u_\lambda|_{1,T} \lesssim \|\boldsymbol{\sigma}_\lambda\| + h_T^{-\frac{1}{2}} \left\|P_T^\partial u_\lambda - \lambda\right\|_{\partial T}. \tag{5.12}$$

Define $\tilde{\lambda} := \lambda - \bar{\lambda}$ and $\bar{u}_\lambda := \frac{1}{|T|}\int_T u_\lambda$, then we have

$$\begin{aligned}
\langle \tilde{\lambda}, u_\lambda\rangle_{\partial T} &= \langle \tilde{\lambda}, u_\lambda - \bar{u}_\lambda\rangle_{\partial T} \\
&\leqslant \left\|\tilde{\lambda}\right\|_{\partial T} \|u_\lambda - \bar{u}_\lambda\|_{\partial T} \\
&\lesssim h_T^{\frac{1}{2}} \left\|\tilde{\lambda}\right\|_{\partial T} |u_\lambda|_{1,T} \\
&\lesssim h_T^{\frac{1}{2}} \left\|\tilde{\lambda}\right\|_{\partial T} \left(\|\boldsymbol{\sigma}_\lambda\|_T + \left\|\alpha_T^{\frac{1}{2}}(P_T^\partial u_\lambda - \lambda)\right\|_{\partial T}\right). \qquad \text{by (5.12)}
\end{aligned}$$

This estimate, together with

$$\begin{aligned}\langle \tilde{\lambda}, \lambda - P_T^{\partial} u_\lambda \rangle_{\partial T} &\leqslant \left\|\tilde{\lambda}\right\|_{\partial T} \left\|\lambda - P_T^{\partial} u_\lambda\right\|_{\partial T} \\ &\lesssim h_T^{\frac{1}{2}} \left\|\tilde{\lambda}\right\|_{\partial T} \left\|\alpha_T^{\frac{1}{2}}(\lambda - P_T^{\partial} u_\lambda)\right\|_{\partial T},\end{aligned}$$

yields

$$\begin{aligned}\left\|\tilde{\lambda}\right\|_{\partial T}^2 &= \langle \tilde{\lambda}, \tilde{\lambda} - P_T^{\partial} u_\lambda \rangle_{\partial T} + \langle \tilde{\lambda}, P_T^{\partial} u_\lambda \rangle_{\partial T} \\ &= \langle \tilde{\lambda}, \lambda - P_T^{\partial} u_\lambda \rangle_{\partial T} + \langle \tilde{\lambda}, u_\lambda \rangle_{\partial T} \\ &\lesssim h_T^{\frac{1}{2}} \left\|\tilde{\lambda}\right\|_{\partial T} \left(\|\boldsymbol{\sigma}_\lambda\|_T + \left\|\alpha_T^{\frac{1}{2}}(P_T^{\partial} u_\lambda - \lambda)\right\|_{\partial T} \right),\end{aligned}$$

which implies (5.10) immediately.

Second, we show

$$\|\boldsymbol{\sigma}_\lambda\|_T + \left\|\alpha_T^{\frac{1}{2}}(P_T^{\partial} u_\lambda - \lambda)\right\|_{\partial T} \lesssim h_T^{-\frac{1}{2}} \left\|\lambda - \bar{\lambda}\right\|_{\partial T}. \tag{5.13}$$

In fact, taking $\boldsymbol{\tau} = \boldsymbol{\sigma}_\lambda$ in (5.3a), $v = u_\lambda - \bar{\lambda}$ in (5.3b), and adding the two resultant equations, we obtain

$$\begin{aligned}&\left\|\boldsymbol{C}^{\frac{1}{2}} \boldsymbol{\sigma}_\lambda\right\|_T^2 + \left\|\alpha_T^{\frac{1}{2}}(P_T^{\partial} u_\lambda - \bar{\lambda})\right\|_{\partial T}^2 \\ =\ & \langle \lambda - \bar{\lambda}, \boldsymbol{\sigma}_\lambda \cdot \boldsymbol{n} \rangle_{\partial T} + \langle \alpha_T (\lambda - \bar{\lambda}), P_T^{\partial} u_\lambda - \bar{\lambda} \rangle_{\partial T} \\ \leqslant\ & \left\|\lambda - \bar{\lambda}\right\|_{\partial T} \|\boldsymbol{\sigma}_\lambda\|_{\partial T} + \left\|\alpha_T^{\frac{1}{2}}(\lambda - \bar{\lambda})\right\|_{\partial T} \left\|\alpha_T^{\frac{1}{2}}(P_T^{\partial} u_\lambda - \bar{\lambda})\right\|_{\partial T} \\ \lesssim\ & h_T^{-\frac{1}{2}} \left\|\lambda - \bar{\lambda}\right\|_{\partial T} \|\boldsymbol{\sigma}_\lambda\|_T + \left\|\alpha_T^{\frac{1}{2}}(\lambda - \bar{\lambda})\right\|_{\partial T} \left\|\alpha_T^{\frac{1}{2}}(P_T^{\partial} u_\lambda - \bar{\lambda})\right\|_{\partial T} \\ \lesssim\ & h_T^{-\frac{1}{2}} \left\|\lambda - \bar{\lambda}\right\|_{\partial T} \left(\|\boldsymbol{\sigma}_\lambda\|_T + \left\|\alpha_T^{\frac{1}{2}}(P_T^{\partial} u_\lambda - \bar{\lambda})\right\|_{\partial T} \right),\end{aligned}$$

which implies

$$\|\boldsymbol{\sigma}_\lambda\|_T + \left\|\alpha_T^{\frac{1}{2}}(P_T^{\partial} u_\lambda - \bar{\lambda})\right\|_{\partial T} \lesssim h_T^{-\frac{1}{2}} \left\|\lambda - \bar{\lambda}\right\|_{\partial T}. \tag{5.14}$$

By noticing that the above estimate also indicates

$$\begin{aligned}\left\|\alpha_T^{\frac{1}{2}}(P_T^{\partial} u_\lambda - \lambda)\right\|_{\partial T} &\leqslant \left\|\alpha_T^{\frac{1}{2}}(P_T^{\partial} u_\lambda - \bar{\lambda})\right\|_{\partial T} + \left\|\alpha_T^{\frac{1}{2}}(\lambda - \bar{\lambda})\right\|_{\partial T} \\ &\lesssim h_T^{-\frac{1}{2}} \left\|\lambda - \bar{\lambda}\right\|_{\partial T},\end{aligned}$$

the estimate (5.13) follows immediately. Then, from (5.10) and (5.13), it follows

$$(\boldsymbol{C}^{-1} \boldsymbol{\sigma}_\lambda, \boldsymbol{\sigma}_\lambda)_T + \langle \alpha_T (P_T^{\partial} u_\lambda - \lambda), P_T^{\partial} u_\lambda - \lambda \rangle_{\partial T} \sim h_T^{-1} \left\|\lambda - \bar{\lambda}\right\|_{\partial T}^2. \tag{5.15}$$

A standard scaling argument shows

$$|\lambda|_{h,\partial T} \sim h_T^{-\frac{1}{2}} \left\|\lambda - \bar{\lambda}\right\|_{\partial T}, \tag{5.16}$$

which, together with (5.15), indicates the desired estimate (5.9). □

Based on Lemma 5.1, it is trivial to derive the proposition below.

Proposition 5.1. *For* **Type 4** *HDG method,* **Assumption 2.1** *holds.*

Remark 5.1. *It has been shown in [17, 18] that, when $\boldsymbol{A}$ is a piecewise constant matrix and $k \geq 1$, the bilinear form $d_h(\cdot,\cdot)$ arising from the hybridized RT mixed element method, i.e.* **Type 1** *HDG method, coincides with that arising from the hybridized BDM mixed element method, i.e.* **Type 2** *HDG method. Then any preconditioner for* **Type 1** *HDG method is also a preconditioner for* **Type 2** *HDG method, and vice versa.*

Remark 5.2. *In [22], a first analysis of multigrid method for* **Type 3** *HDG method was presented. However, it was required there that that the model problem (1.1) admits the regulartiy estimate $\|u\|_{1+\alpha,\Omega} \leqslant C_{\alpha,\Omega} \|f\|_{\alpha-1,\Omega}$ with $\alpha \in (0.5, 1]$ and $C_{\alpha,\Omega}$ being a positive constant that only depends on α and Ω. We note that our analysis in Section 4 for the BPX preconditioner does not require any regularity assumption. In [32], a more general framework for designing multilevel methods for HDG methods were presented and analyzed without any regularity assumption.*

5.2 Weak Galerkin method

At first, we follow [42] to introduce the discrete weak gradients. Let $T \in \mathcal{T}_h$. We define $\boldsymbol{\nabla}_w^i : L^2(T) \to \boldsymbol{W}(T)$ by

$$(\boldsymbol{\nabla}_w^i v, \boldsymbol{q})_T := -(v, \operatorname{div} \boldsymbol{q})_T \quad \text{for all } v \in L^2(T) \text{ and } \boldsymbol{q} \in \boldsymbol{W}(T)\,, \tag{5.17}$$

and define $\boldsymbol{\nabla}_w^b : L^2(\partial T) \to \boldsymbol{W}(T)$ by

$$(\boldsymbol{\nabla}_w^b \mu, \boldsymbol{q})_T := \langle \mu, \boldsymbol{q}\cdot\boldsymbol{n}\rangle_{\partial T} \quad \text{for all } \mu \in L^2(\partial T) \text{ and } \boldsymbol{q} \in \boldsymbol{W}(T). \tag{5.18}$$

Then we define the discrete weak gradients $\boldsymbol{\nabla}_w : L^2(T) \times L^2(\partial T) \to \boldsymbol{W}(T)$ with

$$\boldsymbol{\nabla}_w(v,\mu) := \boldsymbol{\nabla}_w^i v + \boldsymbol{\nabla}_w^b \mu \quad \text{for all } (v,\mu) \in L^2(T) \times L^2(\partial T). \tag{5.19}$$

Hence, the WG discretization reads as follows: Seek $(u_h, \lambda_h) \in V_h \times \mathbb{M}_{h,k}$ such that

$$(\boldsymbol{A}\boldsymbol{\nabla}_w(u_h,\lambda_h), \boldsymbol{\nabla}_w(v_h,\mu_h)) + \sum_{T\in\mathcal{T}_h} \langle \alpha_T (P_T^\partial u_h - \lambda_h), P_T^\partial v_h - \mu_h\rangle_{\partial T} = (f, v_h) \tag{5.20}$$

holds for all $(v_h, \mu_h) \in V_h \times \mathbb{M}_{h,k}$, where α_T denotes a nonnegative penalty function defined on ∂T.

We shall follow the same routine as in the previous subsection to show a new characterization of the WG method. We introduce two local problems as follows.

Local problem 1': For any given $f \in L^2(T)$, seek $u_f \in V(T)$ such that

$$(\boldsymbol{A}\boldsymbol{\nabla}_w^i u_f, \boldsymbol{\nabla}_w^i v)_T + \langle \alpha_T P_T^\partial u_f, P_T^\partial v\rangle_{\partial T} = (f, v)_T \tag{5.21}$$

holds for all $v \in V(T)$.

Local problem 2': For any given $\lambda \in L^2(\partial T)$, seek $u_\lambda \in V(T)$ such that

$$(\boldsymbol{A}\boldsymbol{\nabla}_w^i u_\lambda, \boldsymbol{\nabla}_w^i v)_T + \langle \alpha_T P_T^\partial u_\lambda, P_T^\partial v\rangle_{\partial T} = -(\boldsymbol{A}\boldsymbol{\nabla}_w^b \lambda, \boldsymbol{\nabla}_w^i v)_T + \langle \alpha_T \lambda, P_T^\partial v\rangle_{\partial T} \tag{5.22}$$

holds for all $v \in V(T)$.

Similar to Theorem 5.1, the following conclusion holds.

Theorem 5.2. *Suppose $(u_h, \lambda_h) \in V_h \times \mathbb{M}_{h,k}$ to be the solution to the system (5.20), and suppose, for any $T \in \mathcal{T}_h$, u_f and u_{λ_h} to be the solutions to the local problems (5.21) and (5.22) (by replacing λ with λ_h), respectively. Then it holds*

$$u_h = u_{\lambda_h} + u_f, \tag{5.23}$$

and $\lambda_h \in \mathbb{M}_{h,k}$ is the solution to the system (1.3), i.e.

$$d_h(\lambda_h, \mu_h) = b_h(\mu_h) \quad \textit{for all } \mu_h \in \mathbb{M}_{h,k},$$

where

$$d_h(\lambda_h, \mu_h) := (\boldsymbol{A}\boldsymbol{\nabla}_w(u_{\lambda_h,\lambda_h}), \boldsymbol{\nabla}_w(u_{\mu_h}, \mu_h)) + \sum_{T\in\mathcal{T}_h} \langle \alpha_T (P_T^\partial u_{\lambda_h} - \lambda_h), P_T^\partial u_{\mu_h} - \mu_h\rangle_{\partial T}, \tag{5.24}$$

$$b_h(\mu_h) := (f, u_{\mu_h}). \tag{5.25}$$

We consider two types of WG methods ([42]):

- **Type 1.** $V(T) = P_k(T)$, $\boldsymbol{W}(T) = [P_k(T)]^d + P_K(T)\boldsymbol{x}$ and $\alpha_T = 0$;
- **Type 2.** $V(T) = P_{k-1}(T)$ $(k \geqslant 1)$, $\boldsymbol{W}(T) = [P_k(T)]^d$ and $\alpha_T = 0$.

In both cases, we can prove that **Assumption 2.1** holds.

Theorem 5.3. *For* **Type 1** *WG method,* **Assumption 2.1** *holds.*

Proof. For $T \in \mathcal{T}_h$, define $\boldsymbol{\sigma} := \boldsymbol{\nabla}_w(u_{\lambda_h,\lambda_h})|_T$. Then from (5.22) it follows

$$(\boldsymbol{A}\boldsymbol{\sigma}, \boldsymbol{\nabla}_w^i v)_T = 0 \text{ for all } v \in V(T),$$

which implies

$$\mathrm{div} P_T^{rt}(\boldsymbol{A}\boldsymbol{\sigma}) = 0, \tag{5.26}$$

where $P_T^{tr} : [L^2(T)]^d \to \boldsymbol{W}(T)$ denotes the standard L^2-orthogonal projection operator. By the definition of $\boldsymbol{\nabla}_w$, we have

$$\begin{aligned}
(\boldsymbol{A}\boldsymbol{\sigma}, \boldsymbol{\sigma})_T &= (P_T^{rt}(\boldsymbol{A}\boldsymbol{\sigma}), \boldsymbol{\nabla}_w(u_{\lambda_h}, \lambda_h))_T \\
&= -(\mathrm{div}(P_T^{rt}(\boldsymbol{A}\boldsymbol{\sigma})), u_{\lambda_h})_T + \langle P_T^{rt}(\boldsymbol{A}\boldsymbol{\sigma}) \cdot \boldsymbol{n}, \lambda_h\rangle_{\partial T} \\
&= \langle P_T^{rt}(\boldsymbol{A}\boldsymbol{\sigma}) \cdot \boldsymbol{n}, \lambda_h - m_T(\lambda_h)\rangle_{\partial T} \qquad (\text{by } (5.26)) \\
&\lesssim h_T^{-\frac{1}{2}} \left\|P_T^{rt}(\boldsymbol{A}\boldsymbol{\sigma})\right\|_T \|\lambda_h - m_T(\lambda_h)\|_{\partial T} \\
&\lesssim \|\boldsymbol{A}\boldsymbol{\sigma}\|_T\, |\lambda_h|_{h,\partial T},
\end{aligned}$$

which shows immediately

$$(\boldsymbol{A}\boldsymbol{\sigma},\boldsymbol{\sigma})_T \lesssim |\lambda_h|^2_{h,\partial T}. \tag{5.27}$$

On the other hand, for any $\boldsymbol{\tau}\in\boldsymbol{W}(T)$, from the definition of $\boldsymbol{\nabla}_w$ we have

$$\begin{aligned}(\boldsymbol{\sigma},\boldsymbol{\tau})_T &= (\boldsymbol{\nabla}_w(u_{\lambda_h},\lambda_h),\boldsymbol{\tau})_T\\ &= -(u_{\lambda_h},\mathrm{div}\boldsymbol{\tau})_T + \langle \boldsymbol{\tau}\cdot\boldsymbol{n},\lambda_h\rangle_{\partial T}\\ &= (\boldsymbol{\nabla}u_{\lambda_h},\boldsymbol{\tau})_T + \langle \boldsymbol{\tau}\cdot\boldsymbol{n},\lambda_h-u_{\lambda_h}\rangle_{\partial T},\end{aligned}$$

which yields

$$(\boldsymbol{\sigma}-\boldsymbol{\nabla}u_{\lambda_h},\boldsymbol{\tau})_T = \langle \boldsymbol{\tau}\cdot\boldsymbol{n},\lambda_h-u_{\lambda_h}\rangle_{\partial T}. \tag{5.28}$$

Taking $\boldsymbol{\tau}\in\boldsymbol{W}(T)$ in (5.28) with

$$\left\{\begin{array}{rcll}\int_F \boldsymbol{\tau}\cdot\boldsymbol{n}q &=& \int_F(\lambda_h-u_{\lambda_h})q & \text{for all } F\in\mathcal{F}_T \text{ and } q\in P_k(F),\\ \int_T \boldsymbol{\tau}\cdot\boldsymbol{\nabla}v &=& 0 & \text{for all } v\in V(T),\end{array}\right. \tag{5.29}$$

we have

$$\begin{aligned}\|\lambda_h-u_{\lambda_h}\|^2_{\partial T} = (\boldsymbol{\sigma}-\boldsymbol{\nabla}u_{\lambda_h},\boldsymbol{\tau})_T &= (\boldsymbol{\sigma},\boldsymbol{\tau})_T \leqslant \|\boldsymbol{\sigma}\|_T\,\|\boldsymbol{\tau}\|_T\\ &\lesssim h_T^{\frac{1}{2}}\,\|\boldsymbol{\sigma}\|_T\,\|\lambda_h-u_{\lambda_h}\|_{\partial T}\,,\end{aligned}$$

where we have used the estimate $\|\boldsymbol{\tau}\|_T \lesssim h_T^{\frac{1}{2}}\,\|\lambda_h-u_{\lambda_h}\|_{\partial T}$, which is a trivial result by applying the famous Piola mapping. The above inequality leads to

$$\|\lambda_h-u_{\lambda_h}\|_{\partial T} \lesssim h_T^{\frac{1}{2}}\,\|\boldsymbol{\sigma}\|_T\,. \tag{5.30}$$

Similarly, taking $\boldsymbol{\tau}\in\boldsymbol{W}(T)$ in (5.28) with

$$\left\{\begin{array}{rcll}\int_F \boldsymbol{\tau}\cdot\boldsymbol{n}q &=& 0 & \text{for all } F\in\mathcal{F}_T \text{ and } q\in P_k(F),\\ \int_T \boldsymbol{\tau},\boldsymbol{\nabla}v)_T &=& (\boldsymbol{\nabla}u_{\lambda_h},\boldsymbol{\nabla}v)_T & \text{for all } v\in V(T),\end{array}\right. \tag{5.31}$$

we have

$$|u_{\lambda_h}|_{1,T} \lesssim \|\boldsymbol{\sigma}\|_T\,. \tag{5.32}$$

Since by a standard scaling argument it holds

$$\|\lambda_h-m_T(\lambda_h)\|_{\partial T} \sim \inf_{c\in\mathbb{R}}\|\lambda_h-c\|_{\partial T}\,, \tag{5.33}$$

we have

$$\begin{aligned}\|\lambda_h-m_T(\lambda_h)\|_{\partial T} &\sim \inf_{c\in\mathbb{R}}\|\lambda_h-c\|_{\partial T}\\ &\lesssim \|\lambda_h-u_{\lambda_h}\|_{\partial T} + \inf_{c\in\mathbb{R}}\|u_{\lambda_h}-c\|_{\partial T}\\ &\lesssim \|\lambda_h-u_{\lambda_h}\|_{\partial T} + h_T^{\frac{1}{2}}|u_{\lambda_h}|_{1,T}\\ &\lesssim h_T^{\frac{1}{2}}\,\|\boldsymbol{\sigma}\|_T\,, \qquad \text{(by (5.30) and (5.32))}\end{aligned}$$

which, together with (5.27), yields

$$(\boldsymbol{A}\boldsymbol{\nabla}_w(u_{\lambda_h},\lambda_h),\boldsymbol{\nabla}_w(u_{\lambda_h},\lambda_h))_T \sim |\lambda_h|_{h,\partial T}. \tag{5.34}$$

As a result, the desired estimate (2.2) follows immediately. This completes the proof. □

Remark 5.3. *Similarly, we can show that* **Assumption 2.1** *holds for* **Type 2** *WG method.*

Remark 5.4. *If $\boldsymbol{A}$ is a piecewise constant matrix, the two WG methods are equivalent to the hybridized RT mixed element method and the hybridized BDM mixed element method, respectively. We refer to (Remark 2.1, [30]) for the details.*

5.3 Nonconforming finite element method

In this subsection we take Crouzeix-Raviart element method [23] as an example to show that the theory in Section 4 also applies to nonconforming methods.

At first, we introduce the Crouzeix-Raviart finite element space $\mathcal{L}_h^{CR}$ as follows.

$$\begin{aligned}\mathcal{L}_h^{CR} := \{v_h \in L^2(\Omega) : v_h|_T \in P_1(T), \forall T \in \mathcal{T}_h, v_h \text{ is continuous at the}\\ \text{gravity point of each interior face of } \mathcal{T}_h \text{ and vanishes at the}\\ \text{gravity point of each face of } \mathcal{T}_h \text{ that lies on } \partial\Omega\}.\end{aligned} \tag{5.35}$$

As we know, the standard discretization of the Crouzeix-Raviart element method reads as follows: Seek $u_h \in \mathcal{L}_h^{CR}$ such that

$$(\boldsymbol{A}\boldsymbol{\nabla}_h u_h, \boldsymbol{\nabla}_h v_h) = (f, v_h) \quad \text{for all } v_h \in \mathcal{L}_h^{CR}, \tag{5.36}$$

where $\boldsymbol{\nabla}_h v_h$ is given by

$$\boldsymbol{\nabla}_h v_h|_T := \boldsymbol{\nabla}(v_h|_T) \quad \text{for all } T \in \mathcal{T}_h.$$

We define an operator $\widetilde{\Pi}_h : \mathcal{L}_h^{CR} \to M_{h,0}$ by

$$\widetilde{\Pi}_h v_h|_F := \frac{1}{|F|}\int_F v_h \quad \text{for all } F \in \mathcal{F}_h. \tag{5.37}$$

Obviously, $\widetilde{\Pi}_h$ is a bijective map, and its inverse map $\widetilde{\Pi}_h^{-1} : M_{h,0} \to \mathcal{L}_h^{CR}$ satisfies

$$\int_F \widetilde{\Pi}_h^{-1}\mu_h = \int_F \mu_h \quad \text{for all } F \in \mathcal{F}_h \text{ and } \mu_h \in M_{h,0}. \tag{5.38}$$

By denoting $\mu_h := \widetilde{\Pi}_h v_h$ and $\lambda_h := \widetilde{\Pi}_h u_h$, the system (5.36) is equivalent to the system (1.3), i.e.

$$d_h(\lambda_h, \mu_h) = b_h(\mu_h) \quad \text{for all } \mu_h \in \mathbb{M}_{h,0},$$

where

$$\begin{aligned} d_h(\lambda_h,\mu_h) &:= (\boldsymbol{A}\boldsymbol{\nabla}_h\widetilde{\Pi}_h^{-1}\lambda_h, \boldsymbol{\nabla}_h\widetilde{\Pi}_h^{-1}\mu_h), \\ b_h(\mu_h) &:= (f, \widetilde{\Pi}_h^{-1}\mu_h). \end{aligned} \tag{5.39}$$

By using standard scaling arguments, it is easy to verify that **Assumption 2.1** holds in this case.

Remark 5.5. *We use d_h^{hdg} and d_h^{cr} to denote the bilinear forms defined in (5.7) and (5.39), respectively. When $\boldsymbol{A}$ is a piecewise constant matrix, we can show that for* **Type 4** *HDG method ($k = 0$), $d_h^{hdg} = d_h^{cr}$. In fact, in this case we have, for any $T \in \mathcal{T}_h$,*

$$M(\partial T) := \left\{\mu \in L^2(\partial T) : \mu|_F \in P_0(F) \text{ for each face } F \text{ of } T\right\},$$

$$V(T) = P_1(T), \quad \boldsymbol{W}(T) = [P_0(T)]^d, \quad \alpha_T = O(h_T^{-1}).$$

From (5.3b) it follows

$$\langle \alpha_T (P_T^\partial (u_\lambda - \lambda), P_T^\partial v - \mu\rangle_{\partial T} = 0 \quad \text{for all } (v, \mu) \in V(T) \times M(\partial T).$$

Thus, in view of (5.7), it holds

$$d_h^{hdg}(\lambda_h, \mu_h) = (\boldsymbol{C}\boldsymbol{\sigma}_{\lambda_h}, \boldsymbol{\sigma}_{\mu_h}). \tag{5.40}$$

On the other hand, it follows from (5.3a) and (5.38) that

$$(\boldsymbol{C}\boldsymbol{\sigma}_{\mu_h}, \boldsymbol{\tau})_T = \langle \mu_h, \boldsymbol{\tau}\cdot\boldsymbol{n}\rangle_{\partial T} = \langle \widetilde{\Pi}_h^{-1}\mu_h, \boldsymbol{\tau}\cdot\boldsymbol{n}\rangle_{\partial T} = (\boldsymbol{\nabla}\widetilde{\Pi}_h^{-1}\mu_h, \boldsymbol{\tau})_T$$

holds for all $\boldsymbol{\tau} \in \boldsymbol{W}(T)$. Since $\boldsymbol{C} = \boldsymbol{A}^{-1}$ is a constant matrix on T, the above equality means

$$\boldsymbol{\nabla}_h \widetilde{\Pi}_h^{-1}\mu_h = \boldsymbol{C}\boldsymbol{\sigma}_{\mu_h}. \tag{5.41}$$

Thus, in light of (5.40)-(5.41) and (5.39) we have

$$d_h^{hdg}(\lambda_h, \mu_h) = (\boldsymbol{A}\boldsymbol{\nabla}\widetilde{\Pi}_h^{-1}\lambda_h, \boldsymbol{\nabla}\widetilde{\Pi}_h^{-1}\mu_h) = d_h^{cr}(\lambda_h, \mu_h).$$

Remark 5.6. *As shown in [17, 18], when $\boldsymbol{A}$ is a piecewise constant matrix, the stiffness matrix of $d_h(\cdot,\cdot)$ arising from the lowest order hybridized RT mixed finite element method, i.e.* **Type 1** *HDG method ($k = 0$) in Subsection 5.1, is the same as that arising from the Crouzeix-Raviart element method.*

Remark 5.7. *From Remarks 5.4-5.6, we know that when $\boldsymbol{A}$ is a piecewise constant matrix and $k = 0$, the four methods, namely* **Type 1** *and* **Type 4** *HDG methods in Subsection 5.1,* **Type 1** *WG method in Subsection 5.2, and the Crouzeix-Raviart element method, lead to the same bilinear form $d_h(\cdot,\cdot)$, and hence share the same optimal preconditioners.*

6 Numerical experiments

In this section, we report several numerical examples in two-dimensions to verify the theoretical results of Theorem 4.1 and Theorem 3.1. We only consider the problem (1.1) with the diffusion tensor $\boldsymbol{A} = \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix. We test two types of HDG methods, i.e., **Type 3** HDG method ($k = 0, 1$) with $\alpha_T = 1$ and **Type 4** HDG method ($k = 0, 1$) with $\alpha_T = h_T^{-1}$ for all $T \in \mathcal{T}_h$. We refer to [22, 32] for more numerical results of **Type 3** HDG method.

Example 1. We set $\Omega = (0, 1) \times (0, 1)$ (a square domain) with the initial triangulation $\mathcal{T}_0$ (Figure 1). We produce a sequence of triangulations $\{\mathcal{T}_j : j = 1, 2, \cdots, 10\}$ by a successive refinement procedure: connecting the midpoints of three edges of each triangle.

For each $j = 5, 6, \cdots, 10$, we set $\mathcal{T}_h = \mathcal{T}_j$, and let $\mathcal{D}_h$ and $\mathcal{B}_h$ be defined by (2.5) and (4.2) respectively. Suppose we are to solve the system $\mathcal{D}_h x = b_h$, where b_h is a zero vector. Taking $x_0 = (1, 1, \ldots, 1)^t$ as the initial value, we use the famous preconditioned conjugate gradient method (PCG) to solve this system with the preconditioner $\mathcal{B}_h$. The stopping criterion is that the initial error, i.e. $\sqrt{x_0^T \mathcal{D}_h x_0}$, is reduced by a factor of 10^{-6}.

In the case $k = 0$, the prolongation operators Π_h^1 and Π_h^2 are equivalent, so we have one BPX preconditioner. In the case $k = 1$, we have two different BPX preconditioners since Π_h^1 and Π_h^2 are not equivalent, and we compute both cases. The corresponding numerical results, i.e. the number of iterations in PCG, are listed in Table 1.

Example 2. The only difference between this example and Example 1 is that we set $\Omega = (0, 1) \times (0, 1)/[0, 1) \times \{0.5\}$ (a crack domain) with the initial triangulation $\mathcal{T}_0$ (Figure 2). The corresponding numerical results are presented in Table 2.

Example 3. This example is to verify Theorem 3.1 for graded triangulations. We only consider **Type 3** HDG method. For simplicity we only highlight the difference between this example and the previous two examples.

We set $\Omega = (-1, 1) \times (-1, 1)$ and define $\boldsymbol{A}(x, y) = \mathrm{diag}(a(x, y), a(x, y))$ with

$$a(x, y) := \begin{cases} 1, & -1 < x < 0, \quad -1 < y < 0; \\ 7, & \quad 0 < x < 1, \quad -1 < y < 0; \\ 17, & \quad 0 < x < 1, \quad \quad 0 < y < 1; \\ 3, & -1 < x < 0, \quad \quad 0 < y < 1. \end{cases}$$

We show the first two triangulations $\mathcal{T}_0$ and $\mathcal{T}_1$ in Figure 3 and produce a sequence of graded triangulations $\{\mathcal{T}_j : j = 0, 1, \cdots, 25\}$ in a successive way: $\mathcal{T}_{j+1}$ ($j = 2, 3, \cdots, 24$) is obtained by refining the smallest square containing the origin in $\mathcal{T}_j$ (in $\mathcal{T}_1$, the vertexes of the square to refine is in red color) as what has been done from $\mathcal{T}_0$ to $\mathcal{T}_1$. $\mathcal{T}_{25}$ is shown in Figure 4. For each $j = 5, 10, 15, 20, 25$, we set $\mathcal{T}_h = \mathcal{T}_j$ and, in the definition (3.7) of B_h^G, we set S_h to be the standard symmetric Gauss-Seidel iteration and set $\widetilde{B_h} = A_h^{-1}$, where $A_h^{-1} : V_h^c \to V_h^c$ is defined by

$$(A_h u_h, v_h) := (\boldsymbol{A}\boldsymbol{\nabla} u_h, \boldsymbol{\nabla} v_h) \quad \text{for all } u_h, v_h \in V_h^c.$$

The corresponding numerical results are presented in Table 3.

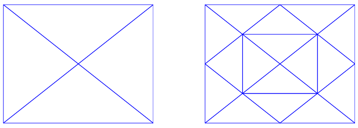

Figure 1: $\mathcal{T}_0$ (left) and $\mathcal{T}_1$ (right) on square domain

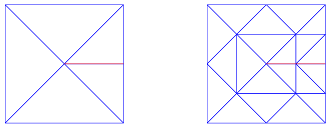

Figure 2: $\mathcal{T}_0$ (left) and $\mathcal{T}_1$ (right) on crack domain

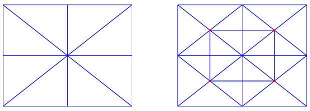

Figure 3: $\mathcal{T}_0$ (left) and $\mathcal{T}_1$ (right)

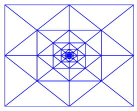

Figure 4: $\mathcal{T}_{25}$

k		$\mathcal{T}_5$	$\mathcal{T}_6$	$\mathcal{T}_7$	$\mathcal{T}_8$	$\mathcal{T}_9$	$\mathcal{T}_{10}$
0	dof	6080	24448	98048	392704	1571840	6289408
		20	20	20	20	20	19
1	dof	12288	49152	196608	786432	3145728	12582912
	Π_h^1	54	57	59	60	61	62
	Π_h^2	31	32	33	34	34	35

Type 3 HDG method

k		$\mathcal{T}_5$	$\mathcal{T}_6$	$\mathcal{T}_7$	$\mathcal{T}_8$	$\mathcal{T}_9$	$\mathcal{T}_{10}$
0	dof	6080	24448	98048	392704	1571840	6289408
		20	20	20	20	20	19
1	dof	12288	49152	196608	786432	3145728	12582912
	Π_h^1	54	57	59	60	61	62
	Π_h^2	31	32	33	34	34	35

Type 4 HDG method

Table 1: Numerical results for the Example 1 (dof denotes the number of degrees of freedom)

k		$\mathcal{T}_5$	$\mathcal{T}_6$	$\mathcal{T}_7$	$\mathcal{T}_8$	$\mathcal{T}_9$	$\mathcal{T}_{10}$
0	dof	7568	30496	122432	490624	1964288	7860732
		26	26	27	27	27	27
1	dof	15360	61440	245760	983040	3932160	15728640
	Π_h^1	55	58	59	61	62	63
	Π_h^2	33	34	35	36	36	37

Type 3 HDG method

k		$\mathcal{T}_5$	$\mathcal{T}_6$	$\mathcal{T}_7$	$\mathcal{T}_8$	$\mathcal{T}_9$	$\mathcal{T}_{10}$
0	dof	7568	30496	122432	490624	1964288	7860732
		26	27	27	27	27	27
1	dof	15360	61440	245760	983040	3932160	15728640
	Π_h^1	55	58	59	61	62	63
	Π_h^2	33	34	35	36	36	37

Type 4 HDG method

Table 2: Numerical results for Example 2

k		$\mathcal{T}_5$	$\mathcal{T}_{10}$	$\mathcal{T}_{15}$	$\mathcal{T}_{20}$	$\mathcal{T}_{25}$
0		14	14	14	14	14
1	Π_h^1	26	26	27	26	26
1	Π_h^2	18	18	18	18	18

Table 3: Numerical results for Example 3

From Tables 1-3 we have the following observations.

- For all the examples, the numbers of iterations in PCG are independent of the mesh size. This means the proposed preconditioners are optimal. Besides, the prolongation operator Π_h^2 behaves better than Π_h^1 in the case $k = 1$.

- Example 1 admits the full elliptic regularity, while Example 2 only admits the regularity estimate $\|u\|_{1+\alpha,\Omega} \leqslant C_{\alpha,\Omega} \|f\|_{\alpha-1,\Omega}$ with $\alpha \in (0, \frac{1}{2})$. These two examples confirm that the proposed BPX preconditioner is optimal. This is conformable to Theorem 4.1.
- Example 3 confirms Theorem 3.1, where the triangulation $\mathcal{T}_h$ is not quasi-uniform.

References

[1] T. ARBOGAST, Z. CHEN, On the implementation of mixed methods as nonconforming methods for second-order elliptic problems, Math. Comp., 64 (1995), 943-972. 3

[2] R. A. ADAMS, J. J. F. FOURNIER, Sobolev Spaces, Academic Press, 2nd ed., 2003. 4

[3] D. N. ARNOLD, F. BREZZI, Mixed and non-conforming finite element methods: implementation, post-processing and error estimates, Modél. Math. Anal. Numér., 19 (1985), 7-35. 2, 14

[4] 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. 2

[5] I. BABUSKA, J. ODEN, J. LEE, Mixed-hybrid finite element approximations of second-order elliptic boundary-value problems, Comput. Methods Appl. Mech. Engrg., 11 (1977), 175-206. 2

[6] D. BRAESS, R. VERFÜRTH, Multigrid methods for nonconforming finite element methods, SIAM J. Numer. Anal., 27(1990), 979-986. 3

[7] J. H. BRAMBLE, Multigrid Methods, Pitman Research Notes in Mathematics, V. 294, John Wiley and Sons, 1993. 3

[8] J. H. BRAMBLE, D.Y. KWAK, J. E. PASCIAK, Uniform convergence of multigrid V-cycle iterations for indefinite and nonsymmetric problems, SIAM J. Numer. Anal., 31 (1994), 1746-1763. 3

[9] J. H. BRAMBLE, J. E. PASCIAK, New estimates for multilevel algorithms including the Vcycle, Math. Comp., 60 (1993) 447-471. 3

[10] J. H. BRAMBLE, J. E. PASCIAK, J. XU, Parallel multilevel preconditioners, Math. Comp., 55 (1990), 1-22. 3

[11] J. H. BRAMBLE, X. ZHANG, Uniform convergence of the multigrid V-cycle for an anisotropic problem, Math. Comp., 70 (2001), 979-986. 3

[12] S. C. BRENNER, An optimal-order multigrid method for P1 nonconforming finite elements, Math. Comp., 52 (1989), 1-16. 3

[13] 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. 3

[14] 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. 2, 14

[15] L. CHEN, Deriving the X-Z identity from auxiliary space method, In: The Proceedings for 19th Conferences for Domain Decomposition Methods, 2010. 9

[16] L. CHEN, J. WANG, Y. WANG, X. YE, An auxiliary space multigrid preconditioner for the weak Galerkin method, Comput. Math. Appl., doi:10.1016/j.camwa.2015.04.016. 2, 6

[17] B. COCKBURN, J. GOPALAKRISHNAN, A characterization of hybridized mixed methods for second order elliptic problems, SIAM J. Numer. Anal., 42 (2004), 283-301. 2, 16, 20

[18] B. COCKBURN, J. GOPALAKRISHNAN, New hybridization techniques, GAMM-Mitt, 2 (2005), 154-183. 2, 16, 20

[19] B. COCKBURN, B. DONG, J. GUZMÁN, A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems, Math. Comp., 77 (2008), 1887-1916. 2

[20] B. COCKBURN, J. GOPALAKRISHNAN, F. J. SAYAS, A projection-based error analysis of HDG methods, Math. Comp., 79 (2010), 1351-1367. 2, 14

[21] 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. 2, 13, 14

[22] B. COCKBURN, O. DUBOIS, J. GOPALAKRISHNAN, Multigrid for an HDG Method, IMA J. Numer. Anal. 2013, doi: 10.1093/imanum/drt024. 3, 10, 14, 16, 21

[23] M. CROUZEIX, P. A. RAVIART, Conforming and nonconforming finite element methods for solving the stationary stokes equations, RAIRO Modél. Math. Anal. Numér., 7 (1973), 33-75. 3, 19

[24] H. DUAN, S. GAO, R. TAN, S. ZHANG, A generalized BPX multigrid framework covering nonnested V-cycle methods, Math. Comp., 76 (2007), 137-152 3

[25] J. GOPALAKRISHNAN, A Schwarz preconditioner for a hybridized mixed method, Comput. Meth. Appl. Math., 3 (2003), 116-134. 3, 4, 5, 14

[26] J. GOPALAKRISHNAN, A convergent multigrid cycle for the hybridized mixed method, Numer. Linear Algebra Appl., 16 (2009), 689-714. 3, 10

[27] J. GOPALAKRISHNAN, G. KANSCHAT, A multilevel discontinuous Galerkin method, Numer. Math., 95 (2003), 527-550. 3

[28] J. KRAUS, S. MARGENOV, J. SYNKA, On the multilevel preconditioning of CrouzeixRaviart elliptic problems, Numer. Linear Algebra Appl., 15(2008), 395416. 3

[29] R.D. LAZAROV, S.D. MARGENOV, On a two-level parallel MIC (0) preconditioning of Crouzeix-Raviart non-conforming FEM systems, Numer. Meth. Appl., Lecture Notes in Computer Science, 2542(2003), 192-201. 3

[30] B. LI, X. XIE, A two-level algorithm for the weak Galerkin discretization of diffusion problems, J. Comput. Appl. Math., 287 (2015), 179-195. 2, 5, 19

[31] B. LI, X. XIE, Analysis of a family of HDG methods for second order elliptic problems, arXiv preprint arXiv:1408.5545, 2014. 2, 14

[32] B. LI, X. XIE, S. ZHANG, Analysis of a two-level algorithm for HDG methods for diffusion problems, arXiv preprint arXiv:1502.04371, 2015. 3, 10, 16, 21

[33] 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. 2

[34] L. MU, J. WANG, X. YE, A weak Galerkin finite element methods with polynomial reduction, arXiv:1304.6481. 2

[35] J. ODEN, J. LEE, Dual-Mixed Hybrid finite element method for second-order elliptic problems, Lecture Notes in Mathematics 606 (1977), 275-291. 2

[36] T. RAHMAN, X. XU, R. HOPPE, Additive Schwarz methods for the Crouzeix-Raviart mortar finite element for elliptic problems with discontinuous coefficients Numer. Math., 101(2005), 551-572. 3

[37] P.-A. RAVIART, J. M. THOMAS, A mixed finite element method for second order elliptic problems, Mathematical Aspects of Finite Element Method (I. Galligani and E. Magenes, eds.), Lecture Notes in Math. 606, Springer-Verlag, New York, 1977, 292-315. 2

[38] P. A. RAVIART, J. M. THOMAS, Primal hybrid finite element methods for 2nd order elliptic equations[J]. Mathematics of computation, 31 (1977), 391-413. 2

[39] P. A. RAVIART, J. M. THOMAS, Dual finite element models for second order elliptic problems[J]. Energy methods in finite element analysis.(A 79-53076 24-39) Chichester, Sussex, England, Wiley-Interscience, 1979, 175-191. 2

[40] F. WANG, J. CHEN, P. HUANG, A multilevel preconditioner for the C-R FEM for elliptic problems with discontinuous coefficients, Sci. China Math., 55(2012),1513-1526. 3

[41] J. WANG, Convergence analysis of multigrid algorithms for nonselfadjoint and indefinite elliptic problems, SIAM J. Numer. Anal., 30 (1993), 275-285. 3

[42] J. WANG AND X.YE, A weak Galerkin finite element method for second-order elliptic problems, J. Comput. Appl. Math., 241 (2013), 103-115. 2, 16, 17

[43] J. XU, The auxiliary space method and optimal multigrid preconditioning techniques for unstructured meshes. Computing, 56 (1996), 215-235. 3, 5, 6

[44] J. XU, L. CHEN, R. H. NOCHETTO, Optimal multilevel methods for H(grad), H(curl), and H(div) systems on graded and unstructured grids, in: Multiscale, Nonlinear and Adaptive Approximation, 2009, 599-659. 3, 11

[45] J. XU, Y. ZHU. Uniform convergent multigrid methods for elliptic problems with strongly discontinuous coefficients, Math. Models Methods Appl. Sci., 18 (2008), 77-105. 3

[46] J. XU, L. ZIKATANOV, The method of alternating projections and the method of subspace corrections in Hilbert space, J. Amer. Math. Soc., 15 (2002), 573-597. 3, 10

[47] X. XU, J. CHEN, Multigrid for the mortar element method for P1 nonconforming element, Numer. Math., 88 (2001), 381-398. 3

[48] Y. ZHU, Analysis of a multigrid preconditioner for Crouzeix-Raviart discretization of elliptic PDE with jump coefficient, Numer. Linear Algebra Appl., 21 (2014), 2438. 3