Abstract. We present the lowest-order hybridizable discontinuous Galerkin schemes with numerical integration (quadrature), denoted as HDG-P0, for the reaction-diffusion equation and the generalized Stokes equations on conforming simplicial meshes in two- and three-dimensions. Here by lowest order, we mean that the (hybrid) finite element space for the global HDG facet degrees of freedom (DOFs) is the space of piecewise constants on the mesh skeleton. A discontinuous piecewise linear space is used for the approximation of the local primal unknowns. We give the optimal a priori error analysis of the proposed HDG-P0 schemes, which hasn’t appeared in the literature yet for HDG discretizations as far as numerical integration is concerned. Moreover, we propose optimal geometric multigrid preconditioners for the statically condensed HDG-P0 linear systems on conforming simplicial meshes. In both cases, we first establish the equivalence of the statically condensed HDG system with a (slightly modified) nonconforming Crouzeix–Raviart (CR) discretization, where the global (piecewise-constant) HDG finite element space on the mesh skeleton has a natural one-to-one correspondence to the nonconforming CR (piecewise-linear) finite element space that live on the whole mesh. This equivalence then allows us to use the well-established nonconforming geometry multigrid theory to precondition the condensed HDG system. Numerical results in two- and three-dimensions are presented to verify our theoretical findings.

Mathematics Subject Classification. 65N30, 65N12, 76S05, 76D07.

Received August 30, 2022. Accepted March 12, 2023.

1. Introduction

Since their first unified introduction for the second order elliptic equation [26], hybridizable discontinuous Galerkin (HDG) methods have been gaining popularity for numerically solving partial differential equations (PDEs), and have been successfully applied in computational fluid dynamics [25, 58], wave propagation [28, 35] and continuum mechanics [37, 52]. Besides succeeding attractive features from the discontinuous Galerkin (DG) schemes including local conservation, allowing unstructured meshes with hanging nodes, and ease of hp-adaptivity, linear systems of HDG schemes can be statically condensed such that only global DOFs on the mesh skeleton remain, resulting in increased sparsity, decreased matrix size, and computational cost [24]. One HDG technique, known as projected jumps ([46], Rem. 1.2.4), further reduces the size of the condensed HDG

Keywords and phrases. HDG, multigrid preconditioner, Crouzeix–Raviart element, reaction-diffusion, generalized Stokes equations.

Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, Notre Dame, IN, USA.
*Corresponding author: wkuang1@nd.edu

© The authors. Published by EDP Sciences, SMAI 2023

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
scheme by making the polynomial spaces of the facet unknowns one order lower than those of the primal variables without loss of accuracy. Thus superconvergence is obtained for the primal variables from the point of view of the globally coupled DOFs. The superconvergence result was rigorously proved for the primal HDG schemes with projected jumps for diffusion and Stokes problems in [55, 56], and also for the mixed HDG schemes with projected jumps for linear elasticity [59], convection-diffusion [57], and incompressible Navier–Stokes [58]. However, no numerical integration effects were considered in these works.

Large-scale simulations based on HDG schemes still face the challenge of constructing scalable and efficient solvers for the condensed linear systems, for which geometric multigrid techniques have been previously explored as either linear system solvers or preconditioners for iterative methods [27, 49]. The main difficulty in designing geometric multigrid algorithms for HDG is the construction of intergrid transfer operators between the coarse and fine meshes, since the spaces of global unknowns on hierarchical meshes live (only) on mesh skeletons and are non-nested.

In the literature, two techniques have been used to overcome this difficulty. The first approach, which is in the spirit of auxiliary space preconditioning [68], uses the conforming piecewise linear finite element method on the same mesh as the coarse grid solver and applies the standard geometric multigrid for conforming finite elements from then on. Hence, the difficulty of constructing integergrid transfer operators between global HDG facet spaces on fine and coarse grids was completely bypassed as the coarse grid HDG space was never used in the multigrid algorithm. This technique was first introduced for HDG schemes in [27] for the diffusion problem, and similarly used in [3, 21, 22, 32] for diffusion and other equations.

The second approach still keeps the HDG facet spaces on coarser meshes to construct the multigrid algorithms, which is referred to as homogeneous multigrid in [49] since it uses the same HDG discretization scheme on all mesh levels. Here the issue of stable intergrid transfer operators between coarse and fine grid HDG facet spaces has to be addressed. Several “obvious” HDG prolongation operators were numerically tested in Tan’s 2009 PhD thesis [64], which, however, failed to be optimal. The work [67] proposed an intergrid transfer operator based on Dirichlet-to-Neumann maps where operators on coarser levels were recursively changed for energy preservation. Numerical results in [67] supported the robustness of their multigrid algorithm, but no theoretical analysis was presented. More recently, a three-step procedure was used in [49] to construct a robust prolongation operator: first, define a continuous extension operator from coarse grid HDG facet space to an \( H^1 \)-conforming finite element space of the same degree on the same mesh by averaging; next, use the natural injection from the coarse grid conforming finite element space to the fine grid conforming finite element space; lastly, restrict the fine grid conforming finite elements data to the fine grid mesh skeleton to recover the fine grid HDG facet data (see more details in [49]). The optimal convergence result of the standard \( V \)-cycle algorithm was proven in [49] for an HDG scheme with stabilization parameter \( \tau = O(1/h) \) for polynomial degree \( k \geq 1 \) for the diffusion problem, where \( h \) is the mesh size. Therein, taking stabilization parameter \( \tau = O(1/h) \) is crucial in the optimal multigrid analysis in [49], which however is not usually preferred in practice as the resulting scheme loses the important property of superconvergence due to too strong stabilization. Nevertheless, the numerical results presented in [49] suggested the multigrid algorithm was still optimal when the stabilization parameter \( \tau = O(1) \), which has the superconvergence property. We note that in these cited works, the finite element spaces in the HDG schemes use the same polynomial degree for all the variables. Intergrid operators based on (superconvergent) post-processing and smoothing was also presented in [29, 30] for the hybrid high-order (HHO) methods. We finally note that the above cited works in the second approach only focus on the pure diffusion problem.

In this study, we work on HDG discretizations for the reaction-diffusion equation and the generalized Stokes equations on conforming simplicial meshes. Specifically, we construct lowest-order HDG schemes with projected jumps and numerical integration, denoted as HDG-P0, for these two sets of equations and present their optimal (superconvergent) \( a \) priori error analysis. We use the space of piecewise constants for the approximation of the (global) hybrid facet unknowns, and piecewise linears for the (local) primal unknowns, and take stabilization parameter \( \tau = O(1/h) \). Due to the use of piecewise linears for the element-wise primal unknowns and the projected jumps in the stabilization, the proposed schemes still enjoy the superconvergence property. We further
provide optimal multigrid preconditioners for the global HDG linear systems with easy-to-implement intergrid transfer operators. The key of our optimal multigrid preconditioner is the establishment of the equivalence of the condensed HDG-P0 schemes with the (slightly modified) nonconforming CR discretizations. Such equivalence enables us to build HDG multigrid preconditioners easily following the rich literature on nonconforming multigrid theory for the reaction-diffusion equation \([4, 6, 8, 14, 16, 20, 31]\) and the generalized Stokes equations \([10, 12, 60, 61, 63, 65]\). We specifically note that equivalence of the lowest order Raviart–Thomas mixed method with certain nonconforming methods was well-known in the literature \([1, 50]\), and multigrid algorithms for hybrid-mixed methods have been designed based on this fact \([11, 19]\). Moreover, an equivalence between the lowest-order primal HDG schemes with projected jumps and the nonconforming CR schemes for pure diffusion and Stokes problems have already been established by Oikawa in \([55, 56]\). Here we use the mixed HDG formulation with projected jumps and establish the stronger equivalence of the two methods for reaction-diffusion and generalized Stokes equations, which has application in time-dependent diffusion and time-dependent incompressible flow problems. As far as the low-order terms are concerned, we find the mixed HDG formulation more robust to formulate and easier to analyze compared with their primal HDG counterparts. In particular, our final algorithm methods rely on conforming and shape-regular simplicial meshes, which can not handle hanging nodes. Besides, the approach of building multigrid methods based on the equivalence to the CR discretizations is specific to HDG-P0, which can not be directly applied to higher order HDG schemes. However, it can be used as a building block when constructing robust \(hp\)-multigrid methods for higher order HDG schemes, which is our ongoing research.

The rest of the paper is organized as follows. Basic notations and the finite element spaces to be applied in HDG-P0 are introduced in Section 2. In Section 3, we propose the HDG-P0 scheme for the mixed formulation of the reaction-diffusion equation, give the optimal \(a\ priori\) error analysis, and present the optimal multigrid preconditioner for the static condensed linear system. We then focus on the HDG-P0 scheme for the generalized Stokes equations in Section 4 and present the optimal \(a\ priori\) error estimates and multigrid preconditioner. Numerical results in two- and three-dimensions are presented in Section 5 to verify the theoretical findings in Sections 3, 4, and we conclude in Section 6.

2. NOTATIONS AND PRELIMINARIES

Let \(\Omega \subset \mathbb{R}^d\), \(d \in \{2, 3\}\), be a bounded polygonal/polyhedral domain with boundary \(\partial \Omega\). We denote \(\mathcal{T}_h\) as a conforming, shape-regular, and quasi-uniform simplicial triangulation of \(\Omega\). Let \(\mathcal{E}_h\) be the collection of the facets (edges in 2D, faces in 3D) of the triangulation \(\mathcal{T}_h\), which is also referred to as the mesh skeleton. We split the mesh skeleton \(\mathcal{E}_h\) into the boundary contribution \(\mathcal{E}^b_h := \{F \in \mathcal{E}_h : F \subset \partial \Omega\}\), and the interior contribution \(\mathcal{E}^i_h := \mathcal{E}_h \setminus \mathcal{E}^b_h\). For any simplex \(K \in \mathcal{T}_h\) with boundary \(\partial K\), we denote the measure of \(K\) by \(|K|\), the \(L_2\)-inner product on \(K\) by \(\langle \cdot, \cdot \rangle_K\), and the \(L_2\)-inner product on \(\partial K\) by \(\langle \cdot, \cdot \rangle_{\partial K}\). We denote \(\mathcal{E}_K\) as the unit normal vector on the element boundaries \(\partial K\) pointing outwards. Furthermore, we denote \(h_K\) as the mesh size of \(K\) defined only on the facets \(\partial K\), whose restriction on a facet \(F \subset \partial K\) is given as \(h_K|_F := |K|/|F|\), where \(|F|\) is the measure of \(F\), and set

\[
h := \max_{K \in \mathcal{T}_h} \max_{F \subset \partial K} h_K|_F
\]

as the maximum mesh size of the triangulation \(\mathcal{T}_h\). To further simplify the notation, we define the discrete \(L_2\)-inner product on the whole mesh as \(\langle \cdot, \cdot \rangle_{\mathcal{T}_h} := \sum_{K \in \mathcal{T}_h} \langle \cdot, \cdot \rangle_K\) and the \(L_2\)-inner product on all element boundaries as \(\langle \cdot, \cdot \rangle_{\partial \mathcal{T}_h} := \sum_{K \in \mathcal{T}_h} \langle \cdot, \cdot \rangle_{\partial K}\).

As usual, we denote \(\|\cdot\|_{m,p,S}\) and \(|\cdot|_{m,p,S}\) as the norm and semi-norm of the Sobolev spaces \(W^{m,p}(S)\) for the domain \(S \subset \mathbb{R}^d\), with \(p\) omitted when \(p = 2\), and \(S\) omitted when \(S = \Omega\) is the whole domain. To this end, we define some numerical integration rules that will be used in this work. Given a simplex \(K \subset \mathbb{R}^d\), we define the following two integration rules for the volume integral \(\int_K g \, dx\):

\[
Q^0_K (g) := |K| g(m_K),
\]
\[ Q^1_K(g) := \frac{|K|}{d+1} \sum_{i=1}^{d+1} g(m^i_K), \]

where \( m_K \) is the barycenter of the element \( K \), and \( m^i_K \) is the barycenter of the \( i \)-th facet of \( K \) for \( i \in [1, d+1] \).

The one-point integration rule \( Q^0_F \) is exact for linear functions, while the \((d+1)\)-point integration rule \( Q^1_K \) is exact for linear functions when \( d = 3 \) and exact for quadratic functions when \( d = 2 \). We also define a one-point quadrature rule for the facet integral \( \int_{F} g \, ds \) over a facet \( F \subset \mathbb{R}^{d-1} \):

\[ Q^0_F(g) := |F| g(m_F), \]

where \( m_F \) is the barycenter of the facet \( F \), which is exact for linear functions. Then we denote the numerical integration on the element boundary \( \partial K \) as

\[ Q^0_{\partial K}(g) := \sum_{i=1}^{d+1} Q^i_F(g) = \sum_{i=1}^{d+1} |F^i| g(m^i_F), \]

where \( \{F^i\}_{i=1}^{d+1} \) is the collection of facets of \( K \) with \( F^i \subset \partial K \).

Moreover, we will use the following finite element spaces:

\begin{align*}
W_h &:= \{ q_h \in L_2(\Omega) : q_h|_K \in P^0(K), \forall K \in \mathcal{T}_h \}, \tag{1a} \\
W^0_h &:= \left\{ q_h \in W_h : \int_{\Omega} q_h \, dx = 0 \right\}, \tag{1b} \\
V_h &:= \{ v_h \in L_2(\Omega) : v_h|_K \in P^1(K), \forall K \in \mathcal{T}_h \}, \tag{1c} \\
V_{h,\text{CR}}^\perp &:= \{ v_h \in V_h : v_h \text{ is continuous at the barycenter } m_F \text{ of } F, \forall F \in \mathcal{E}_h^o \}, \tag{1d} \\
V_{h,\text{CR},0}^\perp &:= \{ v_h \in V_{h,\text{CR}}^\perp : v_h(m_F) = 0, \forall F \in \mathcal{E}_h^o \}, \tag{1e} \\
M_h &:= \{ \hat{v}_h \in L_2(\mathcal{E}_h^o) : \hat{v}_h|_F \in P^0(F), \forall F \in \mathcal{E}_h \}, \tag{1f} \\
M_{h,0}^\perp &:= \{ \hat{v}_h \in M_h : \hat{v}_h|_F = 0, \forall F \subset \partial \Omega \}. \tag{1g}
\end{align*}

We use underline to denote the vector version of the spaces \( \underline{\cdot} := [\cdot]^d \), and double-underline to denote the matrix version \( \underline{\cdot} := [\cdot]^{d \times d} \); e.g., \( \underline{W}_{h,0} := [W_h]^{d \times d} \) is the \( d \times d \) tensor space of piecewise constant functions.

For two positive constants \( a \) and \( b \), we denote \( a \lesssim b \) if there exists a positive constant \( C \) independent of mesh size and model parameters such that \( a \leq Cb \). We denote \( a \simeq b \) when \( a \lesssim b \) and \( b \lesssim a \).

### 3. HDG-P0 FOR THE REACTION-DIFFUSION EQUATION

#### 3.1. The model problem

Let \( f \in L_2(\Omega) \), \( \alpha \in C^1(\overline{\Omega}) \) and \( \beta \in C^0(\overline{\Omega}) \). We assume there exist positive constants \( \alpha_0 \), \( \alpha_1 \) and \( \beta_1 \) with \( \alpha_1 \lesssim \alpha_0 \) such that \( \alpha_1 \geq \alpha \geq \alpha_0 > 0 \) and \( \beta_1 \geq \beta \geq 0 \). We consider the following model problem with a homogeneous Dirichlet boundary condition:

\[ -\nabla \cdot (\alpha \nabla u) + \beta u = f \quad \text{in } \Omega, \]

\[ u = 0 \quad \text{on } \partial \Omega. \tag{2} \]

To define the HDG-P0 scheme, we shall reformulate equation (2) as the following first-order system by introducing the flux \( \underline{\sigma} := -\alpha \nabla u \) as a new variable:

\begin{align*}
\alpha^{-1} \underline{\sigma} + \nabla u &= 0 \quad \text{in } \Omega, \tag{3a} \\
\nabla \cdot \underline{\sigma} + \beta u &= f \quad \text{in } \Omega, \tag{3b} \\
u &= 0 \quad \text{on } \partial \Omega. \tag{3c}
\end{align*}
3.2. The HDG-P0 scheme

With the mesh and finite element spaces given in Section 2, we are ready to present the HDG-P0 scheme for the system (3). A defining feature of hybrid finite element schemes is the use of finite element spaces to approximate the primal variable $u$ both in the mesh $\mathcal{T}_h$ and on the mesh skeleton $\mathcal{E}_h$. We use the (discontinuous) piecewise linear finite element space $V_h$ to approximate $u$ in the mesh $\mathcal{T}_h$, and the (discontinuous) piecewise constant space $M_h^0$ to approximate $u$ on the mesh skeleton $\mathcal{E}_h$, where the homogeneous Dirichlet boundary condition has been built-in to the space $M_h^0$. For the flux variable $\sigma$, we approximate it using the vectorial piecewise constant function space $W_h$. Using these finite element spaces and the numerical integration rules introduced in Section 2, the weak formulation of the HDG-P0 scheme is now given as follows: find $(\sigma_h, u_h, \widehat{u}_h) \in W_h \times V_h \times M_h^0$ such that:

$$\sum_{K \in \mathcal{T}_h} \left( Q_K^0 (\alpha_h^{-1} \sigma_h \cdot \nu_K) + Q_K^0 (\widehat{u}_h \nu_h \cdot \nu_K) \right) = 0,$$

$$\sum_{K \in \mathcal{T}_h} \left( Q_K^0 (\tau_K (u_h - \widehat{u}_h) v_h) + Q_K^0 (\beta u_h v_h) \right) = \sum_{K \in \mathcal{T}_h} Q_K^1 (f v_h),$$

$$\sum_{K \in \mathcal{T}_h} \left( Q_K^0 (\sigma_h \cdot \nu_K \widehat{v}_h) + Q_K^0 (\tau_K (u_h - \widehat{u}_h) \widehat{v}_h) \right) = 0,$$

for all $(\nu_h, v_h, \widehat{v}_h) \in W_h \times V_h \times M_h^0$, where $\alpha_h^{-1} \in W_h$ is the $L_2$-projection of $\alpha^{-1}$ onto the piecewise constant space $W_h$, and $\tau_K := \frac{\beta}{h_K}$ is the stabilization parameter, with $\alpha_h := (\alpha_h^{-1})^{-1}$. Taking test functions $\nu_h = \sigma_h$, $v_h = u_h$, and $\widehat{v}_h = -\widehat{u}_h$ in (4), and adding, we have the following energy identity:

$$\sum_{K \in \mathcal{T}_h} \left( Q_K^0 (\alpha_h^{-1} |\sigma_h|^2) + Q_K^0 (\tau_K (u_h - \widehat{u}_h)^2) + Q_K^1 (\beta u_h^2) \right) = \sum_{K \in \mathcal{T}_h} Q_K^1 (f u_h).$$

We note that the quadratures used in (4) guarantees the scheme is locally mass-conservative with the numerical flux defined as

$$\widehat{\sigma}_h \cdot \nu_K := \sigma_h \cdot \nu_K + \tau_K \Pi_0 (u_h - \widehat{u}_h),$$

where $\Pi_0$ is the $L_2$-projection onto the skeleton space $M_h$, since the equation (4c) is equivalent to

$$\langle \widehat{\sigma}_h \cdot \nu_K, \nu_h \rangle_{\partial \mathcal{T}_h} = 0, \quad \forall \nu_h \in M_h^0.$$

See more discussion in Lemma 3.2 below.

3.3. An a priori error analysis for HDG-P0

To simplify the a priori error analysis in this subsection, we assume the exact solution $(\sigma, u) \in H^1(\Omega) \times (H^2(\Omega) \cap H_0^1(\Omega))$ for the model problem (3), together with the following full elliptic regularity result:

$$\alpha^{-1/2}_0 \|\sigma\|_1 + \left( \alpha_1^{1/2} + \beta_1^{1/2} \right) \|u\|_2 \lesssim c_{\text{reg}} \|f\|_0,$$

which holds when the domain $\Omega$ is convex.

To perform an a priori error analysis of the scheme (4) with numerical integration, we follow the standard convention [23] by assuming the following stronger regularity for $\beta$ and $f$:

$$\beta \in W^{1, \infty} (\bar{\Omega}), \quad f \in W^{1, \infty} (\bar{\Omega}).$$

We use the following approximation results of the quadrature rule $Q_K^1$, which is adapted from Ciarlet ([23], Thms. 4.1.4 and 4.1.5).
Lemma 3.1. For any \( K \in \mathcal{T}_h \), \( f, u \in W^{1,\infty}(K) \), and \( v \in \mathcal{P}^1(K) \), we have:

\[
\begin{align*}
|\langle f, v \rangle_K - Q_K^1(fv)| & \lesssim h^{1+d/2} \|f\|_{1,\infty,K} \|v\|_{1,K}, \\
|\langle fu, v \rangle_K - Q_K^1(fuv)| & \lesssim h \|f\|_{1,\infty,K} \|u\|_{1,K} \|v\|_{0,K}.
\end{align*}
\]

(6a) (6b)

We will compare the solution to (4) with the solution to a similar scheme with exact integration. To this end, let \((\overline{\sigma}_h^1, u_h^1, \hat{u}_h^1) \in W_h \times V_h \times M_h^0\) be the solution to the following system:

\[
\begin{align*}
\sum_{K \in \mathcal{T}_h} (Q_K^0(\alpha_h^{-1}\overline{\sigma}_h^1, \cdot) \cdot \tau_h) + Q_K^0(\hat{u}_h^1, \nabla \cdot \tau_h) &= 0, \\
\sum_{K \in \mathcal{T}_h} (Q_K^0(\tau_K(u_h^1 - \hat{u}_h^1), v_h) + (\beta u_h^1, v_h)_K) &= \sum_{K \in \mathcal{T}_h} (f, v_h)_K, \\
\sum_{K \in \mathcal{T}_h} (Q_K^0(\overline{\sigma}_h^1, \cdot) \cdot \hat{v}_h) + Q_K^0(\tau_K(u_h^1 - \hat{u}_h^1), \hat{v}_h) &= 0,
\end{align*}
\]

for all \((\tau_h, v_h, \hat{v}_h) \in W_h \times V_h \times M_h^0\). Note that the only difference between HDG-P0 (4) and the above system (7) is in equations (4b) and (7b), where the former uses \(Q_K^1\) as the volumetric numerical integration rule and the latter uses exact volumetric integration. The following result shows that the scheme (7) is precisely the lowest-order (mixed) HDG scheme with projected jumps (and exact integration) analyzed in [57, 59] for linear elasticity and convection-diffusion problems, hence its \textit{a priori} error estimates can be readily adapted from [57, 59].

Lemma 3.2. Let \((\overline{\sigma}_h^1, u_h^1, \hat{u}_h^1) \in W_h \times V_h \times M_h^0\) be the solution to (7), then \((\overline{\sigma}_h^1, u_h^1, \hat{u}_h^1)\) satisfies

\[
\begin{align*}
(\alpha^{-1}\overline{\sigma}_h^1, \tau_h)_K - (u_h^1, \nabla \cdot \tau_h)_K + \langle \hat{u}_h^1, \nabla \cdot \tau_h \rangle_{\partial K} &= 0, \\
(\nabla \cdot \overline{\sigma}_h^1, v_h)_K + \langle \tau_K \Pi_0(u_h^1 - \hat{u}_h^1), v_h \rangle_{\partial K} + (\beta u_h^1, v_h)_K &= (f, v_h)_K, \\
\langle \overline{\sigma}_h^1, \cdot \rangle_{\partial K} + \tau_K \Pi_0(u_h^1 - \hat{u}_h^1), \hat{v}_h \rangle_{\partial K} &= 0,
\end{align*}
\]

for all \((\tau_h, v_h, \hat{v}_h) \in W_h \times V_h \times M_h^0\), where \(\Pi_0\) is the \(L_2\)-projection onto piecewise constants on the mesh skeleton.

Proof. Using the fact that \(\overline{\sigma}_h^1, \tau_h \in W_h\) are constant functions on each element \(K\), we have

\[
(\alpha^{-1}\overline{\sigma}_h^1, \tau_h)_K = \overline{\sigma}_h^1(m_K^0) \cdot \tau_h(m_K^0) \int_K \alpha^{-1} \, dx = \overline{\sigma}_h^1(m_K^0) \cdot \tau_h(m_K^0) \alpha_h^{-1} |K| = Q_K^0(\alpha^{-1}\overline{\sigma}_h^1, \tau_h).
\]

Similarly, there holds

\[
\langle \hat{u}_h^1, \tau_h \cdot \hat{v}_h \rangle_{\partial K} = Q_K^0(\hat{u}_h^1, \tau_h \cdot \hat{v}_h).
\]

Using the definition of the projection \(\Pi_0\) and the fact that the restriction of \(\tau_K\) on each facet \(F^i \subset \partial K\) is a constant (denoted as \(\tau_K^i\)), we have

\[
\langle \tau_K \Pi_0(u_h^1 - \hat{u}_h^1), v_h \rangle_{\partial K} = \sum_{i=1}^{d+1} \tau_K^i |F^i| (u_h^1(m_{F^i}) - \hat{u}_h^1(m_{F^i})) v_h(m_{F^i}) = Q_K^0(\tau_K(u_h^1 - \hat{u}_h^1), v_h).
\]

Combining these identities with the fact that \(\nabla \cdot \tau_h = \nabla \cdot \overline{\sigma}_h^1 \equiv 0\), we conclude that the systems (7) and (8) are exactly the same. \(\square\)

We have the following results on the \textit{a priori} error estimates of the scheme (7), whose proof is omitted for simplicity. We refer to the works [57, 59] for a similar analysis.
Corollary 3.1. Let \((\sigma, u)\) be the exact solution to the model problem (3). Let \((\sigma_h, u_h, \tilde{u}_h) \in \mathcal{W}_h \times V_h \times M_h^0\) be the solution to (7). Then there holds
\[
\left\| \alpha^{-1/2} (\sigma - \sigma_h) \right\|_0 + \left\| \beta^{1/2} (u - u_h) \right\|_0 + \left\| \tau_K^{1/2} \Pi_0 (u_h - \tilde{u}_h) \right\|_0,_{\partial \Gamma_h} \lesssim h \Theta, \tag{9a}
\]
\[
\left\| u - u_h \right\|_0 \lesssim \left\| \nabla (u - u_h) \right\|_0 + \left\| h_K^{-1/2} (u_h - \tilde{u}_h) \right\|_0,_{\partial \Gamma_h} \lesssim \alpha_0^{-1/2} h \Theta, \tag{9b}
\]
where
\[
\Theta := \alpha_0^{-1/2} |\sigma|_1 + \left( \alpha_1^{1/2} + \beta_1^{1/2} \right) |u|_2, \tag{9c}
\]
and \(\| \phi \|_{0, \partial \Gamma_h} := \sqrt{\langle \phi, \phi \rangle_{\partial \Gamma_h}}\). Moreover, with the full elliptic regularity assumed in (5), we have the following optimal \(L_2\)-convergence for \(u_h^1\):
\[
\left\| u - u_h^1 \right\|_0 \lesssim c_{\text{reg}} h^2 \Theta. \tag{9d}
\]

Combining the results in this subsection, we are now ready to give the error estimates for the solution to the HDG-P0 scheme (4).

Theorem 3.1. Let \((\sigma, u)\) be the exact solution to the model problem (3). Let \((\sigma_h, u_h, \tilde{u}_h) \in \mathcal{W}_h \times V_h \times M_h^0\) be the solution to (4). Then there holds
\[
\left\| \alpha^{-1/2} (\sigma - \sigma_h) \right\|_0 + \left\| \tau_K^{1/2} \Pi_0 (u_h - \tilde{u}_h) \right\|_0,_{\partial \Gamma_h} \lesssim h (\Theta + \Xi), \tag{10a}
\]
\[
\left\| u - u_h \right\|_0 \lesssim \left\| \nabla (u - u_h) \right\|_0 + \left\| h_K^{-1/2} (u_h - \tilde{u}_h) \right\|_0,_{\partial \Gamma_h} \lesssim \alpha_0^{-1/2} h (\Theta + \Xi), \tag{10b}
\]
where \(\Theta\) is given in (9c) and
\[
\Xi := \alpha_0^{-1/2} \left\| \beta \right\|_{1, \infty} \left( \left\| u \right\|_0 + \alpha_0^{-1/2} h \Theta \right) + \alpha_0^{-1/2} h^{d/2} \left\| f \right\|_{1, \infty}. \tag{10c}
\]

Moreover, assuming full elliptic regularity (5), we have the following optimal \(L_2\)-convergence for \(u_h\):
\[
\left\| u - u_h \right\|_0 \lesssim h^2 \Psi, \tag{10c}
\]
where
\[
\Psi := c_{\text{reg}} (\Theta + \Xi) + \frac{c_{\text{reg}} \alpha_0^{-1/2} \beta_1}{\alpha_1^{1/2} + \beta_1^{1/2} h} \Xi + \left\| \beta \right\|_{1, \infty} \alpha_0^{-3/2} (\Theta + \Xi) + h^{d/2 - 1} \alpha_0^{-1} \left\| f - \beta u \right\|_{1, \infty}. \tag{10c}
\]

Proof. The proof of (10a) and (10b) follows from a standard energy argument, while that of (10c) follows from a duality argument. Here we prove the energy estimates (10a), (10b), and postpone the (slightly more technical) proof of (10c) to the Appendix below.

We denote the difference of the solution \((\sigma_h, u_h, \tilde{u}_h)\) to the system (4) and the solution \((\sigma_h^1, u_h^1, \tilde{u}_h^1)\) to the system (7) as
\[
\varepsilon_\sigma := \sigma_h - \sigma_h^1, \quad e_u := u_h - u_h^1, \quad e_\tilde{u} := \tilde{u}_h - \tilde{u}_h^1.
\]

Subtracting equations (4) from (7) and using Lemma 3.2, we get the following error equation:
\[
(\alpha^{-1} \varepsilon_\sigma, \varepsilon_\sigma)_{\Gamma_h} - (e_u, \nabla \cdot \varepsilon_h)_{\Gamma_h} + (e_u, \varepsilon_h \cdot \Pi_K)_{\partial \Gamma_h} = 0, \tag{11a}
\]
\[
(\nabla \cdot \varepsilon_\sigma, v_h)_{\Gamma_h} + (\tau_K \Pi_0 (e_u - e_\tilde{u}), v_h)_{\partial \Gamma_h} + \sum_{K \in \mathcal{T}_h} \frac{Q_K}{\beta_e} (\beta e_u v_h) = T_1(v_h) - T_2(v_h), \tag{11b}
\]
for all \((\xi_h, v_h, \widehat{v}_h) \in \mathcal{W}_h \times V_h \times M_0^h\) where

\[
T_1(v_h) := (\beta u_h^1, v_h)_{\mathcal{S}_h} - \sum_{K \in \mathcal{S}_h} Q_K^1(\beta u_h^1 v_h),
\]
\[
T_2(v_h) := (f, v_h)_{\mathcal{S}_h} - \sum_{K \in \mathcal{S}_h} Q_K^1(f v_h).
\]

Using Lemma 3.1, we have

\[
T_1(v_h) - T_2(v_h) \lesssim \sum_{K \in \mathcal{S}_h} \left( h \| \beta \|_{1,\infty,K} \| u_h^1 \|_{0,K} + h^{1+d/2} \| f \|_{1,\infty,K} \right) \| v_h \|_{1,K}.
\]

Taking test function \(r_h := \nabla e_u\) in equation (11), by integration by parts, reordering terms, and applying the Cauchy–Schwarz inequality and inverse inequality, we get

\[
(\nabla e_u, \nabla e_u)_{\mathcal{S}_h} = -\left( \alpha^{-1} \xi_h, \nabla e_u \right) + \left( \Pi_0(e_u - e_\widehat{u}), \nabla e_u \right)_{\partial \mathcal{S}_h}
\]
\[
\leq \alpha_0^{-1/2} \left( \| \alpha^{-1/2} \xi_h \|_0 \| \nabla e_u \|_0 + h^{1/2} \| \tau_{K}^{1/2} \Pi_0(e_u - e_\widehat{u}) \|_{0,\partial \mathcal{S}_h} \right)
\]
\[
\lesssim \alpha_0^{-1/2} \left( \| \alpha^{-1/2} \xi_h \|_0 + \| \tau_{K}^{1/2} \Pi_0(e_u - e_\widehat{u}) \|_{0,\partial \mathcal{S}_h} \right) \| \nabla e_u \|_0.
\]

Hence,

\[
\| \nabla e_u \|_0 \lesssim \alpha_0^{-1/2} \left( \| \alpha^{-1/2} \xi_h \|_0 + \| \tau_{K}^{1/2} \Pi_0(e_u - e_\widehat{u}) \|_{0,\partial \mathcal{S}_h} \right).
\]

Moreover, by the definition of \(\Pi_0\), the trace-inverse inequality and Poincaré inequality, we have

\[
\| h_K^{-1/2} (e_u - e_\widehat{u}) \|_{0,\partial \mathcal{S}_h} \leq \left( \| h_K^{-1/2} (e_u - \Pi_0 e_u) \|_{0,\partial \mathcal{S}_h} + \| h_K^{-1/2} \Pi_0 (e_u - e_\widehat{u}) \|_{0,\partial \mathcal{S}_h} \right)
\]
\[
\leq \| h_K^{-1/2} (e_u - P_0 e_u) \|_{0,\partial \mathcal{S}_h} + \| h_K^{-1/2} \Pi_0 (e_u - e_\widehat{u}) \|_{0,\partial \mathcal{S}_h}
\]
\[
\lesssim \| \nabla e_u \|_0 + \| h_K^{-1/2} \Pi_0 (e_u - e_\widehat{u}) \|_{0,\partial \mathcal{S}_h}.
\]
\[
\lesssim \alpha_0^{-1/2} \left( \| \alpha^{-1/2} \xi_h \|_0 + \| \tau_{K}^{1/2} \Pi_0 (e_u - e_\widehat{u}) \|_{0,\partial \mathcal{S}_h} \right),
\]

where \(P_0(e_u)\) is the \(L_2\)-projection of \(e_u\) onto the piecewise constant space \(W_h\). Combining the above two inequalities with the discrete Poincaré inequality for piecewise \(H^1\) functions ([15], Rem. 1.1), we get

\[
\| e_u \|_0 \lesssim \| \nabla e_u \|_0 + \| h_K^{-1/2} (e_u - e_\widehat{u}) \|_{0,\partial \mathcal{S}_h} \lesssim \alpha_0^{-1/2} \left( \| \alpha^{-1/2} \xi_h \|_0 + \| \tau_{K}^{1/2} \Pi_0 (e_u - e_\widehat{u}) \|_{0,\partial \mathcal{S}_h} \right).
\]

Now taking test functions \(r_h = \xi_h, v_h = e_u, \widehat{v}_h = -e_\widehat{u}\) in (11) and adding, we get

\[
\| \alpha^{-1/2} \xi_h \|_0^2 + \| \tau_{K}^{1/2} \Pi_0 (e_u - e_\widehat{u}) \|_{0,\partial \mathcal{S}_h}^2 + \| \beta^{1/2} e_u \|_{0,h} = T_1(e_u) - T_2(e_u)
\]
\[
\lesssim h \left( \alpha_0^{-1/2} \| \beta \|_{1,\infty} \| u_h^1 \|_0 + \alpha_0^{-1/2} h^{d/2} \| f \|_{1,\infty} \right) \left( \| \alpha^{-1/2} \xi_h \|_0 + \| \tau_{K}^{1/2} \Pi_0 (e_u - e_\widehat{u}) \|_{0,\partial \mathcal{S}_h} \right),
\]

where \(|\phi|_{0,h} := \sqrt{\sum_{K \in \mathcal{S}_h} Q_K^1(\phi^2)}\). Combining the above inequality with the triangle inequality

\[
\| u_h^1 \|_0 \leq \| u - u_h^1 \|_0 + \| u \|_0,
\]

and the estimates (9a) and (9b) we get the energy estimate (10a). The estimate (10b) is a direct consequence of (13), (9b), and (10a). \(\square\)
3.4. Equivalence to a nonconforming discretization

We define an interpolation operator $\Pi_h^{\text{CR}} : M_h \rightarrow V_h^{\text{CR}}$ from the piecewise constant facet space $M_h$ to the nonconforming (piecewise linear) CR space $V_h^{\text{CR}}$:

$$
\Pi_h^{\text{CR}}v_h(m_F) = \tilde{v}_h(m_F), \quad \forall v_h \in M_h, F \in \mathcal{E}_h.
$$

(14)

It is clear that $\Pi_h^{\text{CR}}$ is an isomorphic map between these two spaces, which share the same set of DOFs (one DOF per facet). The following result builds a link between the HDG-P0 scheme (4) and a (slightly modified) nonconforming CR discretization.

**Theorem 3.2.** Let $u_h^{\text{CR}} \in V_h^{\text{CR},0}$ be the solution to the following nonconforming scheme

$$
\sum_{K \in \mathcal{T}_h} \left( Q_K^0 (\alpha_h \nabla u_h^{\text{CR}} \cdot \nabla v_h^{\text{CR}}) + Q_K^1 (\gamma_h \beta v_h^{\text{CR}} u_h^{\text{CR}}) \right) = \sum_{K \in \mathcal{T}_h} Q_K^1 (\gamma_f v_h^{\text{CR}}), \quad \forall v_h^{\text{CR}} \in V_h^{\text{CR},0},
$$

(15)

where $\gamma_h := -\alpha_h \nabla \Pi_h^{\text{CR}} \tilde{u}_h$.

Then the solution $(\sigma_h, u_h, \tilde{u}_h)$ to the HDG-P0 scheme (4) satisfies

$$
\sigma_h = -\alpha_h \nabla \Pi_h^{\text{CR}} \tilde{u}_h,
$$

(16a)

$$
u_h^i(m_k) = \gamma_h(m_k) \left( \tilde{u}_h(m_k) + \frac{(h_k^i)^2}{(d+1) \alpha_h} f(m_k^i) \right), \quad \forall i \in \{1, \ldots, d+1\}, \forall K \in \mathcal{T}_h,
$$

(16b)

$$
\Pi_h^{\text{CR}} \tilde{u}_h = u_h^{\text{CR}},
$$

(16c)

where recall that $m_k^i$ is the barycenter of the $i$-th facet $F_i$ of the element $K$, and $h_k^i = |K| |F_i|^{-1}$.

**Proof.** By (4a) and the definition of $\Pi_h^{\text{CR}}$ in (14), we have, for any $\Sigma_h \in \mathcal{P}^0(K)$,

$$
Q_K^0 (\alpha_h^{-1} \Sigma_h \cdot \Sigma_h) = -Q_K^0 (\Pi_h^{\text{CR}} \tilde{u}_h \Sigma_h \cdot \Sigma_h)\K = -Q_K^0 (\Pi_h^{\text{CR}} \tilde{u}_h, \Sigma_h \cdot \Sigma_h)\K = -\left( \nabla \Pi_h^{\text{CR}} \tilde{u}_h, \Sigma_h \right) = -\left( \nabla \Pi_h^{\text{CR}} \tilde{u}_h, \Sigma_h \right)_{\in \mathcal{P}^0(K)}.
$$

Hence,

$$
(\alpha_h^{-1} \Sigma_h + \nabla \Pi_h^{\text{CR}} \tilde{u}_h, \Sigma_h)_{\mathcal{T}_h} = 0,
$$

which implies the identity (16a) thanks to the fact that $\alpha_h |K| \mathcal{P}^0(K)$ and $\Sigma_h, \nabla \Pi_h^{\text{CR}} \tilde{u}_h \in \mathcal{P}^0(K)^d$.

Next, for $i \in \{1, \ldots, d+1\}$, taking test function $v_h$ in (4b) to be supported only on an element $K$, whose values is 1 on $m_k^i$ and zero on $m_k^j$ for $j \neq i$, we get

$$
\tau_K^i (u_h^i(m_k^i) - \tilde{u}_h(m_k^i)) |F_i| + \frac{|K|}{d+1} \beta (m_k^i) u_h(m_k^i) = \frac{|K|}{d+1} \beta (m_k^i) u_h(m_k^i).
$$

Solving for the value $u_h^i(m_k^i)$ and using the definition of $\tau_K$ and $h_K$, i.e., $\tau_K^i = \alpha_h / h_K$ and $h_K = |K| |F_i|^{-1}$, we immediately get the equality (16b).

Finally, let us prove the identity (16c). By the definition of $\Pi_h^{\text{CR}}$ and (16a), (16b), we have

$$
Q_{\partial K}^0 (\sigma_h \cdot n_K \tilde{v}_h) = \left( \sigma_h \cdot n_K, \Pi_h^{\text{CR}} \tilde{v}_h \right)_{\partial K} = (\sigma_h, \nabla \Pi_h^{\text{CR}} \tilde{v}_h)_{K} = -Q_K^0 (\alpha_h \nabla \Pi_h^{\text{CR}} \tilde{u}_h \nabla \Pi_h^{\text{CR}} \tilde{v}_h)_{K},
$$

and

$$
Q_{\partial K}^0 (\tau_K (u_h - \tilde{u}_h) \tilde{v}_h) = \sum_{i=1}^{d+1} |F_i| \tau_K^i (u_h(m_k^i) - \tilde{u}_h(m_k^i)) \tilde{v}_h(m_k^i)
$$
first locally eliminate \( \sigma \) and \( \tau \). The sense that the equality (16c) always holds. \( \tau(4) \) is always equivalent to the original CR discretization for any choice of non-zero stabilization parameter \( \gamma_h \). The asymptotic limit as the mesh size \( h \) approaches zero. Moreover, for the pure diffusion case with \( \beta = 0 \), the scheme (4) is always equivalent to the original CR discretization for any choice of non-zero stabilization parameter \( \tau \) in the sense that the equality (16c) always holds.

\[ \sum_{K \in \mathcal{T}_h} \left( Q^1_K (\alpha_h \nabla \Pi^\text{CR}_h \hat{u}_h \cdot \nabla \Pi^\text{CR}_h \hat{v}_h) + Q^1_K (\gamma_h \beta \Pi^\text{CR}_h \hat{u}_h \Pi^\text{CR}_h \hat{v}_h) \right) = \sum_{K \in \mathcal{T}_h} Q^1_K (\gamma h \beta \Pi^\text{CR}_h \hat{u}_h \Pi^\text{CR}_h \hat{v}_h), \]  

for all \( \hat{v}_h \in M^0_h \), which implies the equivalence (16c).

\[ \square \]

**Remark 3.1** (Equivalence between HDG-P0 and CR discretizations). Theorem 3.2 implies the equivalence between the HDG-P0 scheme (4) and the slightly modified CR discretization (15) with numerical integration, where the modification is the introduction of the scaling parameter \( \gamma_h \). This result is motivated by the equivalence of the lowest order Raviart–Thomas mixed method and the nonconforming CR method [1, 50]. Note that, when \( \beta \neq 0 \), the HDG-P0 scheme (4) converges to the original CR discretization (with \( \gamma_h \) set to 1 in (15)) in the asymptotic limit as the mesh size \( h \) approaches zero. Moreover, for the pure diffusion case with \( \beta = 0 \), the scheme (4) is always equivalent to the original CR discretization for any choice of non-zero stabilization parameter \( \tau \) in the sense that the equality (16c) always holds.

**Remark 3.2** (On HDG-P0 solution procedure). In the practical implementation of the HDG-P0 scheme, we first locally eliminate \( \sigma_h \) and \( u_h \) from (4) to arrive at the global (condensed) linear system (17) for \( \hat{u}_h \). After solving for \( \hat{u}_h \) in the system (17), we then recover \( \sigma_h \) and \( u_h \). The next subsection focuses on the efficient linear system solver for (17) via geometric multigrid. Due to the algebraic equivalence between the condensed HDG system (17) and the (modified) nonconforming system (15), we can simply use the rich multigrid theory for nonconforming methods [4, 6, 8, 14, 16, 20, 31] to precondition the condensed HDG system (17).

### 3.5. Multigrid algorithm

In this subsection, we present the detailed multigrid algorithm for the condensed system (17), using the nonconforming multigrid theory of Brenner [16].

We consider a set of hierarchical meshes for the multigrid algorithm. Let \( \mathcal{T}_l \) be a conforming simplicial triangulation of \( \Omega \) and let \( \mathcal{T}_l \) be obtained by successive mesh refinements for \( l = 2, \ldots, J \), with the final mesh \( \mathcal{T}_J = \mathcal{T}_h \). We denote \( h_l \) as the maximum mesh size of the triangulation \( \mathcal{T}_l \). We assume the triangulation \( \mathcal{T}_l \) is conforming, shape-regular, and quasi-uniform on each level, and the difference of the mesh size between two adjacent mesh levels is bounded, i.e., \( h_l \lesssim h_{l+1} \). Let \( \mathcal{E}_l \) be the mesh skeleton of \( \mathcal{T}_l \). Denote \( W_l, V_l, V_l^\text{CR} \) as the corresponding finite element spaces on the \( l \)-th level mesh \( \mathcal{T}_l \), and \( M_l \) as the corresponding piecewise constant finite element space on the \( l \)-th level mesh skeleton \( \mathcal{E}_l \). We define the following \( L^2 \)-like inner product on the space \( M_l \):

\[ (\hat{u}_l, \hat{v}_l)_{0,l} := \sum_{K \in \mathcal{T}_l} Q^1_K (\Pi^\text{CR}_l \hat{u}_l \Pi^\text{CR}_l \hat{v}_l) = \sum_{K \in \mathcal{T}_l} \sum_{i=1}^{d+1} \frac{|K|}{d+1} \hat{u}_l(m^i_K) \hat{v}_l(m^i_K), \]

with its induced norm as \( \| \cdot \|_{0,l} \), where \( \Pi^\text{CR}_l \) is the interpolation operator (14) from \( M_l \) to \( V_l^\text{CR} \). We define \( A_l : M^0_l \rightarrow M^0_l \) as the linear operator satisfying

\[ (A_l \hat{u}_l, \hat{v}_l)_{0,l} := a_l(\hat{u}_l, \hat{v}_l), \quad \forall \hat{u}_l, \hat{v}_l \in M^0_l, \]

where \( a_l \) is the following bilinear form on the \( l \)-th level mesh:

\[ a_l(\hat{u}_l, \hat{v}_l) := \sum_{K \in \mathcal{T}_l} \left( Q^1_K \left( \alpha_l \nabla \Pi^\text{CR}_l \hat{u}_l \cdot \nabla \Pi^\text{CR}_l \hat{v}_l \right) + Q^1_K \left( \beta_l \Pi^\text{CR}_l \hat{u}_l \Pi^\text{CR}_l \hat{v}_l \right) \right), \quad \forall \hat{u}_l, \hat{v}_l \in M^0_l, \]  

(19)
where \( \alpha_t = (\alpha_t^{-1})^{-1} \) with \( \alpha_t^{-1} \in W_t \) being the \( L_2 \)-projection of \( \alpha_t^{-1} \) onto the piecewise constant space \( W_t \), and \( \beta_t \in V_h \) satisfies

\[
\beta_t(m^t_K) = \gamma_t(m^t_K)\beta(m^t_K), \quad \forall i \in \{1, \ldots, d + 1\}, \ \forall K \in \mathcal{T}_t,
\]

with \( \gamma_t := \frac{\alpha_t}{\alpha_t + h_{K_i}} \) and \( h_{K_i} \) being the mesh size of elements in \( \mathcal{T}_t \). Further denoting \( f_t \in M^0_t \) such that

\[
(f_t, \tilde{v}_t)_{0,t} = \sum_{K \in \mathcal{T}_t} Q^1_K(\gamma_t f_{\Pi_t^{CR}}), \quad \forall \tilde{v}_t \in M^0_t,
\]

the operator form of the linear system \( (17) \) is simply to find \( \tilde{u}_h \in M^0_J \) such that

\[
A_J \tilde{u}_h = f_J. \tag{20}
\]

A multigrid algorithm for the above system \( (20) \) needs two ingredients: an intergrid transfer \( I_{l-1}^t : M^0_{l-1} \rightarrow M^0_l \) operator that connects the spaces \( M^0_{l-1} \) and \( M^0_l \) on two consecutive mesh levels, and a smoothing operator \( R_t : M^0_l \rightarrow M^0_l \) that takes care of high-frequency errors. For the intergrid transfer operator, we use the following well-known nonconforming averaging operator \([4, 8]\):

\[
(I_{l-1}^t v)_{0,l-1} = \begin{cases} 
  (\Pi_{l-1}^{CR} v_{l-1})(m_F), & \text{if } F \in \mathcal{E}_{l-1}^0 \text{ lies inside } \mathcal{T}_{l-1}, \\
  \frac{1}{2} \left( (\Pi_{l-1}^{CR} v_{l-1})^+(m_F) + (\Pi_{l-1}^{CR} v_{l-1})^-(m_F) \right), & \text{if } F \in \mathcal{E}_{l-1}^0 \text{ lies on } \mathcal{E}_{l-1}^a,
\end{cases} \tag{21}
\]

where \((\Pi_{l-1}^{CR} v_{l-1})^+\) and \((\Pi_{l-1}^{CR} v_{l-1})^-\) are the values of \( \Pi_{l-1}^{CR} v_{l-1} \) on two adjacent elements \( K^+, K^- \in \mathcal{T}_{l-1} \) that share the facet \( F \). We further denote the restriction operator \( I_{l-1}^t : M^0_l \rightarrow M^0_{l-1} \) as the transpose of \( I_{l-1}^t \) with respect to the inner product \( (\cdot, \cdot)_{0,l} \):

\[
(I_{l-1}^{i-1} v, \check{w})_{0,l-1} = (\check{v}, I_{l-1}^t \check{w})_{0,l}, \quad \forall \check{v} \in M^0_l, \check{w} \in M^0_{l-1}.
\]

For the smoothing operator \( R_t \), we simply take it to be the classical Jacobi or Gauss–Seidel smoother for the operator \( A_l \). The proof of multigrid convergence requires another operator \( P_{l-1}^t : M^0_{l-1} \rightarrow M^0_{l-1} \), which is the transpose of \( I_{l-1}^t \) with respect to the bilinear form \( a_l \), i.e.,

\[
a_{l-1}(P_l^{-1} u, v_{l-1}) = a_l(\hat{u}_l, I_{l-1}^t v_{l-1}), \quad \forall \hat{u}_l \in M^0_l, v_{l-1} \in M^0_{l-1}.
\]

The operator \( P^{l-1}_l \) is for multigrid analysis only, and never enters the actual multigrid algorithm.

We now follow ([16], Algorithm 2.1) to present the classical symmetric \( V \)-cycle algorithm for the system \( A_l \hat{u}_l = f_l \in M^0_l \).
Algorithm 1. The V-cycle algorithm for $A_l \tilde{u}_l = f_l$.

The $l$-th level symmetric V-cycle algorithm produces $\text{MG}_{A_l}(l, f_l, \tilde{u}^0_l, m)$ as an approximation solution for $A_l \tilde{u}_l = f_l$ with initial guess $\tilde{u}^0_l$, where $m$ denotes the number of pre-smoothing and post-smoothing steps.

if $l = 1$ then
\[ \text{MG}_{A_l}(l, \tilde{u}_l, f_l) = (A_l)^{-1} f_l. \]
else

Perform the following three steps:

1. **pre-smoothing.** For $j = 1, \ldots, m$, compute $\tilde{u}^j_l$ by
\[ \tilde{u}^j_l = \tilde{u}^{j-1}_l + R_l \left( f_l - A_l \tilde{u}^{j-1}_l \right). \]

2. **Coarse grid correction.** Let $\hat{r}_{i-1} = I_{i-1}^l (f_i - A_l \tilde{u}^m_i)$ and compute $\tilde{u}^{m+1}_l$ by
\[ \tilde{u}^{m+1}_l = \tilde{u}^m_l + I_{i-1}^l \text{MG}_{A_{l-1}}(l-1, \hat{r}_{i-1}, 0, m). \]

3. **Post-smoothing.** For $j = m + 2, \ldots, 2m + 1$, compute $\tilde{u}^j_l$ by
\[ \tilde{u}^j_l = \tilde{u}^{j-1}_l + R_l^T \left( f_l - A_l \tilde{u}^{j-1}_l \right), \]

where $R_l^T$ is the transpose of $R_l$ with respect to the inner product $\langle \cdot, \cdot \rangle_{0,l}$.

We then define $\text{MG}_{A_l}(l, f_l, \tilde{u}^0_l, m) = \tilde{u}^{2m+1}_l$.

---

By the algebraic equivalence of the nonconforming system (15) and the condensed HDG system (17), we immediately have the following optimality result of the above multigrid algorithm from [16]. We note that both the equivalence result in Theorem 3.2 and the optimal multigrid theory in [16] do not require the full elliptic regularity assumption (5). Thus our proposed multigrid algorithm for the HDG-P0 scheme is optimal in the low-regularity case, where $u \in H^{1+s}(\Omega) \cap H^s_0(\Omega)$ and $s \in (\frac{1}{2}, 1]$ is the regularity constant for the model problem (3).

**Theorem 3.3** ([16], Thm. 5.2). There exist positive mesh-independent constants $C$ and $m_*$ such that
\[ \| E_{l,m} \hat{v} \|_{1,l} \leq \frac{C}{m^s} \| \hat{v} \|_{1,l}, \quad \forall \hat{v} \in M^0_l, l \geq 1, m \geq m*, \]
where $s \in (\frac{1}{2}, 1]$ is the regularity constant such that the solution $u$ to (2) satisfies
\[ \| u \|_{1+s} \lesssim \| f \|_{-1+s}, \]
$\| \cdot \|_{1,l}$ is the mesh-dependent norm induced by the linear operator $A_l$ (18), i.e., $\| \hat{v} \|_{1,l} := \sqrt{(A_l \hat{v}, \hat{v})_{0,l}}$, and $E_{l,m} : M^0_l \rightarrow M^0_l$ is the operator relating the initial error and the final error of the multigrid V-cycle algorithm, i.e.,
\[ E_{l,m}(\tilde{u}_l - \tilde{u}^0_l) := \tilde{u}_l - \text{MG}_{A_l}(l, f_l, \tilde{u}^0_l, m). \]

**Proof.** This multigrid algorithm is algebraically equivalent to the V-cycle multigrid algorithm for the nonconforming CR scheme (15) with diffusion coefficient $\alpha_l \in W_0$ and reaction coefficient $\beta_l \in V_l$ in each level. It is easy to verify that the assumptions (3.6)–(3.12) of [16] are satisfied for the space $M^0_l$ and operators $I_{l-1}^l$ and $P^{l-1}_l$; see, e.g., Section 6 of [16] and [14]. Hence Theorem 3.3 is simply a restatement of Theorem 5.2 of [16].
4. HDG-P0 for the Generalized Stokes Equations

We follow the same procedures as in the previous section to present the HDG-P0 scheme for the generalized Stokes equations together with its optimal a priori error analysis, and the corresponding multigrid algorithms for the condensed linear system.

4.1. The model problem

We consider the following model problem with a homogeneous Dirichlet boundary condition:

\[
\begin{align*}
\beta u - \nabla \cdot (\mu \nabla u) + \nabla p &= f, & \text{in } \Omega, \\
\nabla \cdot u &= 0, & \text{in } \Omega, \\
u &= 0, & \text{on } \partial \Omega,
\end{align*}
\]

where \( u \) is the velocity, \( p \) is the pressure, \( f \in L^2(\Omega) \) is the source term, \( \mu > 0 \) is the viscosity constant, and \( \beta \geq 0 \) is the low-order term constant, which typically represents the inverse of time step size in an implicit-in-time discretization of the unsteady Stokes flow.

To present the HDG-P0 scheme, we introduce the tensor \( L := -\mu \nabla u \) as a new variable and rewrite the equations (22) into a first-order system:

\[
\begin{align*}
\mu^{-1} L + \nabla u &= 0, & \text{in } \Omega, \\
\nabla \cdot L + \beta u + \nabla p &= f, & \text{in } \Omega, \\
u &= 0, & \text{on } \partial \Omega.
\end{align*}
\]

4.2. The HDG-P0 scheme

The HDG-P0 discretization for the system (23) reads as follows: find \((L_h, u_h, \overline{u}_h, p_h) \in W_h \times V_h \times M_h^0 \times W_0^0\) such that

\[
\sum_{K \in T_h} \left( Q_K^0 \left( \mu^{-1} L_h \cdot \underline{G}_h \right) + Q_K^1 \left( \underline{G}_h u_h \cdot \overline{u}_h \right) \right) = 0,
\]

\[
\sum_{K \in T_h} \left( Q_K^0 (\tau_K (u_h - \overline{u}_h) \cdot \underline{v}_h) + Q_K^1 (\beta u_h \cdot \underline{v}_h) \right) = \sum_{K \in T_h} Q_K^1 (f \cdot \underline{v}_h),
\]

\[
\sum_{K \in T_h} Q_K^0 (\tau_K (u_h - \overline{u}_h) \cdot q_h) = 0,
\]

\[
\sum_{K \in T_h} \left( Q_K^0 \left( \left( \frac{L_h}{\mu} + p_h \right) \underline{v}_h \cdot \overline{v}_h \right) + Q_K^0 (\tau_K (u_h - \overline{u}_h) \cdot \overline{v}_h) \right) = 0
\]

for all \((G_h, v_h, \overline{v}_h, q_h) \in W_h \times V_h \times M_h^0 \times W_0^0\), where \( L \) is the unit diagonal matrix, and \( \tau_K = \frac{\mu}{\mu_K} \) is the stabilization parameter. Taking test functions \((G_h, v_h, \overline{v}_h, q_h) = (L_h, u_h, -\overline{u}_h, p_h)\) in (24) and adding, we have the following energy identity:

\[
\sum_{K \in T_h} \left( Q_K^0 \left( \mu^{-1} |L_h|^2 \right) + Q_K^0 (\tau_K (u_h - \overline{u}_h)^2) + Q_K^1 (\beta u_h^2) \right) = \sum_{K \in T_h} Q_K^1 (f \cdot u_h).
\]

4.3. An a priori error analysis for HDG-P0

Similar to the reaction-diffusion case, to simplify the a priori error analysis in this subsection, we assume \((L, u, p) \in H^1(\Omega) \times (H^2(\Omega) \cap H_0^1(\Omega)) \times (H^1(\Omega) \setminus \mathbb{R})\), together with the following elliptic regularity result:

\[
\mu^{-1/2} |L|_1 + \mu^{-1/2} \|p\|_1 + \left( \mu^{1/2} + \beta^{1/2} \right) \|u\|_2 \lesssim c_{\text{reg}} \|f\|_0,
\]

(25)
which holds when the domain Ω is convex. Furthermore, we assume the following stronger regularity of the source term \( f \) to perform the \textit{a priori} error analysis:

\[
f \in W^{1,\infty}(\Omega),
\]

and compare the solution to the HDG-P0 scheme (24) with the solution to a similar HDG scheme with exact integration. Let \((L_h^1, u_h^1, \hat{u}_h^1, p_h^1) \in W_h \times \mathcal{V}_h \times M_h^0 \times W_h^0\) be the solution to the following scheme:

\[
\begin{align*}
\sum_{K \in \mathcal{T}_h} \left( Q^0_K \left( \mu^{-1} \frac{L_h^1}{K} \cdot G_h \right) + Q^0_K \left( G_h \cdot \hat{u}_h^1 \right) \right) &= 0, \\
\sum_{K \in \mathcal{T}_h} \left( Q^0_K \left( \tau_K \left( u_h^1 - \hat{u}_h^1 \right) \cdot \mathcal{V}_h \right) + \left( \beta_{u_h^1}, \mathcal{V}_h \right)_K \right) &= \sum_{K \in \mathcal{T}_h} \left( f, \mathcal{V}_h \right)_K, \\
\sum_{K \in \mathcal{T}_h} Q^0_K \left( \left( \frac{L_h^1}{K} + p_h^1 I \right) n_K \cdot \mathcal{N}_h \right) + Q^0_K \left( \tau_K \left( u_h^1 - \hat{u}_h^1 \right) \cdot \mathcal{N}_h \right) &= 0, \\
\end{align*}
\]

for all \((G_h, \mathcal{V}_h, \hat{u}_h, q_h) \in W_h \times \mathcal{V}_h \times M_h^0 \times W_h^0\). Note that in two dimension \((d = 2)\), the above two schemes only differ by the right hand side term as \(Q^0_K\) is exact for quadratic polynomial integration, while in three dimensions \((d = 3)\) the quadrature scheme (24) further introduce numerical integration errors in the left hand side volume term in (24b). We use the following lemma to link the above system (26) with the lowest order (mixed) HDG scheme with projected jumps (and exact integration) that has been analyzed in [58] for the incompressible Navier–Stokes equations. The proof procedure is similar to Lemma 3.2 and is omitted here for simplicity.

**Lemma 4.1.** Let \((L_h^1, u_h^1, \hat{u}_h^1, p_h^1) \in W_h \times \mathcal{V}_h \times M_h^0 \times W_h^0\) be the solution to (26), then \((L_h^1, u_h^1, \hat{u}_h^1, p_h^1)\) satisfies

\[
\begin{align*}
\mu^{-1} \left( L_h^1, G_h \right)_{\mathcal{T}_h} - \left( \nabla \cdot \left( \frac{L_h^1}{K} + p_h^1 I \right) \right)_{\mathcal{T}_h} + \left( \nabla \cdot (\hat{u}_h^1), G_h \right)_{\mathcal{T}_h} + \left( \nabla \cdot (\hat{u}_h^1), G_h \right)_{\mathcal{T}_h} &= 0, \\
\left( \nabla \cdot \left( \frac{L_h^1}{K} + p_h^1 I \right), \mathcal{V}_h \right)_{\mathcal{T}_h} + \left( \tau_K \Pi_0 \left( u_h^1 - \hat{u}_h^1 \right), \mathcal{V}_h \right)_{\mathcal{T}_h} + \left( \beta_{u_h^1}, \mathcal{V}_h \right)_{\mathcal{T}_h} &= \left( f, \mathcal{V}_h \right)_{\mathcal{T}_h} \\
\left( \nabla q_h, \mathcal{V}_h \right)_{\mathcal{T}_h} + \left( \nabla q_h, \mathcal{V}_h \right)_{\mathcal{T}_h} &= 0, \\
\end{align*}
\]

for all \((G_h, \mathcal{V}_h, \hat{u}_h, q_h) \in W_h \times \mathcal{V}_h \times M_h^0 \times W_h^0\), where \(\Pi_0\) is the \(L_2\)-projection on to the piecewise constant vector on the mesh skeletons.

With Lemma 4.1, we adapt from [58] and get the following \textit{a priori} error estimates for the HDG scheme (26) with exact integration; see [58] for more details.

**Corollary 4.1.** Let \((L, u, p)\) be the exact solution to the model problem (23). Let \((L_h^1, u_h^1, \hat{u}_h^1, p_h^1) \in W_h \times \mathcal{V}_h \times M_h^0 \times W_h^0\) be the solution to (26). Then there holds

\[
\begin{align}
\| \mu^{-1/2} (L - L_h^1) \|_0 + \| \beta^{1/2} (u - u_h^1) \|_0 + \| \tau_K^{1/2} \Pi_0 \left( u_h^1 - \hat{u}_h^1 \right) \|_{0, \partial \mathcal{T}_h} &\lesssim h \Theta_1, \\
\| u - u_h^1 \|_0 &\lesssim \| \nabla (u - u_h^1) \|_0 + \| h^{-1/2} \left( u_h^1 - \hat{u}_h^1 \right) \|_{0, \partial \mathcal{T}_h} \lesssim \mu^{-1/2} h \Theta_1,
\end{align}
\]

where

\[
\Theta_1 := \mu^{-1/2} |L|_1 + \mu^{-1/2} |p|_1 + \left( \mu^{1/2} + \beta^{1/2} h^2 \right) |u|_2.
\]
Moreover, with the full elliptic regularity assumed in (25), we have the optimal $L_2$-convergence for $u_h^1$:
\[
\|u - u_h^1\|_0 \lesssim h^2(c_{\text{reg}} \Theta_2 + |u|_2),
\]
where
\[
\Theta_2 := \mu^{-1/2} |L|_1 + \mu^{-1/2} |p|_1 + \mu^{1/2} |u|_2.
\]

Now we are ready to present the optimal convergence results of the HDG-P0 scheme (24). We note the proof procedures are essentially the same as in Theorem 3.1 and the detailed proof is omitted here for simplicity.

**Theorem 4.1.** Let $(L, u, p)$ be the exact solution to the model problem (23). Let $(L_h, u_h, p_h) \in \mathcal{W}_h \times V_h \times M_h^0 \times W_h^0$ be the solution to (24). Then there holds
\[
\left\|\mu^{-1/2}(L - L_h)\right\|_0 + \left\|r_K^{-1/2} \Pi_h(u_h - \tilde{u}_h)\right\|_{0, \partial \Omega_h} \lesssim h(\Theta_1 + \Xi),
\]
(32a)
\[
\|u - u_h\|_0 \lesssim \|\nabla (u - u_h)\|_0 + \left\|\mu^{-1/2}(u_h - \tilde{u}_h)\right\|_{0, \partial \Omega_h} \lesssim \mu^{-1/2} h(\Theta_1 + \Xi),
\]
(32b)
where $\Theta_1$ is defined in (29) and
\[
\Xi := \mu^{-1/2} \beta \left(\|u\|_0 + \mu^{-1/2} h \Theta_1\right) + \mu^{-1/2} h^{d/2} \|f\|_{1, \infty}.
\]

Moreover, assuming the full elliptic regularity in (25), we have the following optimal $L_2$-convergence of $u_h$:
\[
\|u - u_h\|_0 \lesssim h^2 \Psi,
\]
(32c)
where
\[
\Psi := c_{\text{reg}} (\Theta_2 + \Xi) + |u|_2 + \frac{c_{\text{reg}} \mu^{-1/2} \beta \Xi}{\mu^{d/2} + \beta^{3/2} h^2} + \|\beta\|_{1, \infty} \mu^{-3/2} (\Theta_1 + \Xi) + h^{d/2 - 1} \mu^{-1} \|f - \beta u\|_{1, \infty},
\]
and $\Theta_2$ is defined in (31).

**Remark 4.1 (On pressure robustness).** Corollary 4.1 and Theorem 4.1 implies the HDG velocity approximation errors depend on the regularity of the pressure, which is not pressure-robust in the sense of [41]. A source modification technique based on discrete Helmholtz decomposition was used in [7, 47] for the nonconforming CR discretization of the Stokes problem ($\beta = 0$) to recover pressure-robust velocity approximations. When $\beta > 0$, a further modification of the mass term using a BDM interpolation [48] was needed to recover pressure-robustness. Due to the equivalence of the proposed HDG-P0 scheme (24) and a nonconforming CR discretization (see Thm. 4.2 below), similar modification can be used for the HDG-P0 scheme to render it pressure-robust by changing the two volume integration terms in (24b) with the following:
\[
Q_K^1(f \cdot \mathbf{v}_h) \rightarrow Q_K^1(f \cdot \Pi_h^{\text{BDM}} \mathbf{v}_h),
\]
\[
Q_K^1(\beta u_h \cdot \mathbf{v}_h) \rightarrow Q_K^1(\beta \Pi_h^{\text{BDM}} \left(\Pi_h^{\text{CR}} \mathbf{v}_h\right) \cdot \Pi_h^{\text{BDM}} \mathbf{v}_h),
\]
where $\Pi_h^{\text{BDM}}$ is the classical BDM interpolation into the space $V_h \cap H(\text{div}; \Omega)$, and $\Pi_h^{\text{CR}}$ is the interpolation from $M_h$ to $V_h^{\text{CR}}$ given in (33) below. Optimal pressure-robust velocity error estimates can be obtained for this modified HDG scheme following the work [7].
4.4. Equivalence to CR discretization

Similar to the reaction-diffusion case, a bijective interpolation operator \( \Pi_h^{\text{CR}} : M_h \rightarrow V_h^{\text{CR}} \) from the facet space \( M_h \) to the CR space \( V_h^{\text{CR}} \) is defined such that:

\[
\Pi_h^{\text{CR}} \hat{u}_h(m_F) = \hat{u}_h(m_F), \quad \forall \hat{u}_h \in M_h, \, F \in \mathcal{E}_h.
\]

Then we have the following equivalence result between the HDG-P0 scheme (24) and a (slightly modified) CR discretization for the generalized Stokes problem:

**Theorem 4.2.** Let \( (u_h^{\text{CR}}, p_h^{\text{CR}}) \in V_h^{\text{CR},0} \times W_h^0 \) be the solution to the following nonconforming scheme:

\[
\sum_{K \in \mathcal{T}_h} \left( Q_K^0 \left( \mu \nabla u_h^{\text{CR}}, \nabla \Pi_h^{\text{CR}} u_h \right) + Q_K^0 \left( \gamma_h \beta u_h^{\text{CR}}, \Pi_h^{\text{CR}} u_h \right) \right) = \sum_{K \in \mathcal{T}_h} Q_K^1 \left( \gamma_h f, \Pi_h^{\text{CR}} u_h \right),
\]

\[
\sum_{K \in \mathcal{T}_h} \left( Q_K^0 \left( q_h \nabla \Pi_h^{\text{CR}} u_h \right) \right) = 0,
\]

for all \( (u_h^{\text{CR}}, q_h) \in V_h^{\text{CR},0} \times W_h^0 \), where \( \gamma_h := \frac{\mu}{\mu + \frac{1}{\gamma_i} \tau_h} \). Then the solution \( (L_h, u_h, \hat{u}_h, p_h) \) to the HDG-P0 scheme (24) satisfies

\[
L_h = -\mu \nabla \Pi_h^{\text{CR}} \hat{u}_h,
\]

\[
u_h(m^i_K) = \gamma_h(m^i_K) \left( \hat{u}_h(m^i_K) + \frac{(h^i_K)^2}{(d+1)\mu} f(m^i_K) \right), \quad \forall i \in \{1, \ldots, d+1\}, \forall K \in \mathcal{T}_h,
\]

\[
\Pi_h^{\text{CR}} \hat{u}_h = u_h^{\text{CR}},
\]

\[
p_h = p_h^{\text{CR}},
\]

where \( m^i_K \) is the barycenter of the \( i \)-th facet \( F^i \) of the element \( K \), and \( h^i_K = \frac{|K|}{|F^i|} \).

**Proof.** The proof procedure is the same as in Theorem 3.2 and we only sketch the main steps here. By the definition of \( \Pi_h^{\text{CR}} \) and integration by parts, we get from (24a), for all \( G \in W_h^1 \),

\[
\sum_{K \in \mathcal{T}_h} \left( Q_K^0 \left( \mu^{-1} L_h \cdot G_h + \nabla \Pi_h^{\text{CR}} \hat{u}_h \right) \right) = \sum_{K \in \mathcal{T}_h} \left( Q_K^0 \left( \mu^{-1} L_h + \nabla \Pi_h^{\text{CR}} \hat{u}_h \right) \cdot G_h \right) = \left( \mu^{-1} L_h + \nabla \Pi_h^{\text{CR}} \hat{u}_h, G_h \right)_{\mathcal{T}_h} = 0,
\]

where we used the fact that \( Q_K^0 \) is exact for \( P^1(K) \) and \( Q_K^0 \) is exact for \( P^1(F) \) for all \( F \in \partial K \). The identity (35a) follows thanks to \( (\mu^{-1} L_h + \nabla \Pi_h^{\text{CR}} \hat{u}_h) \in W_h^1 \).

Next, for \( i \in \{1, \ldots, d+1\} \), taking test function \( u_h \) in (24b) to be supported only on an element \( K \), whose values are 1 on \( m^i_K \) and zero on \( m^j_K \) for \( j \neq i \). By the definition of \( Q_{\partial K}^0 \) and \( Q_K^1 \), we get

\[
\tau_h^i(m^i_K) = \hat{u}_h(m^i_K) ||F^i| + \frac{|K|}{d+1} \beta(m^i_K) u_h(m^i_K) = \frac{|K|}{d+1} f(m^i_K).
\]

We immediately get the equality (35b) by solving for the value \( u_h(m^i_K) \) and using the definition of \( \tau_h \) and \( h_K \).

Finally, by the definition of \( \Pi_h^{\text{CR}} \), \( Q_{\partial K}^0 \) and \( Q_K^1 \), the identities (35a) and (35b), and integration by parts, we have:

\[
Q_{\partial K}^0 \left( \left( L_h + p_h \frac{1}{2} \right) n_K \cdot \hat{u}_h \right) = \left( \left( L_h + p_h \frac{1}{2} \right) \cdot \nabla \Pi_h^{\text{CR}} \hat{u}_h \right)_{\mathcal{T}_h}.
\]
and applies multigrid algorithms for the positive definite system on the divergence-free kernel space $\mathcal{Z}$ for all element schemes in two- and three-dimensions by other researchers; see, e.g., Schöberl’s divergence-free functions, and (ii) a robust block-smoother capturing the divergence free basis functions. This approach proposed by Brenner [12,13] directly works with the saddle point system (with a penalty term), which, however, cannot be applied to the positive definite Schur complement subsystem. The third approach proposed by Schöberl [60,61] also works with a penalty formulation of the saddle point system, where a multigrid theory was applied to the positive definite subsystem for the velocity approximation. The key ingredients in [60,61] are (i) a robust intergrid transfer operator that transfers coarse-grid divergence-free functions to fine-grid (nearly) divergence-free functions, and (ii) a robust block-smoother capturing the divergence free basis functions. This approach is attractive in three dimensions as the intergrid transfer operator does not need to directly work with the divergence-free kernel space, and is much easier to realize in practice than the first approach. Schöberl’s approach was originally introduced for the $P^2/P^0$ discretization on triangles, it has been applied to other finite element schemes in two- and three-dimensions by other researchers; see, e.g., [33,34,39,42,45]. In addition to the aforementioned geometric multigrid methods, Notay [53,54] applied algebraic multigrid methods to an algebraically transformed saddle-point structure where the diagonal blocks dominate. Sharp and robust eigenvalue bounds for the two-grid cases were obtained requiring only mild assumptions.

\[
\begin{align*}
Q^K_{\partial K}(\nabla \cdot n_h q_h) &= (\Pi_h^{\text{CR}} \nabla \cdot \Pi_h^{\text{CR}} \tilde{u}_h, q_h)_{\partial K}, \\
Q^K_{\partial K}(\tilde{u}_h, n_h \cdot \nabla q_h) &= (\Pi_h^{\text{CR}} \tilde{u}_h, n_h, \cdot \nabla q_h)_{\partial K},
\end{align*}
\]

where we used the fact that $Q^K_{\partial K}$ is exact for $\mathcal{P}^0(K)$ and $Q^K_0$ is exact for $\mathcal{P}^1(F)$ for all $F \in \partial K$. By plugging the above identities into (24c) and (24d), we get the condensed HDG-P0 scheme:

\[
\begin{align*}
\sum_{K \in \mathcal{T}_h} Q^K_0 (\gamma h f \cdot \Pi_h^{\text{CR}} \tilde{u}_h) &= \sum_{K \in \mathcal{T}_h} Q^K_1 (\gamma h f \cdot \Pi_h^{\text{CR}} \tilde{u}_h), \\
\sum_{K \in \mathcal{T}_h} Q^K_0 (q_h \nabla \cdot \Pi_h^{\text{CR}} \tilde{u}_h) &= 0,
\end{align*}
\]

for all $(\tilde{u}_h, q_h) \in M_h^0 \times W_h^0$, which implies the result (35c) and (35d).

Same as in Section 3.4 for the reaction-diffusion equation, the above result demonstrates the equivalence between the condensed linear system of HDG-P0 for the generalized Stokes equations (36) and a slightly modified CR discretization (34) with numerical integration, where the modification is introduced by the scaling parameter $\gamma_h$. In practice we first locally eliminate $\tilde{u}_h$ and $\tilde{q}_h$ from the linear system of HDG-P0 (24) to arrive at the condensed system (36), and then recover these local variables after solving for $\tilde{u}_h$ and $p_h$ from (36).

**Remark 4.2** (On multigrid algorithm for (36)). The saddle-point structure (36) creates extra difficulty in design robust multigrid solvers for the system. In the literature, there are three approaches to construct a robust geometric multigrid algorithm for the nonconforming scheme (34), which is equivalent to the condensed system (36). The first approach [10,63,65] explores the fact that the CR discretization produces a cell-wise divergence free velocity approximation

\[
\tilde{u}_h^{\text{CR}} \in \mathcal{Z}_h^{\text{CR},0} := \left\{ \psi_h \in Y_h^{\text{CR},0} : \nabla \cdot \tilde{u}_h|_K = 0, \forall K \in \mathcal{T}_h \right\},
\]

and applies multigrid algorithms for the positive definite system on the divergence-free kernel space $\mathcal{Z}_h^{\text{CR},0}$. This approach is restricted to two dimensions only, which is closely related to multigrid for the nonconforming Morley scheme for the Biharmonic equation [9]. Its extension to three dimensions is highly nontrivial due to the need of constructing (highly complex) intergrid transfer operators between these divergence-free subspaces. The second approach proposed by Brenner [12,13] directly works with the saddle point system (with a penalty term), which, however, cannot be applied to the positive definite Schur complement subsystem. The third approach proposed by Schöberl [60,61] also works with a penalty formulation of the saddle point system, where a multigrid theory was applied to the positive definite subsystem for the velocity approximation. The key ingredients in [60,61] are (i) a robust intergrid transfer operator that transfers coarse-grid divergence-free functions to fine-grid (nearly) divergence-free functions, and (ii) a robust block-smoother capturing the divergence free basis functions. This approach is attractive in three dimensions as the intergrid transfer operator does not need to directly work with the divergence-free kernel space, and is much easier to realize in practice than the first approach. Schöberl’s approach was originally introduced for the $P^2/P^0$ discretization on triangles, it has been applied to other finite element schemes in two- and three-dimensions by other researchers; see, e.g., [33,34,39,42,45]. In addition to the aforementioned geometric multigrid methods, Notay [53,54] applied algebraic multigrid methods to an algebraically transformed saddle-point structure where the diagonal blocks dominate. Sharp and robust eigenvalue bounds for the two-grid cases were obtained requiring only mild assumptions.
In the next subsection, we follow the Schöberl’s geometric multigrid theory to provide a robust geometric multigrid algorithm for the system (36) in combination with an augmented Lagrangian Uzawa iteration method [36,66].

4.5. Multigrid-based augmented Lagrangian Uzawa iteration

We use the same notations for the hierarchical meshes and finite element spaces as in Section 3.5. With slight abuse of notations, we define the following $L_2$-like inner product on the vector facet space $M_l$:

$$\langle \tilde{u}_l, \tilde{v}_l \rangle_{0,l} := \sum_{K \in \mathcal{T}_l} Q^1_K (\Pi_l^CR \tilde{u}_l, \Pi_l^CR \tilde{v}_l) = \sum_{K \in \mathcal{T}_l} \sum_{i=1}^{d+1} \frac{|K|}{d+1} \tilde{u}_l (m^i_K) \tilde{v}_l (m^i_K),$$

with the induced norm $\| \cdot \|_{0,l}$, where $\Pi^CR : M_l \rightarrow V_l^CR$ is the interpolation operator (33) on $l$-th mesh level. The pressure space $W_l$ is equipped with the standard $L_2$-norm, which we denote as

$$[p_l, q_l]_{0,l} := \sum_{K \in \mathcal{T}_l} Q^0_K (p_l, q_l) \quad \forall p_l, q_l \in W_l.$$ 

We define the following linear operators $A_i : M^0_l \rightarrow M^0_l$, and $B_i : M^0_l \rightarrow W_l$

$$(A_i \tilde{u}_l, \tilde{v}_l)_{0,l} := a_i (\tilde{u}_l, \tilde{v}_l), \quad \forall \tilde{u}_l, \tilde{v}_l \in M^0_l,$$

$$[B_i \tilde{u}_l, q_l]_{0,l} := b_i (\tilde{u}_l, q_l), \quad \forall \tilde{u}_l \in M^0_l, q_l \in W_l^0,$$

where $a_i, b_i$ are the following bilinear forms on the $l$-th mesh level:

$$a_i (\tilde{u}_l, \tilde{v}_l) := \sum_{K \in \mathcal{T}_l} (Q^1_K (\mu \nabla \Pi^CR \tilde{u}_l \cdot \nabla \Pi^CR \tilde{v}_l) + Q^1_K (\beta_i \Pi^CR \tilde{u}_l \cdot \Pi^CR \tilde{v}_l)), \quad \forall \tilde{u}_l, \tilde{v}_l \in M^0_l,$$

$$b_i (\tilde{u}_l, q_l) := \sum_{K \in \mathcal{T}_l} Q^0_K (q_l \nabla \cdot \Pi^CR \tilde{u}_l), \quad \forall \tilde{u}_l \in M^0_l, q_l \in W_l,$$

with $\beta_i (m^i_K) := \gamma_i (m^i_K) \beta (m^i_K)$ for all $K \in \mathcal{T}_l$. Further denoting $B_i^* : W_l \rightarrow M^0_l$ as the transpose of $B_i$ with respect to the $L_2$-inner products:

$$(B_i^* p_l, \tilde{v}_l)_{0,l} := [p_l, B_i \tilde{v}_l]_{0,l}, \quad \forall p_l \in W_l, \tilde{v}_l \in M^0_l,$$

and $f_i \in M^0_l$ such that

$$(f_i, \tilde{v}_l)_{0,l} := \sum_{K \in \mathcal{T}_l} Q^0_K (\gamma_i f \cdot \Pi^CR \tilde{v}_l), \quad \forall \tilde{v}_l \in M^0_l.$$

Then the operator form of the system (36) is to find $(\tilde{u}_l, p_l) \in M^0_l \times W^0_l$ satisfying:

$$A_i \tilde{u}_l + B_i^* p_l = f_i, \quad (37a)$$

$$B_i \tilde{u}_l = 0. \quad (37b)$$

Similar to [39,44], we apply the augmented Lagrangian Uzawa iteration method [36,66] to the above saddle-point system (37), which solves the following (equivalent) augmented Lagrangian formulation of (37) iteratively using the Uzawa method

$$\left( \frac{A_i + \epsilon^{-1} B_i^* B_i}{A_i} \right) \tilde{u}_l + B_i^* p_l = f_i, \quad (38a)$$
\[ B_j \tilde{u}_j = 0. \] (38b)

The Uzawa method for (38) with damping parameter \( \epsilon^{-1} \gg 1 \) reads as follows: Start with \( p^0_j = 0 \), for \( k = 1, 2, \cdots \), find \((\tilde{u}^k_j, p^k_j) \in M^0_j \times W^0_j \) such that

\[
\begin{align*}
\mathbf{A}_j^* \tilde{u}^k_j &= f_j - B^*_j p^{k-1}_j, \\
p^k_j &= p^{k-1}_j - \epsilon^{-1} B_j \tilde{u}^k_j.
\end{align*}
\] (39a)

Here the singularly perturbed operator \( \mathbf{A}_j^* \) is associated with the bilinear form

\[
\mathbf{a}_j^*(\tilde{u}_j, \tilde{v}_j) := \mathbf{a}_j(\tilde{u}_j, \tilde{v}_j) + \sum_{k \in \eta_i} Q_k^0(\epsilon^{-1} \nabla \cdot \Pi^I_{CR} \tilde{u}_j \nabla \cdot \Pi^I_{CR} \tilde{v}_j).
\]

We quote a convergence result from Lemma 2.1 of [44] for the above Uzawa method (39).

**Lemma 4.2** ([44], Lem. 2.1). Let \((\tilde{u}_j, p_j) \in M^0_j \times W^0_j \) be the solution of (37), and let \((\tilde{u}^k_j, p^k_j) \) be the \( k \)-th Uzawa iteration solution to (39). Then the following estimate holds:

\[
\| \tilde{u}^k_j - \tilde{u}_j \|_{A_j} \lesssim \sqrt{\epsilon} \| p^k_j - p_j \|_0 \lesssim \sqrt{\epsilon} \left( \frac{\epsilon}{\epsilon + \mu_0} \right)^k \| p_j \|_0,
\]

where \( \mu_0 \) is the minimal eigenvalue of the Schur complement \( S_j = B_j A_j^{-1} B^*_j \), which is independent of the mesh size.

**Remark 4.3** (On practical choice of \( \epsilon \)). The first step Uzawa iteration solution \((\tilde{u}^1_j, p^1_j) \) is simply the penalty method applied to (37) where a small mass term \([\epsilon p_j, q_j]_{0, l}\) is subtracted from the continuity equation (37b). The above convergence result indicates to take \( 0 < \epsilon \ll 1 \) for faster convergence in terms of Uzawa iteration counts. However, taking \( \epsilon \) extremely small will leads to round-off issues as the non-zero matrix entries in \( \mathbf{A}_j \) is of \( \mathcal{O}(1) \) while that of \( \epsilon^{-1} B_j B^*_j \) is of \( \mathcal{O}(\epsilon^{-1}) \). In our numerical experiments, the machine precision is \( 10^{-16} \), and we use \( \epsilon = 10^{-8} \) and perform one Uzawa iteration.

Since the pressure space \( W_j \) is a discontinuous piecewise constant space, there is no equation solve needed for (39b). Hence the major computational cost of a Uzawa iteration is in solving velocity from (39a), for which we use a multigrid following [60, 61]. As mentioned in Remark 4.2, two key ingredients of a robust multigrid algorithm for the system (39a) are a robust intergrid transfer operator and a robust block smoother.

It turns out that a vectorial version of the averaging intergrid transfer operator, denoted as \( \mathbf{L}^-_{l-1} \), used in (21) is not robust for \( \epsilon \ll 1 \). Here we stabilize this averaging operator with a local correction using discrete harmonic extensions [60], which takes care of the coarse-grid divergence. Denoting \( M^0_{l,T} \) as the local subspace of \( M^0_l \) whose DOFs vanishes on the \((l-1)\)-th level mesh skeleton \( E_{l-1} \), the integer grid transfer operator \( \mathbf{L}^-_{l-1} : M^0_{l-1} \rightarrow M^0_{l} \) is defined as follows:

\[
\mathbf{L}^-_{l-1} := \left( \text{id} - P^T_{M^0_l} \right) \mathbf{L}^-_{l-1},
\] (40)

where \( \text{id} \) is the identity operator and the local projection \( P^T_{M^0_l} : M^0_l \rightarrow M^0_{l,T} \) satisfies

\[
\mathbf{a}_l^*\left( P^T_{M^0_l} \tilde{u}_l, \tilde{v}_l^T \right) = \mathbf{a}_l^*(\tilde{u}_l, \tilde{v}_l^T), \quad \forall \tilde{u}_l \in M^0_l, \quad \tilde{v}_l^T \in M^0_{l,T},
\] (41)

which can be solved element-by-element on the coarse mesh \( \mathcal{T}_{l-1} \).

A robust smoother for (39a) needs to take care of the discretely divergence-free kernel space. Here the classical block smoothers for \( H(\text{div}) \)-elliptic problems from Arnold, Falk and Winther are used [2]. In particular, we use...
vertex-patch based damped block Jacobi or block Gauss–Seidel smoother. For completeness, we formulate the vertex-patch damped block Jacobi smoother below. Denoting $V_l$ as the set of vertices of the triangulation $T_l$. We define the subset of facets $E_{l,v}$ meeting at the vertex $v \in V_l$ as

$$E_{l,v} := \bigcup_{F \in E_l, v \in F} F,$$

and decompose the finite element space $M^0_l$ into overlapping subspaces with support on $E_{l,v}$:

$$M^0_l = \sum_{v \in V_l} M^0_{l,v} := \sum_{v \in V_l} \{ \widehat{u}_l \in M^0_l : \text{supp} \, \widehat{u}_l \subset E_{l,v} \}.$$

Further defining $P^u_{A_{l-1}} : M^0_l \to M^0_{l-1}$ as the local projection onto the subspace $M^0_{l-1,v}$ with respect to the bilinear form $a_l^\#$ such that:

$$a_l^\# \left( P^u_{A_{l}} \widehat{u}_l, \widehat{v}_l \right) = a_l^\# \left( \widehat{u}_l, \widehat{v}_l \right), \quad \forall \widehat{u}_l \in M^0_l, \, \widehat{v}_l \in M^0_{l,v}, \, v \in V_l,$$

the damped block Jacobi smoother is then given as:

$$R_l := \zeta \sum_{v \in V_l} P^u_{A_{l}} \left( A_{l}^\# \right)^{-1}, \quad (42)$$

where $\zeta$ is the damping parameter that is small enough (independent of $\epsilon$) to ensure the operator $(id - R_l A_{l}^\#)$ is a positive definite contraction.

With these two ingredients ready, we are ready to present the $W$-cycle and variable $V$-cycle multigrid algorithms for the linear system $A_l^\# u_l = g_l \in M^0_l$ as below:

The analysis of the above multigrid algorithm follows directly from Schöberl’s work [60, 61], which is based on classical multigrid theories [5, 6, 38], and boils down to the verification of three properties of the underlying operators. To perform the analysis, we denote the following parameter-dependent $L_2$-like norm $\| \cdot \|_{\epsilon,l}$ on $M^0_l$:

$$\| \widehat{u}_l \|_{\epsilon,l}^2 := \| \widehat{u}_l \|_{0,l}^2 + \epsilon^{-1} h_l^2 \| \nabla \cdot \Pi_l^{\text{CR}} \widehat{u}_l \|_0^2 + \epsilon^{-2} h_l^2 \| \Pi_{l-1}^W \left( \nabla \cdot \Pi_l^{\text{CR}} \widehat{u}_l \right) \|_0^2,$$

where $\Pi_{l-1}^W$ is the $L_2$-projection onto the piecewise constant space $W_{l-1}$ on the $(l-1)$-th level mesh. We define $\| \cdot \|_{A_l^\#}$ as the norm induced by the SPD operator $A_{l}^\#$, i.e., $\| \cdot \|_{A_l^\#} := \sqrt{A_{l}^\# \cdot \cdot}$, and denote $P^l_{A_{l-1}^\#} : M^0_l \to M^0_{l-1}$ as the transpose of $P_{A_{l-1}^\#}$ with respect to the bilinear form $a_l^\#$ such that:

$$a_{l-1}^\# \left( P^l_{A_{l-1}^\#} \widehat{u}_l, \widehat{v}_{l-1} \right) = a_{l}^\# \left( \widehat{u}_l, P_{A_{l-1}^\#} \widehat{v}_{l-1} \right), \quad \forall \widehat{u}_l \in M^0_l, \, \widehat{v}_{l-1} \in M^0_{l-1}.$$

We quote the abstract multigrid convergence result from Theorem 3.7 of [61].

---

1Edge-patch based block smoother can also be used in three dimensions [2] to reduce memory consumption.
Algorithm 2. The multigrid algorithm for $A_l^j \tilde{u}_l = g_l$.

The $l$-th level multigrid algorithm produces MG$_A^l (l, \tilde{g}_l, \tilde{u}_l, m(l), q)$ as an approximation solution for $A_l^j \tilde{u}_l = g_l$ with initial guess $\tilde{u}_l^0$, where $m(l)$ denotes the number of pre-smoothing and post-smoothing steps on the $l$-th mesh level, and $q \in \{1, 2\}$. Here $q = 1$ corresponding to the $V$-cycle algorithm, and $q = 2$ corresponding to the $W$-cycle algorithm in which constant smoothing steps $m(1) = \cdots = m(l) = m$ is used.

If $l = 1$ then

$$\text{MG}_A^l (l, \tilde{g}_l, \tilde{u}_l, m(l), q) = (A_l^1)^{-1} g_l,$$

else

Perform the following three steps:

(1) Pre-smoothing. For $j = 1, \ldots, m(l)$, compute $\tilde{u}_l^j$ by

$$\tilde{u}_l^j = \tilde{u}_l^{j-1} + R_l (g_l - A_l^j \tilde{u}_l^{j-1}).$$

(2) Coarse grid correction. Let $\delta \tilde{u}_l^{0, -1} = 0$ and $\tilde{v}_l^{-1} = \tilde{v}_l^{-1} (g_l - A_l^j \tilde{u}_l^m(l))$. For $k = 1, \ldots, q$, compute $\delta \tilde{u}_l^{k, -1}$ by

$$\delta \tilde{u}_l^{k, -1} = \text{MG}_A^{-1} \left((l-1, \tilde{v}_l^{-1}, \delta \tilde{u}_l^{k-1, -1}, m(l-1))\right).$$

Then we get

$$\tilde{u}_l^m(l+1) = \tilde{u}_l^m(l) + \tilde{v}_l^{-1} \delta \tilde{u}_l^{m(l)} \delta \tilde{u}_l^{m(l)},$$

where $\delta \tilde{u}_l^{m(l)}$ is the restriction operator satisfying

$$\left(\tilde{v}_l^{-1} \tilde{u}_l, \tilde{v}_l^{-1} \tilde{v}_l^{-1}\right)_{0, l} = \left(\tilde{u}_l, \tilde{v}_l^{-1} \tilde{v}_l^{-1}\right)_{0, l}, \quad \forall \tilde{u}_l \in M^0_l, \quad \forall \tilde{v}_l \in M^0_l.$$

(3) Post-smoothing. For $j = m(l) + 2, \ldots, 2m(l) + 1$, compute $\tilde{u}_l^j$ by

$$\tilde{u}_l^j = \tilde{u}_l^{j-1} + \tilde{R}_l^T (g_l - A_l^j \tilde{u}_l^{j-1}),$$

where $\tilde{R}_l^T$ is the transpose of $\tilde{R}_l$ with respect to the inner product $(\cdot, \cdot)_{0, l}$.

We then define MG$_A (l, \tilde{g}_l, \tilde{u}_l^0, m(l), q) = \tilde{u}_l^{2m(l)+1}$.

Theorem 4.3 ([61], Thm. 3.7). Let the multigrid procedure be as defined in Algorithm 2. Assume there hold the following properties:

- (A0) Scaling property. The spectrum of $\text{id} - \tilde{R}_l A_l^j$ is in the interval $(0, 1)$.
- (A1) Approximation property. There exists a positive constant $C$ independent of mesh size and the parameter $\epsilon$ such that:

$$\left\| \left(\text{id} - \tilde{R}_l A_l^j \right) \tilde{u}_l \right\|_{\epsilon,l} \leq C h_l \| \tilde{u}_l \|_{A_l^j}, \quad \forall \tilde{u}_l \in M^0_l.$$

- (A2) Smoothing property. There exists a positive constant $C$ independent of mesh size $h_l$ and the parameter $\epsilon$ such that:

$$\left\| \left(\text{id} - \tilde{R}_l A_l^j \right)^{m(l)} \tilde{u}_l \right\|_{A_l^j} \leq C m(l) h_l^{-1} \| \tilde{u}_l \|_{\epsilon,l}, \quad \forall \tilde{u}_l \in M^0_l.$$

Then the following multigrid methods lead to convergent solvers:

- The $W$-cycle multigrid algorithm with sufficiently many smoothing steps leads to a convergent method. That is, there exists positive constants $m_*$ and $C$ independent of mesh size $h_l$ and parameter $\epsilon$, such that with $q = 2$ and $m(1) = \cdots = m(l) = m$ in Algorithm 2 we have:

$$\| \tilde{E}_{l,m} \tilde{u}_l \|_{A_l^j} \leq C m^{-1/4} h_l^{-1} \| \tilde{u}_l \|_{\epsilon,l}, \quad \forall \tilde{u}_l \in M^0_l, \quad l \geq 1, \quad m \geq m_*,$$

where $\tilde{E}_{l,m} : M^0_l \rightarrow M^0_l$ is the operator relating the initial error and the final error of the multigrid W-cycle algorithm, i.e.,

$$\tilde{E}_{l,m} (\tilde{u}_l - \tilde{u}_l^0) := \tilde{u}_l - \text{MG}_{A_l^j} (l, \tilde{g}_l, \tilde{u}_l^0, m, 2).$$
The variable $V$-cycle algorithm leads to a robust preconditioner. That is, with $q = 1$ and $\beta_0 m(l) \leq m(l - 1) \leq \beta_1 m(l)$ in Algorithm 2 (1 < $\beta_0 < \beta_1$), there exists positive constant $C$ independent of mesh size $h_l$ and parameter $\epsilon$ such that

$$
\kappa \left( \mathbb{P}_{l,m(l)} A_l^t \right) \leq 1 + C m(l)^{-1/4},
$$

where $\kappa$ is the condition number, and $\mathbb{P}_{l,m(l)} : M_l^0 \rightarrow M_l^0$ is the preconditioning operator relating the residual to the correction of the variable $V$-cycle algorithm with a zero initial guess, i.e.,

$$
\mathbb{P}_{l,m(l)} \mathbb{U}_l := \text{MG}_{A_l^t}(l, \mathbb{U}_l, 0, m(l), 1), \quad \forall \mathbb{U}_l \in M_l^0.
$$

**Remark 4.4** (Verifications of the assumptions in Theorem 4.3). Assumption (A0) is satisfied by taking the damping parameter $\xi$ sufficiently small, which only depends on the (bounded) number of overlapping blocks [2]. Note that if a block Gauss–Seidel smoother is used, there is no need to add damping.

We note that the approximation property (A1), the smoothing property (A2), and thus our proposed multigrid method for the HDG-P0 scheme for the generalized Stokes equations require full elliptic regularity assumed in (25). The verification of the approximation property (A1) essentially follows from Section 4.4 of [60] (see also [60], Sect. 5), which establishes a stable decomposition result for the coarse grid operator $id - \mathbb{P}_l^{l-1} A_l^{-1}$, where the analysis was performed on an equivalent mixed formulation. Here a key step in the proof is to show the averaging operator $I_{l-1}$ preserves the divergence on the coarse-grid:

$$
\int_K \nabla \cdot \Pi_{l-1}^{CR} \mathbb{U}_{l-1} dx = \int_{\partial K} \mathbb{U}_{l-1} \cdot n_K ds = \int_{\partial K} I_{l-1} \mathbb{U}_{l-1} \cdot n_K ds
$$

$$
= \sum_{j=1}^8 \int_{\partial K^j} I_{l-1} \mathbb{U}_{l-1} \cdot n_K ds = \sum_{j=1}^8 \int_{K^j} \nabla \cdot \Pi_{l-1}^{CR} \mathbb{U}_{l-1} dx,
$$

for all $K \in T_{l-1}$, where the elements $K^j$ are children of $K$ such that $\bar{K} = \bigcup_{j=1}^8 K^j$. Hence, the coarse-grid divergence correction only need to be locally performed on the bubble spaces $M_l^0$ in (41).

The verification of the smoothing property (A2) follows from Section 6 of [60], where a key step is to establish the following bound of the interpolation norm ([60], Lem. 7)

$$
\| \mathbb{U}_{l,1} \|_{[D_l^t, A_l^t]} \lesssim \| \mathbb{U}_{l,1} \|_{\epsilon, l}, \quad (43)
$$

where $\mathbb{U}_{l,1} \in Z_l^0 := \{ \mathbb{U}_l \in M_l^0 : \nabla \cdot \Pi_{l}^{CR} \mathbb{U}_l|_K = 0, \forall K \in T_l \}$ is a discretely divergence-free function, $\| \cdot \|_{[D_l^t]}$ is the additive Schwarz norm expressed by

$$
\| \mathbb{U}_l \|^2_{[D_l^t]} := \inf_{u = \sum_{\nu} u_{l, \nu}} \sum_{\nu \in \mathcal{M}_{l, \nu}} \| u_{l, \nu} \|^2_{A_l^t},
$$

and $\| \cdot \|_{[D_l^{t-1}, A_l^{t-1}]}$ is the interpolation norm between $\| \cdot \|_{[D_l^t]}$ and $\| \cdot \|_{[A_l^t]}$. The proof of (43) relies on the explicit characterization of discretely divergence-free functions in $Z_l^0$ using the nonconforming Stokes complex [17, 40], which can be expressed as the discrete curl of the Morley finite element [51] when $d = 2$, and the discrete curl of the $H$(grad curl)-nonconforming finite element [40, 69] when $d = 3$. Using such explicit characterization, one can prove

$$
\| \mathbb{U}_{l,1} \|_{[D_l^t, A_l^t]} \lesssim \| \mathbb{U}_{l,1} \|_{[D_l^t]} \lesssim \| \mathbb{U}_{l,1} \|_{\epsilon, l},
$$

see similar derivation in Lemma 9 of [39].

5. Numerical experiments

In this section, we present numerical experiments to verify the optimal convergence rates of the HDG-P0 schemes and the optimality of the multigrid algorithms. We tested the proposed multigrid algorithms both as
the iteration solvers and as the preconditioners for the preconditioned conjugate gradient (PCG) method to solve the condensed systems (20) and (39a), with a relative tolerance of $10^{-8}$ used as the stopping criterion. The bisection algorithm is used for (local) mesh refinements. All results are obtained by using the NGSolve software [62]. Source code is available at https://github.com/WZKuang/MG4HDG-P0.git.

5.1. Reaction-diffusion equation

5.1.1. Manufactured solution

We first verify the optimal convergence rates of the HDG-P0 scheme (4) for the reaction-diffusion equation with known solution. We set the domain as a unit square/cube $\Omega = [0, 1]^d$ with homogeneous Dirichlet boundary conditions on all sides. We let the coefficients $\alpha = \beta = 1 + \frac{1}{2}\sin(x)\sin(y)$ when $d = 2$, and $\alpha = \beta = 1 + \frac{1}{2}\sin(x)\sin(y)\sin(z)$ when $d = 3$. The exact solution is set as:

$$u = \begin{cases} (x - x^2)(y - y^2), & \text{when } d = 2, \\ (x - x^2)(y - y^2)(z - z^2), & \text{when } d = 3. \end{cases}$$

Table 1 reports the estimated order of convergence (EOC) of the $L_2$ norms $\|u_h - u\|_0$ and $\|\sigma_h - \sigma\|_0$ of the HDG-P0 scheme (4), and optimal convergence rates are obtained as expected.

5.1.2. Non-convex domain and jump diffusion coefficients on the coarsest mesh

In this example, we consider the reaction-diffusion equation with jump diffusion coefficients on the coarsest mesh in a non-convex domain. When $d = 2$, we denote $\Omega_1^{2D}$ as the domain formed by connecting points (0.5, 0.15), (0.65, 0.3), (0.5, 0.45), and (0.35, 0.3) in order. We then define $\Omega_2^{2D} := ([0, 1] \times [0, 0.6]) \setminus \Omega_1^{2D}$, and $\Omega_3^{2D} := [0.2, 0.8] \times [0.6, 0.8]$. The domain $\Omega$ consists of these three subdomains, i.e., $\Omega = \Omega_1^{2D} \cup \Omega_2^{2D} \cup \Omega_3^{2D}$ as illustrated in the left panel of Figure 1. When $d = 3$, $\Omega_1^{3D} := \Omega_1^{2D} \times [0.225, 0.375]$, $\Omega_2^{3D} := \Omega_2^{2D} \times [0, 0.6]$, $\Omega_3^{3D} := \Omega_3^{2D} \times [0.2, 0.4]$, and $\Omega = \Omega_1^{3D} \cup \Omega_2^{3D} \cup \Omega_3^{3D}$. The diffusion coefficient $\alpha$ differs in different subdomains, which is set as 10 in $\Omega_1^{3D}/\Omega_3^{3D}$, 1 in $\Omega_2^{3D}/\Omega_2^{3D}$, and 1000 in $\Omega_3^{3D}/\Omega_3^{3D}$. The source term $f = 1$ in the center.

| Level | $d = 2$ | $d = 3$ |
|-------|--------|--------|
|       | $\|u_h - u\|_0$ | EOC | $\|\sigma_h - \sigma\|_0$ | EOC | $\|u_h - u\|_0$ | EOC | $\|\sigma_h - \sigma\|_0$ | EOC |
| 1     | 1.17e-1 | \_ | 8.29e-1 | \_ | 1.67e-2 | \_ | 1.72e-1 | \_ |
| 2     | 2.96e-2 | 1.98 | 4.22e-1 | 0.97 | 4.25e-3 | 1.97 | 9.22e-2 | 0.90 |
| 3     | 7.44e-3 | 1.99 | 2.12e-1 | 0.99 | 1.07e-3 | 1.99 | 4.68e-2 | 0.98 |
| 4     | 1.86e-3 | 2.00 | 1.06e-1 | 1.00 | 2.66e-4 | 2.00 | 2.35e-2 | 1.00 |
| 5     | 4.66e-4 | 2.00 | 5.31e-2 | 1.00 | 6.64e-5 | 2.00 | 1.17e-2 | 1.00 |
Table 2. Iteration counts of \( V \)-cycle multigrid iteration solver with different smoothers and smoothing steps in Example 5.1.1.

| Parameters | \( d = 2 \) | \( d = 3 \) |
|------------|--------------|--------------|
| \( J \) Facet DOFs | \# It \( \kappa \) | \# It \( \kappa \) | \# It \( \kappa \) | \# It \( \kappa \) | \# It \( \kappa \) |
| 2 | 2.20e2 | 19 | 4.9 | 13 | 2.6 | 9 | 1.5 | 12 | 2.2 | 8 | 1.3 | 6 | 1.1 |
| 3 | 8.48e2 | 22 | 7.1 | 14 | 3.4 | 10 | 1.8 | 13 | 2.8 | 9 | 1.5 | 6 | 1.1 |
| 4 | 3.33e3 | 23 | 8.0 | 15 | 3.8 | 11 | 2.1 | 14 | 3.2 | 9 | 1.7 | 7 | 1.2 |
| 5 | 1.32e4 | 25 | 11 | 16 | 4.7 | 11 | 2.3 | 14 | 3.5 | 10 | 1.8 | 7 | 1.2 |
| 6 | 5.25e4 | 26 | 12 | 16 | 4.9 | 11 | 2.3 | 15 | 3.7 | 10 | 1.9 | 7 | 1.3 |
| 7 | 2.09e5 | 26 | 12 | 16 | 5.1 | 11 | 2.4 | 15 | 4.0 | 10 | 2.0 | 7 | 1.3 |
| 8 | 8.37e5 | 26 | 12 | 16 | 5.1 | 11 | 2.5 | 15 | 4.1 | 10 | 2.0 | 7 | 1.3 |

Table 3. PCG iteration counts for \( V \)-cycle multigrid preconditioner with different smoothers and smoothing steps in Example 5.1.1.

| Parameters | \( d = 2 \) | \( d = 3 \) |
|------------|--------------|--------------|
| \( J \) Facet DOFs | \# It \( \kappa \) | \# It \( \kappa \) | \# It \( \kappa \) | \# It \( \kappa \) | \# It \( \kappa \) | \# It \( \kappa \) |
| 2 | 2.20e2 | 26 | 18 | 4.6 | 15 | 2.4 | 18 | 4.7 | 11 | 1.8 | 7 | 1.2 |
| 3 | 8.48e2 | 37 | 19 | 25 | 8.2 | 17 | 4.2 | 23 | 7.3 | 14 | 2.8 | 9 | 1.6 |
| 4 | 3.33e3 | 38 | 22 | 26 | 10 | 18 | 4.8 | 25 | 10 | 15 | 3.4 | 10 | 1.7 |
| 5 | 1.32e4 | 44 | 25 | 29 | 12 | 20 | 5.5 | 29 | 14 | 16 | 4.0 | 10 | 1.9 |
| 6 | 5.25e4 | 46 | 29 | 30 | 13 | 21 | 6.3 | 31 | 16 | 18 | 4.7 | 11 | 2.1 |
| 7 | 2.09e5 | 49 | 35 | 32 | 18 | 22 | 8.6 | 35 | 22 | 19 | 6.2 | 12 | 2.7 |
| 8 | 8.37e5 | 50 | 35 | 32 | 17 | 22 | 8.1 | 36 | 23 | 19 | 7.0 | 12 | 2.8 |
subdomain $\Omega^D_2/\Omega^D_3$, and $f = 0$ elsewhere. We assume a homogeneous Dirichlet boundary condition on the bottom and homogeneous Neumann boundary conditions on all other boundaries. We use the standard $V$-cycle multigrid algorithms with point Gauss–Seidel as smoother to precondition the conjugate gradient solver. The coarsest mesh is a triangulation of $\Omega$ respecting the subdomain boundaries, with the maximum element diameter not exceeding $1/4$ in both two-dimensional and three-dimensional cases.

We first test the algorithms under uniform mesh refinements, and Table 4 reports the PCG iteration counts with different mesh levels $J$, reaction coefficient $\beta$, and smoothing steps $m$ in both two-dimensional and three-dimensional cases. As observed from the results when $d = 3$, one-step and two-step Gauss–Seidel smoother are not enough to let Theorem 3.3 hold, and the iteration counts grow with increasing facet DOFs. Other results verify the robustness of our preconditioner with respect to mesh size and mesh levels when the smoothing step $m$ is large enough in the low-regularity cases. The iteration counts remain almost unchanged as $\beta$ varies. We note that when smoothing steps are the same, the iteration counts here with jump coefficients in the reaction-diffusion equation are obviously larger than in the previous case with mildly changing coefficients.

We next test with $\beta = 0$ on adaptively refined mesh using a recovery-based error estimator [18]. All other settings are the same. The right panel of Figure 1 demonstrates the finest two-dimensional mesh. Figure 2 reports the obtained PCG iteration counts, and similar results as those on uniformly-refined meshes are observed.

5.1.3. Jump diffusion coefficient on the finest mesh

We consider a reaction-diffusion equation with jump diffusion coefficient on the finest mesh level. For simplicity, we limit ourselves to the two-dimensional case and set the domain as a unit square $\Omega = [0, 1]^2$ with homogeneous Dirichlet boundary conditions on all sides. The source term is set as $f = 1$. The coarsest mesh is a structured triangulation of the domain $\Omega$ with $h = 1/4$ and the hierarchical meshes are obtained by successive uniform refinement. We let the reaction coefficient $\beta = 1$, and the diffusion coefficient $\alpha$ changes alternatively on the finest mesh level $T_J$ between the minimum value $\alpha_0$ ($\alpha_0 = 1$) and the maximum value $\alpha_1$, as demonstrated in the left panel of Figure 3. We define the ratio $\rho := \alpha_1/\alpha_0$. We note that the $\alpha_{l-1}$ in the HDG-P0 scheme (4) requires the projection of $\alpha_{l-1}$ onto piecewise constant space $W_l$ on each mesh level, therefore in this numerical experiment $\alpha_{l-1}$ becomes global constant on coarser mesh levels for $l = 1, \ldots, J - 1$ and equals $2/(\rho + 1)$.

We use the standard $V$-cycle multigrid algorithms with two-step point Gauss–Seidel as smoother to precondition the conjugate gradient solver, and Table 5 reports the PCG iteration counts and the condition number with different ratio $\rho$ and different mesh levels $J$. The robustness of our proposed multigrid method with respect to mesh size and mesh levels is observed. However, as the ratio $\rho$ increases, the iteration counts and the conditioner numbers grow and are not bounded. We refer to [43,70] for further studies on multigrid for CR discretization with jump coefficients, where conforming piecewise linear finite element was used as the auxiliary space to construct multigrid methods.
Table 4. PCG iteration counts for $V$-cycle multigrid preconditioner with Gauss–Seidel smoother and uniform mesh refinement for the reaction-diffusion equation with jump coefficients in Example 5.1.2.

| Parameters | $\beta = 1000$ | $\beta = 1$ | $\beta = 0$ |
|------------|----------------|--------------|--------------|
| $d = 2$    | $m = 1$ | $m = 2$ | $m = 4$ | $m = 1$ | $m = 2$ | $m = 4$ | $m = 1$ | $m = 2$ | $m = 4$ |
| $J$ Facet DOFs | Iteration count | Iteration count | Iteration count | Iteration count | Iteration count | Iteration count |
| 2 2.03e2 | 10 21 14 10 | 21 14 10 | 21 14 10 |
| 3 7.78e2 | 17 34 19 11 | 34 19 11 | 34 19 11 |
| 4 3.04e3 | 24 44 27 13 | 44 27 13 | 44 27 13 |
| 5 1.20e4 | 28 61 31 14 | 61 31 14 | 61 31 14 |
| 6 4.79e4 | 28 72 31 14 | 72 31 14 | 72 31 14 |
| 7 1.91e5 | 28 73 31 14 | 73 31 14 | 73 31 14 |
| 8 7.63e5 | 28 73 31 14 | 73 31 14 | 73 31 14 |

Table 5. PCG iteration counts and condition numbers for $V$-cycle multigrid preconditioner with Gauss–Seidel smoother for the reaction-diffusion equation with jump coefficients on the finest mesh level in Example 5.1.3.

| Parameters | $\rho$ |
|------------|--------|
| $J$ Facet DOFs | # It $\kappa$ | # It $\kappa$ | # It $\kappa$ | # It $\kappa$ | # It $\kappa$ |
| 2 2.03e2 | 9 1.6 16 3.8 22 3.1e1 32 2.6e3 37 2.4e5 |
| 3 7.78e2 | 10 1.7 17 4.1 28 3.3e1 39 2.8e3 49 2.4e5 |
| 4 3.04e3 | 10 1.8 17 4.3 31 3.5e1 47 3.1e3 58 2.4e5 |
| 5 1.20e4 | 11 1.8 17 4.3 31 3.5e1 54 3.2e3 69 2.5e5 |
| 6 4.79e4 | 11 1.8 17 4.3 31 3.5e1 56 3.2e3 71 2.5e5 |
| 7 1.91e5 | 11 1.8 17 4.3 31 3.5e1 56 3.4e3 72 2.5e5 |
| 8 7.63e5 | 11 1.8 17 4.3 31 3.5e1 56 3.4e3 73 2.5e5 |

5.2. Generalized Stokes equations

5.2.1. Manufactured solution

We first verify the optimal convergence rates of the HDG-P0 scheme (24) for the generalized Stokes equations with known solutions. We set the coefficients $\mu = 1$, $\beta = 10$ and the exact solution

$$
\begin{align*}
    u_x &= x^2(x - 1)^22y(1 - y)(2y - 1) \\
    u_y &= y^2(y - 1)^22x(x - 1)(2x - 1) \\
    p &= x(1 - x)(1 - y) - 1/12
\end{align*}
$$

when $d = 2$. 
Figure 2. PCG iteration counts for $V$-cycle multigrid preconditioner with Gauss–Seidel smoother and adaptive mesh refinement for the reaction-diffusion equation with jump coefficients in Example 5.1.2.

Figure 3. The diffusion coefficient on the finest mesh level in Example 5.1.3 (left), and the numerical solution with $J = 3$ and $\rho = 1e4$ (right).

and

$$u_x = x^2(x - 1)^2(2y - 6y^2 + 4y^3)(2z - 6z^2 + 4z^3)$$

$$u_y = y^2(y - 1)^2(2x - 6x^2 + 4x^3)(2z - 6z^2 + 4z^3)$$

$$u_z = -2z^2(z - 1)^2(2x - 6x^2 + 4x^3)(2y - 6y^2 + 4y^3)$$

$$p = x(1 - x)(1 - y)(1 - z) - 1/24$$

when $d = 3$, 
with all other settings the same as in Example 5.1.1. Table 6 reports the values and the corresponding EOC of the discrete $L_2$ norms $∥u_h - u∥_0$, $∥∇ \cdot u_h∥_0$, and $∥L_h - L∥_0$. Optimal convergence rates of $u_h$ and $L_h$ are observed. Since the divergence-free constraint on velocity is imposed weakly in the HDG-P0 scheme, and the normal components of $u$ are not necessarily continuous across mesh facets, the obtained $u_h$ is not exactly divergence-free and the divergence error has first order convergence rate.

We have also tested the $W$-cycle multigrid method in Algorithm 2 as the iteration solver for the problem. The coarsest mesh is a triangulation of $\Omega$ with the maximum element diameter less than $1/4$ when $d = 2$ and less than $1/2$ when $d = 3$, followed by uniform refinement. Table 7 reports the obtained iteration counts with different smoothers and smoothing steps in the multigrid algorithm, where we denote B-JAC as the damped vertex-block Jacobi smoother with the damping parameter set as $0.4$, and B-GS as the vertex-block Gauss–Seidel smoother. As observed from the results, when the smoothing steps are not large enough, the operator $E_{l,m}$ in Theorem 4.3 is no longer a reducer, and the needed smoothing steps in three dimensions for the $W$-cycle iteration to converge are larger than those in two dimensions. Other results verify the robustness of the $W$-cycle iteration solver with respect to the mesh size and mesh level.

### 5.2.2. Lid-driven cavity

In this example, we test the robustness of the proposed multigrid preconditioners for the the lid-driven cavity problem, where the computational domain is a unit square/cube $\Omega = [0, 1]^d$. We assume an inhomogeneous Dirichlet boundary condition $u = |4x(1 - x), 0|^T$ when $d = 2$, or $u = [16x(1 - x)y(1 - y), 0, 0]^T$ when $d = 3$ on the top side, and no-slip boundary conditions on the remaining domain boundaries. The source term $f = 0$. We use the PCG method preconditioned by $W$-cycle and variable $V$-cycle multigrid (with smoothing steps $m(l) = 2^{d-l}m(J)$) in Algorithm 2 to solve the augmented Lagrangian Uzawa iteration for the condensed HDG-P0 scheme (39a). We use vertex-patched block Gauss–Seidel as smoothers in multigrid algorithms to avoid damping parameters. The coarsest mesh is a triangulation of $\Omega$ with the maximum element diameter less than $1/4$ when $d = 2$, or less than $1/2$ when $d = 3$.

Table 8 reports the PCG iteration counts on uniformly refined meshes with different domain dimensions $d$, mesh levels $J$, the low order term coefficient $\beta$, and the smoothing step on the finest mesh $m(J)$. As observed from the results, when $d = 2$ and $\beta = 1000$, the solver preconditioned by the $W$-cycle algorithm with $m(J) = 2$ fails. This is due to the fact that the linear operator of the $W$-cycle multigrid with only two smoothing steps on each level becomes indefinite and cannot be used as a preconditioner, see Section 4 of [6]. On the other hand,

### Table 6. Estimated convergence rates of the HDG-P0 scheme solved by one-step augmented Lagrangian Uzawa iteration in Example 5.2.1.

| Level | $∥h - u∥_0$ | EOC | $∥∇ \cdot u_h∥_0$ | EOC | $∥L_h - L∥_0$ | EOC |
|-------|-------------|-----|-------------------|-----|--------------|-----|
| 1     | 5.87e-3     |     | 1.76e-2           |     | 3.37e-2      |     |
| 2     | 1.67e-3     | 1.81| 9.56e-3           | 0.88| 1.94e-2      | 0.79|
| 3     | 4.42e-4     | 0.95| 4.94e-3           | 0.30| 1.02e-2      | 0.93|
| 4     | 1.13e-4     | 0.99| 2.49e-3           | 0.98| 5.19e-3      | 0.98|
| 5     | 2.83e-5     | 1.00| 1.25e-3           | 1.00| 2.61e-3      | 0.99|

### Table 7. Obtained iteration counts with different settings for the lid-driven cavity problem.

| $d = 2$ | $l$ | $m$ | $Ω$ | $E_{l,m}$ | $∥h - u∥_0$ | EOC | $∥∇ \cdot u_h∥_0$ | EOC | $∥L_h - L∥_0$ | EOC |
|---------|-----|-----|-----|----------|-------------|-----|-------------------|-----|--------------|-----|
| 1       | 1.74e-3 |     | 6.53e-3 |     | 1.34e-2 |     |                   |     |              |     |
| 2       | 5.03e-4 | 1.79| 3.20e-3 | 1.03| 8.05e-3 | 0.74|                   |     |              |     |
| 3       | 1.35e-4 | 1.90| 1.59e-3 | 1.01| 4.29e-3 | 0.91|                   |     |              |     |
| 4       | 3.45e-5 | 1.96| 7.94e-4 | 1.00| 2.19e-3 | 0.97|                   |     |              |     |
| 5       | 8.69e-6 | 1.99| 3.97e-4 | 1.00| 1.10e-3 | 0.99|                   |     |              |     |
Table 7. Iteration counts of $W$-cycle multigrid iteration solver with different smoothers and smoothing steps in Example 5.2.1.

| J  | Facet DOFs | $d = 2$ | $d = 3$ |
|----|------------|---------|---------|
|    |            | B-JAC   | B-GS    |
|    |            | $m = 1$ | $m = 2$ | $m = 4$ | $m = 1$ | $m = 2$ | $m = 4$ |
| 2  | 3.76e2     | 53      | 29      | 20      | 27      | 17      | 12      |
| 3  | 1.57e3     | 22      | 20      | 17      | 17      | 13      | 11      |
| 4  | 6.40e3     | 46      | 36      | 25      | 30      | 16      | 11      |
| 5  | 2.59e4     | N/A     | 62      | 31      | 26      | 14      | 11      |
| 6  | 1.04e5     | N/A     | 78      | 36      | 33      | 15      | 10      |
| 7  | 4.17e5     | N/A     | 72      | 31      | 30      | 11      | 8       |
| 8  | 1.67e6     | N/A     | 86      | 24      | 21      | 11      | 7       |

variable $V$-cycle multigrid is more robust and works with $m(J) = 1$. Other results verify the robustness of our multigrid preconditioner with respect to mesh size and mesh levels.

5.2.3. Backward-facing step flow

Finally we test the robustness of the proposed multigrid preconditioners for the backward-facing step flow problem, with the non-convex L-shaped domain $\Omega = ([0, 5] \times [0, 0.5]) \cup ([0, 5] \times [0.5, 1])$ when $d = 2$, or $\Omega = ((([0.5, 5] \times [0, 0.5]) \cup ([0, 5] \times [0.5, 1])) \times [0, 1]$ when $d = 3$. We assume an inhomogeneous Dirichlet boundary condition $u = [16(1-y)(y-0.5), 0]^T$ when $d = 2$, or $u = [64(1-y)(y-0.5)z(1-z), 0, 0]^T$ when $d = 3$ for the inlet flow on $\{x = 0\}$, with do-nothing boundary condition on $\{x = 5\}$ and no-slip boundary conditions on the remaining sides. The maximum element diameter is less than $1/2$ in both two-dimensional and three-dimensional cases. Other settings are the same as in Example 5.2.2.

Table 9 reports the PCG iteration counts under uniform mesh refinements. Similar results are observed as in the previous example. In particular, variable $V$-cycle algorithm with $m(J) = 1$ leads to a robust preconditioner. Next we test with $\beta = 0$ on adaptively refined meshes using a recovery-based error estimator as in Example 5.1.2. Figure 4 reports the obtained PCG iteration counts, and similar results as those on the uniformly-refined meshes are observed. We note that although the multigrid analysis in Theorem 4.3 requires full elliptic regularity, our proposed multigrid method works well in this low-regularity case.

6. Conclusion

In this study, we present the lowest order HDG schemes with projected jumps and numerical integration (HDG-P0) for the reaction-diffusion equation and the generalized Stokes equations, prove their optimal $a$ priori error analysis, and construct optimal multigrid algorithms for the condensed HDG-P0 schemes for the two sets of equations on conforming simplicial meshes. The key idea of constructing the optimal multigrid is the equivalence between the condensed HDG-P0 scheme and the (slightly modified) CR discretization, which enables us to follow
Table 8. Iteration counts of the preconditioned conjugate gradient solver for the generalized Stokes equations in lid-driven cavity problems in Example 5.2.2.

| $d = 2$ | Multigrid Variable $V$ cycle | $W$ cycle |
|--------|-------------------------------|------------|
| $\beta$ | $m(J)$ | 1000 | 1 | 0 | 1000 | 1 | 0 |
| $m(J)$ | Facet DOFs | Iteration counts | |
| 2 | 3.76e2 | 13 | 10 | 12 | 10 | 12 | 10 | N/A | 8 | 10 | 8 | 10 | 8 |
| 3 | 5.17e3 | 18 | 14 | 15 | 12 | 15 | 12 | N/A | 12 | 10 | 9 | 10 | 9 |
| 4 | 6.40e3 | 20 | 16 | 17 | 13 | 17 | 13 | N/A | 11 | 11 | 9 | 11 | 9 |
| 5 | 2.59e4 | 20 | 15 | 18 | 14 | 18 | 14 | N/A | 9 | 11 | 10 | 11 | 10 |
| 6 | 1.04e5 | 20 | 15 | 19 | 15 | 19 | 15 | N/A | 9 | 11 | 9 | 11 | 9 |
| 7 | 4.17e5 | 20 | 15 | 15 | 20 | 15 | 20 | N/A | 9 | 11 | 9 | 11 | 9 |
| 8 | 1.67e6 | 21 | 15 | 21 | 15 | 21 | 15 | N/A | 9 | 11 | 9 | 11 | 9 |

| $d = 3$ | Multigrid Variable $V$ cycle | $W$ cycle |
|--------|-------------------------------|------------|
| $m(J)$ | Facet DOFs | Iteration counts | |
| 2 | 1.41e3 | 10 | 7 | 11 | 8 | 11 | 8 | 7 | 5 | 8 | 6 | 8 | 6 |
| 3 | 4.61e3 | 11 | 9 | 9 | 9 | 11 | 9 | 8 | 7 | 8 | 7 | 8 | 7 |
| 4 | 1.44e4 | 12 | 10 | 10 | 9 | 12 | 9 | 9 | 8 | 9 | 8 | 9 | 8 |
| 5 | 3.89e4 | 13 | 11 | 12 | 10 | 12 | 10 | 8 | 8 | 8 | 7 | 8 | 7 |
| 6 | 1.06e5 | 14 | 12 | 12 | 10 | 13 | 10 | 8 | 8 | 8 | 7 | 8 | 7 |
| 7 | 2.74e5 | 14 | 12 | 12 | 10 | 12 | 10 | 8 | 7 | 8 | 7 | 8 | 7 |
| 8 | 7.00e5 | 14 | 12 | 13 | 10 | 13 | 10 | 8 | 8 | 9 | 7 | 9 | 7 |

Table 9. Iteration counts of the preconditioned conjugate gradient solver for the generalized Stokes equations in backward-facing step flow problems with uniformly refined mesh in Example 5.2.3.

| $d = 2$ | Multigrid Variable $V$ cycle | $W$ cycle |
|--------|-------------------------------|------------|
| $m(J)$ | Facet DOFs | Iteration counts | |
| 2 | 3.92e2 | 9 | 7 | 9 | 7 | 9 | 7 | N/A | 4 | 6 | 5 | 6 | 6 |
| 3 | 1.65e3 | 13 | 10 | 10 | 8 | 10 | 8 | N/A | 7 | 6 | 6 | 6 | 6 |
| 4 | 6.75e3 | 16 | 13 | 11 | 8 | 11 | 8 | N/A | 10 | 6 | 6 | 6 | 6 |
| 5 | 2.73e4 | 18 | 14 | 11 | 9 | 11 | 9 | N/A | 7 | 6 | 6 | 6 | 6 |
| 6 | 1.10e5 | 17 | 12 | 12 | 9 | 12 | 9 | N/A | 6 | 6 | 6 | 6 | 6 |
| 7 | 4.41e5 | 15 | 11 | 12 | 9 | 12 | 9 | N/A | 6 | 6 | 6 | 6 | 6 |
| 8 | 1.77e6 | 13 | 10 | 12 | 9 | 12 | 9 | N/A | 6 | 6 | 6 | 6 | 6 |

| $d = 3$ | Multigrid Variable $V$ cycle | $W$ cycle |
|--------|-------------------------------|------------|
| $m(J)$ | Facet DOFs | Iteration counts | |
| 2 | 3.02e3 | 7 | 5 | 6 | 5 | 6 | 5 | 4 | 3 | 4 | 4 | 4 | 4 |
| 3 | 8.55e3 | 7 | 6 | 6 | 5 | 6 | 5 | 5 | 4 | 5 | 4 | 4 | 4 |
| 4 | 2.47e4 | 8 | 7 | 8 | 6 | 8 | 6 | 6 | 5 | 5 | 5 | 5 | 5 |
| 5 | 6.88e4 | 9 | 7 | 7 | 6 | 7 | 6 | 6 | 5 | 5 | 5 | 5 | 5 |
| 6 | 1.82e5 | 10 | 8 | 8 | 7 | 8 | 7 | 6 | 6 | 5 | 5 | 5 | 5 |
| 7 | 4.78e5 | 10 | 8 | 8 | 7 | 8 | 7 | 6 | 6 | 5 | 5 | 5 | 5 |
| 8 | 1.22e6 | 10 | 9 | 8 | 7 | 8 | 7 | 5 | 5 | 5 | 5 | 5 | 5 |
Figure 4. Iteration counts of the preconditioned conjugate gradient solver for the generalized Stokes equations in backward-facing step flow problems with adaptively refined mesh in Example 5.2.3.

the rich literature on multigrid algorithms for CR discretizations to design robust multigrid schemes for HDG-P0. Numerical experiments are presented with the proposed multigrid algorithms as both iterative solvers and preconditioners to solve the condensed HDG-P0 schemes, where optimal $L_2$ norm convergence rates are obtained, and the condition number of the preconditioned operators are bounded independent of mesh level and mesh size. We further note that the proposed $h$-multigrid algorithms for the condensed HDG-P0 schemes in this study can be used as a building block to construct robust preconditioners for higher-order HDG schemes using the auxiliary space preconditioning technique [68], which will be our forthcoming research.

Appendix A. Proof of (10c)

Let $(\psi, \phi) \in H^1(\Omega) \times (H^2(\Omega) \cap H^1_0(\Omega))$ be the solution to the following dual problem together with the regularity assumption as in (5):

$$\alpha^{-1} \psi + \nabla \phi = 0, \quad \nabla \cdot \psi + \beta \phi = e_u, \quad \text{in } \Omega. \quad (A.1)$$

We define $\Pi_V$ and $\Pi_W$ as $L_2$-projection onto the finite element space $V_h^0$ and $W_h^0$. To further simplify notation, we denote

$$\phi_h := \Pi_V \phi, \quad \psi_h := \Pi_W \psi.$$ 

Moreover, by continuity of the $L_2$-projection and the continuous dependency result on $\phi$, there holds:

$$\|\phi_h\|_{1,h} \lesssim \|\phi\|_1 \lesssim \alpha_0^{-1} \|e_u\|_0, \quad (A.2)$$
where the norm \( \| \phi_h \|_{1,h} := \sqrt{\sum_{K \in \mathcal{T}_h} \| \phi_h \|^2_{0,K} + \| \nabla \phi_h \|^2_{0,K}} \).

Using the definition of the dual problem (A.1), we have, for any \( \psi_h \in W_h \) and \( \phi_h \in V_h \),

\[
\| e_u \|^2_0 = \left( e_u, \nabla \cdot \psi - \beta \phi \right)_{\mathcal{T}_h} = \left( e_u, \nabla \cdot \left( \frac{1}{2} \psi - \psi_h \right) \right)_{\mathcal{T}_h} + \left( e_u, \beta (\phi - \phi_h) \right)_{\mathcal{T}_h} + \left( e_u, \nabla \cdot \psi_h \right)_{\mathcal{T}_h} + \left( \beta e_u, \phi_h \right)_{\mathcal{T}_h}.
\]

Taking \( T_h = \psi_h \) in (11a), we get

\[
\left( e_u, \nabla \cdot \psi_h \right)_{\mathcal{T}_h} = \left( \alpha^{-1} \xi_{\sigma}, \psi_h \right)_{\mathcal{T}_h} + \left( e \tilde{a}, \psi_h, u \right)_{\mathcal{T}_h} = \left( \alpha^{-1} \xi_{\sigma}, \psi_h - \psi \right)_{\mathcal{T}_h} - \left( \xi_{\sigma}, \nabla \phi \right)_{\mathcal{T}_h} + \left( e \tilde{a}, \left( \psi_h - \psi \right) \cdot u \right)_{\mathcal{T}_h},
\]

where we used the first equation in (A.1), single-valuedness of \( e \tilde{a} \) on the interior mesh skeleton \( E_h^o \) and \( e \tilde{a} |_{\partial \Omega} = 0 \) in the second step. Taking \( v_h = \phi_h \) in (11b), we have

\[
\left( \beta e_u, \phi_h \right)_{\mathcal{T}_h} = \left( \beta e_u, \phi_h \right)_{\mathcal{T}_h} - \sum_{K \in \mathcal{T}_h} Q^1_K (\beta e_u \phi_h)_{\mathcal{T}_h}
\]

where in the second step we used the single-valuedness of \( e \tilde{a} \cdot u \) on the interior mesh skeleton \( E_h^o \) and \( e \tilde{a} |_{\partial \Omega} = 0 \). Combing the above three equalities and simplifying, we get

\[
\| e_u \|^2_0 \leq I_1 + I_2 + I_3 + I_4
\]

Using approximation properties of the \( L_2 \)-projections and (A.2), we can bound each of the above right hand side terms by the following:

\[
I_1 \leq \left( \| \nabla e_u \|_0 + \alpha_0^{-1/2} \| \nabla^{-1/2} \xi_{\sigma} \|_0 \right) \| \psi - \psi_h \|_0 \leq \alpha_0^{-1/2} h^2 \Xi \| \psi \|_1,
\]

\[
I_2 \leq \| h^{-1/2} (e_u - \tilde{e}_u) \|_{0,\partial \mathcal{T}_h} \left( h^{-1/2} \left( \psi - \psi_h \right) \right)_{0,\partial \mathcal{T}_h} \leq \alpha_0^{-1/2} h^2 \Xi \| \psi \|_1,
\]

\[
I_3 \leq (\beta_1 \| e_u \|_0 + \| \nabla \cdot e_{\sigma} \|_0) \| \phi - \phi_h \|_0 \leq (\beta_1 \| e_u \|_0 + \alpha_1^{-1/2} \Xi) h^2 \| \phi \|_2,
\]

\[
I_4 \leq \| h^{-1/2} \Pi_0 (e_u - \tilde{e}_u) \|_{0,\partial \mathcal{T}_h} \| h^{-1/2} (\phi - \phi_h) \|_{0,\partial \mathcal{T}_h} \leq \alpha_1^{-1/2} \Xi \| \phi \|_2,
\]

\[
T_3 (\phi_h) \leq h \| \beta \|_{1,\infty} \| e_u \|_{1,h} \| \phi_h \|_0 \leq \| \beta \|_{1,\infty} \alpha_0^{-3/2} h^2 \Xi \| e_u \|_0,
\]

\[
T_1 (\phi_h) - T_2 (\phi_h) \leq h \| \beta \|_{1,\infty} \| u - u_h \|_{1,h} \| \phi_h \|_0 + h^{1+d/2} \| f - \beta u \|_{1,\infty} \| \phi_h \|_1 \leq \left( h \| \beta \|_{1,\infty} \alpha_0^{-3/2} h^2 \Theta + h^{1+d/2} \alpha_0^{-1} \| f - \beta u \|_{1,\infty} \right) \| e_u \|_0.
\]

Combining these inequalities with the regularity assumption (5), we finally get

\[
\| e_u \|^2_0 \leq h^2 \Xi \left( \alpha_0^{-1/2} \| \psi \|_1 + \alpha_1^{1/2} \| \phi \|_2 + \left( \beta_1 \| \phi \|_2 + \| \beta \|_{1,\infty} \alpha_0^{-3/2} (\Theta + \Xi) + h^{1+d/2} \alpha_0^{-1} \| f - \beta u \|_{1,\infty} \right) h^2 \| e_u \|_0.
\]
\[
\lesssim \left( c_{\text{reg}}\Xi + \frac{c_{\text{reg}}}{\alpha_0^{1/2}} \beta_1^{1/2} \Xi + \|\beta\|_{1,\infty} \alpha_0^{-3/2} (\Theta + \Xi) + h^{d/2-1} \alpha_0^{-1} \|f - \beta u\|_{1,\infty} \right) h^2 \|e_u\|_0.
\]

The estimate (10c) now following from the above inequality and (9d) by the triangle inequality. This completes the proof of (10c). \qed

Acknowledgements. We gratefully acknowledge the partial support of this work by the U.S. National Science Foundation through grant DMS-2012031.

References

[1] D.N. Arnold and F. Brezzi, Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. RAIRO Model. Math. Anal. Numér. 19 (1985) 7–32.
[2] D.N. Arnold, R.S. Falk and R. Winther, Multigrid in \(H(\text{div})\) and \(H(\text{curl})\). Numer. Math. 85 (2000) 197–217.
[3] J. Betteridge, T.H. Gibson, I.G. Graham and E.H. Müller, Multigrid preconditioners for the hybridised discontinuous Galerkin discretisation of the shallow water equations. J. Comput. Phys. 426 (2021) 34.
[4] D. Braess and R. Verfürth, Multigrid methods for nonconforming finite element methods. SIAM J. Numer. Anal. 27 (1990) 979–996.
[5] J.H. Bramble and J.E. Pasciak, New convergence estimates for multigrid algorithms. Math. Comp. 49 (1987) 311–329.
[6] J.H. Bramble, J.E. Pasciak and J. Xu, The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms. Math. Comput. 56 (1991) 1–34.
[7] C. Brennecke, A. Linke, C. Merdon and J. Schöberl, Optimal and pressure-independent \(L^2\) velocity error estimates for a modified Crouzeix–Raviart element with BDM reconstructions, in Finite volumes for Complex Applications VII. Methods and Theoretical Aspects. Vol. 77 of Springer Proc. Math. Stat. Springer, Cham (2014) 159–167.
[8] S.C. Brenner, An optimal-order multigrid method for \(P1\) nonconforming finite elements. Math. Comput. 52 (1989) 1–15.
[9] S.C. Brenner, An optimal-order nonconforming multigrid method for the biharmonic equation. SIAM J. Numer. Anal. 26 (1989) 1124–1138.
[10] S.C. Brenner, A nonconforming multigrid method for the stationary Stokes equations. Math. Comput. 55 (1990) 411–437.
[11] 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.
[12] S.C. Brenner, A nonconforming mixed multigrid method for the pure displacement problem in planar linear elasticity. SIAM J. Numer. Anal. 30 (1993) 116–135.
[13] S.C. Brenner, A nonconforming mixed multigrid method for the pure traction problem in planar linear elasticity. Math. Comput. 63 (1994) 435–460, S1–S5.
[14] S.C. Brenner, Convergence of nonconforming multigrid methods without full elliptic regularity. Math. Comput. 68 (1999) 25–53.
[15] S.C. Brenner, Poincaré–Friedrichs inequalities for piecewise \(H^1\) functions. SIAM J. Numer. Anal. 41 (2003) 306–324.
[16] S.C. Brenner, Convergence of nonconforming V-cycle and F-cycle multigrid algorithms for second order elliptic boundary value problems. Math. Comput. 73 (2004) 1041–1066.
[17] S.C. Brenner, Forty years of the Crouzeix–Raviart element. Numer. Methods Part. Differ. Equ. 31 (2015) 367–396.
[18] Z. Cai and S. Zhang, Recovery-based error estimators for interface problems: mixed and nonconforming finite elements. SIAM J. Numer. Anal. 48 (2010) 30–52.
[19] Z. Chen, Equivalence between and multigrid algorithms for nonconforming and mixed methods for second-order elliptic problems. East-West J. Numer. Math. 4 (1996) 1–33.
[20] Z. Chen and D.Y. Kwak, The analysis of multigrid algorithms for nonconforming and mixed methods for second order elliptic problems, in IMA Preprints Series, Institute for Mathematics and Its Applications, University of Minnesota (1994).
[21] H. Chen, P. Lu and X. Xu, A robust multilevel method for hybridizable discontinuous Galerkin method for the Helmholtz equation. J. Comput. Phys. 264 (2014) 133–151.
[22] L. Chen, J. Wang, Y. Wang and X. Ye, An auxiliary space multigrid preconditioner for the weak Galerkin method. Comput. Math. Appl. 70 (2015) 330–344.
[23] P.G. Ciarlet, The Finite Element Method for Elliptic Problems. Vol. 40 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2002). Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (80 #15001)].
[24] B. Cockburn, Static condensation, hybridization, and the devising of the HDG methods, in Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations. Vol. 114 of Lect. Notes Comput. Sci. Eng. Springer, Cham (2016) 129–177.
[25] B. Cockburn, Discontinuous Galerkin methods for computational fluid dynamics, in Encyclopedia of Computational Mechanics, 2nd edition. John Wiley & Sons, Ltd (2017) 1–63.
[26] B. Cockburn, J. Gopalakrishnan and R. Lazarov, Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.* **47** (2009) 1319–1365.

[27] B. Cockburn, O. Duque, J. Gopalakrishnan and S. Tan, Multigrid for an HDG method. *IMA J. Numer. Anal.* **34** (2014) 1386–1425.

[28] B. Cockburn, N.C. Nguyen and J. Peraire, HDG methods for hyperbolic problems, in *Handbook of Numerical Methods for Hyperbolic Problems*, Vol. 17 of *Handb. Numer. Anal.* Elsevier/North-Holland, Amsterdam (2016) 173–197.

[29] D.A. Di Pietro, F. Hülsemann, P. Matalon, P. Mycek, U. Rüde and D. Ruiz, An h-multigrid method for hybrid high-order discretizations. *SIAM J. Sci. Comput.* **43** (2021) S839–S861.

[30] D.A. Di Pietro, F. Hülsemann, P. Matalon, P. Mycek, U. Rüde and D. Ruiz. Towards robust, fast solutions of elliptic equations on complex domains through hybrid high-order discretizations and non-nested multigrid methods. *Int. J. Numer. Methods Eng.* **122** (2021) 6576–6595.

[31] H.-Y. Duan, S.-Q. Gao, R.C.E. Tan and S. Zhang, A generalized BPX multigrid framework covering nonnested V-cycle methods. *Math. Comput.* **76** (2007) 137–152.

[32] M.S. Fabien, M.G. Knepley, R.T. Mills and B.M. Rivière, Manycore parallel computing for a hybridizable discontinuous Galerkin nested multigrid method. *SIAM J. Sci. Comput.* **41** (2019) C73–C96.

[33] P.E. Farrell, L. Mitchell and F. Wechsung, An augmented Lagrangian preconditioner for the 3D stationary incompressible Navier–Stokes equations at high Reynolds number. *SIAM J. Sci. Comput.* **41** (2019) A3073–A3096.

[34] P.E. Farrell, L. Mitchell, L.R. Scott and F. Wechsung, Robust multigrid methods for nearly incompressible elasticity using macro elements. *IMA J. Numer. Anal.* **42** (2021) 3386–3399.

[35] P. Fernández, A. Christophe, S. Terrana, N.C. Nguyen and J. Peraire, Hybridized discontinuous Galerkin methods for wave propagation. *J. Sci. Comput.* **77** (2018) 1566–1604.

[36] M. Fortin and R. Glowinski, Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary Value Problems, translated from the French by B. Hunt and D. C. Spicer. Vol. 15 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam (1983).

[37] G. Fu, C. Lehrenfeld, A. Linke and T. Streckenbach, Locking-free and gradient-robust $H(div)$-conforming HDG methods for linear elasticity. *J. Sci. Comput.* **86** (2021) 30.

[38] W. Hackbusch, Multigrid Methods and Applications. Vol. 4 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin (1985).

[39] Q. Hong, J. Kraus, J. Xu and L. Zikatanov, A robust multigrid method for discontinuous Galerkin discretizations of Stokes and linear elasticity equations. *Numer. Math.* **132** (2016) 23–49.

[40] X. Huang, Nonconforming finite element stokes complexes in three dimensions. *Sci. China Math.* (2023). DOI: 10.1007/s11425-021-2026-7.

[41] V. John, A. Linke, C. Merdon, M. Neilan and L.G. Rebholz, On the divergence constraint in mixed finite element methods for incompressible flows. *SIAM Rev.* **59** (2017) 492–544.

[42] G. Kanschat and Y. Mao, Multigrid methods for $H(div)$-conforming discontinuous Galerkin methods for the Stokes equations. *J. Numer. Math.* **23** (2015) 51–66.

[43] T.V. Kolev, J. Xu and Y. Zhu, Multilevel preconditioners for reaction-diffusion problems with discontinuous coefficients. *J. Sci. Comput.* **67** (2016) 324–350.

[44] Y.-J. Lee, J. Wu, J. Xu and L. Zikatanov, Robust subspace correction methods for nearly singular systems. *Math. Models Methods Appl. Sci.* **17** (2007) 1937–1963.

[45] Y.-J. Lee, J. Wu and J. Chen, Robust multigrid method for the planar linear elasticity problems. *Numer. Math.* **113** (2009) 473–496.

[46] C. Lehrenfeld, *Hybrid Discontinuous Galerkin methods for solving incompressible flow problems*. Ph.D. thesis, RWTH Aachen University (2010).

[47] A. Linke, On the role of the Helmholtz decomposition in mixed methods for incompressible flows and a new variational crime. *Comput. Methods Appl. Mech. Eng.* **268** (2014) 782–800.

[48] A. Linke and C. Merdon, Pressure-robustness and discrete Helmholtz projectors in mixed finite element methods for the incompressible Navier–Stokes equations. *Comput. Methods Appl. Mech. Eng.* **311** (2016) 304–326.

[49] P. Lu, A. Rupp and G. Kanschat, Homogeneous multigrid for HDG. *IMA J. Numer. Anal.* **42** (2021) 3135–3153.

[50] L.D. Marini, An inexpensive method for the evaluation of the solution of the lowest order Raviart–Thomas mixed method. *SIAM J. Numer. Anal.* **22** (1985) 493–496.

[51] L.S.D. Morley, The triangular equilibrium element in the solution of plate bending problems. *Aeronaut. Q.* **19** (1968) 149–169.

[52] N.C. Nguyen and J. Peraire, Hybridizable discontinuous Galerkin methods for partial differential equations in continuum mechanics. *J. Comput. Phys.* **231** (2012) 5955–5988.

[53] Y. Notay, A new algebraic multigrid approach for Stokes problems. *Numer. Math.* **132** (2016) 51–84.

[54] Y. Notay, Algebraic multigrid for Stokes equations. *SIAM J. Sci. Comput.* **39** (2017) S88–S111.

[55] I. Oikawa, A hybridized discontinuous Galerkin method with reduced stabilization. *J. Sci. Comput.* **65** (2015) 327–340.

[56] I. Oikawa, Analysis of a reduced-order HDG method for the Stokes equations. *J. Sci. Comput.* **67** (2016) 475–492.

[57] W. Qiu and K. Shi, An HDG method for convection diffusion equation. *J. Sci. Comput.* **66** (2016) 346–357.

[58] W. Qiu and K. Shi, A superconvergent HDG method for the incompressible Navier–Stokes equations on general polyhedral meshes. *IMA J. Numer. Anal.* **36** (2016) 1943–1967.
[59] W. Qiu, J. Shen and K. Shi, An HDG method for linear elasticity with strong symmetric stresses. *Math. Comput.* 87 (2018) 69–93.

[60] J. Schöberl, Multigrid methods for a parameter dependent problem in primal variables. *Numer. Math.* 84 (1999) 97–119.

[61] J. Schöberl, *Robust multigrid methods for parameter dependent problems*. Ph.D. thesis, Johannes Kepler University Linz (1999).

[62] J. Schöberl, *C++11 implementation of finite elements in NGSolve*. ASC Report 30/2014, Institute for Analysis and Scientific Computing, Vienna University of Technology (2014).

[63] R. Stevenson, Nonconforming finite elements and the cascadic multi-grid method. *Numer. Math.* 91 (2002) 351–387.

[64] S. Tan, *Iterative solvers for hybridized finite element methods*. Ph.D. thesis, University of Florida (2009).

[65] S. Turek, Multigrid techniques for a divergence-free finite element discretization. *East-West J. Numer. Math.* 2 (1994) 229–255.

[66] H. Uzawa, Iterative methods for concave programming, in Studies in Linear and Non-Linear Programming, Stanford University Press, Stanford, CA (1958) 154–165.

[67] T. Wildey, S. Muralikrishnan and T. Bui-Thanh, Unified geometric multigrid algorithm for hybridized high-order finite element methods. *SIAM J. Sci. Comput.* 41 (2019) S172–S195.

[68] J. Xu, The auxiliary space method and optimal multigrid preconditioning techniques for unstructured grids. *Computing* 56 (1996) 215–235.

[69] B. Zheng, Q. Hu and J. Xu, A nonconforming finite element method for fourth order curl equations in $\mathbb{R}^3$. *Math. Comput.* 80 (2011) 1871–1886.

[70] Y. Zhu, Analysis of a multigrid preconditioner for Crouzeix–Raviart discretization of elliptic partial differential equation with jump coefficients. *Numer. Linear Algebra Appl.* 21 (2014) 24–38.

---

Please help to maintain this journal in open access!

This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at https://edpsciences.org/en/subscribe-to-open-s2o.