Available online at www.sciencedirect.com
ScienceDirect
Computer Methods in Applied Mechanics and Engineering 350 (2019) 315–334
Computer methods in applied mechanics and engineering

A Hybridized Discontinuous Galerkin Method for A Linear Degenerate Elliptic Equation Arising from Two-Phase Mixtures

Shinhoo Kang[a], Tan Bui-Thanh[b], Todd Arbogast[c]

[a]*Department of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin, Austin, TX 78712, USA.*
[b]*Department of Aerospace Engineering and Engineering Mechanics, and Institute for Computational Engineering & Sciences, The University of Texas at Austin, Austin, TX 78712, USA.*
[c]*Department of Mathematics, and Institute for Computational Engineering & Sciences, The University of Texas at Austin, Austin, TX 78712*

Abstract

We develop a high-order hybridized discontinuous Galerkin (HDG) method for a linear degenerate elliptic equation arising from a two-phase mixture of mantle convection or glacier dynamics. We show that the proposed HDG method is well-posed by using an energy approach. We derive *a priori* error estimates for the method on simplicial meshes in both two- and three-dimensions. The error analysis shows that the convergence rates are optimal for both the scaled pressure and the scaled velocity for non-degenerate problems and are sub-optimal by half order for degenerate ones. Several numerical results are presented to confirm the theoretical estimates. Degenerate problems with low regularity solutions are also studied, and numerical results show that high-order methods are beneficial in terms of accuracy.

Keywords: discontinuous Galerkin methods, hybridization, degenerate elliptic equation, two-phase mixtures, error estimates

1. Introduction

The Earth's core is hotter than the Earth's surface, which leads to thermal convection in which the cold mantle is dense and sinks while the hot mantle is light and rises to the surface. The induced current, i.e., mantle convection, moves slowly and cools gradually. The evolution and circulation of the mantle induce plate tectonics, volcanic activity, and variation in crustal chemical composition. Therefore, the study of mantle dynamics is critical to understanding how the planet functions [1]. Glacier dynamics, on the other hand, describes the movement of glaciers and ice sheets. Glaciers and ice sheets interact with the atmosphere, the oceans, and the landscape [2], which could lead to a large impact on weather and climate change [3]. Though mantle convection and glacier dynamics are different in nature, their dynamics can be mathematically modeled by the Stokes equations combined with a Darcy equation accounting for melt.

In this paper, we are interested in developing numerical methods for both glacial dynamics and mantle convection described by a similar two-phase mathematical model, as we now briefly discuss. In glacial dynamics, the mixture of ice and water is observed near the temperature at the pressure-melting point, which is in a phase-transition process [4, 5]. In mantle dynamics, a partially molten rock is generated by supplying heat or reducing the pressure. In both cases, the relative motion between the melt and the solid matrix is modeled by two-phase flow [6].

We adopt the mathematical model in [7, 6, 8, 9, 10, 11, 12]. In particular, the mixture parameter of fluid melt and solid matrix is described by the porosity ϕ—the relative volume of fluid melt with respect to the bulk volume—which separates the solid single-phase ($\phi = 0$) and fluid-solid two-phase ($\phi > 0$) regions [13]. The partially molten regions are governed by Darcy flow through a deformable solid matrix which is modeled as a highly viscous Stokes fluid

[14, 12]. We use the subscript f and s to distinguish between the fluid melt and solid matrix, and boldface lowercase letters for vector-valued functions. We denote by $\mathbf{v}_f$ and $\mathbf{v}_s$ the velocities of fluid and solid, $\tilde{p}_f$ and $\tilde{p}_s$ the pressures, ρ_f and ρ_s the densities, $\hat{q}_f$ and $\hat{q}_s$ the viscosities, and σ_f and σ_s the stresses. Darcy's law [15, 6] states

$$\phi\left(\mathbf{v}_f - \mathbf{v}_s\right) = -\frac{\kappa(\phi)}{\hat{q}_f}\left(\nabla\tilde{p}_f - \rho_f\mathbf{g}\right), \tag{1}$$

where $\kappa(\phi)$ is the permeability with $\kappa(0) = 0$ and $\mathbf{g}$ is the gravity. We assume that the solid matrix is more viscous than the fluid melt ($\hat{q}_f \ll \hat{q}_s$) so that the fluid and the solid stresses can be modeled as

$$\sigma_f := -\tilde{p}_f\mathcal{I}, \tag{2}$$

$$\sigma_s := -\tilde{p}_s\mathcal{I} + \hat{q}_s\left(\nabla\mathbf{v}_s + \nabla\mathbf{v}_s^T\right) - \frac{2}{3}\hat{q}_s\nabla\cdot\mathbf{v}_s\mathcal{I}, \tag{3}$$

where $\mathcal{I}$ is the second order identity tensor. The mixture of the melt and the solid matrix obeys the Stokes equation [10, 12]

$$\nabla\cdot\left(\phi\sigma_f + (1-\phi)\sigma_s\right) + \left(\phi\rho_f + (1-\phi)\rho_s\right)\mathbf{g} = 0. \tag{4}$$

The mass conservations of the fluid melt and the solid matrix are given as [6, 10]

$$\frac{\partial\left(\rho_f\phi\right)}{\partial t} + \nabla\cdot\left(\rho_f\phi\mathbf{v}_f\right) = 0, \tag{5}$$

$$\frac{\partial(\rho_s(1-\phi))}{\partial t} + \nabla\cdot(\rho_s(1-\phi)\mathbf{v}_s) = 0. \tag{6}$$

Applying a Boussinesq approximation (constant and equal densities for non-buoyancy terms) to (5)–(6), the total mass conservation of the mixture can be written as

$$\nabla\cdot\left(\phi\mathbf{v}_f + (1-\phi)\mathbf{v}_s\right) = 0. \tag{7}$$

The pressure jump between the melt and the matrix phases (the compaction relation) is given by [16, 11]

$$\left(\tilde{p}_s - \tilde{p}_f\right) = -\frac{\hat{q}_s}{\phi}\nabla\cdot\mathbf{v}_s. \tag{8}$$

The coupled Darcy-Stokes system (1), (4), (7) and (8) describes the motion of the mantle flow (and glacier dynamics). The challenge is when $\phi = 0$. Since solid matrix always exists, the Stokes part is well-posed, but the Darcy part is degenerate when $\phi = 0$.

In this paper, we shall focus on addressing the challenge of solving the linear degenerate elliptic equation arising from the Darcy part of the system. With a change of variables, (1) and a combination of (7)–(8) become

$$\tilde{\mathbf{v}} + d(\phi)^2(\nabla\tilde{p} - \tilde{\mathbf{g}}) = 0, \qquad \text{in } \Omega, \tag{9a}$$

$$\nabla\cdot\tilde{\mathbf{v}} + \phi\tilde{p} = \phi^{\frac{1}{2}}f, \qquad \text{in } \Omega, \tag{9b}$$

$$\phi\tilde{p} = \phi^{\frac{1}{2}}g_D, \qquad \text{on } \Gamma_D, \tag{9c}$$

where $\Omega \subset \mathbb{R}^{dim}, dim = 2$ or 3, is an open and bounded domain, $\Gamma_D = \partial\Omega$ the Dirichlet boundary, g_D the Dirichlet data, $\mathbf{n}$ the outward unit normal vector, $\tilde{\mathbf{v}} = \phi(\mathbf{v}_f - \mathbf{v}_s)$ the Darcy velocity, $\tilde{p} = \tilde{p}_f$ the fluid pressure, $\tilde{\mathbf{g}} = \rho_f\mathbf{g}$, $f = \phi^{\frac{1}{2}}\tilde{p}_s$, $\hat{q}_s = 1$ and $d(\phi) = \sqrt{\frac{\kappa(\phi)}{\hat{q}_f}}$. Though d is a function of ϕ, we shall write d instead of $d(\phi)$ for the simplicity of the exposition.

The boundary value problem (9) has been studied in [17], where the scaled velocity and pressure were proposed in order to obtain well-posedness. For numerical implementation, a cell-centered finite difference method [18] and a mixed finite element method [17] have been studied. The results showed that the numerical schemes are stable and have an optimal convergence rate for smooth solutions. However, these schemes are low order accurate approaches.

Meanwhile, the high-order discontinuous Galerkin (DG) method—originally developed [19, 20, 21] for the neutron transport equation—has been studied extensively for virtually all types of partial differential equations (PDEs) [22, 23, 24, 25, 26]. This is due to the fact that DG combines advantages of finite volume and finite element methods. As such, it is well-suited to problems with large gradients including shocks and with complex geometries, and large-scale simulations demanding parallel implementations. In particular, for numerical modeling of magma dynamics, the DG methods have been used to study the interaction between the fluid melt and the solid matrix [27, 28], and to include a porosity-dependent bulk viscosity and a solid upwelling effect [29]. In spite of these advantages, DG methods for steady state and/or time-dependent problems that require implicit time-integrators are more expensive in comparison to other existing numerical methods, since DG typically has many more (coupled) unknowns.

As an effort to mitigate the computational expense associated with DG methods, the hybridized (also known as hybridizable) discontinuous Galerkin (HDG) methods are introduced for various types of PDEs including Poisson-type equation [30, 31, 32, 33, 34, 35], Stokes equation [36, 37], Euler and Navier-Stokes equations, wave equations [38, 39, 40, 41, 42, 43, 44, 45], to name a few. In [46, 47, 48], one of the authors has proposed an upwind HDG framework that provides a unified and systematic construction of HDG methods for a large class of PDEs. We note that the weak Galerkin (WG) methods in [49, 50, 51, 52] share many similar advantages with HDG. In fact, HDG and WG are the same for the degenerate elliptic problem in this paper.

The main goal of this paper is to develop a high-order HDG scheme for the linear degenerate elliptic equation (9). In section 2, we briefly discuss the scaled system for (9). In section 3, we derive the HDG formulation for the scaled system based on the upwind HDG framework. The key feature is that we have modified the upwind HDG flux to accommodate the degenerate regions. When the porosity vanishes, the resulting HDG system becomes ill-posed because the upwind parameter associated with the HDG flux disappears. To overcome the difficulty, we introduce a generalized stabilization parameter that is an extension of the upwind based stabilization parameter. It has positive values on the degenerate interfaces. Next, we show the well-posedness and error analysis of the HDG system under the assumption that the grid well matches with the intersection between the fluid melt and the solid matrix. In section 4, various numerical results for the scaled system will be presented to confirm the accuracy and robustness of the proposed HDG scheme. Finally, we conclude the paper and discuss future research directions in section 5.

2. Handling the degeneracy

Let $(\cdot,\cdot)_\Omega$ be the L^2 inner-product on Ω, and $\langle\cdot,\cdot\rangle_{\partial\Omega}$ be the L^2 inner-product on $\partial\Omega$. We denote the L^2 norm by $\|\cdot\|_\Omega = (\cdot,\cdot)_\Omega^{\frac{1}{2}}$ on Ω and by $\|\cdot\|_{\partial\Omega} = \langle\cdot,\cdot\rangle_{\partial\Omega}^{\frac{1}{2}}$ on $\partial\Omega$. We also define the weighted L^2 norm on $\partial\Omega$ by $\|\cdot\|_{\partial\Omega,\tau} = \langle|\tau|\cdot,\cdot\rangle_{\partial\Omega}^{\frac{1}{2}} = \left(\int_{\partial\Omega}|\tau|(\cdot)^2 dx\right)^{\frac{1}{2}}$. For any $s \neq 0$, we denote the $H^s(D)$-norm as $\|\cdot\|_{s,D}$, for example, $\|\cdot\|_{\frac{1}{2},\partial\Omega}$ is the norm of $H^{\frac{1}{2}}(\partial\Omega)$.

2.1. The scaled system

When the porosity becomes zero, the system (9) degenerates. However, we can still investigate how the solutions behave as the porosity vanishes. According to [17], a priori energy estimates for the system (9) read as

$$\|d^{-1}\tilde{\mathbf{v}}\|_\Omega + \|\phi^{\frac{1}{2}}\tilde{p}\|_\Omega + \|\phi^{-\frac{1}{2}}\nabla\cdot\tilde{\mathbf{v}}\|_\Omega \le c\left(\|g_D\|_{H^{\frac{1}{2}}(\partial\Omega)} + \|d\tilde{\mathbf{g}}\|_\Omega + \|f\|_\Omega\right), \tag{10}$$

for some constant $c > 0$. Note that we may lose control of the pressure $\tilde{p}$ as the porosity approaches zero.

To have the control of the pressure, following [17], we define the scaled velocity and the scaled pressure as $\mathbf{u} = d^{-1}\tilde{\mathbf{v}}$ and $p = \phi^{\frac{1}{2}}\tilde{p}$, respectively. The system (9) becomes

$$\mathbf{u} + d\nabla\left(\phi^{-\frac{1}{2}}p\right) = d\tilde{\mathbf{g}}, \quad \text{in } \Omega, \tag{11a}$$

$$\phi^{-\frac{1}{2}}\nabla\cdot(d\mathbf{u}) + p = f, \quad \text{in } \Omega, \tag{11b}$$

$$p = g_D, \quad \text{on } \Gamma_D. \tag{11c}$$

Here, we interpret the differential operators in (11) as

$$d\nabla(\phi^{-\frac{1}{2}}p) = -\frac{1}{2}\phi^{-\frac{3}{2}}d\nabla\phi p + \phi^{-\frac{1}{2}}d\nabla p, \tag{12}$$

$$\phi^{-\frac{1}{2}}\nabla\cdot(d\mathbf{u}) = \phi^{-\frac{1}{2}}\nabla d\cdot\mathbf{u} + \phi^{-\frac{1}{2}}d\nabla\cdot\mathbf{u}, \tag{13}$$

where we assume that

$$\phi^{-\frac{1}{2}} d \in L^\infty(\Omega), \tag{14a}$$

$$\phi^{-\frac{1}{2}} \nabla d \in (L^\infty(\Omega))^{dim}, \tag{14b}$$

$$\phi^{-\frac{3}{2}} d \nabla \phi \in (L^\infty(\Omega))^{dim}. \tag{14c}$$

With the assumption (14), the scaled system does not degenerate. If the porosity vanishes, then $d(\phi) = 0$, which leads to $\mathbf{u} = 0$ and $p = f$. The energy estimates for the scaled system (11) read as

$$\|u\|_\Omega + \|p\|_\Omega + \|\phi^{-\frac{1}{2}} \nabla \cdot (d\mathbf{u})\|_\Omega \leq c \left(\|g_D\|_{H^{\frac{1}{2}}(\partial\Omega)} + \|d\tilde{\mathbf{g}}\|_\Omega + \|f\|_\Omega \right), \tag{15}$$

for some constant $c > 0$ [17]. We clearly see that we have control of the scaled pressure p even when the porosity becomes zero.

2.2. Upwind-based HDG flux

With some simple manipulation, the scaled system (11) can be rewritten as

$$\mathbf{u} - \phi^{-\frac{1}{2}} \nabla d p + \nabla \cdot \left(\phi^{-\frac{1}{2}} d p \mathcal{I} \right) = d\tilde{\mathbf{g}}, \quad \text{in } \Omega, \tag{16a}$$

$$\frac{1}{2} \phi^{-\frac{3}{2}} d \nabla \phi \cdot \mathbf{u} + p + \nabla \cdot \left(\phi^{-\frac{1}{2}} d \mathbf{u} \right) = f, \quad \text{in } \Omega. \tag{16b}$$

We cast the scaled system (16) into the conservative form

$$\nabla \cdot \mathcal{F}(\mathbf{r}) + \mathcal{G}\mathbf{r} = \mathbf{f}, \quad \text{in } \Omega, \tag{17}$$

where we have defined the solution vector $\mathbf{r} := (u_1, u_2, u_3, p)$, the source vector $\mathbf{f} := (d\tilde{g}_1, d\tilde{g}_2, d\tilde{g}_3, f)$, the flux tensor

$$\mathcal{F} := (F_1, F_2, F_3) := \mathcal{F}(\mathbf{r}) := \phi^{-\frac{1}{2}} d \begin{pmatrix} p & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & p \\ u_1 & u_2 & u_3 \end{pmatrix} \tag{18}$$

and

$$\mathcal{G} := \begin{pmatrix} 1 & 0 & 0 & -\phi^{-\frac{1}{2}} \frac{\partial d}{\partial x} \\ 0 & 1 & 0 & -\phi^{-\frac{1}{2}} \frac{\partial d}{\partial y} \\ 0 & 0 & 1 & -\phi^{-\frac{1}{2}} \frac{\partial d}{\partial z} \\ \frac{1}{2}\phi^{-\frac{3}{2}} d \frac{\partial \phi}{\partial x} & \frac{1}{2}\phi^{-\frac{3}{2}} d \frac{\partial \phi}{\partial y} & \frac{1}{2}\phi^{-\frac{3}{2}} d \frac{\partial \phi}{\partial z} & 1 \end{pmatrix}. \tag{19}$$

We define

$$\mathcal{A} = \sum_{k=1}^{3} n_k \frac{\partial F_k}{\partial \mathbf{r}} = \phi^{-\frac{1}{2}} d \begin{pmatrix} 0 & 0 & 0 & n_1 \\ 0 & 0 & 0 & n_2 \\ 0 & 0 & 0 & n_3 \\ n_1 & n_2 & n_3 & 0 \end{pmatrix}, \tag{20}$$

which has four eigenvalues $\left(-\phi^{-\frac{1}{2}} d, 0, 0, \phi^{-\frac{1}{2}} d \right)$ and distinct eigenvectors

$$W_1 = \begin{pmatrix} -n_1 \\ -n_2 \\ -n_3 \\ 1 \end{pmatrix}, W_2 = \begin{pmatrix} -n_2 \\ n_1 \\ 0 \\ 0 \end{pmatrix}, W_3 = \begin{pmatrix} -n_3 \\ 0 \\ n_1 \\ 0 \end{pmatrix}, \text{ and } W_4 = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ 1 \end{pmatrix}. \tag{21}$$

The system (17) can be considered as a steady state hyperbolic system [53]. Finally, following the upwind HDG framework in [46] we can construct the upwind HDG flux with scalar $\hat{p}$ and vector $\hat{\mathbf{u}}$ trace unknowns as

$$\hat{\mathcal{F}}(\hat{\mathbf{r}})\cdot\mathbf{n} = \phi^{-\frac{1}{2}} d \begin{pmatrix} n_1\hat{p} \\ n_2\hat{p} \\ n_3\hat{p} \\ \hat{\mathbf{u}}\cdot\mathbf{n} \end{pmatrix} := \mathcal{F}(\mathbf{r})\cdot\mathbf{n} + |\mathcal{A}|(\mathbf{r}-\hat{\mathbf{r}}) = \phi^{-\frac{1}{2}} d \begin{pmatrix} n_1\left(p+(\mathbf{u}-\hat{\mathbf{u}})\cdot\mathbf{n}\right) \\ n_2\left(p+(\mathbf{u}-\hat{\mathbf{u}})\cdot\mathbf{n}\right) \\ n_3\left(p+(\mathbf{u}-\hat{\mathbf{u}})\cdot\mathbf{n}\right) \\ \mathbf{u}\cdot\mathbf{n}+(p-\hat{p}) \end{pmatrix}, \tag{22}$$

where $\hat{\mathbf{r}} = (\hat{u}_1, \hat{u}_2, \hat{u}_3, \hat{p})$, $|\mathcal{A}| := W|D|W^{-1}$, D is the diagonal matrix of eigenvalues of W_1, W_2, W_3 and W_4, and W is the matrix of corresponding eigenvectors. Following [46], we can compute $\hat{\mathbf{u}}\cdot\mathbf{n}$ as a function of $\hat{p}$, and hence $\hat{\mathbf{u}}\cdot\mathbf{n}$ can be eliminated. The upwind HDG flux can be then written in terms of $\hat{p}$ as

$$\hat{\mathcal{F}}(\hat{\mathbf{r}})\cdot\mathbf{n} = \phi^{-\frac{1}{2}} d \begin{pmatrix} n_1\hat{p} \\ n_2\hat{p} \\ n_3\hat{p} \\ \mathbf{u}\cdot\mathbf{n}+(p-\hat{p}) \end{pmatrix}. \tag{23}$$

3. HDG formulation

We denote by $\Omega_h := \cup_{i=1}^{N_e} K_i$ the mesh containing a finite collection of non-overlapping elements, K_i, that partition Ω. Here, h is defined as $h := \max_{j\in\{1,\dots,N_e\}} diam\left(K_j\right)$. Let $\partial\Omega_h := \{\partial K : K \in \Omega_h\}$ be the collection of the boundaries of all elements. Let us define $\mathcal{E}_h := \mathcal{E}_h^o \cup \mathcal{E}_h^\partial$ as the skeleton of the mesh which consists of the set of all uniquely defined faces/interfaces, where $\mathcal{E}_h^\partial$ is the set of all boundary faces on $\partial\Omega$, and $\mathcal{E}_h^o = \mathcal{E}_h \setminus \mathcal{E}_h^\partial$ is the set of all interior interfaces. For two neighboring elements K^+ and K^- that share an interior interface $e = K^+ \cap K^-$, we denote by $q^\pm$ the trace of the solutions on e from $K^\pm$. We define $\mathbf{n}^-$ as the unit outward normal vector on the boundary ∂K^- of element K^-, and $\mathbf{n}^+ = -\mathbf{n}^-$ the unit outward normal of a neighboring element K^+. On the interior interfaces $e \in \mathcal{E}_h^o$, we define the mean/average operator $\{\!\{\mathbf{v}\}\!\}$, where $\mathbf{v}$ is either a scalar or a vector quantify, as $\{\!\{\mathbf{v}\}\!\} := \left(\mathbf{v}^- + \mathbf{v}^+\right)/2$, and the jump operator $[\![\mathbf{v}\cdot\mathbf{n}]\!] := \mathbf{v}^+\cdot\mathbf{n}^+ + \mathbf{v}^-\cdot\mathbf{n}^-$ On the boundary faces $e \in \mathcal{E}_h^\partial$, we define the mean and jump operators as $\{\!\{\mathbf{v}\}\!\} := \mathbf{v}, \quad [\![\mathbf{v}]\!] := \mathbf{v}$.

Let $\mathcal{P}^k(D)$ denote the space of polynomials of degree at most k on a domain D. Next, we introduce discontinuous piecewise polynomial spaces for scalars and vectors as

$$\begin{aligned}
V_h(\Omega_h) &:= \left\{v \in L^2(\Omega_h) :\ v|_K \in \mathcal{P}^k(K), \forall K \in \Omega_h\right\}, \\
\Lambda_h(\mathcal{E}_h) &:= \left\{\lambda \in L^2(\mathcal{E}_h) :\ \lambda|_e \in \mathcal{P}^k(e), \forall e \in \mathcal{E}_h\right\}, \\
\mathbf{V}_h(\Omega_h) &:= \left\{\mathbf{v} \in \left[L^2(\Omega_h)\right]^m :\ \mathbf{v}|_K \in \left[\mathcal{P}^k(K)\right]^m, \forall K \in \Omega_h\right\}, \\
\mathbf{\Lambda}_h(\mathcal{E}_h) &:= \left\{\boldsymbol{\lambda} \in \left[L^2(\mathcal{E}_h)\right]^m :\ \boldsymbol{\lambda}|_e \in \left[\mathcal{P}^k(e)\right]^m, \forall e \in \mathcal{E}_h\right\}.
\end{aligned}$$

and similar spaces $V_h(K)$, $\Lambda_h(e)$, $\mathbf{V}_h(K)$, and $\mathbf{\Lambda}_h(e)$ by replacing Ω_h with K and $\mathcal{E}_h$ with e. Here, m is the number of components of the vector under consideration.

We define the broken inner products as $(\cdot,\cdot)_{\Omega_h} := \sum_{K\in\Omega_h}(\cdot,\cdot)_K$ and $\langle\cdot,\cdot\rangle_{\partial\Omega_h} := \sum_{\partial K\in\partial\Omega_h}\langle\cdot,\cdot\rangle_{\partial K}$, and on the mesh skeleton as $\langle\cdot,\cdot\rangle_{\mathcal{E}_h} := \sum_{e\in\mathcal{E}_h}\langle\cdot,\cdot\rangle_e$. We also define the associated norms as $\|\cdot\|_{\Omega_h} := \left(\sum_{K\in\Omega_h}\|\cdot\|_K^2\right)^{\frac{1}{2}}$, $\|\cdot\|_{\partial\Omega_h} := \left(\sum_{K\in\Omega_h}\|\cdot\|_{\partial K}^2\right)^{\frac{1}{2}}$, and the weighted norm $\|\cdot\|_{\partial\Omega_h,\tau} := \left(\sum_{K\in\Omega_h}\|\cdot\|_{\partial K,\tau}^2\right)^{\frac{1}{2}}$ (recall $\|\cdot\|_{\partial K,\tau} = \||\tau|^{\frac{1}{2}}(\cdot)\|_{\partial K}$).

3.1. Weak form

From now on, *we conventionally use $\mathbf{u}^e$, p^e and $\hat{p}^e$ for the exact solution while $\mathbf{u}$, p and $\hat{p}$ are used to denote the HDG solution.* Unlike the DG approach, in which $\hat{p}$ on an interface is computed using information from neighboring elements that share that interface, i.e.,

$$\hat{p} = \frac{1}{2}[\![\mathbf{u}\cdot\mathbf{n}]\!] + \{\!\{p\}\!\}, \tag{24}$$

the idea behind HDG is to treat $\hat{p}$ as a new unknown. Testing (16) or (17) with $(\mathbf{v}, q)$ and integrating by parts we obtain the local solver for each element by replacing the flux $\langle \mathcal{F}\cdot\mathbf{n}, (\mathbf{v},q)\rangle_{\partial K}$ with the HDG numerical flux $\left\langle \hat{\mathcal{F}}\cdot\mathbf{n}, (\mathbf{v},q)\right\rangle_{\partial K}$. The local solver reads: find $(\mathbf{u}, p, \hat{p}) \in \mathbf{V}_h(K)\times V_h(K)\times \Lambda_h(\partial K)$ such that

$$(\mathbf{u},\mathbf{v})_K - \left(\phi^{-\frac{1}{2}}\nabla d p, \mathbf{v}\right)_K - \left(\phi^{-\frac{1}{2}} d p, \nabla\cdot\mathbf{v}\right)_K + \left\langle \phi^{-\frac{1}{2}} d\hat{p}, \mathbf{v}\cdot\mathbf{n}\right\rangle_{\partial K} = (d\tilde{\mathbf{g}}, \mathbf{v})_K\,, \tag{25a}$$

$$(p,q)_K + \left(\frac{1}{2}\phi^{-\frac{3}{2}} d\nabla\phi\cdot\mathbf{u}, q\right)_K - \left(\phi^{-\frac{1}{2}} d\mathbf{u}, \nabla q\right)_K + \left\langle \phi^{-\frac{1}{2}} d\left(\mathbf{u}\cdot\mathbf{n} + (p-\hat{p})\right), q\right\rangle_{\partial K} = (f,q)_K\,, \tag{25b}$$

for all $(\mathbf{v}, q) \in \mathbf{V}_h(K)\times V_h(K)$.

Clearly we need an additional equation to close the system since we have introduced an additional trace unknown $\hat{p}$. The natural condition is the conservation, that is, the continuity of the HDG flux. For the HDG method to be conservative, it is sufficient to weakly enforce the continuity of the last component of the HDG flux (22) on each face e of the mesh skeleton, i.e.,

$$\left\langle [\![\phi^{-\frac{1}{2}} d\mathbf{u}\cdot\mathbf{n} + \phi^{-\frac{1}{2}} d\,(p-\hat{p})]\!], \hat{q}\right\rangle_e = 0, \quad \forall e \in \mathcal{E}_h^o. \tag{26}$$

On degenerate faces, where $\phi = 0$, the conservation condition (26) is trivially satisfied. These faces would need to be sorted out and removed from the system. However, this creates implementation issues. To avoid this, we introduce a more general HDG flux

$$\hat{\mathcal{F}}\cdot\mathbf{n} := \begin{pmatrix} n_1\phi^{-\frac{1}{2}} d\hat{p} \\ n_2\phi^{-\frac{1}{2}} d\hat{p} \\ n_3\phi^{-\frac{1}{2}} d\hat{p} \\ \phi^{-\frac{1}{2}} d\mathbf{u}\cdot\mathbf{n} + \tau(p-\hat{p}) \end{pmatrix}, \tag{27}$$

where τ is a positive function on the edge. For example, we can take $\tau = \phi^{-\frac{1}{2}} d$ for non-degenerate faces (i.e., faces with $\phi > 0$), and for degenerate ones (i.e. faces with $\phi = 0$) we take $\tau = \gamma > 0$. Alternatively, we can take a single value $\tau = O(1/h)$ over the entire mesh skeleton. We shall compare these choices in Section 4. With this HDG flux, the conservation condition (26) becomes

$$\left\langle [\![\phi^{-\frac{1}{2}} d\mathbf{u}\cdot\mathbf{n} + \tau\,(p-\hat{p})]\!], \hat{q}\right\rangle_e = 0, \quad \forall e \in \mathcal{E}_h^o, \quad \forall \hat{q} \in \Lambda_h(e). \tag{28}$$

On the Dirichlet boundary Γ_D, we impose the boundary data g_D to $\hat{p}$ through the weak form of

$$\langle \tau\hat{p}, \hat{q}\rangle_{\Gamma_D} = \langle \tau g_D, \hat{q}\rangle_{\Gamma_D}\,, \quad \forall \hat{q} \in \Lambda_h(\Gamma_D). \tag{29}$$

With the general HDG flux (27) and Dirichlet boundary condition (11c), the local equation (25) now becomes

$$\begin{aligned}(\mathbf{u},\mathbf{v})_K - \left(\phi^{-\frac{1}{2}}\nabla d p, \mathbf{v}\right)_K - \left(\phi^{-\frac{1}{2}} d p, \nabla\cdot\mathbf{v}\right)_K + \left\langle \phi^{-\frac{1}{2}} d\hat{p}, \mathbf{v}\cdot\mathbf{n}\right\rangle_{\partial K\setminus\Gamma_D} \\ + \left\langle \phi^{-\frac{1}{2}} d g_D, \mathbf{v}\cdot\mathbf{n}\right\rangle_{\partial K\cap\Gamma_D} = (d\tilde{\mathbf{g}}, \mathbf{v})_K\,,\end{aligned} \tag{30a}$$

$$\begin{aligned}(p,q)_K + \left(\frac{1}{2}\phi^{-\frac{3}{2}} d\nabla\phi\cdot\mathbf{u}, q\right)_K - \left(\phi^{-\frac{1}{2}} d\mathbf{u}, \nabla q\right)_K + \left\langle \phi^{-\frac{1}{2}} d\mathbf{u}\cdot\mathbf{n} + \tau(p-\hat{p}), q\right\rangle_{\partial K\setminus\Gamma_D} \\ + \left\langle \phi^{-\frac{1}{2}} d\mathbf{u}\cdot\mathbf{n} + \tau(p-g_D), q\right\rangle_{\partial K\cap\Gamma_D} = (f,q)_K\,.\end{aligned} \tag{30b}$$

The HDG comprises the local solver (30), the global equation (28) and the boundary condition (29). By summing (30) over all elements and (28) over the mesh skeleton, we obtain the complete HDG system with the weakly imposed Dirichlet boundary condition (29): find $(\mathbf{u}, p, \hat{p}) \in \mathbf{V}_h(\Omega_h)\times V_h(\Omega_h)\times \Lambda_h(\mathcal{E}_h)$ such that

$$\begin{aligned}(\mathbf{u},\mathbf{v})_{\Omega_h} - \left(\phi^{-\frac{1}{2}}\nabla d p, \mathbf{v}\right)_{\Omega_h} - \left(\phi^{-\frac{1}{2}} d p, \nabla\cdot\mathbf{v}\right)_{\Omega_h} + \left\langle \phi^{-\frac{1}{2}} d\hat{p}, \mathbf{v}\cdot\mathbf{n}\right\rangle_{\partial\Omega_h\setminus\Gamma_D} \\ = (d\tilde{\mathbf{g}}, \mathbf{v})_{\Omega_h} - \left\langle g_D, \phi^{-\frac{1}{2}} d\mathbf{v}\cdot\mathbf{n}\right\rangle_{\Gamma_D}\,,\end{aligned} \tag{31a}$$

$$\begin{aligned}(p,q)_{\Omega_h} + \left(\frac{1}{2}\phi^{-\frac{3}{2}} d\nabla\phi\cdot\mathbf{u}, q\right)_{\Omega_h} - \left(\phi^{-\frac{1}{2}} d\mathbf{u}, \nabla q\right)_{\Omega_h} + \left\langle \phi^{-\frac{1}{2}} d\mathbf{u}\cdot\mathbf{n} + \tau\,(p-\hat{p})\,, q\right\rangle_{\partial\Omega_h\setminus\Gamma_D} \\ + \left\langle \phi^{-\frac{1}{2}} d\mathbf{u}\cdot\mathbf{n} + \tau p, q\right\rangle_{\Gamma_D} = (f,q)_{\Omega_h} + \langle \tau g_D, q\rangle_{\Gamma_D}\,,\end{aligned} \tag{31b}$$

$$-\left\langle [\![\phi^{-\frac{1}{2}} d\mathbf{u}\cdot\mathbf{n} + \tau\,(p-\hat{p})]\!], \hat{q}\right\rangle_{\mathcal{E}_h\setminus\Gamma_D} + \langle \tau\hat{p}, \hat{q}\rangle_{\Gamma_D} = \langle \tau g_D, \hat{q}\rangle_{\Gamma_D}\,, \tag{31c}$$

for all $(\mathbf{v}, q, \hat{q}) \in \mathbf{V}_h(\Omega_h) \times V_h(\Omega_h) \times \Lambda_h(\mathcal{E}_h)$. Note that this form resembles the weak Galerkin framework [49, 50, 51, 52]. Indeed, HDG and the weak Galerkin method are equivalent in this case.

The HDG computation consists of three steps: first, solve the local solver for $(\mathbf{u}, p)$ as a function of $\hat{p}$ element-by-element, completely independent of each other; second, substitute $(\mathbf{u}, p)$ into the global equation (28) to solve for $\hat{p}$ on the mesh skeleton; and finally recover the local volume unknown $(\mathbf{u}, p)$ in parallel.

3.2. Well-posedness

Let us denote the bilinear form on the left hand side of (31) as $a\left((\mathbf{u}, p, \hat{p}); (\mathbf{v}, q, \hat{q})\right)$ and the linear form on right hand side as $\ell\left((\mathbf{v}, q, \hat{q})\right)$. We begin with an energy estimate for the HDG solution.

Proposition 1 (Discrete energy estimate)**.** *Suppose $g_D \in L^2(\Gamma_D)$, $f \in L^2(\Omega_h)$, and $d(\phi)\tilde{\mathbf{g}} \in L^2(\Omega_h)$. If $\tau = O(1/h)$, then it holds that*

$$a\left((\mathbf{u}, p, \hat{p}); (\mathbf{u}, p, \hat{p})\right) = \|\mathbf{u}\|^2_{\Omega_h} + \|p\|^2_{\Omega_h} + \|\hat{p}\|^2_{\Gamma_D,\tau} + \|p\|^2_{\Gamma_D,\tau} + \|p - \hat{p}\|^2_{\partial\Omega_h\setminus\Gamma_D,\tau} \tag{32}$$

$$\leq c\left(\|g_D\|^2_{\Gamma_D,\tau} + \|\tilde{\mathbf{g}}\|^2_{\Omega_h} + \|f\|^2_{\Omega_h}\right), \tag{33}$$

for some positive constant $c = c(\phi, d, \tau, h, k)$. In particular, there is a unique solution $(\mathbf{u}, p, \hat{p})$ to the HDG system (31)*.*

Proof. We start with the following identities

$$-\left(\phi^{-\frac{1}{2}}\nabla dp, \mathbf{v}\right)_K - \left(\phi^{-\frac{1}{2}} dp, \nabla\cdot\mathbf{v}\right)_K = -\left(p, \phi^{-\frac{1}{2}}\nabla\cdot(d\mathbf{v})\right)_K, \tag{34a}$$

$$\left(\frac{1}{2}\phi^{-\frac{3}{2}} d\nabla\phi\cdot\mathbf{u}, q\right)_K - \left(\phi^{-\frac{1}{2}} d\mathbf{u}, \nabla q\right)_K = \left(\phi^{-\frac{1}{2}}\nabla\cdot(d\mathbf{u}), q\right)_K - \left\langle \phi^{-\frac{1}{2}} d\mathbf{u}\cdot\mathbf{n}, q\right\rangle_{\partial K}. \tag{34b}$$

Now taking $\mathbf{v} = \mathbf{u}$, $q = p$, and $\hat{q} = \hat{p}$ in (31) and (34), and then adding all equations in (31) gives

$$\begin{aligned} a\left((\mathbf{u}, p, \hat{p}); (\mathbf{u}, p, \hat{p})\right) = \|\mathbf{u}\|^2_{\Omega_h} + \|p\|^2_{\Omega_h} + \|\hat{p}\|^2_{\Gamma_D,\tau} + \|p\|^2_{\Gamma_D,\tau} + \|p - \hat{p}\|^2_{\partial\Omega_h\setminus\Gamma_D,\tau} = \\ -\left\langle g_D, \phi^{-\frac{1}{2}} d\mathbf{u}\cdot\mathbf{n}\right\rangle_{\Gamma_D} + \langle \tau g_D, p\rangle_{\Gamma_D} + \langle \tau g_D, \hat{p}\rangle_{\Gamma_D} + (d\tilde{\mathbf{g}}, \mathbf{u})_{\Omega_h} + (f, p)_{\Omega_h}, \end{aligned}$$

which, after invoking the Cauchy-Schwarz and Young inequalities, becomes

$$\begin{aligned} a\left((\mathbf{u}, p, \hat{p}); (\mathbf{u}, p, \hat{p})\right) \leq \frac{\left\|\phi^{-\frac{1}{2}} d\right\|_\infty}{2\varepsilon_1}\|g_D\|^2_{\Gamma_D,\tau} + \frac{\varepsilon_1}{2}\|\mathbf{u}\|^2_{\Gamma_D,\tau^{-1}} \\ + \frac{1}{2\varepsilon_2}\|g_D\|^2_{\Gamma_D,\tau} + \frac{\varepsilon_2}{2}\|p - \hat{p}\|^2_{\Gamma_D,\tau} + \frac{1}{2\varepsilon_3}\|d\tilde{\mathbf{g}}\|^2_{\Omega_h} + \frac{\varepsilon_3}{2}\|\mathbf{u}\|^2_{\Omega_h} + \frac{1}{2\varepsilon_4}\|f\|^2_{\Omega_h} + \frac{\varepsilon_4}{2}\|p\|^2_{\Omega_h}, \end{aligned}$$

which yields the desired energy estimate after applying an inverse trace inequality (c.f. Lemma (A.1)) for the second term on right hand side and choosing sufficiently small values for $\varepsilon_1, \varepsilon_2, \varepsilon_3$ and ε_4. □

Since the HDG system (31) is linear and square in terms of the HDG variables $(\mathbf{u}, p, \hat{p})$, the uniqueness result in Proposition 1 implies existence and stability, and hence the well-posedness of the HDG system.

Lemma 1 (Consistency)**.** *Suppose $(\mathbf{u}^e, p^e)$ is a weak solution of* (11)*, which is sufficiently regular. Then $\left(\mathbf{u}^e, p^e, p^e|_{\mathcal{E}_h}\right)$ satisfies the HDG formulation* (31)*. In particular, the Galerkin orthogonality holds, i.e.,*

$$a\left(\left(\mathbf{u}^e - \mathbf{u}, p^e - p, p^e|_{\mathcal{E}_h} - \hat{p}\right); (\mathbf{v}, q, \hat{q})\right) = 0, \quad \forall (\mathbf{v}, q, \hat{q}) \in \mathbf{V}_h(\Omega_h) \times V_h(\Omega_h) \times \Lambda_h(\mathcal{E}_h). \tag{35}$$

The proof is a simple application of integration by parts and hence omitted.

3.3. Error analysis

We restrict the analyis for simplicial meshes and adopt the projection-based error analysis in [31]. To begin, we define $\widehat{p^e}$ as the trace of p^e. For any element K, $e \in \mathcal{E}_h$, $e \subset \partial K$, we denote by $\boldsymbol{P}\left(\mathbf{u}^e, p^e, \widehat{p^e}\right) := \left(\mathbb{P}\mathbf{u}^e, \mathbb{P}p^e, \Pi\widehat{p^e}\right)$, where Π is the standard L^2-projection, a collective projection of the exact solution. Let us define

$$\boldsymbol{\varepsilon}_{\mathbf{u}}^I := \mathbf{u}^e - \mathbb{P}\mathbf{u}^e, \quad \boldsymbol{\varepsilon}_{\mathbf{u}}^h := \mathbf{u} - \mathbb{P}\mathbf{u}^e, \tag{36}$$

$$\varepsilon_p^I := p^e - \mathbb{P}p^e, \quad \varepsilon_p^h := p - \mathbb{P}p^e, \tag{37}$$

$$\varepsilon_{\hat{p}}^I := \widehat{p^e} - \Pi\widehat{p^e}, \quad \varepsilon_{\hat{p}}^h := \hat{p} - \Pi\widehat{p^e}, \tag{38}$$

and then the projections $\mathbb{P}\mathbf{u}^e$ and $\mathbb{P}p^e$ are defined by

$$\left(\boldsymbol{\varepsilon}_{\mathbf{u}}^I, \mathbf{v}\right)_K = 0, \quad \mathbf{v} \in [\mathcal{P}_{k-1}(K)]^{dim}, \tag{39a}$$

$$\left(\varepsilon_p^I, q\right)_K = 0, \quad q \in \mathcal{P}_{k-1}(K), \tag{39b}$$

$$\left\langle \alpha\boldsymbol{\varepsilon}_{\mathbf{u}}^I \cdot \mathbf{n} + \tau\varepsilon_p^I, \hat{q}\right\rangle_e = 0, \quad \hat{q} \in \mathcal{P}_k(e), \tag{39c}$$

for each $K \in \Omega_h$, $e \in \mathcal{E}_h$ and $e \subset \partial K$. Here α, to be defined below, is a positive constant on each face e of element K.

Lemma 2. *Let $\tau_K := \tau/\alpha$. The projections $\mathbb{P}\mathbf{u}^e$ and $\mathbb{P}p^e$ are well-defined, and*

$$\left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^I\right\|_K + h\left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^I\right\|_{1,K} \le ch^{k+1}\left\|\mathbf{u}^e\right\|_{k+1,K} + ch^{k+1}\tau_K^*\left\|p^e\right\|_{k+1,K},$$

$$\left\|\varepsilon_p^I\right\|_K + h\left\|\varepsilon_p^I\right\|_{1,K} \le c\frac{h^{k+1}}{\tau_K^{\max}}\left\|\nabla\cdot\mathbf{u}^e\right\|_{k,K} + ch^{k+1}\left\|p^e\right\|_{k+1,K},$$

where $\tau_K^{\max} := \max \tau_K|_{\partial K}$ and $\tau_K^* := \tau_K|_{\partial K\setminus e^*}$, where e^* is the edge on which τ_K is maximum.

The proof can be obtained from [31].

Since the interpolation errors $\boldsymbol{\varepsilon}_{\mathbf{u}}^I, \varepsilon_p^I$ and $\varepsilon_{\hat{p}}^I$ have optimal convergence order, by the triangle inquality, the convergent rates of the total errors $\boldsymbol{\varepsilon}_{\mathbf{u}} = \boldsymbol{\varepsilon}_{\mathbf{u}}^I + \boldsymbol{\varepsilon}_{\mathbf{u}}^h, \varepsilon_p = \varepsilon_p^I + \varepsilon_p^h$, and $\varepsilon_{\hat{p}} = \varepsilon_{\hat{p}}^I + \varepsilon_{\hat{p}}^h$ are determined by those of the discretization errors $\boldsymbol{\varepsilon}_{\mathbf{u}}^h, \varepsilon_p^h$ and $\varepsilon_{\hat{p}}^h$. We use an energy approach to estimate the discretization errors. To begin, let us define

$$\mathcal{E}_h^2 := \left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^h\right\|_{\Omega_h}^2 + \left\|\varepsilon_p^h\right\|_{\Omega_h}^2 + \left\|\varepsilon_{\hat{p}}^h\right\|_{\Gamma_D,\tau}^2 + \left\|\varepsilon_p^h\right\|_{\Gamma_D,\tau}^2 + \left\|\varepsilon_p^h - \varepsilon_{\hat{p}}^h\right\|_{\partial\Omega_h\setminus\Gamma_D,\tau}^2.$$

Lemma 3 (Error equation)**.** *It holds that*

$$\mathcal{E}_h^2 = \underbrace{-\left(\phi^{-\frac{1}{2}}\varepsilon_p^I, \nabla\cdot\left(d\boldsymbol{\varepsilon}_{\mathbf{u}}^h\right)\right)_{\Omega_h}}_{A} + \underbrace{\left\langle\phi^{-\frac{1}{2}}d\varepsilon_{\hat{p}}^I, \boldsymbol{\varepsilon}_{\mathbf{u}}^h\cdot\mathbf{n}\right\rangle_{\partial\Omega_h\setminus\Gamma_D}}_{B} \underbrace{-\left(d\boldsymbol{\varepsilon}_{\mathbf{u}}^I, \nabla\left(\phi^{-\frac{1}{2}}\varepsilon_p^h\right)\right)_{\Omega_h}}_{C}$$
$$+ \underbrace{\left\langle\left(\phi^{-\frac{1}{2}}d-\alpha\right)\boldsymbol{\varepsilon}_{\mathbf{u}}^I\cdot\mathbf{n}, \varepsilon_p^h - \varepsilon_{\hat{p}}^h\right\rangle_{\partial\Omega_h\setminus\Gamma_D} + \left\langle\left(\phi^{-\frac{1}{2}}d-\alpha\right)\boldsymbol{\varepsilon}_{\mathbf{u}}^I\cdot\mathbf{n}, \varepsilon_p^h\right\rangle_{\Gamma_D}}_{D} \tag{40}$$

Proof. The proof is straightforward by first adding and subtracting appropriate projections in the Galerkin orthogonality equation (35), second using the definition of the projections (39), and finally taking $\mathbf{v} = \boldsymbol{\varepsilon}_{\mathbf{u}}^h$, $q = \varepsilon_p^h$, and $\hat{q} = \varepsilon_{\hat{p}}^h$. □

The next step is to estimate A, B, C and D. To that end, we define α on faces of an element K as

$$\alpha := \begin{cases} \overline{\phi^{-\frac{1}{2}}d} & \text{if } \overline{\phi^{-\frac{1}{2}}d} \neq 0 \\ 1 & \text{otherwise} \end{cases}, \tag{41}$$

where $\overline{\phi^{-\frac{1}{2}}d}$ is the average of $\phi^{-\frac{1}{2}}d$ on the element K.

Lemma 4 (Estimation for A)**.** *There exists a positive constant $c = c(\phi, d)$ such that*

$$|A| \le c \left\|\varepsilon_p^I\right\|_{\Omega_h} \left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^h\right\|_{\Omega_h}.$$

Proof. We have

$$|A| \le \left|\left(\varepsilon_p^I, \phi^{-\frac{1}{2}} \nabla d \cdot \boldsymbol{\varepsilon}_{\mathbf{u}}^h\right)_{\Omega_h}\right| + \left|\left(\varepsilon_p^I, \phi^{-\frac{1}{2}} d \nabla \cdot \boldsymbol{\varepsilon}_{\mathbf{u}}^h\right)_{\Omega_h}\right|$$

Bounding the first term is straightforward:

$$\left|\left(\varepsilon_p^I, \phi^{-\frac{1}{2}} \nabla d \cdot \boldsymbol{\varepsilon}_{\mathbf{u}}^h\right)_{\Omega_h}\right| \le c \left\|\phi^{-\frac{1}{2}} \nabla d\right\|_{\infty} \left\|\varepsilon_p^I\right\|_{\Omega_h} \left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^h\right\|_{\Omega_h}.$$

For the second term, we have

$$\begin{aligned}\left|\left(\varepsilon_p^I, \phi^{-\frac{1}{2}} d \nabla \cdot \boldsymbol{\varepsilon}_{\mathbf{u}}^h\right)_{\Omega_h}\right| = \left|\left(\varepsilon_p^I, \left(\phi^{-\frac{1}{2}} d - \overline{\phi^{-\frac{1}{2}} d}\right) \nabla \cdot \boldsymbol{\varepsilon}_{\mathbf{u}}^h\right)_{\Omega_h}\right| &\le ch \left\|\varepsilon_p^I\right\|_{\Omega_h} \left\|\phi^{-\frac{1}{2}} d\right\|_{W^{1,\infty}(\Omega_h)} \left\|\nabla \cdot \boldsymbol{\varepsilon}_{\mathbf{u}}^h\right\|_{\Omega_h} \\ &\le c \left\|\varepsilon_p^I\right\|_{\Omega_h} \left\|\phi^{-\frac{1}{2}} d\right\|_{W^{1,\infty}(\Omega_h)} \left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^h\right\|_{\Omega_h},\end{aligned}$$

where we have used (39b) in the first equality, the Cauchy-Schwarz inequality and the Bramble–Hilbert lemma (see, e.g., [54]) in the first inequality, and Lemma 8 (in the appendix) in the last inequality. Here, $W^{1,\infty}$ is a standard Sobolev space. □

Lemma 5 (Estimation for B)**.** *There exists a positive constant $c = c(\phi, d)$ such that*

$$|B| \le ch^{\frac{1}{2}} \left\|\varepsilon_p^I\right\|_{\partial\Omega_h} \left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^h\right\|_{\Omega_h}.$$

Proof. We have

$$\begin{aligned}|B| = \left|\left\langle \varepsilon_{\hat{p}}^I, \left(\phi^{-\frac{1}{2}} d - \overline{\phi^{-\frac{1}{2}} d}\right) \boldsymbol{\varepsilon}_{\mathbf{u}}^h \cdot \mathbf{n} \right\rangle_{\partial\Omega_h \setminus \Gamma_D}\right| &\le \left\|\varepsilon_{\hat{p}}^I\right\|_{\partial\Omega_h} \left\|\phi^{-\frac{1}{2}} d - \overline{\phi^{-\frac{1}{2}} d}\right\|_{L^\infty(\partial\Omega_h)} \left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^h\right\|_{\partial\Omega_h} \\ &\le ch \left\|\phi^{-\frac{1}{2}} d\right\|_{W^{1,\infty}(\Omega_h)} \left\|\varepsilon_{\hat{p}}^I\right\|_{\partial\Omega_h} \left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^h\right\|_{\partial\Omega_h},\end{aligned}$$

where we have used the property of L^2-projection $\Pi\widehat{p^e}$ in the first equality, the Cauchy-Schwarz inequality in the first inequality, and the Bramble–Hilbert lemma in the last inequality. Now the best approximation of $\Pi\widehat{p^e}$ implies $\|\varepsilon_{\hat{p}}^I\|_{\partial\Omega_h} \le \|\varepsilon_p^I\|_{\partial\Omega_h}$ and (A.4) gives the result. □

Lemma 6 (Estimation for C)**.** *There exists a positive constant $c = c(\phi, d)$ such that*

$$|C| \le c \left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^I\right\|_{\Omega_h} \left\|\varepsilon_p^h\right\|_{\Omega_h}.$$

Proof. We have

$$|C| \le \left|\left(\frac{1}{2}\phi^{-\frac{3}{2}} d \nabla\phi \cdot \boldsymbol{\varepsilon}_{\mathbf{u}}^I, \varepsilon_p^h\right)_{\Omega_h}\right| + \left|\left(\boldsymbol{\varepsilon}_{\mathbf{u}}^I, \phi^{-\frac{1}{2}} d \nabla \varepsilon_p^h\right)_{\Omega_h}\right|.$$

The rest of the proof is similar to that of Lemma 4 by using (39a). □

Lemma 7 (Estimation for D)**.** *There exists a positive constant $c = c(\phi, d)$ such that*

$$|D| \le c\beta \left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^I\right\|_{\partial\Omega_h, \tau^{-1}} \left(\left\|\varepsilon_p^h - \varepsilon_{\hat{p}}^h\right\|_{\partial\Omega_h \setminus \Gamma_D, \tau} + \left\|\varepsilon_p^h\right\|_{\Gamma_D, \tau}\right),$$

where

$$\beta := \begin{cases} h & \text{if } \overline{\phi^{-\frac{1}{2}} d} \ne 0 \quad \forall K \in \Omega_h \\ 1 & \text{otherwise} \end{cases}.$$

Proof. Employing similar techniques as in estimating B, we have

$$|D| \le \left\|\phi^{-\frac{1}{2}} d - \alpha\right\|_{L^\infty(\partial\Omega_h)} \left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^I\right\|_{\partial\Omega_h,\tau^{-1}} \left(\left\|\varepsilon_p^h - \varepsilon_{\hat{p}}^h\right\|_{\partial\Omega_h\setminus\Gamma_D,\tau} + \left\|\varepsilon_p^h\right\|_{\Gamma_D,\tau}\right).$$

Now using the definition of α in (41) and the Bramble–Hilbert lemma,

$$\left\|\phi^{-\frac{1}{2}} d - \alpha\right\|_{L^\infty(\partial\Omega_h)} \le \left\|\phi^{-\frac{1}{2}} d - \alpha\right\|_{L^\infty(\Omega_h)} \le c\beta,$$

and this ends the proof. □

Now comes the main result of this section.

Theorem 1. *Suppose* $\mathbf{u}^e \in \left[H^{k+1}(\Omega_h)\right]^{dim}$ *and* $p^e \in H^{k+1}(\Omega_h)$*. Then*

$$\begin{aligned}\left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^h\right\|_{\Omega_h} + \left\|\varepsilon_p^h\right\|_{\Omega_h} + \left\|\varepsilon_{\hat{p}}^h\right\|_{\Gamma_D,\tau} + \left\|\varepsilon_p^h\right\|_{\Gamma_D,\tau} + \left\|\varepsilon_p^h - \varepsilon_{\hat{p}}^h\right\|_{\partial\Omega_h\setminus\Gamma_D,\tau} \\ \le c\left(\|\mathbf{u}^e\|_{k+1,\Omega_h} + \|p^e\|_{k+1,\Omega_h}\right) \times \begin{cases} h^{k+1} & \text{if } \overline{\phi^{-\frac{1}{2}} d} \ne 0 \quad \forall K \in \Omega_h \\ h^{k+\frac{1}{2}} & \text{otherwise} \end{cases},\end{aligned}$$

where $c = c(\phi, d, \tau)$ *is a positive constant independent of* h*.*

Proof. Using the results in Lemmas 3—7 and the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\mathcal{E}_h^2 \le c\left(\left\|\varepsilon_p^I\right\|_{\Omega_h}^2 + \beta h\left\|\varepsilon_p^I\right\|_{\partial\Omega_h}^2 + \left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^I\right\|_{\Omega_h}^2 + \beta\left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^I\right\|_{\partial\Omega_h,\tau^{-1}}^2\right)^{\frac{1}{2}} \times \\ \left(\left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^h\right\|_{\Omega_h}^2 + \left\|\varepsilon_p^h\right\|_{\Omega_h}^2 + \left\|\varepsilon_p^h - \varepsilon_{\hat{p}}^h\right\|_{\partial\Omega_h\setminus\Gamma_D,\tau}^2 + \left\|\varepsilon_p^h\right\|_{\Gamma_D,\tau}^2\right)^{\frac{1}{2}} \\ \le c\left(\left\|\varepsilon_p^I\right\|_{\Omega_h}^2 + \beta h\left\|\varepsilon_p^I\right\|_{\partial\Omega_h}^2 + \left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^I\right\|_{\Omega_h}^2 + \beta\left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^I\right\|_{\partial\Omega_h,\tau^{-1}}^2\right)^{\frac{1}{2}} \times \mathcal{E}_h. \quad (42)\end{aligned}$$

The estimate for $\left\|\varepsilon_p^I\right\|_{\Omega_h}^2$ and $\left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^I\right\|_{\Omega_h}^2$ can be obtained directly from Lemma 2. Now using Lemma 9 in the appendix and approximation properties of $\mathbb{P}\mathbf{u}^e, \mathbb{P}p^e$ in Lemma 2 gives

$$\left\|\varepsilon_p^I\right\|_{\partial\Omega_h}^2 \le c\sum_K \left(\left\|\nabla\varepsilon_p^I\right\|_{0,K} + h^{-1}\left\|\varepsilon_p^I\right\|_{0,K}\right)\left\|\varepsilon_p^I\right\|_{0,K} \le ch^{2k+1}\left(\max_K \frac{1}{\tau_K^{\max}}\|\mathbf{u}^e\|_{k+1,\Omega_h} + \|p^e\|_{k+1,\Omega_h}\right)^2. \quad (43)$$

Similarly we can obtain

$$\left\|\boldsymbol{\varepsilon}_{\mathbf{u}}^I\right\|_{\partial\Omega_h,\tau^{-1}}^2 \le ch^{2k+1}\max_K \frac{1}{\tau}\left(\|\mathbf{u}^e\|_{k+1,\Omega_h} + \max_K \tau_K^*\|p^e\|_{k+1,\Omega_h}\right)^2. \quad (44)$$

The assertion is now ready by combing the inequalities (42)—(44), the definition of β, and the Cauchy-Schwarz inequality. □

Remark 1. *When the system is degenerate, but the exact solution is piecewise smooth, the convergence rate is sub-optimal by half order. The above proof, especially inequality* (44)*, shows that this suboptimality may not be improved by using* $\tau = O\left(h^{-1}\right)$*. The reason is that the gain by half order from* $\max_K \frac{1}{\tau}$ *is taken away by the loss of half order from* $\max_K \tau_K^*$*. This will be confirmed in our numerical studies of the sensitivity of* τ *on the convergence rate in Section 4.4.*

4. Numerical results

In this section, we present numerical examples to support the HDG approach and its convergence analysis. For a non-degenerate case, we consider a sine solution test, while for degenerate cases, we choose smooth and non-smooth solution tests [17]. We take the upwind based parameter $\tau = \phi^{-\frac{1}{2}} d$ for a non-degenerate case, and the generalized parameter

$$\tau = \begin{cases} \phi^{-\frac{1}{2}} d & \text{for } \phi > 0, \\ 1/h & \text{for } \phi = 0 \end{cases}$$

for degenerate cases. We also conduct several numerical computations to understand if the stabilization parameter τ can affect the accuracy of the HDG solution and its convergence rate. We assume that porosity ϕ is known and $d = \phi$ in all the numerical examples. The domain Ω is chosen as $\Omega = (0,1)^{dim}$ or $\Omega = (-1,1)^{dim}$, which is either uniformly discretized with n_e rectangular tensor product elements in each dimension (so that the total number of elements is $N_e = n_e^{dim}$), or N_e triangular elements. Though we have rigorous optimal convergence theory for only simplicial meshes (see Theorem 1), a similar result is expected for quadrilateral/hexahedral meshes (see the numerical results in the following sections). Since rectangular meshes are convenient for all problems in this paper with simple interfaces between the fluid melt and the solid matrix, we use rectangular meshes hereafter, except for the test in Section 4.1.

4.1. Non-degenerate case

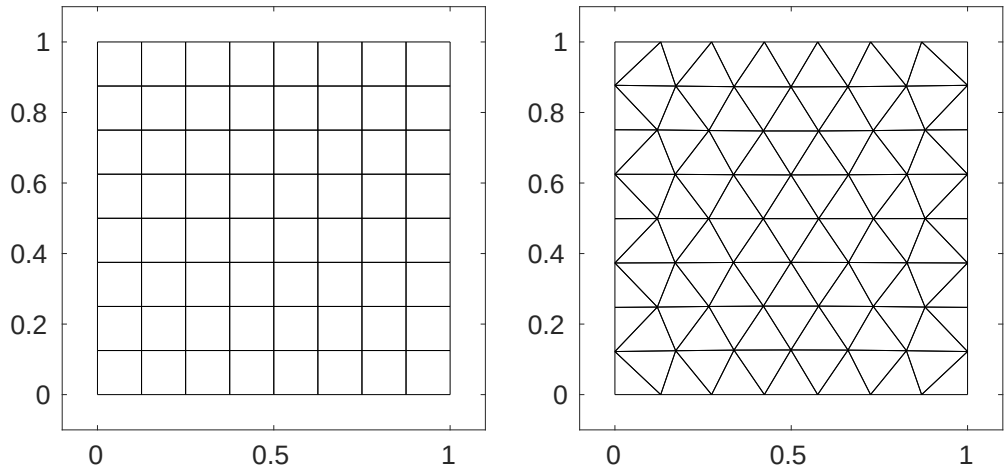
We consider a non-degenerate case on $\Omega = (0,1)^2$ with the porosity given by $\phi = \exp(2(x+y))$. We choose the pressure to be $\tilde{p}^e = \exp(-(x+y))\sin(m_x\pi x)\sin(m_y\pi y)$. The corresponding manufactured scaled solutions are given as

$$p^e = \sin(m_x\pi x)\sin(m_y\pi y), \tag{45a}$$

$$u_x^e = \exp(x+y)\sin(m_y\pi y)\left(\sin(m_x\pi x) - m_x\pi\cos(m_x\pi x)\right), \tag{45b}$$

$$u_y^e = \exp(x+y)\sin(m_x\pi x)\Big(\sin(m_y\pi y) - m_y\pi\cos(m_y\pi y)\Big). \tag{45c}$$

Here, we take $m_x = 2$ and $m_y = 3$.



(a) Rectangular elements (b) Triangular elements

Figure 1. Coarse grids for non-degenerate case with (a) rectangular and (b) triangular elements.

Table 1 shows h-convergence results in the $L^2(\Omega_h)$-norm using a sequence of nested meshes with $N_e = \{8^2, 32^2, 128^2\}$ for rectangular and $N_e = \{104, 416, 1664\}$ for triangular elements, respectively. The corresponding coarse meshes are shown in Figure 1. We observe approximately the optimal convergence rates of $(k+1)$ for both scaled pressure p and scaled velocity $\mathbf{u}$ for both mesh types.

Table 1. Non-degenerate case: the results show that the HDG solutions for scaled pressure p and scaled velocity $\mathbf{u}$ converge to the exact solution with optimal order of $k+1$ for both triangular and rectangular meshes. The upwind based parameter $\tau = \phi^{-\frac{1}{2}} d$ is used.

		Rectangular elements				Triangular elements				
k	h	$\|p^e - p\|_{\Omega_h}$ error	order	$\|\mathbf{u}^e - \mathbf{u}\|_{\Omega_h}$ error	order	h	$\|p^e - p\|_{\Omega_h}$ error	order	$\|\mathbf{u}^e - \mathbf{u}\|_{\Omega_h}$ error	order
1	0.0312	3.628E-02	–	8.546E-01	1.395	0.1400	6.521E-01	–	5.021E+00	–
	0.0156	1.159E-02	1.646	2.878E-01	1.570	0.0700	1.813E-01	1.847	1.436E+00	1.806
	0.0078	3.389E-03	1.775	8.804E-02	1.709	0.0350	4.726E-02	1.940	3.794E-01	1.920
2	0.0312	1.067E-03	–	2.734E-02	–	0.1400	9.445E-02	–	8.312E-01	–
	0.0156	1.597E-04	2.741	4.272E-03	2.678	0.0700	1.245E-02	2.924	1.071E-01	2.956
	0.0078	2.226E-05	2.843	6.170E-04	2.791	0.0350	1.587E-03	2.971	1.354E-02	2.984
3	0.0312	1.970E-05	–	4.691E-04	–	0.1400	8.929E-03	–	6.658E-02	–
	0.0156	1.405E-06	3.809	3.478E-05	3.753	0.0700	5.865E-04	3.928	4.570E-03	3.865
	0.0078	9.480E-08	3.890	2.414E-06	3.848	0.0350	3.728E-05	3.976	2.942E-04	3.957
4	0.0312	3.327E-07	–	8.829E-06	–	0.1400	8.015E-04	–	7.305E-03	–
	0.0156	1.168E-08	4.832	3.211E-07	4.781	0.0700	2.571E-05	4.963	2.292E-04	4.995
	0.0078	3.906E-10	4.903	1.115E-08	4.848	0.0350	8.124E-07	4.984	7.199E-06	4.992

4.2. Degenerate case with a smooth solution

Following [17] we consider the smooth pressure of the form $\tilde{p}^e = \cos(6xy^2)$ on $\Omega = (-1,1)^2$ and the following degenerate porosity

$$\phi = \begin{cases} 0, & x \le -\frac{3}{4} \text{ or } y \le -\frac{3}{4}, \\ (x+\frac{3}{4})^\alpha (y+\frac{3}{4})^{2\alpha}, & \text{otherwise.} \end{cases} \tag{46}$$

We note that $\phi^{-\frac{1}{2}} \nabla \phi \in [L^\infty(\Omega)]^2$ for $\alpha \ge 2$, and we take $\alpha = 2$. The one-phase region is denoted as $\Omega_1 := \{(x,y) : x \le -\frac{3}{4} \text{ or } y \le -\frac{3}{4}\}$ with $\phi = 0$, and the two-phase region is given by $\Omega_2 := \{(x,y) : -\frac{3}{4} < x < 1 \text{ and } -\frac{3}{4} < y < 1\}$ with $\phi > 0$. We define the intersection of these two regions by $\partial\Omega_{12} := \overline{\Omega}_1 \cap \overline{\Omega}_2$. In $\overline{\Omega}_1$, the exact scaled pressure and scaled velocity vanish. In $\overline{\Omega}_2$, the exact solutions are given by

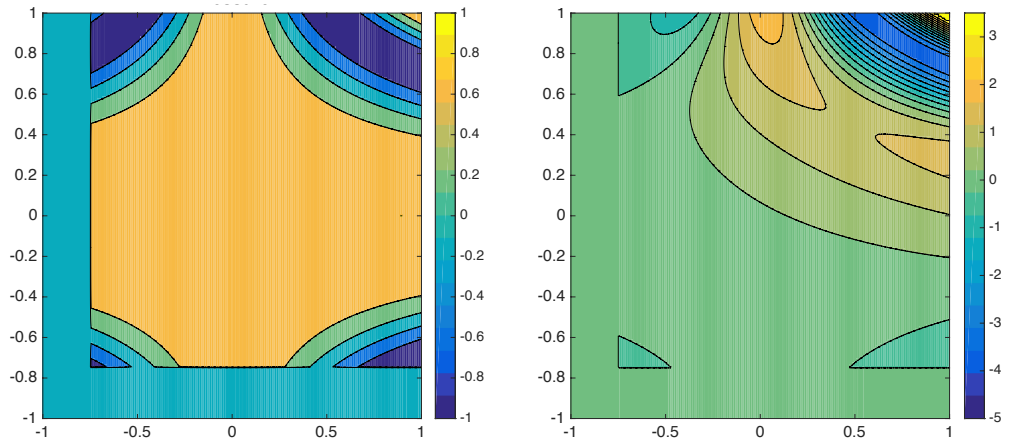
$$p^e = \left(x+\frac{3}{4}\right)^{\frac{\alpha}{2}} \left(y+\frac{3}{4}\right)^{\alpha} \cos(6xy^2), \tag{47a}$$

$$u_x^e = 6y^2 \left(x+\frac{3}{4}\right)^{\alpha} \left(y+\frac{3}{4}\right)^{2\alpha} \sin\left(6xy^2\right), \tag{47b}$$

$$u_y^e = 12xy \left(x+\frac{3}{4}\right)^{\alpha} \left(y+\frac{3}{4}\right)^{2\alpha} \sin(6xy^2). \tag{47c}$$

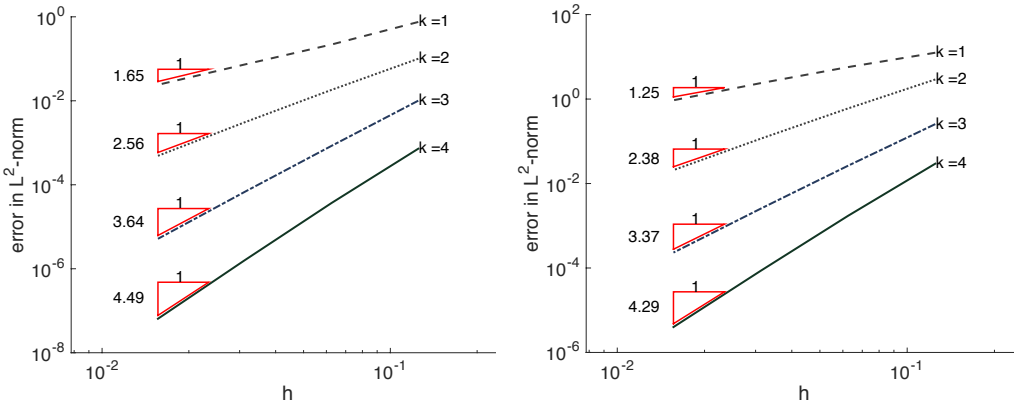
In Figure 2 are the contours of the pressure $\tilde{p}$ and the scaled pressure p computed from our HDG method using $N_e = 64^2$ rectangular elements and solution order $k = 4$. We observe that the pressure $\tilde{p}$ changes smoothly in the two-phase region Ω_2, but abruptly becomes zero in the one-phase region Ω_1. The sudden pressure jump on the intersection Ω_{12} is alleviated with the use of the scaled pressure p.

For a convergence study, we use a sequence of meshes with $n_e = \{16, 32, 64, 128\}$ and with $k = \{1, 2, 3, 4\}$. Here we choose an even number of elements so that the mesh skeleton aligns with the intersection $\partial\Omega_{12}$. As can be seen in Figure 3, the convergence rate of $\left(k+\frac{1}{2}\right)$ is observed more or less for both the scaled pressure p and the scaled velocity $\mathbf{u}$, and this agrees with Theorem 1 for the degenerate case with a piecewise smooth solution.



(a) Pressure $\tilde{p}$ (b) Scaled pressure p

Figure 2. Degenerate case with a smooth solution: (a) contour plot of the pressure $\tilde{p}$ field and (b) contour plot of the scaled pressure p with $N_e = 64^2$ and $k = 4$. The pressure field changes smoothly in the two-phase region Ω_2, but suddenly becomes zero in the one-phase region Ω_1. The abrupt change near the intersection Ω_{12} between the one- and two-phase regions is alleviated with the use of the scaled pressure p.



(a) p (b) $\mathbf{u}$

Figure 3. Degenerate case with a smooth solution: convergence study for (a) the scaled pressure p field and (b) the scaled velocity $\mathbf{u}$ field. The $\left(k + \frac{1}{2}\right)$ convergence rates are obtained approximately for both the scaled pressure p and the scaled velocity $\mathbf{u}$.

4.3. Degenerate case with low solution regularity

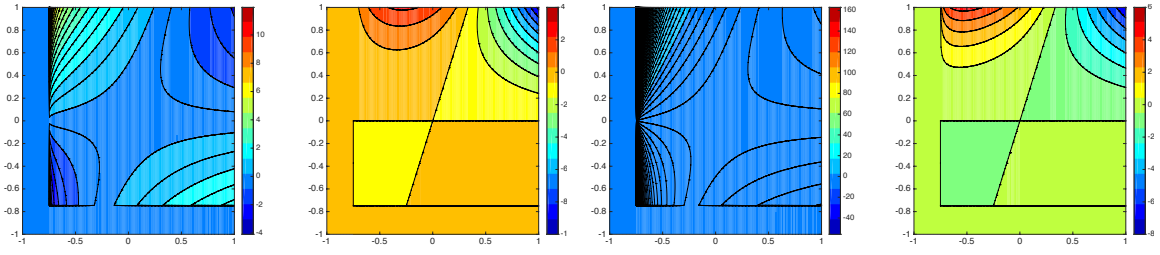
Similar to [17], we choose the exact pressure to be $\tilde{p}^e = y(y - 3x)(x + \frac{3}{4})^\beta$ with $\beta = -\frac{1}{4}$ or $-\frac{3}{4}$, and the porosity ϕ is defined in (46). Similar to Section 4.2, we take $\alpha = 2$. The exact solutions then read

$$p^e = y(y - 3x)\left(x + \frac{3}{4}\right)^{\frac{\alpha}{2}+\beta}\left(y + \frac{3}{4}\right)^{\alpha}, \tag{48a}$$

$$u_x^e = y\left(\beta(3x - y) + 3\left(x + \frac{3}{4}\right)\right)\left(x + \frac{3}{4}\right)^{\alpha+\beta-1}\left(y + \frac{3}{4}\right)^{2\alpha}, \tag{48b}$$

$$u_y^e = (3x - 2y)\left(x + \frac{3}{4}\right)^{\alpha+\beta}\left(y + \frac{3}{4}\right)^{2\alpha}. \tag{48c}$$

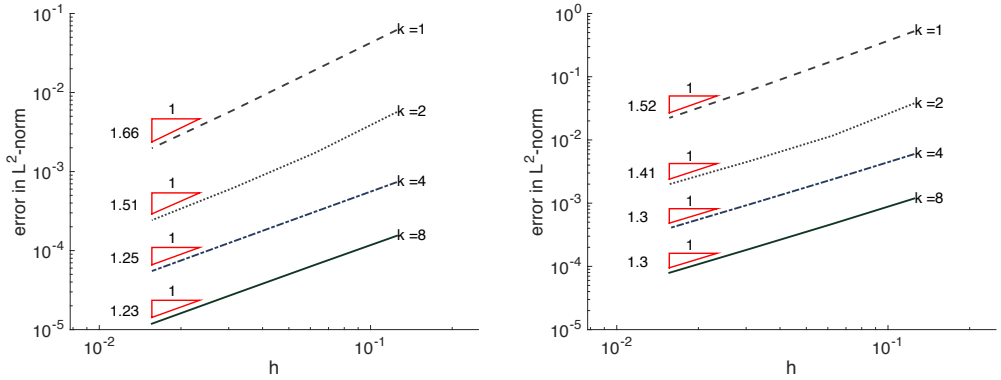
The pressure and the scaled pressure fields are simulated with $N_e = 64^2$ and $k = 4$ for the two different cases: $\beta = -\frac{1}{4}$ and $\beta = -\frac{3}{4}$ in Figure 4. As can be seen from (48) and Figure 4 that smaller β implies lower solution regularity. The pressure field with $\beta = -\frac{3}{4}$ is less regular than that with $\beta = -\frac{1}{4}$. For both cases, we also observe that the pressure $\tilde{p}$ fields become stiffer (stiff "boundary layer") near the intersection at $x = -\frac{3}{4}$, while the scaled pressure p fields are much less stiff.



(a) $\tilde{p}$ with $\beta = -\frac{1}{4}$ (b) p with $\beta = -\frac{1}{4}$ (c) $\tilde{p}$ with $\beta = -\frac{3}{4}$ (d) p with $\beta = -\frac{3}{4}$

Figure 4. Degenerate case with low solution regularity: simulated with $N_e = 64^2$ and $k = 4$ are (a) pressure $\tilde{p}$ for $\beta = -\frac{1}{4}$, (b) scaled pressure p for $\beta = -\frac{1}{4}$, (c) pressure $\tilde{p}$ with $\beta = -\frac{3}{4}$, and (d) scaled pressure p for $\beta = -\frac{3}{4}$. The pressure field with $\beta = -\frac{3}{4}$ is less regular than that with $\beta = -\frac{1}{4}$. In both the cases, the pressure fields have low regularity near the intersection Ω_{12}.

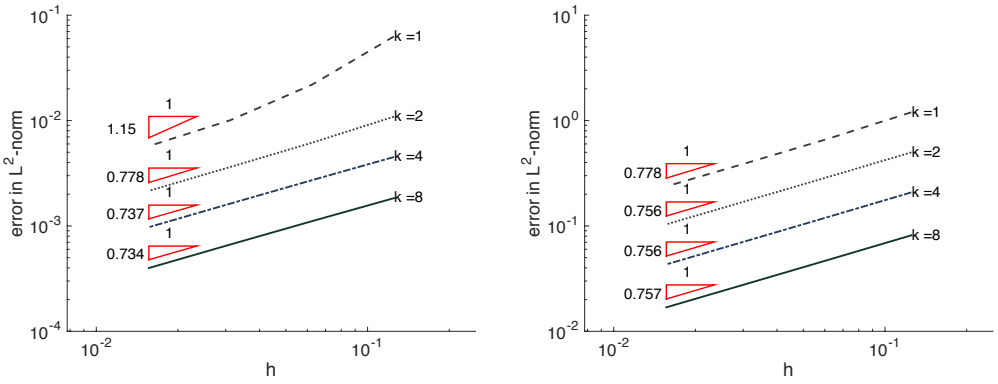
When $\beta = -\frac{1}{4}$, the scaled pressure p and the scaled velocity $\mathbf{u}$ reside in $H^{1.25-\varepsilon}$ for $\varepsilon > 0$ [17]. In order to see how the HDG solution behaves for this case, we perform a convergence study with $n_e = \{16, 32, 64, 128\}$ and $k = \{1, 2, 4, 8\}$. As shown in Figure 5, the scaled pressure p and the scaled velocity $\mathbf{u}$ converge to the exact counterparts with the rate of about 1.25. Note that though our error analysis in Section 3.3 considers exact solutions residing in standard Sobolev spaces with integer powers, it can be straightforwardly extended to solutions in fractional Sobolev spaces. For this example, the convergence rate is bounded above by $1.25 - \varepsilon$ regardless of the solution order. However, the high order HDG solutions are still beneficial in terms of accuracy, for example, the HDG solution with $k = 8$ is 4.5 times more accurate than that with $k = 4$.



(a) Convergence of p (b) Convergence of $\mathbf{u}$

Figure 5. Degenerate case with low solution regularity: a convergence study with $\beta = -\frac{1}{4}$ for (a) the scaled pressure p field and (b) the scaled velocity $\mathbf{u}$ field.

When $\beta = -\frac{3}{4}$, the scaled pressure p and the scaled velocity $\mathbf{u}$ lie in $H^{0.75-\varepsilon}$ for $\varepsilon > 0$ [17]. We conduct a convergence study with $n_e = \{16, 32, 64, 128\}$ and $k = \{1, 2, 4, 8\}$. As shown in Figure 6, the convergence rate of about 0.75 is observed for both the scaled pressure p and the scaled velocity $\mathbf{u}$. Similar to the case of $\beta = -\frac{1}{4}$, high order HDG solutions, in spite of more computational demand, are beneficial from an accuracy standpoint. For instance, the HDG solution with $k = 8$ is 2.5 times more accurate than that with $k = 4$.



(a) Convergence of p (b) Convergence of $\mathbf{u}$

Figure 6. Degenerate case with low solution regularity: a convergence study with $\beta = -\frac{3}{4}$ for (a) the scaled pressure p field and (b) the scaled velocity $\mathbf{u}$ field.

4.4. Sensitivity of τ for the degenerate case with smooth solution

In this section we assess numerically whether the sub-optimality in Theorem 1 is sharp. To that end, we consider the degenerate example with smooth solution in Section 4.2 again here. Recall that the generalized parameter τ is given by

$$\tau := \begin{cases} \phi^{-\frac{1}{2}} d & \text{for } \phi > 0, \\ \gamma & \text{for } \phi = 0. \end{cases} \tag{49}$$

We first compare the convergence rates for three different values of γ, namely $\gamma \in \{1/h, 1, 10\}$, and the numerical results (not shown here) show that the convergence rates are the same and are sub-optimal by half order. For that reason we show only the case when $\gamma = 1/h$ in the fourth column of Tables 2 and 3, in which we report the convergence rates of p and $\mathbf{u}$, respectively.

We now present convergence rates for the cases where we use a single value for τ over the entire mesh skeleton $\mathcal{E}_h$. We consider three cases: $\tau = (1/h, 1, 10)$. The convergence rates of p and $\mathbf{u}$ for these parameters are shown in the sixth, eighth, and tenth columns of Tables 2 and 3. The results for $\tau = 1$ and $\tau = 10$ show the convergence rate of about $\left(k + \frac{1}{2}\right)$. The cases with $\tau = \frac{1}{h}$ initially have the convergence rate of $(k + 1)$ for both the scaled pressure p and the scaled velocity $\mathbf{u}$, then approach the predicted asymptotic rate of $\left(k + \frac{1}{2}\right)$ as the grid is refined. If we look at the value of the errors at any grid level, the cases with $\tau = 1/h$ have the smallest errors compared to the other cases (including the cases with τ given in (49)). It could be due to the initial higher-order convergence and/or smaller error constants. We thus recommend that $\tau = 1/h$ should be used.

4.5. Non-degenerate case in three dimensions

We consider finally a non-degenerate case on $\Omega = (0, 1)^3$ with the positive porosity $\phi = \exp(2(x + y + z))$. Let the pressure $\tilde{p}^e = \sin(m_x\pi x)\sin(m_y\pi y)\sin(m_z\pi z)\exp(-(x + y + z))$. The corresponding manufactured scaled solutions are given as

$$q^e = \sin(m_x\pi x)\sin(m_y\pi y)\sin(m_z\pi z), \tag{50a}$$

$$u_x^e = \exp(x + y + z)\sin(m_y\pi y)\sin(m_z\pi z)\left(\sin(m_x\pi x) - m_x\pi\cos(m_x\pi x)\right), \tag{50b}$$

$$u_y^e = \exp(x + y + z)\sin(m_x\pi x)\sin(m_z\pi z)\left(\sin(m_y\pi y) - m_y\pi\cos(m_y\pi y)\right), \tag{50c}$$

$$u_z^e = \exp(x + y + z)\sin(m_x\pi x)\sin(m_y\pi y)\left(\sin(m_z\pi z) - m_z\pi\cos(m_z\pi z)\right). \tag{50d}$$

Table 4 shows h-convergence results in the $L^2(\Omega_h)$-norm using a sequence of nested meshes with $n_e = \{8, 12, 16, 20\}$. Here we take $m_x = m_y = m_z = 1$ and use the upwind based parameter $\tau = \phi^{-\frac{1}{2}} d$. We observe the convergence rates

Table 2. Degenerate case with a smooth solution: the errors $\|p^e - p\|_{\Omega_h}$ and the convergence rates for the scaled pressure. Four cases are presented: τ given (49), $\tau = \frac{1}{h}$, $\tau = 1$ and $\tau = 10$.

k	h	$\tau = \begin{cases} \phi^{-\frac{1}{2}} d & \text{for } \phi > 0 \\ 1/h & \text{for } \phi = 0 \end{cases}$		$\tau = \frac{1}{h}$		$\tau = 1$		$\tau = 10$	
		error	order	error	order	error	order	error	order
1	0.1250	7.534E-01	–	1.752E-01	–	2.606E+00	–	3.827E-01	–
	0.0625	2.188E-01	1.784	4.107E-02	2.093	4.809E-01	2.438	1.228E-01	1.639
	0.0312	7.323E-02	1.579	9.138E-03	2.168	1.424E-01	1.756	4.098E-02	1.584
	0.0156	2.403E-02	1.608	2.306E-03	1.987	5.002E-02	1.509	1.312E-02	1.643
2	0.1250	1.004E-01	–	3.313E-02	–	1.642E-01	–	6.442E-02	–
	0.0625	1.819E-02	2.465	3.315E-03	3.321	3.283E-02	2.322	1.115E-02	2.531
	0.0312	3.083E-03	2.561	3.887E-04	3.092	6.316E-03	2.378	1.791E-03	2.638
	0.0156	4.907E-04	2.651	6.161E-05	2.658	1.154E-03	2.453	2.713E-04	2.723
3	0.1250	1.016E-02	–	2.815E-03	–	1.823E-02	–	5.781E-03	–
	0.0625	8.531E-04	3.574	1.651E-04	4.092	1.705E-03	3.419	4.683E-04	3.626
	0.0312	6.857E-05	3.637	1.098E-05	3.910	1.540E-04	3.469	3.630E-05	3.690
	0.0156	5.239E-06	3.710	8.459E-07	3.698	1.317E-05	3.547	2.694E-06	3.752
4	0.1250	7.243E-04	–	2.445E-04	–	1.177E-03	–	4.613E-04	–
	0.0625	3.700E-05	4.291	7.196E-06	5.086	6.741E-05	4.126	2.298E-05	4.327
	0.0312	1.615E-06	4.518	2.170E-07	5.051	3.288E-06	4.358	9.600E-07	4.581
	0.0156	6.562E-08	4.621	8.533E-09	4.669	1.508E-07	4.447	3.730E-08	4.686

between $\left(k + \frac{1}{2}\right)$ and $(k + 1)$ for both scaled pressure p and scaled velocity $\mathbf{u}$. Recall that the optimal convergence rate of $k + 1$ is proved for only simplices, though similar results for quadrilaterals and hexahedra are expected. Indeed, Table 4 shows that as the solution order increases, the convergence rate is above $k + \frac{1}{2}$.

5. Conclusions and future work

In this paper, we developed numerical methods for both glacier dynamics and mantle convection. Both phenomena can be described by a two-phase mixture model, in which the mixture of the fluid and the solid is described by the porosity ϕ (i.e., $\phi > 0$ implies the fluid-solid two-phase and $\phi = 0$ means the solid single-phase region). The challenge is when the porosity vanishes because the system degenerates, which make the problem difficult to solve numerically. To address the issue, following [17], we start by scaling variables to obtain the well-posedness. Then we spatially discretize the system using the upwind HDG framework. The key feature is that we have modified the upwind HDG flux to accommodate the degenerate (one-phase) region. When the porosity vanishes, the unmodified HDG system becomes ill-posed because the stabilization parameter associated with the HDG flux disappears. For this reason, we introduce the generalized stabilization parameter that is composed of the upwind based parameter $\tau = \phi^{-\frac{1}{2}} d$ in the two-phase region and a positive parameter $\tau = \frac{1}{h} > 0$ in the one-phase region. This enabled us to develop a high-order HDG method for a linear degenerate elliptic equation arising from a two-phase mixture of both glacier dynamics and mantle convection.

We have shown the well-posedness and the convergence analysis of our HDG scheme. The rigorous theoretical results tell us that our HDG method has the convergence rates of $(k + 1)$ for a non-degenerate case and $\left(k + \frac{1}{2}\right)$ for a degenerate case with a piecewise smooth solution.

Several numerical results confirm that our proposed HDG method works well for linear degenerate elliptic equations. For the non-degenerate case, we obtain the $(k + 1)$ convergence rates of both the scaled pressure p and the

Table 3. Degenerate case with a smooth solution: the errors $\|\mathbf{u}^e - \mathbf{u}\|_{\Omega_h}$ and the convergence rates for the scaled velocity. Four cases are presented: τ given (49), $\tau = \frac{1}{h}$, $\tau = 1$ and $\tau = 10$.

k	h	$\tau = \begin{cases} \phi^{-\frac{1}{2}} d & \text{for } \phi > 0 \\ 1/h & \text{for } \phi = 0 \end{cases}$		$\tau = \frac{1}{h}$		$\tau = 1$		$\tau = 10$	
		error	order	error	order	error	order	error	order
1	0.1250	1.251E+01	–	7.416E+00	–	1.515E+01	–	1.028E+01	–
	0.0625	5.714E+00	1.130	2.229E+00	1.734	7.362E+00	1.041	4.479E+00	1.199
	0.0312	2.386E+00	1.260	5.664E-01	1.977	3.334E+00	1.143	1.784E+00	1.329
	0.0156	9.371E-01	1.348	1.505E-01	1.912	1.449E+00	1.202	6.660E-01	1.421
2	0.1250	2.911E+00	–	1.656E+00	–	3.649E+00	–	2.361E+00	–
	0.0625	5.996E-01	2.279	2.372E-01	2.804	8.361E-01	2.126	4.566E-01	2.371
	0.0312	1.154E-01	2.378	3.575E-02	2.730	1.821E-01	2.199	8.247E-02	2.469
	0.0156	2.080E-02	2.472	5.949E-03	2.587	3.770E-02	2.272	1.408E-02	2.550
3	0.1250	2.551E-01	–	1.237E-01	–	3.595E-01	–	1.876E-01	–
	0.0625	2.635E-02	3.275	9.486E-03	3.705	4.001E-02	3.168	1.885E-02	3.315
	0.0312	2.542E-03	3.374	6.954E-04	3.770	4.232E-03	3.241	1.763E-03	3.418
	0.0156	2.316E-04	3.456	5.436E-05	3.677	4.236E-04	3.321	1.565E-04	3.494
4	0.1250	2.951E-02	–	1.710E-02	–	3.615E-02	–	2.437E-02	–
	0.0625	1.717E-03	4.103	6.531E-04	4.710	2.348E-03	3.944	1.344E-03	4.181
	0.0312	8.585E-05	4.322	2.274E-05	4.844	1.321E-04	4.152	6.358E-05	4.402
	0.0156	3.994E-06	4.426	8.741E-07	4.701	6.999E-06	4.238	2.813E-06	4.498

scaled velocity $\mathbf{u}$ in two dimensions, whereas in three dimensions we observe the convergence rates above $\left(k + \frac{1}{2}\right)$. For the degenerate case with a smooth solution, the convergence rate of $\left(k + \frac{1}{2}\right)$ is observed for both the scaled pressure p and the scaled velocity $\mathbf{u}$. For the degenerate case with low solution regularity, the convergence rates of the numerical solutions are bounded by the solution regularity, but the high-order method still shows a benefit in terms of accuracy. Through a parameter study, we found that using a positive parameter on the one-phase region does not affect the accuracy of a numerical solution. We also found that $\tau = 1/h$ showed slightly better performance in terms of error levels and convergence rates for the degenerate case with smooth solution.

In order for our proposed method to work in two-phase flows, the interfaces between matrix solid and fluid melt need to be identified and grids should be aligned with the interfaces. In other words, the degeneracies are always required to lie on a set of measure zero. Note that we do not consider the full set of dynamical equations (1), (4), (7), (8) yet. We will tackle this challenge in a future work.

Appendix A. Auxiliary results

In this appendix we collect some technical results that are useful for our analysis.

Lemma 8 (Inverse Inequality [55, Lemma 1.44])**.** *For $v \in \mathcal{P}_k(K)$ with $K \in \Omega_h$, there exists $c > 0$ independent of h such that*

$$\|\nabla v\|_{0,K} \leq c h_K^{-1} \|v\|_{0,K} . \tag{A.1}$$

Lemma 9 (Trace inequality [55, Lemma 1.49])**.** *For $v \in H^1(\Omega_h)$ and for $K \in \Omega_h$ with $e \subset \partial K$, there exists $c > 0$ independent of h such that*

$$\|v\|_{0,e}^2 \leq c \left(\|\nabla v\|_{0,K} + h_K^{-1} \|v\|_{0,K} \right) \|v\|_{0,K} . \tag{A.2}$$

Table 4. Non-degenerate case in three dimensions: the results show that the HDG solutions for scaled pressure p and scaled velocity $\mathbf{u}$ converge to the exact solutions with the rate in above $k + \frac{1}{2}$.The upwind based parameter $\tau = \phi^{-\frac{1}{2}} d$ is used.

k	h	$\|p^e - p\|_2$ error	order	$\|\mathbf{u}^e - \mathbf{u}\|_2$ error	order
1	0.1250	2.553E-02	–	5.895E-01	–
	0.0833	1.473E-02	1.356	3.527E-01	1.267
	0.0625	9.754E-03	1.433	2.399E-01	1.340
	0.0500	6.994E-03	1.491	1.757E-01	1.396
2	0.1250	1.405E-03	–	4.408E-02	–
	0.0833	5.274E-04	2.417	1.712E-02	2.333
	0.0625	2.576E-04	2.491	8.579E-03	2.402
	0.0500	1.460E-04	2.545	4.961E-03	2.455
3	0.1250	4.872E-05	–	1.477E-03	–
	0.0833	1.183E-05	3.491	3.684E-04	3.425
	0.0625	4.234E-06	3.572	1.348E-04	3.495
	0.0500	1.885E-06	3.626	6.105E-05	3.550
4	0.1250	1.100E-06	–	2.741E-05	–
	0.0833	1.705E-07	4.598	4.428E-06	4.496
	0.0625	4.447E-08	4.672	1.188E-06	4.573
	0.0500	1.551E-08	4.720	4.235E-07	4.622

Applying the arithmetic-geometric mean inequality to the right side, we can derive

$$\|v\|_{0,e} \le c\left(h_k^{\frac{1}{2}} \|\nabla v\|_{0,K} + h_K^{-\frac{1}{2}} \|v\|_{0,K}\right). \tag{A.3}$$

If $v \in H^1(\Omega_h)$ is in a piecewise polynomial space, we can derive the following inequality from Lemma 9 and the inverse inequality (Lemma 8):

$$\|v\|_{0,e} \le c h_K^{-\frac{1}{2}} \|v\|_{0,K}. \tag{A.4}$$

Acknowledgements

The first and the second authors are partially supported by the NSF Grant NSF-DMS1620352. The third author is partially supported by NSF-DMS1720349. We are grateful for the support.

References

[1] T. Herring, Geodesy: treatise on geophysics, Elsevier, 2010.
[2] R. J. Chorley, B. A. Kennedy, Physical geography: a systems approach, Prentice Hall, 1971.
[3] G. Kaser, Glacier-climate interaction at low latitudes, Journal of Glaciology 47 (157) (2001) 195–204.
[4] A. Fowler, On the transport of moisture in polythermal glaciers, Geophysical & Astrophysical Fluid Dynamics 28 (2) (1984) 99–140.
[5] A. Aschwanden, E. Bueler, C. Khroulev, H. Blatter, An enthalpy formulation for glaciers and ice sheets, Journal of Glaciology 58 (209) (2012) 441–457.
[6] D. McKenzie, The generation and compaction of partially molten rock, Journal of Petrology 25 (3) (1984) 713–765.
[7] D. P. McKenzie, J. M. Roberts, N. O. Weiss, Convection in the earth's mantle: towards a numerical simulation, Journal of Fluid Mechanics 62 (3) (1974) 465–538.
[8] D. R. Scott, D. J. Stevenson, Magma solitons, Geophysical Research Letters 11 (11) (1984) 1161–1164.
[9] D. R. Scott, D. J. Stevenson, Magma ascent by porous flow, Journal of Geophysical Research: Solid Earth 91 (B9) (1986) 9283–9296.

[10] D. Bercovici, Y. Ricard, G. Schubert, A two-phase model for compaction and damage: 3. applications to shear localization and plate boundary formation, Journal of Geophysical Research: Solid Earth 106 (B5) (2001) 8925–8939.
[11] D. Bercovici, Y. Ricard, Energetics of a two-phase model of lithospheric damage, shear localization and plate-boundary formation, Geophysical Journal International 152 (3) (2003) 581–596.
[12] O. Šrámek, Y. Ricard, D. Bercovici, Simultaneous melting and compaction in deformable two-phase media, Geophysical Journal International 168 (3) (2007) 964–982.
[13] T. Arbogast, M. A. Hesse, A. L. Taicher, Mixed methods for two-phase darcy–stokes mixtures of partially melted materials with regions of zero porosity, SIAM Journal on Scientific Computing 39 (2) (2017) B375–B402.
[14] I. Hewitt, A. Fowler, Partial melting in an upwelling mantle column, in: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, Vol. 464, The Royal Society, 2008, pp. 2467–2491.
[15] H. P. G. Darcy, Les Fontaines publiques de la ville de Dijon. Exposition et application des principes à suivre et des formules à employer dans les questions de distribution d'eau, etc, Victor Dalmont, 1856.
[16] N. H. Sleep, Tapping of melt by veins and dikes, Journal of Geophysical Research: Solid Earth 93 (B9) (1988) 10255–10272.
[17] T. Arbogast, A. L. Taicher, A linear degenerate elliptic equation arising from two-phase mixtures, SIAM Journal on Numerical Analysis 54 (5) (2016) 3105–3122.
[18] T. Arbogast, A. L. Taicher, A cell-centered finite difference method for a degenerate elliptic equation arising from two-phase mixtures, Computational Geosciences 21 (4) (2017) 700–712.
[19] W. H. Reed, T. R. Hill, Triangular mesh methods for the neutron transport equation, Tech. Rep. LA-UR-73-479, Los Alamos Scientific Laboratory (1973).
[20] P. LeSaint, P. A. Raviart, On a finite element method for solving the neutron transport equation, in: C. de Boor (Ed.), Mathematical Aspects of Finite Element Methods in Partial Differential Equations, Academic Press, 1974, pp. 89–145.
[21] C. Johnson, J. Pitkäranta, An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation, Mathematics of Computation 46 (173) (1986) 1–26.
[22] J. Douglas, T. Dupont, Interior penalty procedures for elliptic and parabolic Galerkin methods, in: Computing methods in applied sciences, Springer, 1976, pp. 207–216.
[23] M. F. Wheeler, An elliptic collocation-finite element method with interior penalties, SIAM Journal on Numerical Analysis 15 (1) (1978) 152–161.
[24] D. N. Arnold, An interior penalty finite element method with discontinuous elements, SIAM journal on numerical analysis 19 (4) (1982) 742–760.
[25] B. Cockburn, G. E. Karniadakis, C.-W. Shu, The development of discontinuous Galerkin methods, in: Discontinuous Galerkin Methods, Springer, 2000, pp. 3–50.
[26] D. N. Arnold, F. Brezzi, B. Cockburn, L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM journal on numerical analysis 39 (5) (2002) 1749–1779.
[27] S. Tirupathi, J. S. Hesthaven, Y. Liang, M. Parmentier, Multilevel and local time-stepping discontinuous Galerkin methods for magma dynamics, Computational Geosciences 19 (4) (2015) 965–978.
[28] S. Tirupathi, J. S. Hesthaven, Y. Liang, Modeling 3d magma dynamics using a discontinuous Galerkin method, Communications in Computational Physics 18 (1) (2015) 230–246.
[29] A. R. Schiemenz, M. A. Hesse, J. S. Hesthaven, Modeling magma dynamics with a mixed fourier collocation-discontinuous Galerkin method, Communications in Computational Physics 10 (2) (2011) 433–452.
[30] B. Cockburn, J. Gopalakrishnan, 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.
[31] B. Cockburn, J. Gopalakrishnan, F.-J. Sayas, A projection-based error analysis of HDG methods, Mathematics Of Computation 79 (271) (2010) 1351–1367.
[32] R. M. Kirby, S. J. Sherwin, B. Cockburn, To CG or to HDG: A comparative study, J. Sci. Comput. 51 (2012) 183–212.
[33] N. C. Nguyen, J. Peraire, B. Cockburn, An implicit high-order hybridizable discontinuous Galerkin method for linear convection-diffusion equations, Journal Computational Physics 228 (2009) 3232–3254.
[34] B. Cockburn, B. Dong, J. Guzman, M. Restelli, R. Sacco, A hybridizable discontinuous Galerkin method for steady state convection-diffusion-reaction problems, SIAM J. Sci. Comput. 31 (2009) 3827–3846.
[35] H. Egger, J. Schoberl, A hybrid mixed discontinuous Galerkin finite element method for convection-diffusion problems, IMA Journal of Numerical Analysis 30 (2010) 1206–1234.
[36] B. Cockburn, J. Gopalakrishnan, The derivation of hybridizable discontinuous Galerkin methods for Stokes flow, SIAM J. Numer. Anal 47 (2) (2009) 1092–1125.
[37] N. C. Nguyen, J. Peraire, B. Cockburn, A hybridizable discontinous Galerkin method for Stokes flow, Comput Method Appl. Mech. Eng. 199 (2010) 582–597.
[38] N. C. Nguyen, J. Peraire, B. Cockburn, An implicit high-order hybridizable discontinuous Galerkin method for the incompressible Navier-Stokes equations, Journal Computational Physics 230 (2011) 1147–1170.
[39] D. Moro, N. C. Nguyen, J. Peraire, Navier-Stokes solution using hybridizable discontinuous Galerkin methods, American Institute of Aeronautics and Astronautics 2011-3407.
[40] N. C. Nguyen, J. Peraire, B. Cockburn, Hybridizable discontinuous Galerkin method for the time harmonic Maxwell's equations, Journal Computational Physics 230 (2011) 7151–7175.
[41] L. Li, S. Lanteri, R. Perrrussel, A hybridizable discontinuous Galerkin method for solving 3D time harmonic Maxwell's equations, in: Numerical Mathematics and Advanced Applications 2011, Springer, 2013, pp. 119–128.
[42] N. C. Nguyen, J. Peraire, B. Cockburn, High-order implicit hybridizable discontinuous Galerkin method for acoustics and elastodynamics, Journal Computational Physics 230 (2011) 3695–3718.
[43] R. Griesmaier, P. Monk, Error analysis for a hybridizable discontinous Galerkin method for the Helmholtz equation, J. Sci. Comput. 49 (2011)

291–310.
[44] J. Cui, W. Zhang, An analysis of HDG methods for the Helmholtz equation, IMA J. Numer. Anal. 34 (1) (2014) 279–295.
[45] S. Rhebergen, G. N. Wells, A hybridizable discontinuous Galerkin method for the navier–stokes equations with pointwise divergence-free velocity field, arXiv preprint arXiv:1704.07569.
[46] T. Bui-Thanh, From Godunov to a unified hybridized discontinuous Galerkin framework for partial differential equations, Journal of Computational Physics 295 (2015) 114–146.
[47] T. Bui-Thanh, From Rankine-Hugoniot Condition to a Constructive Derivation of HDG Methods, Lecture Notes in Computational Sciences and Engineering, Springer, 2015, pp. 483–491.
[48] T. Bui-Thanh, Construction and analysis of HDG methods for linearized shallow water equations, SIAM Journal on Scientific Computing 38 (6) (2016) A3696–A3719.
[49] J. Wang, X. Ye, A weak Galerkin finite element method for second-order elliptic problems, Journal of Computational and Applied Mathematics 241 (2013) 103–115.
[50] J. Wang, X. Ye, A weak Galerkin mixed finite element method for second order elliptic problems, Math. Comp. 83 (2014) 2101–2126.
[51] Q. Zhai, R. Zhang, X. Wang, A hybridized weak Galerkin finite element scheme for the Stokes equations, Science China Mathematics 58 (11) (2015) 2455–2472.
[52] L. Mu, J. Wang, X. Ye, A new weak Galerkin finite element method for the Helmholtz equation, IMA Journal of Numerical Analysis 35 (3) (2014) 1228–1255.
[53] E. F. Toro, Riemann solvers and numerical methods for fluid dynamics: a practical introduction, Springer Science & Business Media, 2013.
[54] S. Brenner, R. Scott, The mathematical theory of finite element methods, Vol. 15, Springer Science & Business Media, 2007.
[55] D. A. D. Pietro, A. Ern, Mathematical aspects of Discontinuous Galerkin methods, Springer, 2012.