

A UNIFIED HP-HDG FRAMEWORK FOR FRIEDRICHS' PDE SYSTEMS*

JAU-UEI CHEN[†], SHINHOO KANG[‡], TAN BUI-THANH[§], AND JOHN N. SHADID[¶]

This paper is dedicated to Professor Leszek Demkowicz on the occasion of his 70th birthday.

Abstract. This work proposes a unified hp -adaptivity framework for hybridized discontinuous Galerkin (HDG) method for a large class of partial differential equations (PDEs) of Friedrichs' type. In particular, we present unified hp -HDG formulations for abstract one-field and two-field structures and prove their well-posedness. In order to handle non-conforming interfaces we simply take advantage of HDG built-in mortar structures. With split-type mortars and the approximation space of trace, a numerical flux can be derived via Godunov approach and be naturally employed without any additional treatment. As a consequence, the proposed formulations are parameter-free. We perform several numerical experiments for time-independent and linear PDEs including elliptic, hyperbolic, and mixed-type to verify the proposed unified hp -formulations and demonstrate the effectiveness of hp -adaptation. Two adaptivity criteria are considered: one is based on a simple and fast error indicator, while the other is rigorous but more expensive using an adjoint-based error estimate. The numerical results show that these two approaches are comparable in terms of convergence rate even for problems with strong gradients, discontinuities, and singularities.

Key words. Hybridized Discontinuous Galerkin; Friedrichs' system; Discontinuous Galerkin; Hybridization; hp -adaptation

AMS subject classifications. 68Q25, 68R10, 68U05

1. Introduction. The hybridized discontinuous Galerkin (HDG) methods were first introduced in [32] and they inherit many benefits of discontinuous Galerkin (DG) methods including the applicability to a wide variety of partial differential equations (PDEs), the capability of handling complex geometries, and high-order accuracy support, to name a few. In addition, HDG methods improve computational efficiency [30] by condensing out the local unknowns, and the linear system to be solved for the trace unknowns on the mesh skeleton is smaller than DG counterparts. With these favorable advantages, HDG methods indeed have great success solving various kinds of PDEs such as Poisson equation [30, 34, 70], convection-diffusion equations [86, 87, 55], Stokes equations [31, 33, 88, 37, 68], Navier-Stokes equations [91, 24], Maxwell equation [90, 79], acoustics and elastodynamics equations [89], Helmholtz equation [61, 39], and magneto-hydrodynamic equations [74], to mention a few. A constructive and unified HDG framework for a large class of physics governed by elliptic, parabolic, hyperbolic, and mixed-typed PDEs has been developed in [20] that not only rediscovers most of the existing HDG methods but also discovers new ones.

As with any numerical discretization method, standard HDG could be inefficient in some crucial situations where high gradient, discontinuous, and/or singular features are present. Unfortunately, these extreme features are not uncommon in almost all engineering/physics applications. A cure to this issue is to employ hp -adaptivity. The idea is first proposed in [4] and is systematically studied in [62, 63]. It consists of two key findings. The first one is that an exponential convergence rate can be attained by uniformly increasing the degree of approximation (p -refinement) if the solution is regular enough [5]. The second one is that a low degree of approximation along with refined mesh (h -refinement) is desired if the solution is non-smooth. In brief, the ideal situation is to locally execute either h - or p - refinement according to the local behavior of the solution.

In fact, the adaptive feature has been routinely applied in the context of HDG methods either through h -adaptivity [23, 85, 46, 40, 107, 27, 35, 96, 2, 52, 77, 83, 101, 76, 75, 99, 7, 38, 78], p -adaptivity [58, 57, 65, 97, 56, 82], or hp -adaptivity [8, 105, 106, 9]. To drive the adaptation process, some indicator is necessary. There are three popular approaches: *a posteriori* error estimator, adjoint-based error estimate, and heuristic indicator. Although the reliability and efficiency of an *a posteriori* error estimator sometimes can be guaranteed [46, 27, 2], the derivation is problem-dependent and is typically non-trivial, especially for nonlinear problems, [23, 46, 27, 2, 77, 76, 75, 99]. In addition, a post-process may be required in this

*Submitted to the editors DATE.

Funding: This work was funded by some funds

[†]Department of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin, Austin, TX 78712, USA (chenju@utexas.edu).

[‡]Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, IL 60439, USA.

[§]Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX 78712, USA

[¶]Computational Mathematics Department, Sandia National Laboratories, P.O. Box 5800, MS 1321, Albuquerque, NM 87185, US

type of error estimator [46, 2, 99, 58, 57]. On the other hand, an adjoint-based error estimate is popular in engineering applications [8, 40, 107, 106, 9, 52, 82, 38] since, in this scenario, one is usually more interested in some specific quantities instead of the solution itself. An adaptation process driven by an adjoint-based error estimate has been developed for computing accurate values for such quantities of interest. Finally, some heuristic indicator can also be employed in driving adaptation [66] and is typically inexpensive to be computed. For example, in [96, 65], a measure of jump of flux is used as an error indicator. However, an error indicator is not necessarily associated with “error”. For instance, authors in [85, 7] take advantage of artificial viscosity as an indicator and the work [83] uses damage-field as an indicator. In this paper, we study both adjoint-based error estimate and error indicator. The first option is easier to be derived compared to *a posteriori* error estimator but it still possesses certain robustness [8, 40, 107, 106, 9]. The second option is more ad-hoc and lacks robustness. The discussion of both approaches will be addressed in Section 4 and the numerical comparison will be made in Section 5.

The adaptation procedure can involve either h -refinement or p -refinement or both for the mesh under consideration. The local mesh refinement can be achieved without having any hanging node at the element boundaries for simplex meshes. In this case, no special treatment is required. The classic algorithms without re-meshing are bisection [94, 93, 100] and red-green procedures [10]. Another approach is to simply re-generate the mesh where the layout of the small and large elements depends on some metric [106, 52, 77, 82, 7, 78]. This mechanism is usually more computational expensive but the resulting mesh is more economical. We would like to mention that, however, h -nonconforming interfaces are typically involved in local h -refinement for quadrilateral and hexahedral meshes. Thanks to natural built-in mortars in HDG methods, the relevant techniques can be easily utilized to treat nonconforming interfaces. In addition, the issue of p -nonconforming interfaces due to local enrichment of approximation space can also be addressed by the mortar techniques. Nonetheless, special attention is needed for curved boundaries [97, 56]. In this work, though our hp -HDG approaches are valid for triangular/tetrahedral/quadrilateral/hexahedral elements with straight edges/faces in both 2D and 3D, our numerical results are only for two-dimensional problems with triangle elements.

So far, we have reviewed various HDG schemes for solving different physical problems. Since each physics has a unique characteristic, it is natural to develop different numerical scheme for different problem. However, the PDEs of Friedrichs’ type [54] embraces a large class of PDEs with similar mathematical structure and this provides an opportunity of developing a single unified framework. This idea is first adopted in a series of papers [49, 50, 51] in the analysis of DG methods. Friedrichs’ system is also the basis to unify various discontinuous Petrov-Galerkin methods [21]. In the work [20], the author uses Friedrichs’ system to propose a unified and constructive framework for HDG schemes via a Godunov approach, with the assumption that the interfaces are conforming.

This paper extends the work in [20] in two important directions. First, our extension now provides a unified HDG framework for PDEs with two-field structure (to be defined). Second we develop two unified hp -HDG frameworks: one for one-field PDE structure and another for two-field PDE structure. In particular, we consider Friedrichs’ systems with more general assumptions that cover one- and two-field structures. For two-field structures, both full and partial coercivities are examined. The resulting approaches thus cover a wide range PDEs including hyperbolic, elliptic, or mixed-type PDEs. We then propose two hp -HDG formulations: one for one-field PDEs and the other for two-field PDEs. The derivation heavily relies on the Godunov approach. For the two-field formulation, we further exploit its intrinsic structure to obtain the corresponding reduced trace system. A few assumptions are identified to guarantee the existence of the numerical flux, and this is also a key to prove the well-posedness. Several numerical experiments are carried out to verify the effectiveness of the abstract hp -HDG formulations when applied to specific PDEs. In order to drive the adaptivity, an adhoc error indicator and an adjoint-based error estimation are implemented and their performance are compared. As shall be shown, using either of these criteria, numerically polluted areas induced by high gradient/discontinuity/singular can indeed decrease through the hp -adaptation process, and acceptable convergence rates can be attained in many cases.

The paper is organized as follows. Section 2 briefly reviews Friedrichs’ systems and outlines important assumptions that will be used in the well-posedness analysis. In Section 3, key concepts about mortar techniques are discussed. In addition, HDG numerical fluxes and the corresponding hp -HDG formulations for one-field and two-field Friedrichs’ systems are derived. The well-posedness of these formulations is then proved. The hp -adaptation strategy with adaptive criteria based on ad-hoc and adjoint-based error indicator

is presented in Section 4. Several numerical examples for elliptic PDEs (with corner singularity, anisotropic diffusion with discontinuous boundary condition, heterogeneous anisotropic with discontinuous diffusivity field), linear hyperbolic PDE (with variable speed and discontinuous boundary condition), and convection-diffusion PDE (with boundary layer and discontinuous boundary condition) are presented in Section 5. Section 6 concludes the paper with future work.

2. Linear PDEs of Friedrichs' type. The main idea of Friedrichs' unification of PDEs [54] is to cast wide classes of PDEs into the first order systems which share the same mathematical structure. In this section, we outline one-field and two-field PDE of Friedrichs' type. The following notations are used in the paper. Boldface lowercases are reserved for (column) vectors, uppercase letters are for matrices, and boldface uppercase letters are for third order tensors. Considering the following general system of linear PDEs defined in a Lipschitz domain $\Omega \subset \mathbb{R}^d$, where d refers to the spatial dimension:

$$(2.1) \quad \sum_{k=1}^d \partial_k F_k(\mathbf{z}) + G\mathbf{z} := \sum_{k=1}^d \partial_k (A_k \mathbf{z}) + G\mathbf{z} = \mathbf{f} \text{ in } \Omega,$$

where $F_k(\mathbf{z}) := A_k \mathbf{z}$ is the k -th component of the flux tensor $\mathbf{F}(\mathbf{z})$, $\mathbf{z}$ the unknown solution with values in $\mathbb{R}^m$, and $\mathbf{f} \in [\mathcal{L}^2(\Omega)]^d$ the forcing term. Here, $\mathcal{L}^2(\Omega)$ is the space of square-integrable functions on Ω . Additionally, ∂_k stands for the k -th (component-wise) partial derivative. Different types of constraints imposed on A_k and G will result in different types of Friedrichs' systems, and we shall discuss each case separately in the following sub-sections.

2.1. One-field Friedrichs' systems. One-field Friedrichs' systems come with the following standard assumptions [54, 67, 49]:

$$(A.1) \quad G \in [\mathcal{L}^\infty(\Omega)]^{m,m}$$

$$(A.2) \quad \forall k \in \{1, \dots, d\}, A_k \in [\mathcal{L}^\infty(\Omega)]^{m,m} \text{ and } \sum_{k=1}^d \partial_k A_k \in [\mathcal{L}^\infty(\Omega)]^{m,m}.$$

$$(A.3) \quad \forall k \in \{1, \dots, d\}, A_k = (A_k)^T \text{ a.e. in } \Omega.$$

$$(A.4) \quad \exists \mu_0 > 0, G + G^T + \sum_{k=1}^d \partial_k A_k \geq 2\mu_0 I_m \text{ a.e. in } \Omega.$$

where I_m is the $m \times m$ identity matrix. In this paper, the inequality of the type (A.4) stands for the semi-positive definiteness of the difference between the matrices on the left-hand side and the right-hand side. Inequality (A.4) is also known as full coercivity [49]. Here, $[\mathcal{L}^\infty(\Omega)]^{m,m}$ denotes the space of $m \times m$ matrix-valued essentially bounded functions on Ω . It turns out that any symmetric and strictly hyperbolic PDE system is an example of one-field Friedrichs' system. The advection equation, for example, falls into this category and it is discussed in Section 5.2. More examples can be found in Sections 3.1, 3.2 in [49].

2.2. Two-field Friedrichs' systems. Let m^σ and m^u be two positive integers such that $m = m^\sigma + m^u$. Denote $\mathcal{L}_\sigma := [\mathcal{L}^2(\Omega)]^{m^\sigma}$, $\mathcal{L}_u := [\mathcal{L}^2(\Omega)]^{m^u}$, and $\mathcal{L} := \mathcal{L}_\sigma \times \mathcal{L}_u$. Suppose we have the decomposition $\mathbf{z} = (\boldsymbol{\sigma}, \mathbf{u})$ for all $\mathbf{z} \in \mathcal{L}$, and

$$(2.2) \quad G = \begin{bmatrix} G^{\sigma\sigma} & G^{\sigma u} \\ G^{u\sigma} & G^{uu} \end{bmatrix}, \quad A_k = \begin{bmatrix} A_k^{\sigma\sigma} & B_k \\ B_k^T & C_k \end{bmatrix}, \quad k \in \{1, \dots, d\}.$$

Note that C_k is symmetric owing to (A.3). Two additional key assumptions on which the two-field theory is based are [50]:

$$(A.5) \quad A_k^{\sigma\sigma} = 0, \forall k \in \{1, \dots, d\},$$

$$(A.6) \quad G^{\sigma\sigma} \geq k_0 I_{m^\sigma} \text{ for some } k_0 > 0,$$

where I_{m^σ} is the identity matrix in $\mathbb{R}^{m^\sigma, m^\sigma}$. Assumptions (A.5)-(A.6) allow us to eliminate the $\boldsymbol{\sigma}$ -component of $\mathbf{z}$ in the PDE system and the resulting differential equation is an elliptic-like PDE for the $\mathbf{u}$ -component. The two-field Friedrichs' systems that satisfy assumptions (A.1)-(A.6) cover a wide variety of PDEs including convection-diffusion-reaction equation, compressible linear continuum mechanics with a reaction term, and simplified MHD. These examples are studied in Section 3 in [50].

We note that the positivity condition (A.4) can be further relaxed to account for systems that have two-field structures with partial coercivity. This class includes convection-diffusion, anisotropic diffusion, and typical compressible linear continuum mechanics (e.g., linearized compressible elasticity or linearized compressible Navier-Stokes) equations, to name a few. This can be accomplished (see [51]) by replacing assumption (A.4) with the following:

(A4.a) $\exists \mu_0 > 0, G + G^T + \sum_{k=1}^d A_k \geq 2\mu_0 I_{m^\sigma}^\sigma$ a.e. in Ω , where $I_{m^\sigma}^\sigma$ is an $m \times m$ matrix defined as

$$I_{m^\sigma}^\sigma := \begin{bmatrix} I_{m^\sigma} & 0 \\ 0 & 0 \end{bmatrix}.$$

(A4.b) $G^{\sigma u} = (G^{u\sigma})^T = 0$ and B_k are constant over Ω .

The diffusion equations are discussed in Section 5.1 and the convection-diffusion equation is discussed in Section 5.3. In addition, compressible linear continuum mechanics is discussed in Section 3.4 in [51].

REMARK 1. *Here, we omit one additional inequality ((A3B'') in [51]) required for two-field Friedrichs' systems with partial coercivity since it will not be used in our analysis of HDG. However, the inequality is critical in the proof of well-posedness of the continuous PDE stated in (2.1). Such an inequality can be viewed as a generalized form of Friedrichs–Poincaré [18] or Korn's [19] inequality, and the discrete version of it can actually be used in our analysis which, however, will lead to a mesh-dependent HDG scheme.*

2.3. Boundary conditions. Though the numerical results in Section 5 use non-homogeneous boundary conditions, it is sufficient to show the well-posedness of the one-field setting (A.1)-(A.4) and the two-field setting (A.1)-(A.6) (or (A.1)-(A.3), (A4.a)-(A4.b), (A.5)-(A.6)) with homogeneous boundary condition. Similar to [49, 50, 51], we consider a general homogeneous boundary condition of the following form

$$(2.3a) \quad (A - M)z = 0, \quad \text{on } \partial\Omega,$$

where $M : \partial\Omega \rightarrow \mathbb{R}^{m,m}$ and $A := \sum_{k=1}^d n_k A_k$ with $n = (n_1, \dots, n_d)^T$ being the unit outward vector of $\partial\Omega$. In addition, we assume that

$$(2.3b) \quad M \geq 0,$$

$$(2.3c) \quad \mathcal{N}(A - M) + \mathcal{N}(A + M) = \mathbb{R}^m,$$

with $\mathcal{N}(\cdot)$ denoting the nullspace of its argument. It should be noted that the definition of M depends on the boundary, that is, different choices of M associate with different boundary conditions.

3. hp -HDG Formulations. In this section, we are going to derive hp -HDG formulations for Friedrichs' systems outlined in Section 2. Toward formulating an hp -HDG scheme, it is essential to derive a numerical flux due to discontinuous approximation space(s) used for the volume unknown(s). The well-known Godunov approach, which involves solving the Riemann problem either exactly or approximately, is one of the most popular methods to construct numerical fluxes. The key is to realize that the Godunov flux can be hybridized [20]. In other words, the Godunov flux¹ can be defined implicitly along with trace unknown(s) and thus can be employed as an HDG numerical flux. In addition, such an approach is desirable since it can lead to a parameter-free scheme. The idea of hybridizing the upwind flux to constructively and systematically derive HDG methods for abstract (and particular) PDEs is thoroughly discussed in [20]. In that work [20], we only discussed conforming HDG approaches. In this paper, we extended this idea to derive the upwind HDG methods for Friedrichs' system with hp -nonconforming meshes. The key ingredient to handle hp -nonconforming interfaces is to construct such flux directly on the mortars which are naturally built-in HDG methods. As we shall show, this can be achieved with the specific choice of the configuration of the mortars and the approximation space(s) of trace unknown(s). Further, we will show that the construction is quite straightforward for one-field Friedrichs' system. For two-field systems, it is less so especially when we would like to reduce the number of trace unknowns in our hp -HDG formulations for the sake of efficiency. As pointed out in Lemma 3.5, in order to accomplish this goal, we require a few more assumptions to derive such efficient upwind HDG schemes on an abstract level.

3.1. Nomenclatures. This section collects notations and conventions for the rest of the paper. Again, boldface lowercases are reserved for (column) vectors, uppercase letters are for matrices, and boldface uppercase letters are for third order tensors. A partition $\mathcal{T}_h$ of the domain $\Omega \subset \mathbb{R}^d$ is a finite collection of disjoint elements K such that $\cup_{K \in \mathcal{T}_h} \bar{K} = \bar{\Omega}$ where the mesh size h is defined as $\max_{K \in \mathcal{T}_h} \text{diam}(K)$. For the simplicity of the exposition, we will use two-dimensional simplex elements to convey our idea, though our approach is valid for three-dimensional settings as well. The set of elemental boundaries is denoted

¹It should be noted that the Godunov flux is simply an upwind flux if the problem of interest is linear and thus we may use these two terms interchangeably in this paper without confusion.

by $\partial\mathcal{T}_h = \{\partial K \mid K \in \mathcal{T}_h\}$ each of which comes with unit outward normal vector $\mathbf{n}^K$. We conventionally identify $\mathbf{n}$ as the normal vector on the boundary ∂K of element K (also denoted as K^-) and $\mathbf{n}^+ = -\mathbf{n}^-$ as the normal vector of the boundary of a neighboring element (also denoted as K^+). An element K^+ is said to be a neighbor of the element K^- when $\partial K^+ \cap \partial K^-$ has a positive $d-1$ Lebesgue measure. For an element K of the partition $\mathcal{T}_h$, we define a face of the element $K \in \mathcal{T}_h$ by $F \subset \partial K$. For an interior interface (nonconforming or not), we introduce a *mortar* e as $e = \partial K^+ \cap \partial K^-$, and $e = \partial K \cap \partial\Omega$ on the boundary of Ω . Note that $e = F^+ = F^-$ if we have a conforming interface (i.e., $F^+ = F^-$), otherwise, they are different (see Figure 1 and Figure 2). For any conforming interface the mortar e is clearly defined and for any nonconforming interface the mortar is defined in Section 3.2.1. The collection of mortars, called *mesh skeleton*, is denoted by $\mathcal{E}_h$, $\mathcal{E}_h = \mathcal{E}_h^\circ \cup \mathcal{E}_h^\partial$ with $\mathcal{E}_h^\partial = \{e \in \mathcal{E}_h \mid e \subset \partial\Omega\}$ and $\mathcal{E}_h^\circ = \mathcal{E}_h \setminus \mathcal{E}_h^\partial$. The derivation of an HDG scheme is centered around the HDG numerical flux which typically comes with the newly introduced unknowns residing on the skeleton. Such unknowns are usually termed trace unknowns while the usual unknowns defined within elements, such as the ones in DG methods, are termed as volume unknowns.

For the quantity f that is possibly double-valued on the mesh skeleton, we define the jump of f on $e \in \mathcal{E}_h$ as:

$$[[f]] = f^- + f^+, \quad \text{for } \forall e \in \mathcal{E}_h^\circ, \quad [[f]] = f, \quad \text{for } \forall e \in \mathcal{E}_h^\partial,$$

where $f^\pm(x) = \lim_{y \rightarrow x, y \in K^\pm} f(y)$.

We define $\mathcal{P}^p(D)$ as the space of polynomials of degree at most p on a domain D . In particular, we denote the degree of polynomials in an element K by p_K and on a mortar e by p_e . Next, we introduce discontinuous piecewise polynomial spaces

$$\begin{aligned} \mathbf{W}_h &= [P_h]^m, & P_h &:= \{z_h \in \mathcal{L}^2(\mathcal{T}_h) : z_h|_K \in \mathcal{P}^{p_K}(K), \forall K \in \mathcal{T}_h\}, \\ \widehat{\mathbf{W}}_h &= [\widehat{P}_h]^m, & \widehat{P}_h &:= \{\widehat{z}_h \in \mathcal{L}^2(\mathcal{E}_h) : \widehat{z}_h|_e \in \mathcal{P}^{p_e}(e), \forall e \in \mathcal{E}_h\}. \end{aligned}$$

To account for various boundary conditions, we denote $\partial\Omega_D$ as Dirichlet type of boundary, $\partial\Omega_N$ as Neumann type of boundary, and $\partial\Omega_R$ as Robin type boundary. The boundary now can be decomposed as $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N \cup \partial\Omega_R$ where the intersections of any two types of boundaries are empty set. To facilitate the discussion of the two-field Friedrichs' system later, we further introduce some additional approximation spaces:

$$\begin{aligned} \Sigma_h &= [P_h]^{m^\sigma}, & U_h &= [P_h]^{m^u}, \\ \widehat{\Sigma}_h &= [\widehat{P}_h]^{m^\sigma}, & \widehat{U}_h &= [\widehat{P}_h]^{m^u}. \end{aligned}$$

Finally, we define the inner product for the aforementioned finite element spaces. $(\cdot, \cdot)_D$ is defined as the $\mathcal{L}^2$ -inner product on a domain $D \subset \mathbb{R}^d$ and $\langle \cdot, \cdot \rangle_D$ as the $\mathcal{L}^2$ -inner product on a domain D if $D \subset \mathbb{R}^{d-1}$. To make our presentation more concise, we introduce the following definitions:

$$\begin{cases} (\cdot, \cdot)_{\mathcal{T}_h} := \sum_{K \in \mathcal{T}_h} (\cdot, \cdot)_K, \\ \langle \cdot, \cdot \rangle_{\partial\mathcal{T}_h} := \sum_{\partial K \in \partial\mathcal{T}_h} \sum_{F \subset \partial K} \sum_{e \subseteq F} \langle \cdot, \cdot \rangle_e, \\ \langle \cdot, \cdot \rangle_{\partial K} := \sum_{F \subset \partial K} \sum_{e \subseteq F} \langle \cdot, \cdot \rangle_e, \\ \langle \cdot, \cdot \rangle_{\mathcal{E}_h} := \sum_{e \in \mathcal{E}_h} \langle \cdot, \cdot \rangle_e, \\ \langle \cdot, \cdot \rangle_{\mathcal{E}_h^\partial} := \sum_{e \in \mathcal{E}_h^\partial} \langle \cdot, \cdot \rangle_e. \end{cases}$$

3.2. Mortar-based technique. A mortar technique is characterized by the introduction of mortars, finite element spaces on the mortars, and the method that uses mortars to patch the subdomains/elements. Our mortar approach is built upon four mortar approaches, all of which share the aforementioned three steps, and the key difference is the way they compute the mortar unknowns. The first approach is due to [81, 1, 15, 16, 17, 11], originally developed for elliptic PDEs, that uses the mortar unknowns to weakly maintain the continuity of the solution across the mortars. In this case, the mortar unknowns are solved together with the volume unknowns on subdomains or elements. The second approach was developed in [71, 72, 69, 22, 53] for hyperbolic PDEs in the context of spectral element and DG approaches. The upwind

states, which can be considered as the mortar unknowns [20], are computed on the mortars to construct the numerical fluxes to ensure flux conservation across the mortars. The third approach by [14, 104, 3, 92], originally developed in the context of mixed finite element methods for elliptic-type PDEs, calls the mortar unknowns as Lagrange multipliers and, similar to the first approach, they are solved together with the volume unknowns. The key difference is that the weak continuity of the flux is enforced instead of the weak continuity of the solution. Finally, the HDG approach [32, 30, 34, 70, 86, 87, 55, 31, 33, 88, 37, 68, 74, 20], in which the mortar unknowns are called trace unknowns, uses mortar unknowns to enforce the weak continuity of the flux similar to the third approach. The mortar unknowns are also solved together with the volume unknowns.

In this paper, we extend the HDG built-in mortars to fully account for hp -nonconforming interfaces. To that end, two ingredients are required: i) the appropriate choice of mortar configuration, and ii) the finite element space defined over the mortars. We will show that our choice, without any additional interpolation or projection, can lead to a setting where the Riemann problem is well-defined and the numerical flux can be derived via the Godunov approach.

3.2.1. h -nonconforming interfaces. In h -nonconforming interfaces (see Figure 1a), F^+ is not necessarily equal to F^- . We hence need to carefully consider the definition of mortar. There are two options for constructing a mortar as shown in Figure 1. In the first option, a set of *split-sided* mortars (i.e., Figure 1b) are deployed to conform to the smaller sides of the adjacent elements, while in the second option a *full-sided* mortar (i.e., Figure 1c) is used to conform to the larger side of the adjacent element. In the context of HDG methods, the first option is used in [28, 46, 29] and the second one is used in [96, 40, 83]. Although the usage of the full-sided mortar can be less computationally intensive, the split-sided mortars (Figure 1b) are chosen in this work to facilitate the implementation of the Godunov approach.

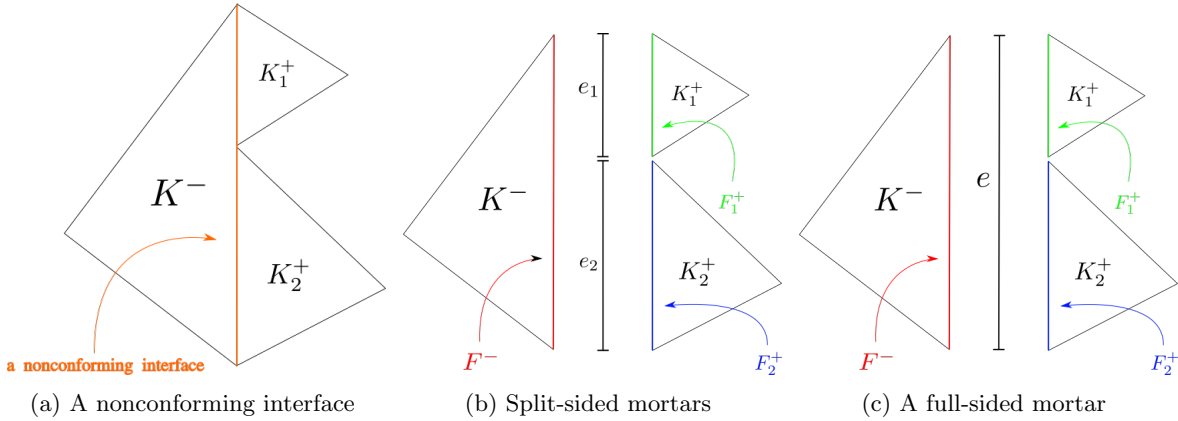


Fig. 1: Two options of mortars on a nonconforming interface.

REMARK 2. It should be noted that in [73] the authors showed theoretically that either full- or split-sided mortar can lead to a stable DG scheme on a discrete level for a time-dependent linear elasticity problem. Additionally, they also showed numerically that both type of mortars can lead to conservative schemes. However, split-sided mortar is still suggested in [73] in the sense that

1. it is the most natural approach for DG methods,
2. full-sided mortar has a spectral radius more than twice as large as the split-sided mortar (hence, more restrictive time step size for explicit methods).

3.2.2. p -nonconforming interfaces. For p -nonconforming interfaces, $p_{K^+} = p_{K^-}$ does not hold in general. Furthermore, the degree of approximation of trace unknowns p_e could differ from p_{K^+} or p_{K^-} . In this work, we choose:

$$(3.1) \quad \begin{cases} p_e = \max \{p_{K^+}, p_{K^-}\}, & \text{for } \forall e \in \mathcal{E}_h^o, \\ p_e = p_K, & \text{for } \forall e \subset \partial K \cap \mathcal{E}_h^\partial, \end{cases}$$

to facilitate the Godnuov approach and stability. This choice is also suggested in [32].

3.2.3. hp -nonconforming interfaces. By combining the setting presented in Section 3.2.1 and Section 3.2.2, we can now handle hp -nonconforming interfaces and construct an hp -HDG scheme for the Friedrichs' system using the Godnuov approach. In addition, neither projection nor interpolation is required: thanks to our specific selection of the configuration of mortars and of the degree of approximation of the trace unknowns.

To illustrate the idea, we consider the nonconforming interface shown in Figure 1b and focus on the segment e_1 . It is not clear how to implement the Godnuov approach since the Riemann problem is not well-defined. The left state and right state are defined on the domains that do not conform to each other (i.e $F^- \neq F_1^+$). To resolve this issue, we can either project and interpolate the states onto the mortar e_1 . Then, the Godnuov approach can be applied by solving the Riemann problem that is properly defined by these intermediate states. This methodology is already proven to be successful in the context of DG methods [71, 72, 69, 22, 53]. *Throughout the paper we assume that all the edges/faces are straight, that is, the meshes are affine. Curved elements are more delicate to treat and this will be part of our future work.* Owing to the natural built-in mortar in HDG methods and the way we handle the nonconforming interfaces, both projection, and interpolation are actually implicitly implied. To see this, we consider the following piecewise polynomial functions $\mathbf{z}_h \in \mathbf{W}_h$, $(\mathbf{z}_h^*)^\pm \in \widehat{\mathbf{W}}_h$ and $\widehat{\mathbf{w}}_h \in \widehat{\mathbf{W}}_h$. Moreover, we define a projection operator $\mathbb{P}(\cdot)$ that is the $\mathcal{L}^2$ -projection into the space $\widehat{\mathbf{W}}_h$. The projection from left state defined on F^- to the left intermediate state $(\mathbf{z}_h^*)^-$ can be stated as $\mathbb{P}(\mathbf{z}_h|_{F^- \cap e_1}) =: (\mathbf{z}_h^*)^-|_{e_1}$ and the equality holds in the sense that

$$(3.2) \quad \langle \mathbf{z}_h, \widehat{\mathbf{w}}_h \rangle_{e_1} = \langle (\mathbf{z}_h^*)^-, \widehat{\mathbf{w}}_h \rangle_{e_1} \quad \forall \widehat{\mathbf{w}}_h \in \widehat{\mathbf{W}}_h.$$

Since we use split-sided mortars and choose degree approximation of the trace test space $\widehat{\mathbf{W}}_h$ by Eq. (3.1), it is obvious that for any polynomial function $f \in \mathcal{P}^{p_K-}(K^-)$ it has to satisfy that $f|_{F^- \cap e_1} \subseteq \mathcal{P}^{p_e}(e_1)$. Due to the unique representation of polynomials, the projection actually does nothing here, and hence $\mathbb{P}(\mathbf{z}_h|_{F^- \cap e_1}) = \mathbf{z}_h|_{F^- \cap e_1}$. As consequence, the left intermediate state is nothing but just the restriction of the left state: $(\mathbf{z}_h^*)^-|_{e_1} = \mathbf{z}_h|_{F^- \cap e_1}$. The same argument can also be made in terms of interpolation. Similarly, we have the right intermediate state $(\mathbf{z}_h^*)^+|_{e_1} = \mathbf{z}_h|_{F^+ \cap e_1}$. Now the upwind numerical flux can be constructed by solving the Riemann problem locally along the normal $\mathbf{n}$ of the segment $F^- \cap e_1$. *Given that being aligned with a single direction is one-dimension in nature, a line along the normal direction $\mathbf{n}$ can be parameterized by some scalar x where $x = 0$ corresponds to the location of the mortar e_1 (see also Figure 2 for the illustration). By extending the definition of the coefficient matrix $A := \sum_{k=1}^d n_k A_k$ with $\mathbf{n} = (n_1, \dots, n_d)^T$ being unit outward vector of ∂K for $\forall K \in \mathcal{T}_h$, the statement of the Riemann problem [102] reads: find $\mathbf{z}_h(x, t)$ such that*

$$(3.3) \quad \frac{\partial \mathbf{z}_h}{\partial t} + \frac{\partial (A \mathbf{z}_h)}{\partial x} = \mathbf{0},$$

with initial condition $\mathbf{z}_h(x, 0) = (\mathbf{z}_h^*)^-$ for $x < 0$, $\mathbf{z}_h(x, 0) = (\mathbf{z}_h^*)^+$ for $x > 0$. Here, (artificial) time t is introduced to help understand the Godunov flux via the Riemann problem, but it is otherwise not necessary in the derivation. Figure 2 illustrates the idea of how the Riemann problem is defined in direction $\mathbf{n}$ that is parametrized by x . With the well-defined problem (3.3), we are now in the position to derive upwinding HDG flux by following the procedure outlined in [20]. In this paper, the coefficient matrix A will be assumed to be continuous across the mesh skeleton². In particular, A is symmetric according to (A.3) and hence its eigen-decomposition is guaranteed to exist. We thus can also define $|A| := R|\Lambda|R^{-1}$ where $\Lambda := \text{diag}(\lambda_1, \dots, \lambda_m)$ and λ_i are eigenvalues of A , and R is the matrix composed by the corresponding eigenvectors.

3.3. A constructive derivation of an hp -HDG formulation.

3.3.1. Friedrichs' system with one-field structure. In this section, we derive hp -HDG formulation for linear PDE in Eq. (2.1) that satisfies one-field Friedrichs' system assumptions (A.1)-(A.4). To begin, we apply Galerkin approximation to Eq. (2.1) on an element $K \in \mathcal{T}_h$ together with integration by parts. The

²This condition can be relaxed, but we will use it to keep the presentation concise.

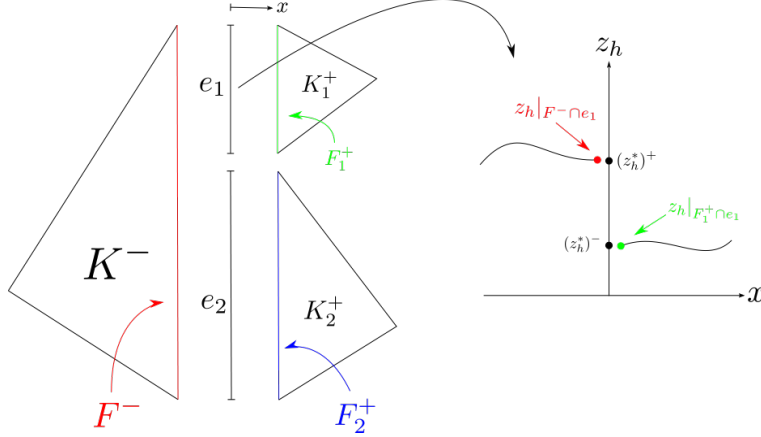


Fig. 2: Illustration of how the Riemann problem is defined along a normal direction $\mathbf{n}$ where a line aligned in this direction can be defined as $\mathbf{x}_0 + x\mathbf{n}$ with $\mathbf{x}_0 \in e_1$.

resulting local problem reads: seek $\mathbf{z}_h \in \mathbf{W}_h$ such that

$$(3.4) \quad -\sum_{k=1}^d (A_k \mathbf{z}_h, \partial_k \mathbf{w}_h)_K + (G \mathbf{z}_h, \mathbf{w}_h)_K + \langle \mathbf{F}(\mathbf{z}_h) \mathbf{n}, \mathbf{w}_h \rangle_{\partial K} = (\mathbf{f}, \mathbf{w}_h)_K, \quad \forall K \in \mathcal{T}_h,$$

for all $\mathbf{w}_h \in \mathbf{W}_h$ and the flux $\mathbf{F}(\mathbf{z}_h)$ is a tensor in which each component is a $m \times d$ matrix. As a result, $\mathbf{F}(\mathbf{z}_h) \mathbf{n}$ is a vector with the dimension m . By treating nonconforming interfaces in the fashion presented in Sec. 3.2.3, the normal flux $\mathbf{F}(\mathbf{z}_h) \mathbf{n}$ on $e \in \mathcal{E}_h$ is still not well-defined since the traces of both $\mathbf{z}_h^-$ of element K^- and $\mathbf{z}_h^+$ of element K^+ co-exist on e . However, it can be resolved by Godunov-type methods [59] through first solving, either exactly or approximately, the Riemann problem (3.3) for the upwind state $\mathbf{z}_h^*$ at the mortar e and then introducing the upwind numerical flux³ $\mathbf{F}^*(\mathbf{z}_h^-, \mathbf{z}_h^*) \mathbf{n}$. Furthermore, as reported in [20], such flux is hybridizable. The upwind-based HDG flux can then be constructed by replacing the upwind state with the designated trace unknown. Following this procedure, the upwind HDG flux reads:

$$(3.5) \quad \hat{\mathbf{F}}(\mathbf{z}_h, \hat{\mathbf{z}}_h) \mathbf{n} := A \mathbf{z}_h + |A|(\mathbf{z}_h - \hat{\mathbf{z}}_h).$$

Note that $|A|$ can also be replaced by some other stability parameter matrix T and that will result in different numerical fluxes (also see the discussion in [20]). By replacing $\mathbf{F}(\mathbf{z}_h) \mathbf{n}$ by $\hat{\mathbf{F}}(\mathbf{z}_h, \hat{\mathbf{z}}_h) \mathbf{n}$, we arrive at the so-called *local equations*: seek $(\mathbf{z}_h, \hat{\mathbf{z}}_h) \in \mathbf{W}_h \times \hat{\mathbf{W}}_h$ such that

$$(3.6) \quad -\sum_{k=1}^d (A_k \mathbf{z}_h, \partial_k \mathbf{w}_h)_K + (G \mathbf{z}_h, \mathbf{w}_h)_K + \langle A \mathbf{z}_h + |A|(\mathbf{z}_h - \hat{\mathbf{z}}_h), \mathbf{w}_h \rangle_{\partial K} = (\mathbf{f}, \mathbf{w}_h)_K, \quad \forall K \in \mathcal{T}_h,$$

for all $\mathbf{w}_h \in \mathbf{W}_h$. To close the system, we still require one more constraint. This can be achieved by weakly enforcing the continuity of the normal numerical flux (3.5) across the mortars: for $(\mathbf{z}_h, \hat{\mathbf{z}}_h) \in \mathbf{W}_h \times \hat{\mathbf{W}}_h$,

$$(3.7) \quad \langle [A \mathbf{z}_h + |A|(\mathbf{z}_h - \hat{\mathbf{z}}_h)], \hat{\mathbf{w}}_h \rangle_e = 0, \quad \forall e \in \mathcal{E}_h^\circ,$$

is enforced for $\forall \hat{\mathbf{w}}_h \in \hat{\mathbf{W}}_h$. Equation (3.7) is called as *conservativity condition* [32] since it will guarantee that the scheme is locally conservative. In addition, it couples all volume unknowns and hence is referred to as a *global equation*. On the boundary, it is clearer to weakly enforce non-homogeneous version of boundary conditions (2.3a) directly through the trace unknown which is already defined on $\mathcal{E}_h^\partial$:

$$(3.8) \quad \langle (A - M) \hat{\mathbf{z}}_h, \hat{\mathbf{w}}_h \rangle_e = \langle (A - M) \mathbf{g}, \hat{\mathbf{w}}_h \rangle_e, \quad \forall \hat{\mathbf{w}}_h \in \hat{\mathbf{W}}_h \text{ and } \forall e \in \mathcal{E}_h^\partial,$$

³Here, we consider the upwind numerical flux in the one-sided form $\mathbf{F}^*(\mathbf{z}_h^-, \mathbf{z}_h^*) \mathbf{n} = A \mathbf{z}_h^- + |A|(\mathbf{z}_h^- - \mathbf{z}_h^*)$. Typically, such a flux is defined as a function of states from adjacent elements in the DG setting. That is, $\mathbf{F}^* = \mathbf{F}^*(\mathbf{z}_h^-, \mathbf{z}_h^+)$.

where we set $M := |A|$ and the function $\mathbf{g} : \partial\Omega \rightarrow \mathbb{R}^m$ is defined as

$$(3.9) \quad \mathbf{g} := \begin{cases} \mathbf{g}_D & \text{if } (A - M) \neq 0, \\ \mathbf{0} & \text{if } (A - M) = 0, \end{cases}$$

where $\mathbf{g}_D$ is the Dirichlet data and is set to be zero for homogeneous boundary condition. It should be noted that equation (3.8) corresponds to the "inflow" boundary condition [20]. In addition, (3.8) is analogous to its continuous version stated in (2.3a). Since (3.8) only specify inflow condition, it is clear that $\widehat{\mathbf{z}}_h$ cannot be uniquely determined on the outflow. Thus, we further require that

$$(3.10) \quad \langle A\mathbf{z}_h + |A|(\mathbf{z}_h - \widehat{\mathbf{z}}_h), \widehat{\mathbf{w}}_h \rangle_e = \langle A\widehat{\mathbf{z}}_h, \widehat{\mathbf{w}}_h \rangle_e, \quad \forall \widehat{\mathbf{w}}_h \in \widehat{\mathbf{W}}_h \text{ and } \forall e \in \mathcal{E}_h^\partial.$$

Equation (3.10) is resulted from maintaining consistency of the numerical flux, and corresponds to outflow conditions. In fact, (3.8) and (3.10) can be incorporated into a single equation as:

$$(3.11) \quad \langle \llbracket A\mathbf{z}_h + |A|(\mathbf{z}_h - \widehat{\mathbf{z}}_h) \rrbracket, \widehat{\mathbf{w}}_h \rangle_{\mathcal{E}_h^\partial} = - \left\langle \frac{1}{2}(A - M)\mathbf{g}, \widehat{\mathbf{w}}_h \right\rangle_{\mathcal{E}_h^\partial} + \left\langle \frac{1}{2}(A + M)\widehat{\mathbf{z}}_h, \widehat{\mathbf{w}}_h \right\rangle_{\mathcal{E}_h^\partial}.$$

In this paper, we will work with the general form of boundary condition (3.11) for one-field Friedrichs' system. The complete hp -HDG formulation for the one-field Friedrichs' system is established by combining Eq. (3.6), Eq. (3.7), and Eq. (3.11) together: seek $(\mathbf{z}_h, \widehat{\mathbf{z}}_h) \in \mathbf{W}_h \times \widehat{\mathbf{W}}_h$ such that ⁴

$$(3.12a) \quad - \sum_{k=1}^d (A_k \mathbf{z}_h, \partial_k \mathbf{w}_h)_{\mathcal{T}_h} + (G\mathbf{z}_h, \mathbf{w}_h)_{\mathcal{T}_h} + \langle A\mathbf{z}_h + |A|(\mathbf{z}_h - \widehat{\mathbf{z}}_h), \mathbf{w}_h \rangle_{\partial\mathcal{T}_h} = (\mathbf{f}, \mathbf{w}_h)_{\mathcal{T}_h},$$

$$(3.12b) \quad \langle \llbracket A\mathbf{z}_h + |A|(\mathbf{z}_h - \widehat{\mathbf{z}}_h) \rrbracket, \widehat{\mathbf{w}}_h \rangle_{\mathcal{E}_h} = - \left\langle \frac{1}{2}(A - M)\mathbf{g}, \widehat{\mathbf{w}}_h \right\rangle_{\mathcal{E}_h^\partial} + \left\langle \frac{1}{2}(A + M)\widehat{\mathbf{z}}_h, \widehat{\mathbf{w}}_h \right\rangle_{\mathcal{E}_h^\partial},$$

for all $(\mathbf{w}_h, \widehat{\mathbf{w}}_h) \in \mathbf{W}_h \times \widehat{\mathbf{W}}_h$. We now show that the numerical scheme in (3.12) is trivially locally and globally conservative, and furthermore well-posed.

LEMMA 3.1 (Local conservation). *The hp -HDG scheme in (3.12) is locally conservative.*

Proof. Taking $\mathbf{w}_h = \mathbf{1}$ in the local equations (3.6), we obtain

$$(3.13) \quad (G\mathbf{z}_h, \mathbf{1})_K + \langle A\mathbf{z}_h + |A|(\mathbf{z}_h - \widehat{\mathbf{z}}_h), \mathbf{1} \rangle_{\partial K} = (\mathbf{f}, \mathbf{1})_K, \quad \forall K \in \mathcal{T}_h,$$

and thus

$$(3.14) \quad (G\mathbf{z}_h, \mathbf{1})_K + \sum_{F \subset \partial K} \langle A\mathbf{z}_h + |A|(\mathbf{z}_h - \widehat{\mathbf{z}}_h), \mathbf{1} \rangle_F = (\mathbf{f}, \mathbf{1})_K, \quad \forall K \in \mathcal{T}_h,$$

which indicates the scheme is locally conservative. In particular, the amount of flux entering an element K is equal to the amount of flux leaving the element if both the reaction term and forcing term vanish (i.e., $G = 0$ and $\mathbf{f} = \mathbf{0}$). $\square$

As we will show later, the locally conservative property can also be easily proven for Friedrichs' system with two-field structure. A similar result is presented in [46] as well for an hp -HDG method used to solve the problem of Stokes flow.

LEMMA 3.2 (Global conservation). *The hp -HDG scheme in (3.12) is globally conservative.*

Proof. Taking $(\mathbf{w}_h, \widehat{\mathbf{w}}_h) = (\mathbf{1}, \mathbf{1})$ in Eq. (3.12), we obtain

$$(3.15a) \quad (G\mathbf{z}_h, \mathbf{1})_{\mathcal{T}_h} + \langle \llbracket A\mathbf{z}_h + |A|(\mathbf{z}_h - \widehat{\mathbf{z}}_h) \rrbracket, \mathbf{1} \rangle_{\mathcal{E}_h} = (\mathbf{f}, \mathbf{1})_{\mathcal{T}_h},$$

$$(3.15b) \quad \langle \llbracket A\mathbf{z}_h + |A|(\mathbf{z}_h - \widehat{\mathbf{z}}_h) \rrbracket, \mathbf{1} \rangle_{\mathcal{E}_h} = - \left\langle \frac{1}{2}(A - M)\mathbf{g}, \mathbf{1} \right\rangle_{\mathcal{E}_h^\partial} + \left\langle \frac{1}{2}(A + M)\widehat{\mathbf{z}}_h, \mathbf{1} \right\rangle_{\mathcal{E}_h^\partial}.$$

⁴Although $\mathbf{g}$ is set to be zero, we still keep it in the right-hand side of the global equation so that reader can easily observe that inflow and outflow boundaries are switched in the adjoint hp -HDG formulation.

Substitute Eq. (3.15b) into Eq. (3.15a), we arrive at

$$(3.16) \quad (G\mathbf{z}_h, \mathbf{1})_{\mathcal{T}_h} - \left\langle \frac{1}{2}(A - M)\mathbf{g}, \mathbf{1} \right\rangle_{\varepsilon_h^\partial} + \left\langle \frac{1}{2}(A + M)\widehat{\mathbf{z}}_h, \mathbf{1} \right\rangle_{\varepsilon_h^\partial} = (\mathbf{f}, \mathbf{1})_{\mathcal{T}_h},$$

which implies the scheme is globally conservative. $\square$

LEMMA 3.3 (Well-posedness of the local equation). *Suppose that the assumptions (A.1)-(A.4) hold, the local solver (3.12a) is well-posed, that is, given $(\widehat{\mathbf{z}}_h, \mathbf{f})$, there exists a unique solution $\mathbf{z}_h$ of the local system.*

Proof. Thanks to the hp -nonconforming treatment, the proof is the same as the proof of [20, Lemma 6.1], and hence omitted. $\square$

THEOREM 3.4 (Well-posedness of the hp -HDG formulation). *Suppose that*

1. *the assumptions (A.1)-(A.4) and (2.3b) hold,*

2. $\mathcal{N}(A) = \{\mathbf{0}\}$ ⁵,

the hp -HDG formulation in (3.12) is well-posed in the sense that, given $\mathbf{f}$ and the homogeneous Dirichlet data, there exists a unique solution $(\mathbf{z}_h, \widehat{\mathbf{z}}_h)$.

Proof. Following the discussion presented in Theorem 6.2 in [20], we first take $(\mathbf{w}_h, \widehat{\mathbf{w}}_h) = (\mathbf{z}_h, \widehat{\mathbf{z}}_h)$ and assume $\mathbf{f} = \mathbf{0}$. We then perform integration by part for (3.12a), subtract (3.12b) from the resulting equations, and substitute the following identity

$$(3.17) \quad \langle A(\widehat{\mathbf{z}}_h - \mathbf{z}_h), \widehat{\mathbf{z}}_h \rangle_e = \frac{1}{2} [\langle A\widehat{\mathbf{z}}_h, \widehat{\mathbf{z}}_h \rangle_e + \langle A(\mathbf{z}_h - \widehat{\mathbf{z}}_h), (\mathbf{z}_h - \widehat{\mathbf{z}}_h) \rangle_e - \langle A\mathbf{z}_h, \mathbf{z}_h \rangle_e],$$

which is valid for $\forall e \in \mathcal{E}_h$. Together with homogeneous Dirichlet boundary condition, we obtain

$$(3.18) \quad \frac{1}{2} \left(\left[G + G^T + \sum_{k=1}^d \partial_k A_k \right] \mathbf{z}_h, \mathbf{z}_h \right)_{\mathcal{T}_h} + \frac{1}{2} \langle M\widehat{\mathbf{z}}_h, \widehat{\mathbf{z}}_h \rangle_{\varepsilon_h^\partial} + \left\langle \left(\frac{1}{2}A + |A| \right) (\mathbf{z}_h - \widehat{\mathbf{z}}_h), (\mathbf{z}_h - \widehat{\mathbf{z}}_h) \right\rangle_{\partial\mathcal{T}_h} = 0$$

whose left-hand side is non-negative owing to the coercivity condition (A.4), semi-positiveness of boundary operator (2.3b) and semi-positiveness of $\frac{1}{2}A + |A| \geq 0$. Therefore, we conclude that $\mathbf{z}_h = \mathbf{0}$ in K for $\forall K \in \mathcal{T}_h$. Furthermore, the last two terms in (3.18) is equal to zero since our assumption $\mathcal{N}(A) = \{\mathbf{0}\}$ implies that $\mathcal{N}(|A|) = \{\mathbf{0}\}$. We thus can conclude that $\widehat{\mathbf{z}}_h = \mathbf{0}$ as well. $\square$

3.3.2. Friedrichs' system with two-field structure. In this section, we derive hp -HDG formulation for two-field Friedrichs' system in (2.1) where the coefficient matrices G and A_k can be decomposed into block matrices as presented in Eq. (2.2). In addition, a set of assumptions (A.1)-(A.6) are assumed to hold. However, the strong coercivity is not necessarily required and can be further weakened by replacing (A.4) with (A4.a)-(A4.b).

Through the Galerkin approximation along with integration by part, we will again obtain Eq. (3.4). Moreover, the two-field structure can be further exploited by taking advantage of the decomposition where we can introduce $\mathbf{z}_h = (\boldsymbol{\sigma}_h, \mathbf{u}_h)$, $\mathbf{w}_h = (\mathbf{s}_h, \mathbf{v}_h)$, $\mathbf{F} = (\mathbf{F}^\sigma, \mathbf{F}^u)$, and $\mathbf{f} = (\mathbf{f}^\sigma, \mathbf{f}^u)$. Indeed, with the aid of Eq.(2.2), the local equation can be rewritten as: seek $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \boldsymbol{\Sigma}_h \times \mathbf{U}_h$ such that

$$(3.19a) \quad - \sum_{k=1}^d (B_k \mathbf{u}_h, \partial_k \mathbf{s}_h)_K + (G^{\sigma\sigma} \boldsymbol{\sigma}_h + G^{\sigma u} \mathbf{u}_h, \mathbf{s}_h)_K + \langle \mathbf{F}^\sigma(\boldsymbol{\sigma}_h, \mathbf{u}_h) \mathbf{n}, \mathbf{s}_h \rangle_{\partial K} = (\mathbf{f}^\sigma, \mathbf{s}_h)_K, \quad \forall K \in \mathcal{T}_h,$$

$$(3.19b) \quad - \sum_{k=1}^d (B_k^T \boldsymbol{\sigma}_h + C_k \mathbf{u}_h, \partial_k \mathbf{v}_h)_K + (G^{u\sigma} \boldsymbol{\sigma}_h + G^{uu} \mathbf{u}_h, \mathbf{v}_h)_K + \langle \mathbf{F}^u(\boldsymbol{\sigma}_h, \mathbf{u}_h) \mathbf{n}, \mathbf{v}_h \rangle_{\partial K} = (\mathbf{f}^u, \mathbf{v}_h)_K, \quad \forall K \in \mathcal{T}_h,$$

for all $(\mathbf{s}_h, \mathbf{v}_h) \in \boldsymbol{\Sigma}_h \times \mathbf{U}_h$. Now the upwind flux $\mathbf{F}^*(\boldsymbol{\sigma}_h^-, \mathbf{u}_h^-, \boldsymbol{\sigma}_h^+, \mathbf{u}_h^+) \mathbf{n}$ for the two-field system can also be derived by solving the Riemann problem stated in Eq. (3.3). To that end, it is required to compute the eigen-decomposition of the coefficient matrix A . The two-field structure can be exploited again by decomposing

⁵This condition actually implies the condition (2.3c) which is a key to make the exact solution unique (see also Remark 6.3 in [20]).

409 $|A|$ as follows:

$$410 \quad (3.20) \quad |A| = \begin{bmatrix} \mathcal{A}^{\sigma\sigma} & \mathcal{A}^{\sigma u} \\ \mathcal{A}^{u\sigma} & \mathcal{A}^{uu} \end{bmatrix},$$

411 where $\mathcal{A}^{\sigma\sigma}$, $\mathcal{A}^{\sigma u}$, $\mathcal{A}^{u\sigma}$, and $\mathcal{A}^{uu}$ are $m^\sigma \times m^\sigma$, $m^\sigma \times m^u$, $m^u \times m^\sigma$, and $m^u \times m^u$ sub-block matrices of $|A|$.
 412 In addition, the $m \times d$ matrix of upwind flux $\mathbf{F}^* \mathbf{n}$ can also be decomposed into $m^\sigma \times d$ and $m^u \times d$ block
 413 matrices $(\mathbf{F}^\sigma)^* \mathbf{n}$ and $(\mathbf{F}^u)^* \mathbf{n}$, respectively. By introducing the upwind states σ_h^* and $\mathbf{u}_h^*$, the numerical
 414 flux can be expressed in one-sided form:

$$415 \quad \mathbf{F}^* (\sigma_h^-, \mathbf{u}_h^-, \sigma_h^*, \mathbf{u}_h^*) \mathbf{n} = \begin{bmatrix} (\mathbf{F}^\sigma)^* (\sigma_h^-, \mathbf{u}_h^-, \sigma_h^*, \mathbf{u}_h^*) \mathbf{n} \\ (\mathbf{F}^u)^* (\sigma_h^-, \mathbf{u}_h^-, \sigma_h^*, \mathbf{u}_h^*) \mathbf{n} \end{bmatrix} = \sum_{k=1}^d n_k \begin{bmatrix} A_k^{\sigma\sigma} & B_k \\ B_k^T & C_k \end{bmatrix} \begin{bmatrix} \sigma_h^- \\ \mathbf{u}_h^- \end{bmatrix} + \begin{bmatrix} \mathcal{A}^{\sigma\sigma} & \mathcal{A}^{\sigma u} \\ \mathcal{A}^{u\sigma} & \mathcal{A}^{uu} \end{bmatrix} \begin{bmatrix} \sigma_h^- - \sigma_h^* \\ \mathbf{u}_h^- - \mathbf{u}_h^* \end{bmatrix},$$

416 where $A_k^{\sigma\sigma} = 0$ by the assumption (A.5). At this point, we could replace the upwind states $(\sigma_h^*, \mathbf{u}_h^*)$ with
 417 the trace unknowns $(\hat{\sigma}_h, \hat{\mathbf{u}}_h)$ and obtain an upwind-based HDG flux. However, one of the upwind states
 418 (and hence one of the trace unknowns) can be eliminated, and it is desirable since the system becomes even
 419 cheaper to solve. Such reduction can be achieved since $\mathbf{u}_h^*$ and σ_h^* are linearly dependent. To see this, we
 420 know that the numerical flux $\mathbf{F}^* \mathbf{n}$ must match the physical flux $\mathbf{F} \mathbf{n}$ evaluated at the upwind states. That
 421 is,⁶

$$422 \quad (3.22) \quad \mathbf{F}^* (\sigma_h^-, \mathbf{u}_h^-, \sigma_h^*, \mathbf{u}_h^*) \mathbf{n} = \mathbf{F} (\sigma_h^*, \mathbf{u}_h^*) \mathbf{n},$$

423 where

$$424 \quad (3.23) \quad \mathbf{F} (\sigma_h^*, \mathbf{u}_h^*) \mathbf{n} = \sum_{k=1}^d n_k \begin{bmatrix} A_k^{\sigma\sigma} & B_k \\ B_k^T & C_k \end{bmatrix} \begin{bmatrix} \sigma_h^* \\ \mathbf{u}_h^* \end{bmatrix} \text{ and } A_k^{\sigma\sigma} = 0 \text{ by the assumption (A.5).}$$

425 By substituting (3.23) into (3.21) we can remove either σ_h^* or $\mathbf{u}_h^*$. In this work, we eliminate σ_h^* . For the
 426 other elimination possibilities, one can consult with [98].

427 LEMMA 3.5 (The reduced upwind flux). *The upwind numerical flux can be expressed as a function of*
 428 *$\mathbf{u}_h^*$ only:*

$$429 \quad (3.24) \quad \mathbf{F}^* (\sigma_h, \mathbf{u}_h, \mathbf{u}_h^*) \mathbf{n} = \begin{bmatrix} B \mathbf{u}_h^* \\ B^T \sigma_h + C \mathbf{u}_h + T (\mathbf{u}_h - \mathbf{u}_h^*) \end{bmatrix}$$

430 where $B := \sum_{k=1}^d n_k B_k$ and $C := \sum_{k=1}^d n_k C_k$ are with $\mathbf{n} = (n_1, \dots, n_d)^T$ being unit outward vector of ∂K
 431 for $\forall K \in \mathcal{T}_h$. The value of T depends on either of the following assumptions:

432 (F.1)

433 (F1.a) *There exists an invertible matrix $\Phi \in \mathbb{R}^{m^u, m^u}$ such that $B \Phi \mathcal{A}^{u\sigma} = \mathcal{A}^{\sigma\sigma}$, and*

434 (F1.b) *there exists a matrix $\Psi \in \mathbb{R}^{m^u, m^u}$ such that $\mathcal{A}^{\sigma u} = B \Psi$, and*

435 (F1.c) $\bigcap_{k=1}^d \text{Range}(B_k) = \{\mathbf{0}\}$ and $\mathcal{N}(B_k) = \{\mathbf{0}\}$ for $\forall k = 1, \dots, d$.

436 (F.2) $\mathcal{A}^{\sigma u} = 0$ (note that $\mathcal{A}^{\sigma u} = (\mathcal{A}^{u\sigma})^T$ owing to the assumption (A.3)).

437 In particular, if assumption (F.1) holds then:

$$438 \quad (3.25) \quad T := -(\Phi^T B^T B \Phi)^{-1} \Phi^T B^T B (\Psi + I_{m^u}) + \mathcal{A}^{uu},$$

439 where I_{m^u} is the $m^u \times m^u$ identity matrix. On the other hand, if assumption (F.2) holds then,

$$440 \quad (3.26) \quad T := \mathcal{A}^{uu}.$$

441 *Proof.* It can be proved by simple algebraic manipulation along with the corresponding assumption. See
 442 Appendix B for detailed proof. $\square$

⁶ It should be noted that both of upwind states σ_h^* and $\mathbf{u}_h^*$ are the function of states from adjacent elements. That is,
 $\sigma_h^* = \sigma_h^*(\sigma_h^-, \mathbf{u}_h^-, \sigma_h^+, \mathbf{u}_h^+)$ and $\mathbf{u}_h^* = \mathbf{u}_h^*(\sigma_h^-, \mathbf{u}_h^-, \sigma_h^+, \mathbf{u}_h^+)$

We would like to mention that the assumption (F.1) and (F.2) are mutually exclusive from each other. For example, the derivation of HDG numerical flux for a convection-diffusion equation rely on (F.1) while an elliptic equation uses (F.2) (See Section 5 for more detail). According to Lemma 3.5, we can construct the upwind-based HDG flux by replacing $\mathbf{u}_h^*$ with $\hat{\mathbf{u}}_h$:

$$(3.27) \quad \hat{\mathbf{F}}(\boldsymbol{\sigma}_h, \mathbf{u}_h, \hat{\mathbf{u}}_h) \mathbf{n} := \begin{bmatrix} B\hat{\mathbf{u}}_h \\ B^T \boldsymbol{\sigma}_h + C\mathbf{u}_h + T(\mathbf{u}_h - \hat{\mathbf{u}}_h) \end{bmatrix},$$

where the definitions of B and C follow ones introduced in Lemma 3.5 (these notations will be used in the rest of the paper as well). In fact, the numerical flux (3.27) can represent the larger class of HDG family other than just upwind-based HDG. It is possible to obtain different HDG schemes by setting the stability matrix T to be different from (3.25) and (3.26). Such an exploration is also studied in [20].

Now the local equation of an hp -HDG scheme for the two-field Friedrichs system can be constructed by substituting the upwind-based HDG flux (3.27) back into (3.19): seek $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \hat{\mathbf{u}}_h) \in \boldsymbol{\Sigma}_h \times \mathbf{U}_h \times \hat{\mathbf{U}}_h$ such that

$$(3.28a) \quad -\sum_{k=1}^d (B_k \mathbf{u}_h, \partial_k \mathbf{s}_h)_K + (G^{\sigma\sigma} \boldsymbol{\sigma}_h + G^{\sigma u} \mathbf{u}_h, \mathbf{s}_h)_K + \langle B\hat{\mathbf{u}}_h, \mathbf{s}_h \rangle_{\partial K} = (\mathbf{f}^\sigma, \mathbf{s}_h)_K, \quad \forall K \in \mathcal{T}_h,$$

$$(3.28b) \quad -\sum_{k=1}^d (B_k^T \boldsymbol{\sigma}_h + C_k \mathbf{u}_h, \partial_k \mathbf{v}_h)_K + (G^{u\sigma} \boldsymbol{\sigma}_h + G^{uu} \mathbf{u}_h, \mathbf{v}_h)_K + \langle B^T \boldsymbol{\sigma}_h + C\mathbf{u}_h + T(\mathbf{u}_h - \hat{\mathbf{u}}_h), \mathbf{v}_h \rangle_{\partial K} = (\mathbf{f}^u, \mathbf{v}_h)_K, \quad \forall K \in \mathcal{T}_h,$$

for all $(\mathbf{s}_h, \mathbf{v}_h) \in \boldsymbol{\Sigma}_h \times \mathbf{U}_h$. Again, we close the system with a conservative condition. Since the first component in $\hat{\mathbf{F}}(\boldsymbol{\sigma}_h, \mathbf{u}_h, \hat{\mathbf{u}}_h) \mathbf{n}$ is already uniquely defined, we weakly enforce the continuity in the second component: for $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \hat{\mathbf{u}}_h) \in \boldsymbol{\Sigma}_h \times \mathbf{U}_h \times \hat{\mathbf{U}}_h$,

$$(3.29) \quad \langle [B^T \boldsymbol{\sigma}_h + C\mathbf{u}_h + T(\mathbf{u}_h - \hat{\mathbf{u}}_h)], \hat{\mathbf{v}}_h \rangle_e = \mathbf{0}, \quad \forall e \in \mathcal{E}_h^\circ, \quad \forall \hat{\mathbf{v}}_h \in \hat{\mathbf{U}}_h$$

Finally, the boundary conditions are specified in a similar way as in Eq. (3.8). The difference is that instead of taking $M := T$, we make use of the characteristic of the two-field structure and choose M in a special way

$$(3.30) \quad M := \begin{bmatrix} 0 & -\alpha B \\ \alpha B^T & M^{uu} \end{bmatrix},$$

where $M^{uu} : \partial\Omega \rightarrow \mathbb{R}^{m^u, m^u}$, $M^{uu} \geq 0$, and $\alpha \in \{-1, +1\}$. With this specific setting, the boundary can then be enforced through Eq. (3.8) with the boundary data $\mathbf{g} : \partial\Omega \rightarrow \mathbb{R}^{m^\sigma \times m^u}$, $\mathbf{g} = (\mathbf{g}^\sigma, \mathbf{g}^u)$ in which $\mathbf{g}^\sigma : \partial\Omega \rightarrow \mathbb{R}^{m^\sigma}$ and $\mathbf{g}^u : \partial\Omega \rightarrow \mathbb{R}^{m^u}$. Again, for clarity and for the numerical results, we use nonhomogeneous boundary conditions, but in the well-posedness analysis it is sufficient to consider homogeneous boundary conditions. We further set $M^{uu} = 2\rho I_{m^u} + C$, where I_{m^u} is the $m^u \times m^u$ identity matrix and ρ is chosen on the case-by-case basis (see Section 5). Again, the boundary operator M is not unique but must satisfy assumptions (2.3b) and (2.3c). In particular, (2.3b) requires:

$$(3.31) \quad 2\rho I_{m^u} + C \geq 0.$$

For Dirichlet type of boundary we set $\alpha = 1$ and $\rho = \frac{1}{2}$, the boundary condition (3.8) now is restated as⁷:

$$(3.32) \quad \langle \hat{\mathbf{u}}_h, \hat{\mathbf{v}}_h \rangle_e = 0, \quad \forall \hat{\mathbf{v}}_h \in \hat{\mathbf{U}}_h, \quad \forall e \in \mathcal{E}_h^\partial \cap \partial\Omega_D.$$

⁷In fact, with $\alpha = 1$ and $\rho = \frac{1}{2}$ we obtain two equations from (3.8):

$$\langle B\hat{\mathbf{u}}_h, \hat{\mathbf{s}}_h \rangle_e = 0, \quad \forall \hat{\mathbf{s}}_h \in \hat{\boldsymbol{\Sigma}}_h, \quad \forall e \in \mathcal{E}_h^\partial \cap \partial\Omega_D, \quad \text{and} \quad \langle \hat{\mathbf{u}}_h, \hat{\mathbf{v}}_h \rangle_e = 0, \quad \forall \hat{\mathbf{v}}_h \in \hat{\mathbf{U}}_h, \quad \forall e \in \mathcal{E}_h^\partial \cap \partial\Omega_D.$$

These two equations are re equivalent and we use the latter formulation in this paper since:

1. Given that only the trace unknown $\hat{\mathbf{u}}_h$ is introduced, no test function in $\hat{\boldsymbol{\Sigma}}_h$ should be involved, and
2. The latter option is more economical.

On the other hand, for Neumann or Robin type of boundary condition, we set $\alpha = -1$ and the variable ϱ depends on problems to be solved. In this case, (3.8) now becomes:

$$(3.33) \quad \left\langle B^T \sigma_h + C u_h + T(u_h - \hat{u}_h), \hat{v}_h \right\rangle_e = \langle (\varrho I_m u + C) \hat{u}_h, \hat{v}_h \rangle_e, \quad \forall \hat{v}_h \in \hat{U}_h, \forall e \in \mathcal{E}_h^\partial \cap (\partial\Omega_N \cup \partial\Omega_R).$$

Again, the stabilization parameter T is set to be (3.25) if the assumption (F.1) holds or to be (3.26) if the assumption (F.2) holds. Finally, the consistency condition like (3.10) is not needed here since the equation (3.8) itself along with the set-up (3.30) is sufficient to determine the trace unknown $\hat{u}_h$ on the boundary. Combining (3.28), (3.29), (3.32), and (3.33), we can obtain the complete the hp -HDG formulation for the two-field Friedrichs' system: seek $(\sigma_h, u_h, \hat{u}_h) \in \Sigma_h \times U_h \times \hat{U}_h$ such that

$$(3.34a) \quad -\sum_{k=1}^d (B_k u_h, \partial_k s_h)_{\mathcal{T}_h} + (G^{\sigma\sigma} \sigma_h + G^{\sigma u} u_h, s_h)_{\mathcal{T}_h} + \langle B \hat{u}_h, s_h \rangle_{\partial\mathcal{T}_h} = (f^\sigma, s_h)_{\mathcal{T}_h},$$

$$(3.34b) \quad -\sum_{k=1}^d \left(B_k^T \sigma_h + C_k u_h, \partial_k v_h \right)_{\mathcal{T}_h} + (G^{u\sigma} \sigma_h + G^{uu} u_h, v_h)_{\mathcal{T}_h} + \left\langle B^T \sigma_h + C u_h + T(u_h - \hat{u}_h), v_h \right\rangle_{\partial\mathcal{T}_h} = (f^u, v_h)_{\mathcal{T}_h},$$

$$(3.34c) \quad \left\langle \llbracket B^T \sigma_h + C u_h + T(u_h - \hat{u}_h) \rrbracket, \hat{v}_h \right\rangle_{\mathcal{E}_h \setminus \partial\Omega_D} = \langle (\varrho I_m u + C) \hat{u}_h, \hat{v}_h \rangle_{\mathcal{E}_h \cap (\partial\Omega_N \cup \partial\Omega_R)},$$

$$(3.34d) \quad \langle \hat{u}_h, \hat{v}_h \rangle_{\mathcal{E}_h^\partial \cap \partial\Omega_D} = 0,$$

for all $(s_h, v_h, \hat{v}_h) \in \Sigma_h \times U_h \times \hat{U}_h$. We now show that the numerical scheme in (3.34) is both locally and globally conservative, and well-posed. For the well-posedness proof, both full and partial coercivity will be discussed. It turns out that a few extra assumptions are needed for the well-posedness and they are different for full and partial coercivity cases.

LEMMA 3.6 (Local conservation). *The hp -HDG scheme in (3.34) is both locally and globally conservative.*

Proof. The proofs are the same as the proof of Lemma 3.1 and Lemma 3.2, and hence omitted. $\square$

LEMMA 3.7 (Well-posedness of the local equation-with full coercivity). *Suppose*

1. *the assumptions (A.1)-(A.6) hold, and*
2. *$\frac{1}{2}C + T \geq 0$, and*
3. *B_k is a constant and is non-zero for $k = 1, \dots, d$.*

Then, the local solver composed by (3.34a) and (3.34b) is well-posed, that is, given $(\hat{u}_h, f^\sigma, f^u)$, there exists a unique solution (σ_h, u_h) of the local solver.

Proof. Since the formulation is linear and $z_h = (\sigma_h, u_h)$ is in finite dimensional space W_h , it is sufficient to restrict to a single element K and show that the solution $(\sigma_h, u_h) = 0$ is a unique solution in K for $K \in \mathcal{T}_h$ provided that $\hat{u}_h$ and f are zero. Let $\hat{u}_h$ and f be zero and (s_h, v_h) be (σ_h, u_h) in (3.34a) and (3.34b). Adding the equations yields

$$(3.35) \quad -\sum_{k=1}^d (B_k u_h, \partial_k \sigma_h)_K + (G^{\sigma\sigma} \sigma_h + G^{\sigma u} u_h, \sigma_h)_K - \sum_{k=1}^d \left(B_k^T \sigma_h + C_k u_h, \partial_k u_h \right)_K + (G^{u\sigma} \sigma_h + G^{uu} u_h, u_h)_K + \left\langle B^T \sigma_h + C u_h + T u_h, u_h \right\rangle_{\partial K} = 0.$$

By invoking the assumption that B_k is a constant for $k = 1, \dots, d$, the term $\frac{1}{2} \sum_{k=1}^d ((\partial_k B_k^T) \sigma_h, u_h)_K$ contribute nothing and can be freely added into (3.35). It gives

$$(3.36) \quad -\sum_{k=1}^d (B_k u_h, \partial_k \sigma_h)_K + (G^{\sigma\sigma} \sigma_h + G^{\sigma u} u_h, \sigma_h)_K - \sum_{k=1}^d \left(B_k^T \sigma_h + C_k u_h, \partial_k u_h \right)_K + \frac{1}{2} \sum_{k=1}^d \left((\partial_k B_k^T) \sigma_h, u_h \right)_K + (G^{u\sigma} \sigma_h + G^{uu} u_h, u_h)_K + \left\langle B^T \sigma_h + C u_h + T u_h, u_h \right\rangle_{\partial K} = 0.$$

509 The first term in (3.36) can be further expanded as:

$$510 \quad (3.37) \quad - \sum_{k=1}^d (B_k \mathbf{u}_h, \partial_k \boldsymbol{\sigma}_h)_K = \sum_{k=1}^d (B_k \partial_k \mathbf{u}_h, \boldsymbol{\sigma}_h)_K + \frac{1}{2} \sum_{k=1}^d ((\partial_k B_k) \mathbf{u}_h, \boldsymbol{\sigma}_h)_K - \langle B \mathbf{u}_h, \boldsymbol{\sigma}_h \rangle_{\partial K}.$$

511 The second term on the right-hand side of (3.37) is zero owing to our assumption and hence can be multiplied
512 by an arbitrary constant. Similarly, it is easy to show that the following identity holds:

$$513 \quad (3.38) \quad - \sum_{k=1}^d (C_k \mathbf{u}_h, \partial_k \mathbf{u}_h)_K = \frac{1}{2} \sum_{k=1}^d ((\partial_k C_k) \mathbf{u}_h, \mathbf{u}_h)_K - \frac{1}{2} \langle C \mathbf{u}_h, \mathbf{u}_h \rangle_{\partial K}.$$

514 Substituting (3.37) and (3.38) back into (3.36), and combining (undo the decomposition) the volume integrals,
515 we arrive at

$$516 \quad (3.39) \quad \frac{1}{2} \left(\left(G + G^T + \sum_{k=1}^d \partial_k A_k \right) \mathbf{z}_h, \mathbf{z}_h \right)_K + \left\langle \left(\frac{1}{2} C + T \right) \mathbf{u}_h, \mathbf{u}_h \right\rangle_{\partial K} = 0.$$

517 With the assumption of full-coercivity (A.4) and the assumption of semi-positiveness $\frac{1}{2}C + T \geq 0$, we can
518 conclude that $\mathbf{z}_h = (\boldsymbol{\sigma}_h, \mathbf{u}_h) = \mathbf{0}$ in K for any $K \in \mathcal{T}_h$. $\square$

519 **THEOREM 3.8** (Well-posedness of the hp -HDG formulation-with full coercivity). *Suppose:*

- 520 1. the assumptions (A.1)-(A.6) and (2.3b) hold, and
- 521 2. $\frac{1}{2}C + T \geq 0$, and
- 522 3. $\bar{B}_k$ is constant and is nonzero for $k = 1, \dots, d$, and
- 523 4. $\bigcap_{k=1}^d \text{Range}(B_k) = \{\mathbf{0}\}$ and $\mathcal{N}(B_k) = \{\mathbf{0}\}$ for $\forall k = 1, \dots, d$.

524 Then, the hp -HDG formulation in (3.34) is well-posed, that is, given $\mathbf{f}^\sigma$, $\mathbf{f}^u$, and the homogeneous Dirichlet
525 data, there exists a unique solution $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \hat{\mathbf{u}}_h)$.

526 *Proof.* Due to the finite-dimensional nature and the linearity of the global system, it is sufficient to show
527 that the solution $\hat{\mathbf{u}}_h = 0$ is the unique solution if $\mathbf{f} = 0$ along with homogeneous boundary data $\mathbf{g} = \mathbf{0}$. We
528 first let $\mathbf{f} = \mathbf{0}$ and $(s_h, \mathbf{v}_h, \hat{\mathbf{v}}_h) = (\boldsymbol{\sigma}_h, \mathbf{u}_h, \hat{\mathbf{u}}_h)$. The boundary condition (3.34d) now reads

$$529 \quad (3.40) \quad \langle \hat{\mathbf{u}}_h, \hat{\mathbf{u}}_h \rangle_{\mathcal{E}_h^\partial \cap \partial \Omega_D} = 0,$$

530 which implies that $\hat{\mathbf{u}}_h = \mathbf{0}$ at e for $\forall e \in \mathcal{E}_h^\partial \cap \partial \Omega_D$. Adding (3.34a) and (3.34b) together, and then subtracting
531 (3.34c) from the resulting equation, we obtain

$$\begin{aligned} & - \sum_{k=1}^d (B_k \mathbf{u}_h, \partial_k \boldsymbol{\sigma}_h)_{\mathcal{T}_h} + (G^{\sigma\sigma} \boldsymbol{\sigma}_h + G^{\sigma u} \mathbf{u}_h, \boldsymbol{\sigma}_h)_{\mathcal{T}_h} + \langle B \hat{\mathbf{u}}_h, \boldsymbol{\sigma}_h \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega_D} \\ & - \sum_{k=1}^d \left(B_k^T \boldsymbol{\sigma}_h + C_k \mathbf{u}_h, \partial_k \mathbf{u}_h \right)_{\mathcal{T}_h} + (G^{u\sigma} \boldsymbol{\sigma}_h + G^{uu} \mathbf{u}_h, \mathbf{u}_h)_{\mathcal{T}_h} \\ 532 \quad (3.41) \quad & + \left\langle B^T \boldsymbol{\sigma}_h + C \mathbf{u}_h + T(\mathbf{u}_h - \hat{\mathbf{u}}_h), \mathbf{u}_h \right\rangle_{\partial \mathcal{T}_h \setminus \partial \Omega_D} \\ & - \left\langle B^T \boldsymbol{\sigma}_h + C \mathbf{u}_h + T(\mathbf{u}_h - \hat{\mathbf{u}}_h), \hat{\mathbf{u}}_h \right\rangle_{\partial \mathcal{T}_h \setminus \partial \Omega_D} \\ & + \langle (\varrho I_{m^u} + C) \hat{\mathbf{u}}_h, \hat{\mathbf{u}}_h \rangle_{\mathcal{E}_h \cap (\partial \Omega_N \cup \partial \Omega_R), I} = 0, \end{aligned}$$

533 where the result of $\hat{\mathbf{u}}_h = \mathbf{0}$ at e for $\forall e \in \mathcal{E}_h^\partial \cap \partial \Omega_D$ is already applied. We can add the additional term

534 $\frac{1}{2} \sum_{k=1}^d ((\partial_k B_k^T) \sigma_h, \mathbf{u}_h)_{\mathcal{T}_h}$ as B_k is constant for $k = 1, \dots, d$ to obtain

$$\begin{aligned}
 & - \sum_{k=1}^d (B_k \mathbf{u}_h, \partial_k \sigma_h)_{\mathcal{T}_h} + (G^{\sigma\sigma} \sigma_h + G^{\sigma\mathbf{u}} \mathbf{u}_h, \sigma_h)_{\mathcal{T}_h} + \langle B \hat{\mathbf{u}}_h, \sigma_h \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega_D} \\
 & - \sum_{k=1}^d (B_k^T \sigma_h + C_k \mathbf{u}_h, \partial_k \mathbf{u}_h)_{\mathcal{T}_h} + (G^{\mathbf{u}\sigma} \sigma_h + G^{\mathbf{u}\mathbf{u}} \mathbf{u}_h, \mathbf{u}_h)_{\mathcal{T}_h} + \frac{1}{2} \sum_{k=1}^d ((\partial_k B_k^T) \sigma_h, \mathbf{u}_h)_{\mathcal{T}_h} \\
 & + \langle B^T \sigma_h + C \mathbf{u}_h + T(\mathbf{u}_h - \hat{\mathbf{u}}_h), \mathbf{u}_h \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega_D} \\
 & - \langle B^T \sigma_h + C \mathbf{u}_h + T(\mathbf{u}_h - \hat{\mathbf{u}}_h), \hat{\mathbf{u}}_h \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega_D} \\
 & + \langle (\varrho I_{m^u} + C) \hat{\mathbf{u}}_h, \hat{\mathbf{u}}_h \rangle_{\mathcal{E}_h \cap (\partial \Omega_N \cup \partial \Omega_R), I} = 0.
 \end{aligned}
 \tag{3.42}$$

536 We have the following identity by inspection:

$$\begin{aligned}
 & - \langle C \mathbf{u}_h, \hat{\mathbf{u}}_h \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega_D} = \frac{1}{2} \langle C(\mathbf{u}_h - \hat{\mathbf{u}}_h), (\mathbf{u}_h - \hat{\mathbf{u}}_h) \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega_D} \\
 & - \frac{1}{2} \langle C \mathbf{u}_h, \mathbf{u}_h \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega_D} - \frac{1}{2} \langle C \hat{\mathbf{u}}_h, \hat{\mathbf{u}}_h \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega_D}.
 \end{aligned}
 \tag{3.43}$$

538 Note that $\langle C \hat{\mathbf{u}}_h, \hat{\mathbf{u}}_h \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} = 0$ since C is assumed to be continuous across the mesh skeleton, and the trace
 539 unknown $\hat{\mathbf{u}}_h$ is uniquely defined on the mortar e for all $e \in \mathcal{E}_h^\circ$. As a consequence, the last term in (3.43)
 540 can be rewritten as $-\frac{1}{2} \langle C \hat{\mathbf{u}}_h, \hat{\mathbf{u}}_h \rangle_{\partial \mathcal{T}_h \cap (\partial \Omega_N \cup \partial \Omega_R), I}$. Substituting equality (3.37), (3.38), and (3.43) back into
 541 (3.42), and combining (undo the decomposition) the volume integrals, we arrive at the following

$$\begin{aligned}
 & \frac{1}{2} \left\langle \left(G + G^T + \sum_{k=1}^d \partial_k A_k \right) \mathbf{z}_h, \mathbf{z}_h \right\rangle_{\mathcal{T}_h} + \left\langle \left(\frac{1}{2} C + T \right) (\mathbf{u}_h - \hat{\mathbf{u}}_h), (\mathbf{u}_h - \hat{\mathbf{u}}_h) \right\rangle_{\partial \mathcal{T}_h \setminus \partial \Omega_D} \\
 & + \left\langle \left(\frac{1}{2} C + T \right) \mathbf{u}_h, \mathbf{u}_h \right\rangle_{\partial \mathcal{T}_h \cap \partial \Omega_D} + \left\langle \left(\frac{1}{2} C + \varrho I_{m^u} \right) \hat{\mathbf{u}}_h, \hat{\mathbf{u}}_h \right\rangle_{\partial \mathcal{T}_h \cap (\partial \Omega_N \cup \partial \Omega_R)} = 0.
 \end{aligned}
 \tag{3.44}$$

543 With full-coercivity (A.4), semi-positiveness of the boundary operator (2.3b) (hence inequality (3.31)), and
 544 semi-positiveness $\frac{1}{2} C + T \geq 0$ assumptions, we can conclude that $\mathbf{z}_h = (\sigma_h, \mathbf{u}_h) = \mathbf{0}$ in K for all $K \in \mathcal{T}_h$.
 545 Now substituting $(\sigma_h, \mathbf{u}_h) = \mathbf{0}$ back to the sub-equation (3.34a) in the local solver along with $\mathbf{f} = \mathbf{0}$ and
 546 $\hat{\mathbf{u}}_h = \mathbf{0}$ at e for $\forall e \in \mathcal{E}_h^\partial \cap \partial \Omega_D$, we get:

$$\langle B \hat{\mathbf{u}}_h, \mathbf{s}_h \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega_D} = 0 \quad \forall \mathbf{s}_h \in \Sigma_h,
 \tag{3.45}$$

548 which implies that $B \hat{\mathbf{u}}_h = 0$. By invoking our assumption that $\bigcap_{k=1}^d \text{Range}(B_k) = \{\mathbf{0}\}$ and $\mathcal{N}(B_k) = \{\mathbf{0}\}$
 549 for $\forall k = 1, \dots, d$, the condition $\mathcal{N}(B) = \{\mathbf{0}\}$ can be concluded. We then conclude $\hat{\mathbf{u}}_h = 0$ in e for all
 550 $e \in \mathcal{E}_h \setminus \partial \Omega_D$. $\square$

551 **LEMMA 3.9** (Well-posedness of the local equation-with partial coercivity). *Assume:*

- 552 1. the assumption (A.1)-(A.3), (A4.a)-(A4.b) and (A.5)-(A.6) hold, and
- 553 2. $\frac{1}{2} C + T > 0$, and
- 554 3. $\bigcap_{k=1}^d \text{Range}(B_k) = \{\mathbf{0}\}$ and $\mathcal{N}(B_k) = \{\mathbf{0}\}$ for $\forall k = 1, \dots, d$.

555 Then, the local solver composed by (3.34a) and (3.34b) is well-posed, that is, given $(\hat{\mathbf{u}}_h, \mathbf{f}^\sigma, \mathbf{f}^u)$, there exists
 556 a unique solution $(\sigma_h, \mathbf{u}_h)$ of the local solver.

557 *Proof.* Essentially, the hp -HDG formulation for the two-field Friedrichs' system with partial coercivity is
 558 the same as the one with full coercivity. Hence, we can obtain the equation (3.39) as well following the same
 559 arguments discussed in the proof of Lemma 3.7. By applying the assumption of partial coercivity (A4.a) and
 560 of positiveness of $\frac{1}{2} C + T > 0$, we can conclude that $\sigma_h = \mathbf{0}$ on K for any $K \in \mathcal{T}_h$ and $\mathbf{u}_h = \mathbf{0}$ on $F \subset \partial K$
 561 for all $K \in \mathcal{T}_h$. By applying integration by part to the first term in (3.34a), and substituting the result that
 562 we just obtained into it along with $\hat{\mathbf{u}}_h = \mathbf{0}$ and $\mathbf{f} = \mathbf{0}$, we have

$$\sum_{k=1}^d (\partial_k (B_k \mathbf{u}_h), \mathbf{s}_h)_K = 0 \quad \forall \mathbf{s}_h \in \Sigma_h,
 \tag{3.46}$$

which implies that $\sum_{k=1}^d \partial_k(B_k \mathbf{u}_h) = \mathbf{0}$ in K . Furthermore, it can be rewritten as $\sum_{k=1}^d B_k \partial_k(\mathbf{u}_h) = \mathbf{0}$ in K owing to assumption (A4.b). Based on our assumption that $\bigcap_{k=1}^d \text{Range}(B_k) = \{\mathbf{0}\}$ and $\mathcal{N}(B_k) = \{\mathbf{0}\}$ for $\forall k = 1, \dots, d$, we can further conclude that $\mathbf{u}_h = \mathbf{0}$ on K for any $K \in \mathcal{T}_h$. $\square$

THEOREM 3.10 (Well-posedness of the hp -HDG formulation -with partial coercivity). *Suppose:*

1. the assumptions (A.1)-(A.3), (A4.a)-(A4.b), (A.5)-(A.6), and (2.3b) hold, and
2. $\frac{1}{2}C + T > 0$,

3. $\bigcap_{k=1}^d \text{Range}(B_k) = \{\mathbf{0}\}$ and $\mathcal{N}(B_k) = \{\mathbf{0}\}$ for $\forall k = 1, \dots, d$.

Then, the hp -HDG formulation stated in (3.34) is well-posed in the sense that given $\mathbf{f}^\sigma$, $\mathbf{f}^u$, and the homogeneous Dirichlet data, there exists a unique solution $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \hat{\mathbf{u}}_h)$.

Proof. Given that the hp -HDG formulation for the two-field Friedrichs' system with partial coercivity is the same as the one with full coercivity, we can directly follow the same arguments used in the proof of Theorem 3.8 and it should lead us to equation (3.44) as well. With the stated assumptions we can conclude that $\boldsymbol{\sigma}_h = \mathbf{0}$ in K for all $K \in \mathcal{T}_h$, and $\mathbf{u}_h = \hat{\mathbf{u}}_h$ in $F \cap e$ for all $F \subset \partial K$ for $\forall \partial K \in \partial \mathcal{T}_h \setminus \partial \Omega_D$ and for all $e \in \mathcal{E}_h \setminus \partial \Omega_D$. Now we perform integration by part to the first term in (3.34a), transfer all integration over ∂K to the summation of the integration over e where $e \in \mathcal{E}_h$, and apply the conclusion we just obtained into the resultant equation along with $\mathbf{f} = \mathbf{0}$, we get:

$$(3.47) \quad \sum_{k=1}^d (\partial_k(B_k \mathbf{u}_h), \mathbf{s}_h)_{\mathcal{T}_h} = 0 \quad \forall \mathbf{s}_h \in \boldsymbol{\Sigma}_h,$$

which implies that $\sum_{k=1}^d \partial_k(B_k \mathbf{u}_h) = \mathbf{0}$ in K for all $K \in \mathcal{T}_h$. Furthermore, it can be rewritten as $\sum_{k=1}^d B_k \partial_k(\mathbf{u}_h) = \mathbf{0}$ owing to assumption (A4.b). Given that we assume $\bigcap_{k=1}^d \text{Range}(B_k) = \{\mathbf{0}\}$ and $\mathcal{N}(B_k) = \{\mathbf{0}\}$ for $\forall k = 1, \dots, d$, we can conclude that $\mathbf{u}_h = \mathbf{0}$ in K for all $K \in \mathcal{T}_h$ but $\mathbf{u}_h = \hat{\mathbf{u}}_h$ in $F \cap e$ for all $F \subset \partial \mathcal{T}_h \setminus \partial \Omega_D$ for all $e \in \mathcal{E}_h \setminus \partial \Omega_D$. This leads to $\hat{\mathbf{u}}_h = \mathbf{0}$ in $\mathcal{E}_h \setminus \partial \Omega_D$. $\square$

4. Strategy for hp -adaptation. The formulations stated in (3.12) and (3.34) provide us with HDG schemes that can be carried out on hp -nonconforming meshes. As a result, we have a lot of flexibility when constructing finite element spaces. It is well-known that a smooth solution can be well resolved using a high degree of approximation even on a coarse mesh, whereas a solution with a sharp gradient is more suitable for low degree approximations on a fine mesh. Given that these different behaviors may occur locally, it is beneficial to use an adaptation procedure that allows us to improve the numerical results with a reasonable computational cost. This process can be achieved by refining elements locally via either dividing them into smaller ones (h -adaptation), or enriching their approximation spaces (p -adaptation). To that end, two essential ingredients are needed: an error indication for each element and a method to define a new spatial discretization [66]. For the first ingredient, two different approaches are adopted in this work. One is to use an error indicator while the other is to use an adjoint-based error estimate. For the second ingredient, the regularity indicator proposed in [43] is applied. In the following discussion, we will discuss the error indicator obtained by two different approaches and then outline the algorithm for hp -adaptation.

4.1. Doleji's approach. By denoting $\mathbf{q}_h$ as an approximate solution, h_F as a length of a face of an element and $\mathbf{g}_D$ as Dirichlet data, a local error estimator is defined as the following [43]:

$$(4.1) \quad \mathcal{E}_{h,K}^{\text{Doleji}}(\mathbf{q}_h) := \left(\sum_{F \subset \partial K \setminus \partial \Omega} \sum_{e \subset F} \frac{1}{h_F} \langle [\![\mathbf{q}_h]\!] \rangle_e + \sum_{F \subset \partial K \cap \partial \Omega_D} \frac{1}{h_F} \langle \mathbf{q}_h - \mathbf{g}_D, \mathbf{q}_h - \mathbf{g}_D \rangle_F \right)^{\frac{1}{2}} \quad \forall K \in \mathcal{T}_h,$$

which, originally, is derived in the context of the interior-penalty DG methods. We simply use it as a local error indicator to probe errors in our work. It is inexpensive since only the computation of the jump between adjacent elements is needed. In particular, $\mathbf{q}_h := \mathbf{z}_h$ and $\mathbf{q}_h := \mathbf{u}_h$ are picked for the one and two-field Friedrichs system, respectively.

4.2. Adjoint approach. The main idea of the adjoint approach is to measure the error in the output functional of interest. The error arises when the output functional is evaluated by a numerical solution. Based on the pioneering work [47], the dual-weight-residual (DWR) method have been developed for error control and mesh optimization within the context of finite element methods [13]. In this method, an additional

linear system formed by an adjoint equation is needed to be solved, which then induces an estimate of the error in the target functional. This estimate can be used as a criterion to drive adaptation so that the error in the target functional is reduced. Recently, this method is adopted in solving the elliptic equations by using HDG method in [38] along with h -adaptivity, and the evaluation of error estimate is completed with the aid of the postprocessing technique [36].

In our work, we develop a discrete weak adjoint approach where the procedure outlined in the DWR method is still largely followed but the primal problem considered here is already in a discretized weak form (i.e., Eq. (3.12) or (3.34)). For HDG methods, the discrete adjoint approach had been studied and implemented in [8, 40, 107, 106, 9, 52, 82]. To proceed with the discussion, some additional notations are needed. Let $\mathcal{J}(\cdot)$ be a (Gâteaux or Fréchet) differentiable output functional and, for a clearer exposition, we further decompose it into two differential functionals $\mathcal{J}(\cdot) = \mathcal{J}^{adjoint}(\cdot) + \mathcal{J}^{boundary}(\cdot)$ where $\mathcal{J}^{adjoint}(\cdot)$ is a user-defined functional and $\mathcal{J}^{boundary}(\cdot)$ is a boundary-associated functional. Examples of a user-defined functional could be a drag coefficient, a lift coefficient, an energy across the entire domain, and so on. On the other hand, the boundary-associated functional is also defined by a user and closely related to the boundary conditions of the adjoint problem (such as adjoint hp -HDG formulations in this paper). A more detailed discussion about the boundary-associated functional is addressed in Appendix A. We shall use subscript H to denote the approximation computed at a coarse discretization while h is for a finer level. We then define the operator $\mathbf{I}_H^h$ as the injection from level H to level h and this operation can be done by interpolation. In addition, the interpolated quantity is denoted with a subscript H along with a superscript h . For example $\mathbf{z}_H^h = \mathbf{I}_H^h \mathbf{z}_H$ is obtained by interpolating the approximate solution $\mathbf{z}_H$ that is solved at the coarser level (i.e., a lower degree of approximation or a coarser mesh or the combination.) to the finer level. Moreover, we define the lumped variables $\mathcal{Z}_h$ and $\mathcal{W}_h$ as

$$(4.2a) \quad \mathcal{Z}_h := \begin{cases} (\mathbf{z}_h, \widehat{\mathbf{z}}_h) & \text{for one-field Friedrichs' system,} \\ (\boldsymbol{\sigma}_h, \mathbf{u}_h, \widehat{\mathbf{u}}_h) & \text{for two-field Friedrichs' system,} \end{cases}$$

$$(4.2b) \quad \mathcal{W}_h := \begin{cases} (\mathbf{w}_h, \widehat{\mathbf{w}}_h) & \text{for one-field Friedrichs' system,} \\ (\mathbf{s}_h, \mathbf{v}_h, \widehat{\mathbf{v}}_h) & \text{for two-field Friedrichs' system,} \end{cases}$$

where $(\mathbf{z}_h, \widehat{\mathbf{z}}_h), (\mathbf{w}_h, \widehat{\mathbf{w}}_h) \in \mathbf{W}_h \times \widehat{\mathbf{W}}_h$ and $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \widehat{\mathbf{u}}_h), (\mathbf{s}_h, \mathbf{v}_h, \widehat{\mathbf{v}}_h) \in \boldsymbol{\Sigma}_h \times \mathbf{U}_h \times \widehat{\mathbf{U}}_h$. Furthermore, we introduce the bilinear form $\mathcal{R}_h^{\text{one}}(\cdot, \cdot)$ to denote the residual of the hp -HDG formulation for the one-field Friedrichs' system (3.12). It is the sum of other bilinear forms $\mathcal{R}_{h,K}^{\mathbf{z}}(\cdot, \cdot)$ and $\mathcal{R}_{h,e}^{\widehat{\mathbf{z}}}(\cdot, \cdot)$:

$$(4.3) \quad \mathcal{R}_h^{\text{one}}(\mathcal{Z}_h, \mathcal{W}_h) := \sum_{K \in \mathcal{T}_h} \mathcal{R}_{h,K}^{\mathbf{z}}(\mathcal{Z}_h, \mathcal{W}_h) + \sum_{e \in \mathcal{E}_h} \mathcal{R}_{h,e}^{\widehat{\mathbf{z}}}(\mathcal{Z}_h, \mathcal{W}_h),$$

where

$$(4.4a) \quad \mathcal{R}_{h,K}^{\mathbf{z}}(\mathcal{Z}_h, \mathcal{W}_h) := - \sum_{k=1}^d (A_k \mathbf{z}_h, \partial_k \mathbf{w}_h)_K + (G \mathbf{z}_h, \mathbf{w}_h)_K + \langle A \mathbf{z}_h + T(\mathbf{z}_h - \widehat{\mathbf{z}}_h), \mathbf{w}_h \rangle_{\partial K} - (\mathbf{f}, \mathbf{w}_h)_K,$$

$$(4.4b) \quad \mathcal{R}_{h,e}^{\widehat{\mathbf{z}}}(\mathcal{Z}_h, \mathcal{W}_h) := \langle \llbracket A \mathbf{z}_h + T(\mathbf{z}_h - \widehat{\mathbf{z}}_h) \rrbracket, \widehat{\mathbf{w}}_h \rangle_e + \left\langle \frac{1}{2} (A - M) \mathbf{g}, \widehat{\mathbf{w}}_h \right\rangle_{e \cap \mathcal{E}_h^\partial} - \left\langle \frac{1}{2} (A + M) \widehat{\mathbf{z}}_h, \widehat{\mathbf{w}}_h \right\rangle_{e \cap \mathcal{E}_h^\partial}.$$

On the other hand, the residual $\mathcal{R}_h^{\text{two}}(\cdot, \cdot)$ of the hp -HDG formulation for the two-field Friedrichs' system is the bilinear form composed by three bilinear forms $\mathcal{R}_{h,K}^{\boldsymbol{\sigma}}(\cdot, \cdot)$, $\mathcal{R}_{h,K}^{\mathbf{u}}(\cdot, \cdot)$, and $\mathcal{R}_{h,e}^{\widehat{\mathbf{u}}}(\cdot, \cdot)$:

$$(4.5) \quad \mathcal{R}_h^{\text{two}}(\mathcal{Z}_h, \mathcal{W}_h) := \sum_{K \in \mathcal{T}_h} \mathcal{R}_{h,K}^{\boldsymbol{\sigma}}(\mathcal{Z}_h, \mathcal{W}_h) + \sum_{K \in \mathcal{T}_h} \mathcal{R}_{h,K}^{\mathbf{u}}(\mathcal{Z}_h, \mathcal{W}_h) + \sum_{e \in \mathcal{E}_h} \mathcal{R}_{h,e}^{\widehat{\mathbf{u}}}(\mathcal{Z}_h, \mathcal{W}_h),$$

645 where

$$646 \quad (4.6a) \quad \mathcal{R}_{h,K}^{\sigma}(\mathcal{Z}_h, \mathcal{W}_h) := - \sum_{k=1}^d (B_k \mathbf{u}_h, \partial_k \mathbf{s}_h)_K + (G^{\sigma\sigma} \boldsymbol{\sigma}_h + G^{\sigma\mathbf{u}} \mathbf{u}_h, \mathbf{s}_h)_K + \langle B \hat{\mathbf{u}}_h, \mathbf{s}_h \rangle_{\partial K} - (\mathbf{f}^{\sigma}, \mathbf{s}_h)_K,$$

$$647 \quad (4.6b) \quad \begin{aligned} \mathcal{R}_{h,K}^{\mathbf{u}}(\mathcal{Z}_h, \mathcal{W}_h) &:= - \sum_{k=1}^d \left(B_k^T \boldsymbol{\sigma}_h + C_k \mathbf{u}_h, \partial_k \mathbf{v}_h \right)_K + (G^{\mathbf{u}\sigma} \boldsymbol{\sigma}_h + G^{\mathbf{u}\mathbf{u}} \mathbf{u}_h, \mathbf{v}_h)_K \\ &\quad + \left\langle B^T \boldsymbol{\sigma}_h + C \mathbf{u}_h + T(\mathbf{u}_h - \hat{\mathbf{u}}_h), \mathbf{v}_h \right\rangle_{\partial K} - (\mathbf{f}^{\mathbf{u}}, \mathbf{v}_h)_K, \\ 648 \quad (4.6c) \quad \mathcal{R}_{h,e}^{\hat{\mathbf{u}}_e}(\mathcal{Z}_h, \mathcal{W}_h) &:= \left\langle \llbracket B^T \boldsymbol{\sigma}_h + C \mathbf{u}_h + T(\mathbf{u}_h - \hat{\mathbf{u}}_h) \rrbracket, \hat{\mathbf{v}}_h \right\rangle_{e \setminus \partial \Omega_D} - \langle (\varrho I_m^{\mathbf{u}} + C) \hat{\mathbf{u}}_h, \hat{\mathbf{v}}_h \rangle_{e \cap (\partial \Omega_N \cup \partial \Omega_R)} \\ 649 \quad &\quad + \langle C \hat{\mathbf{u}}_h, \hat{\mathbf{v}}_h \rangle_{e \cap \partial \Omega_D}. \end{aligned}$$

650 At this point, we can further define a more general residual based on (4.3) and (4.5) as:

$$651 \quad (4.7) \quad \mathcal{R}_h(\mathcal{Z}_h, \mathcal{W}_h) := \begin{cases} \mathcal{R}_h^{\text{one}}(\mathcal{Z}_h, \mathcal{W}_h) & \text{if only identified as a one-field Friedrichs' system,} \\ \mathcal{R}_h^{\text{two}}(\mathcal{Z}_h, \mathcal{W}_h) & \text{if identified as a two-field Friedrichs' system.} \end{cases}$$

652 Obviously, the residual is always zero if it is evaluated by using the correct solution while it is generally
653 non-zero when using the interpolated solution. That is, $\mathcal{R}_h(\mathcal{Z}_h, \mathcal{W}_h) = 0$ but in general $\mathcal{R}_h(\mathcal{Z}_H^h, \mathcal{W}_h) \neq 0$.
654 Finally, the error of the output functional $\mathcal{J}(\cdot)$ can now be approximated as [40, 9, 105]:

$$655 \quad (4.8) \quad \mathcal{J}(\mathcal{Z}_H) - \mathcal{J}(\mathcal{Z}_h) \approx -\mathcal{R}_h(\mathcal{Z}_H^h, \mathcal{W}_h).$$

656 Here, $\mathcal{W}_h$ is also referred to as an adjoint variable and serves as a detection of the sensitivity of output
657 functional error induced by a less accurate solution. Further, it has to satisfy the adjoint hp -HDG formulation
658 that is either (a.1) or (a.4) with the given right-hand sides depending on an output functional $\mathcal{J}(\cdot)$ and on
659 an interpolated solution $\mathcal{Z}_H^h$. The derivation of the adjoint hp -HDG formulation and well-posedness analysis
660 are discussed in Appendix A.

661 From (4.8), it can be seen that two different approximation spaces (at the level h and at the level H)
662 are required. In this work, we construct the finer space by enriching the degree of approximation without
663 refining the mesh. That is, the meshes used in solving the primal and adjoint hp -HDG formulations are
664 the same (i.e., $\mathcal{T}_h = \mathcal{T}_H$) but the finite element spaces on each element for the primal and adjoint hp -
665 HDG formulations differ by one degree. The benefits are twofold: reasonable computational cost and easy
666 implementation. Toward the adaptation, we need to localize the error approximation presented in (4.8). By
667 defining the localized residual $\mathcal{R}_{h,K}$ as:

$$668 \quad (4.9) \quad \mathcal{R}_{h,K}(\mathcal{Z}_h, \mathcal{W}_h) := \begin{cases} \mathcal{R}_{h,K}^z(\mathcal{Z}_h, \mathcal{W}_h) & \text{for one-field Friedrichs' system,} \\ \mathcal{R}_{h,K}^{\sigma}(\mathcal{Z}_h, \mathcal{W}_h) + \mathcal{R}_{h,K}^{\mathbf{u}}(\mathcal{Z}_h, \mathcal{W}_h) & \text{for two-field Friedrichs' system,} \end{cases}$$

669 and following the works in [40, 105, 9], the local error indicator based on the adjoint approach can be defined
670 as:

$$671 \quad (4.10) \quad \mathcal{E}_{H,K}^{\text{adjoint}}(\mathcal{Z}_H^h, \mathcal{W}_h) := |\mathcal{R}_{h,K}(\mathcal{Z}_H^h, \mathcal{W}_h)|.$$

672 It should be emphasized that the error indicator (4.10) does not include the contribution from the trace
673 unknowns (i.e., $\mathcal{R}_{h,e}^{\hat{\mathbf{u}}_e}$ and $\mathcal{R}_{h,e}^{\hat{\mathbf{u}}_e}$ are neglected) owing to its insignificant influence [40, 105].

674 **REMARK 3.** We would also like to point out that the output error stated in (4.8) can directly be computed
675 by evaluating the difference between $\mathcal{J}(\mathcal{Z}_H)$ and $\mathcal{J}(\mathcal{Z}_h)$, where we have to solve the hp -HDG formulation
676 (3.12) or (3.34) at two different levels of approximation. However, in this work, we stick to the approximation
677 given by the DWR method (i.e., evaluation of the right-hand side of (4.8)). This method is more general in
678 the sense that the adjoint problem is always linear and is the only problem that needs to be solved at the fine
679 level of approximation. It holds true regardless of whether the primal problem is linear or not.

4.3. An adaptation algorithm. The algorithm used in this work is the simplified version of the strategy proposed in [43], which provides all necessary keys for carrying out hp -adaption. Combining the previous discussion on error indicators, we denote a general local and global error indicator as:

$$(4.11a) \quad \mathcal{E}_{H,K} := \mathcal{E}_{H,K}^{\text{Doleji}}(\mathbf{q}_H) \text{ or } \mathcal{E}_{H,K}^{\text{adjoint}}(\mathcal{Z}_H^h, \mathcal{W}_h),$$

$$(4.11b) \quad \mathcal{E}_H := \left(\sum_{K \in \mathcal{T}_h} \mathcal{E}_{H,K}^2 \right)^{\frac{1}{2}}.$$

To drive full hp -adapation, a method to decide how to construct a new spatial discretization is also necessary. In this work, it is desirable that the spatial discretization can be constructed according to the smoothness of the solution. To this end, a local regularity indicator is needed and the one proposed in [43] is deployed in this work. By denoting $|K|$ as the area of an element, the indicator reads:

$$(4.12) \quad \mathcal{G}_K(\mathbf{q}_h) := \frac{\sum_{F \subset \partial K} \sum_{e \subseteq F \setminus \partial \Omega} \langle \llbracket \mathbf{q}_h \rrbracket, \llbracket \mathbf{q}_h \rrbracket \rangle_e}{|K| h_K^{2p_K-3}},$$

where $\mathbf{q}_h := \mathbf{z}_h$ and $\mathbf{q}_h := \mathbf{u}_h$ are one-field and two-field Friedrichs' system, respectively. Once error and regularity indicators are computed, one or a couple of the following operations will be performed:

- h -refinement⁸: to split a given *mother element* K into four *child elements* K' by connecting centers of its edges.
- p -refinement: to increase the degree of polynomial approximation for a given element K , i.e., we set $p_K = p_K + 1$.
- p -coarsening: to decrease the degree of polynomial approximation for a given element K , i.e., we set $p_K = p_K - 1$.

In the original strategy presented in [43], there are two additional operations called h -coarsening and hp -substitution. They merge elements that have arisen in a previous adaptation cycle along with p -refinement, p -coarsening, or nothing. However, according to our numerical experiments, this action only slightly increased efficiency, and sometimes the performance seems to be degrading. For this reason, we remove these operations from our adaptation strategy. Given the user-defined tolerance $0 \leq \omega \leq 1$ and the maximum cycle number, the hp -adaption procedure can now be performed by following the strategy outlined in Algorithm 4.1.

5. Numerical experiments. In this section, we are going to present several numerical experiments for different kinds of PDEs. The numerical solution is obtained by solving hp -HDG formulations (3.12) or (3.34) along with the adaptivity strategy discussed in Section 4. The main goal is to demonstrate the validity of the unified hp formulations and examine the performance of our proposed approaches. We point out that the output functional employed in the adjoint approach is given as

$$(5.1) \quad \mathcal{J}(\mathbf{q}_h) := \left[\sum_{K \in \mathcal{T}_h} \left(\mathcal{E}_{h,K}^{\text{Doleji}}(\mathbf{q}_h) \right)^2 \right]^{\frac{1}{2}}, \quad \text{where } \mathbf{q}_h := \begin{cases} \mathbf{z}_h, & \text{for one-field Friedrichs' system} \\ \mathbf{u}_h, & \text{for two-field Friedrichs' system} \end{cases},$$

so that we can fairly compare the computational performance of Doleji's approach and the adjoint approach, as the same quantity is minimized through the hp -adaptation process. From now on let us denote by $\delta \mathcal{J}_{\mathbf{q}_h}(\cdot; \delta \mathbf{q}_h)$ the directional derivative of a functional $\mathcal{J}(\cdot)$ with respect to some variable $\mathbf{q}_h$ in the direction $\delta \mathbf{q}_h$. As a result, the directional derivative of the output functional given in (5.1) reads:

$$(5.2) \quad \delta \mathcal{J}_{\mathbf{q}_h}(\mathbf{q}_h; \delta \mathbf{q}_h) = \mathcal{J}(\mathbf{q}_h)^{-1} \sum_{K \in \mathcal{T}_h} \left(\sum_{F \subset \partial K \setminus \partial \Omega_D} \sum_{e \subset F} \frac{1}{h_F} \langle \llbracket \mathbf{q}_h \rrbracket, \delta \mathbf{q}_h \rangle_e + \sum_{F \subset \partial K \cap \partial \Omega_D} \frac{1}{h_F} \langle \mathbf{q}_h - \mathbf{g}_D, \delta \mathbf{q}_h \rangle_e \right),$$

⁸We also enforce the number of hanging nodes resulting from local h -refinement to be always one in each interface within a mesh.

Algorithm 4.1 An hp -adaptation algorithm

```

1:  $\mathcal{E}_{H,K} \leftarrow 1 \quad \forall K \in \mathcal{T}_H$  and  $\mathcal{E}_H \leftarrow 0$  ▷ The initialization for starting the adaptation cycle(s).
2: while  $\max_{K \in \mathcal{T}_H} \mathcal{E}_{H,K} \geq \omega \mathcal{E}_H$  or cycle number  $\leq$  max. cycle number do
3:   Solve the  $hp$ -HDG formulation stated in (3.12) or (3.34) on a coarse (current) level.
4:   (Solve the adjoint  $hp$ -HDG formulation stated in (a.1) or (a.4))
5:   on a fine (by enriching degree of approximation) level if the adjoint approach is applied).
6:   Compute and update local and global error indicator presented in (4.11).
7:   for  $K \in \mathcal{T}_H$  do
8:     if  $\mathcal{E}_{H,K} \geq \omega \max_{K \in \mathcal{T}_H} \mathcal{E}_{H,K}$  then
9:       if  $\mathcal{G}_K(\mathbf{q}_H) \leq h_K^{-2}$  then
10:        Tag the element as  $p$ -refinement
11:       else if  $h_K^{-2} < \mathcal{G}_K(\mathbf{q}_H) \leq h_K^{-4}$  then
12:        Tag the element as  $h$ -refinement
13:       else
14:        Tag the element as  $h$ -refinement along with  $p$ -coarsening
15:       end if
16:     end if
17:   end for
18:   Perform adaption and construct the new corresponding finite element space
19: end while

```

717 where

$$718 \quad \delta \mathbf{q}_h := \begin{cases} \delta \mathbf{z}_h \in \mathbf{W}_h, & \text{for one-field Friedrichs' system} \\ \delta \mathbf{u}_h \in \mathbf{U}_h, & \text{for two-field Friedrichs' system} \end{cases}.$$

719 The directional derivative $\delta \mathcal{J}_{\mathbf{q}_h}(\mathbf{q}_h; \delta \mathbf{q}_h)$ will appear in the right-hand side of the adjoint hp -HDG formulation
720 (see (a.1) and (a.4)), and hence needs to be computed when solving the adjoint system. Instead of exactly
721 computing $\delta \mathcal{J}_{\mathbf{q}_h}(\mathbf{q}_h; \delta \mathbf{q}_h)$, an approximation $\delta \mathcal{J}_{\mathbf{q}_h}(\mathbf{I}_H^h \mathbf{q}_H; \delta \mathbf{q}_h)$ is applied. We would like to point out that
722 the boundary-associated functional $\mathcal{J}^{boundary}(\cdot)$ is set to zero in this paper. That is, we have homogeneous
723 boundary conditions for the adjoint hp -HDG formulation. To measure the computational performance, we
724 plot the convergence rate of the error in the $\mathcal{L}^2$ -norm versus the number of degrees of freedoms (DOFs)
725 resulting from the statically condensed hp -HDG formulations. It should be noted that the required DOFs
726 for the adjoint approach include DOFs needed by hp -HDG formulations and DOFs needed by the adjoint hp -
727 HDG formulation since we additionally have to solve for the adjoint solution to evaluate the error indicator
728 (4.10).

729 The PDEs under consideration in the experiments can be classified as elliptic, hyperbolic, and mixed
730 equations. In the following subsections, we will briefly discuss each type of PDEs and justify the well-
731 posedness of their hp -HDG formulation by using the results in Section 3. We use subscript h to denote the
732 numerical solution and this should not be confused with the notations used in Section 4 where h and H refer
733 to different refinement levels.

734 **5.1. Elliptic PDEs.** For elliptic PDEs, we consider:

- 735 (E.1) Poisson's problem (isotropic diffusion) with a corner singularity,
736 (E.2) anisotropic diffusion problem with discontinuous Dirichlet boundary condition, and
737 (E.3) heterogeneous anisotropic diffusion problem with discontinuous field $\hat{\mathbf{k}}$.

738 We analyze these problems by using the two-field Friedrichs' system with partial coercivity. The problem
739 reads: find a function $u : \Omega \rightarrow \mathbb{R}$ such that:

$$\begin{aligned}
& -\nabla \cdot (\hat{\mathbf{k}} \nabla u) = f, \quad \text{in } \Omega, \\
& u = g^u, \quad \text{on } \partial\Omega_D, \\
& \hat{\mathbf{k}} \nabla u \cdot \mathbf{n} + \lambda u = g^u, \quad \text{on } \partial\Omega_N \cup \partial\Omega_R \text{ where } \lambda = 0 \text{ when on the } \partial\Omega_N,
\end{aligned}$$

740 (5.3)

741 where the boundary data $g^u : \partial\Omega \mapsto \mathbb{R}$ is in $\mathcal{L}^2(\partial\Omega)$ and is defined as

$$742 \quad (5.4) \quad g^u = \begin{cases} g_D & \text{on } \partial\Omega_D, \\ g_N & \text{on } \partial\Omega_N, \\ g_R & \text{on } \partial\Omega_R. \end{cases}$$

743 Here, $f \in \mathcal{L}^2(\Omega)$ is a source term; and $\hat{\kappa} \in [\mathcal{L}^\infty(\Omega)]^{d,d}$ is a symmetric positive-definite diffusivity coefficient
 744 with its lowest eigenvalue uniformly bounded away from zero. To be able to interpret the numerical result
 745 later, we briefly review some physical aspects of the PDE stated in (5.3). At each location within Ω , we
 746 have the principal direction of anisotropy denoted by X and the direction of weak diffusion denoted by Y .
 747 As shown in Figure 3 along with coordinate of physical domain (x_1, x_2) , it is possible to align x_1 to X by
 748 rotating the system with the angle θ_m so that the equation (5.3) becomes:

$$749 \quad \kappa_X \frac{\partial^2 u}{\partial X^2} + \kappa_Y \frac{\partial^2 u}{\partial Y^2} = f \quad \text{in } \Omega,$$

750 where κ_X and κ_Y are referred to the diffusivity in X -direction and in Y -direction, respectively. Since κ_X
 751 represents the principal direction of anisotropy, we always have $\kappa_X \geq \kappa_Y$. At this point, we can define the
 752 anisotropy ratio $\mathcal{A}_\kappa := \kappa_X / \kappa_Y$ which indicates the strength of the anisotropy. The case $\mathcal{A}_\kappa = 1$ corresponds
 753 to isotropic diffusion (i.e., a Laplace's or Poisson's equation). Now the diffusivity coefficient $\hat{\kappa}$ can be
 754 expressed as:

$$755 \quad \hat{\kappa} = \begin{bmatrix} \kappa_X \cos^2(\theta_m) + \kappa_Y \sin^2(\theta_m) & (\kappa_X - \kappa_Y) \sin(\theta_m) \cos(\theta_m) \\ (\kappa_X - \kappa_Y) \sin(\theta_m) \cos(\theta_m) & \kappa_X \sin^2(\theta_m) + \kappa_Y \cos^2(\theta_m) \end{bmatrix},$$

756 or

$$757 \quad \hat{\kappa} = \begin{bmatrix} \mathcal{A}_\kappa \cos^2(\theta_m) + \sin^2(\theta_m) & (\mathcal{A}_\kappa - 1) \sin(\theta_m) \cos(\theta_m) \\ (\mathcal{A}_\kappa - 1) \sin(\theta_m) \cos(\theta_m) & \mathcal{A}_\kappa \sin^2(\theta_m) + \cos^2(\theta_m) \end{bmatrix}.$$

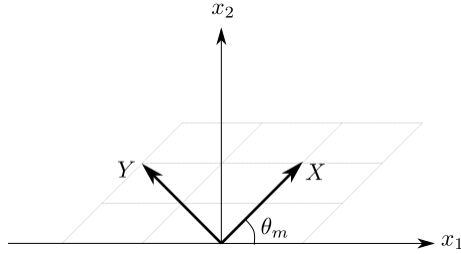


Fig. 3: The skewed domain of anisotropic field with X parallel to the anisotropic principal direction.

758 To cast the problems into the Friedrichs' framework, we rewrite the original PDE stated in (5.3) into
 759 the first order form by introducing the *auxiliary variable* $\sigma := -\hat{\kappa} \nabla u$:

$$760 \quad (5.5a) \quad \nabla u + \hat{\kappa}^{-1} \sigma = 0, \quad \text{in } \Omega,$$

$$761 \quad (5.5b) \quad \nabla \cdot \sigma = f, \quad \text{in } \Omega,$$

$$762 \quad (5.5c) \quad \begin{aligned} u &= g^u, & \text{on } \partial\Omega_D, \\ -\sigma \cdot \mathbf{n} + \lambda u &= g^u, & \text{on } \partial\Omega_N \cup \partial\Omega_R \text{ where } \lambda = 0 \text{ when on the } \partial\Omega_N. \end{aligned}$$

764 Thus, the size of the system is given as: $m = d + 1$, $m^\sigma = d$, and $m^u = 1$. The corresponding two-field
 765 Friedrichs' system reads:

$$766 \quad (5.6) \quad G = \begin{bmatrix} \hat{\kappa}^{-1} & 0_{d \times 1} \\ 0_{1 \times d} & 0 \end{bmatrix}, \quad A_k = \begin{bmatrix} 0_{d \times d} & \mathbf{e}_k \\ \mathbf{e}_k^T & 0 \end{bmatrix},$$

where $\mathbf{e}_k$ stands for the k -th canonical basis in $\mathbb{R}^{m^\sigma}$ and 0 with subscript indicates the zero matrix with its dimension specified by the subscripts. To enforce boundary conditions properly, the boundary operator M defined in (3.30) can be specified as:

$$(5.7) \quad \begin{cases} \alpha = +1, M^{\mathbf{u}\mathbf{u}} = 2\rho I_{m^\mathbf{u}}, & \text{on } \partial\Omega_D \text{ where } \rho := \frac{1}{2}, \\ \alpha = -1, M^{\mathbf{u}\mathbf{u}} = 2\rho I_{m^\mathbf{u}}, & \text{on } \partial\Omega_N \cup \partial\Omega_R \text{ where } \rho := \lambda, \end{cases}$$

where we require $\lambda = 0$ on $\partial\Omega_N$ and $\lambda > 0$ on $\partial\Omega_R$ in order for the conditions in (2.3b) and (2.3c) to hold.

LEMMA 5.1. *The hp-HDG formulation for the PDE stated in (5.5) is well-posed both locally and globally.*

Proof. The assumptions (A.1)-(A.3), (A4.a)-(A4.b) and (A.5)-(A.6) is obviously satisfied by substituting (5.6) into each conditions.

On the other hand, the numerical flux falls into the category (F.2) where we have:

$$(5.8) \quad |A| = \begin{bmatrix} \frac{\mathbf{n}\mathbf{n}^T}{\|\mathbf{n}\|_2} & 0_{d \times 1} \\ 0_{1 \times d} & \|\mathbf{n}\|_2^2 \end{bmatrix}$$

in which $\|\cdot\|_2$ is a standard Euclidean norm. Thus, the stabilization parameter reads $T = \|\mathbf{n}\|_2^2 = 1$. It is evident that:

- $\frac{1}{2}C + T = T = 1 > 0$, and
- $\bigcap_{k=1}^d \text{Range}(\mathbf{e}_k) = \{\mathbf{0}\}$ and $\mathcal{N}(\mathbf{e}_k) = \{\mathbf{0}\}$ for $\forall k = 1, \dots, d$.

Hence, by Lemma 3.9 and Theorem 3.10 we can conclude that the hp-HDG formulation for the elliptic PDE (5.5) is well-posed locally and globally. $\square$

5.2. Hyperbolic PDE. We consider the following hyperbolic PDE:

(HP.1) steady-state linear advection with variable speed and discontinuous inflow condition.

The PDE for steady-state linear advection reads: find a function $u : \Omega \rightarrow \mathbb{R}$ such that:

$$(5.9) \quad \begin{aligned} \nabla \cdot (\boldsymbol{\beta}u) &= f, & \text{in } \Omega, \\ u &= g_D, & \text{on } \partial\Omega^- \end{aligned}$$

with $\boldsymbol{\beta} \in [\mathcal{L}^\infty(\Omega)]^d$, $\nabla \cdot \boldsymbol{\beta} \in \mathcal{L}^\infty(\Omega)$, $f \in \mathcal{L}^2(\Omega)$ and $g_D \in \mathcal{L}^2(\partial\Omega^-)$. Here, we adopt the convention $\partial\Omega^- = \{x \in \partial\Omega : \boldsymbol{\beta} \cdot \mathbf{n} < 0\}$ to denote the inflow boundaries, and they are essentially Dirichlet boundaries in this problem set. It is well-known that singularity (or discontinuity) can be propagated by linear advection. Hence, we can expect that there is a shock within the domain Ω if a discontinuity is specified at the inflow boundary $\partial\Omega^-$. The problem can be analyzed by the one-field Friedrichs' system. The size of the system is $m = 1$ and the corresponding system reads:

$$(5.10) \quad G = 0, \quad A_k = \beta_k \text{ for } k = 1, \dots, d.$$

It is clear that that assumptions (A.1)-(A.3) are valid. To have coercivity (A.4), we further assume that

$$(5.11) \quad \text{ess inf}_\Omega \frac{1}{2} \nabla \cdot \boldsymbol{\beta} \geq 0.$$

Finally, we also require the following conditions to obtain a well-posed HDG scheme:

$$(5.12) \quad \boldsymbol{\beta} \cdot \mathbf{n} \neq 0 \text{ on } e, \quad \forall e \in \mathcal{E}_h, \quad \llbracket \boldsymbol{\beta} \cdot \mathbf{n} \rrbracket = 0 \text{ on } \partial K, \quad \forall K \in \mathcal{T}_h,$$

where we assume $\boldsymbol{\beta} \cdot \mathbf{n}$ is always continuous across element edges and does not vanish at edges (or mortars). Note that the condition for continuity can be relaxed and the resulting numerical flux has the weight-average type of stabilization parameter [20].

LEMMA 5.2. *The hp-HDG formulation for the PDE stated in (5.9) is well-posed both locally and globally if the assumptions (5.11) and (5.12) hold.*

Proof. Consider the following transformation:

$$(5.13) \quad u = \chi \tilde{u}, \quad \text{where } \chi := e^{-\gamma(\mathbf{x} - \mathbf{x}_0) \cdot \boldsymbol{\beta}},$$

in which $\gamma \in \mathbb{R}$, $\mathbf{x}_0 \in \Omega$, and $\tilde{u} : \Omega \rightarrow \mathbb{R}$. Substituting (5.13) back into (5.9) gives us:

$$(5.14) \quad \begin{aligned} \nabla \cdot (\tilde{\boldsymbol{\beta}} \tilde{u}) &= f, \quad \text{in } \Omega, \\ \tilde{u} &= \tilde{g}_D, \quad \text{on } \partial\Omega^-, \end{aligned}$$

where $\tilde{\boldsymbol{\beta}} = \chi \boldsymbol{\beta}$ and $\tilde{g}_D = \chi^{-1} g_D$. Note that $\chi > 0$ and hence its inverse always exists. The PDE (5.14) can also be identified as a one-field Friedrichs' system where:

$$(5.15) \quad G = 0, \quad A_k = \tilde{\boldsymbol{\beta}}_k.$$

It is obvious that (A.1)-(A.3) are valid and

$$(5.16) \quad G + G^T + \sum_{k=1}^d \partial_k A_k = \sum_{k=1}^d \partial_k \tilde{\boldsymbol{\beta}}_k = \chi \nabla \cdot \boldsymbol{\beta} - \|\boldsymbol{\beta}\|_2^2 \gamma \chi > 0,$$

where the last inequality will hold by the assumption (5.11) and by taking $\gamma < 0$. Therefore, condition (A.4) is also satisfied. Finally, $\mathcal{N}(A) = \mathcal{N}(\tilde{\boldsymbol{\beta}} \cdot \mathbf{n}) = \{0\}$ along all surfaces of the elements since the continuity of $\boldsymbol{\beta} \cdot \mathbf{n}$ is assumed and the mapping χ is diffeomorphism. With the aid of Lemma 3.3 and Theorem 3.4, we can conclude that the hp -HDG formulation for (5.14) is well-posed both locally and globally. Given that the mapping χ is bijective, this conclusion is also valid for (5.9). $\square$

REMARK 4. *It is possible to extend the Friedrichs' framework discussed in this paper to time-dependent problems. One way to achieve this is to treat one of the spatial variables as the time. For example, the model (5.9) is readily to be rewritten as one-dimensional unsteady linear advection by changing x_1 as t and specifying $\boldsymbol{\beta}$ as $(1, a)$ where the scalar a is advection velocity. However, this way may only be straightforward for linear scalar problems. In particular, it is difficult for PDEs with vector states (i.e., the first-order form of the unsteady heat equation).*

The more general and easier extension is to employ Rothe's method [95]. By applying Rothe's method, the time derivative term becomes a reaction term and the rest of the terms can still easily be written in the general form outlined in Eq. (2.1). In addition, due to the positiveness of the time variable, the newly introduced reaction term induced by the time derivative term would not pose a negative effect on the (partial or full) coercivity condition ((A.4) or (A4.a)).

5.3. Mixed PDE. For mixed PDE, we consider:

(HB.1) steady-state convection-diffusion problem with discontinuous inflow condition,

$$(5.17) \quad \begin{aligned} \nabla \cdot (\boldsymbol{\beta} u - \hat{\boldsymbol{\kappa}} \nabla u) &= f, \quad \text{in } \Omega, \\ \begin{cases} u = g^u, & \text{on } \partial\Omega^+, \\ (\boldsymbol{\beta} u - \hat{\boldsymbol{\kappa}} \nabla u) \cdot \mathbf{n} = g^u, & \text{on } \partial\Omega^- \cup \partial\Omega_0, \end{cases} \end{aligned}$$

where the boundary data $g^u : \partial\Omega \mapsto \mathbb{R}$ is in $\mathcal{L}^2(\partial\Omega)$ and is defined as

$$(5.18) \quad g^u = \begin{cases} g_D & \text{on } \partial\Omega^+, \\ g_{N,R} & \text{on } \partial\Omega_0 \cup \partial\Omega^-. \end{cases}$$

Here, $\boldsymbol{\beta} \in [\mathcal{L}^\infty(\Omega)]^d$, $\nabla \cdot \boldsymbol{\beta} \in \mathcal{L}^\infty(\Omega)$, $f \in \mathcal{L}^2(\Omega)$, and $\hat{\boldsymbol{\kappa}}$ is a symmetric positive definite matrix-valued defined on Ω with lowest eigenvalue uniformly bounded away from zero. In addition, $\partial\Omega^- \cup \partial\Omega^+ \cup \partial\Omega_0 = \partial\Omega$ where $\partial\Omega^- = \{x \in \partial\Omega : \boldsymbol{\beta} \cdot \mathbf{n} < 0\}$ is an inflow boundary; $\partial\Omega^+ = \{x \in \partial\Omega : \boldsymbol{\beta} \cdot \mathbf{n} > 0\}$ is an outflow boundary; and $\partial\Omega_0 = \{x \in \partial\Omega : \boldsymbol{\beta} \cdot \mathbf{n} = 0\}$ is a zero-flow boundary. It is evident that $\partial\Omega^+ = \partial\Omega_D$ and $\partial\Omega_0 \cup \partial\Omega^- = \partial\Omega_N \cup \partial\Omega_R$. The problem can be analyzed by a two-field Friedrichs' system with partial

coercivity. The size of the system is given as $m = d + 1$, $m^\sigma = d$, and $m^u = 1$. The two-field Friedrichs' system for this problem reads:

$$(5.19) \quad G = \begin{bmatrix} \hat{\kappa}^{-1} & 0_{d \times 1} \\ 0_{1 \times d} & 0 \end{bmatrix} \quad \text{and} \quad A_k = \begin{bmatrix} 0_{d \times d} & \mathbf{e}_k \\ \mathbf{e}_k^T & \beta_k \end{bmatrix}.$$

We further assume that:

$$(5.20) \quad \operatorname{ess\,inf}_{\Omega} \left(\frac{1}{2} \nabla \cdot \boldsymbol{\beta} \right) \geq 0,$$

to gain partial coercivity. The boundary conditions can be enforced by specifying the boundary operator M as

$$(5.21) \quad \begin{cases} \alpha = +1, M^{uu} = 2\varrho I_{m^u} + \boldsymbol{\beta} \cdot \mathbf{n}, & \text{on } \partial\Omega_D, \\ \alpha = -1, M^{uu} = 2\varrho I_{m^u} + \boldsymbol{\beta} \cdot \mathbf{n}, & \text{on } \partial\Omega_N \cup \partial\Omega_R. \end{cases}$$

In addition, we set $\varrho = \frac{1}{2}$ on $\partial\Omega_D$, $\varrho = 0$ on $\partial\Omega_N$ and $\varrho = -\boldsymbol{\beta} \cdot \mathbf{n}$ on $\partial\Omega_R$. Thus, conditions (2.3b) and (2.3c) are satisfied. Finally, the following condition is also assumed:

$$(5.22) \quad \boldsymbol{\beta} \cdot \mathbf{n} \neq 0 \text{ on } e, \quad \forall e \in \mathcal{E}_h, \quad \llbracket \boldsymbol{\beta} \cdot \mathbf{n} \rrbracket = 0 \text{ on } \partial K, \quad \forall K \in \mathcal{T}_h.$$

That is, we assume $\boldsymbol{\beta} \cdot \mathbf{n}$ is always continuous across element edges and does not vanish at edges (or mortars). As mentioned in the previous example, this condition can be relaxed by modifying the derivation of the upwind flux.

LEMMA 5.3. *The hp-HDG formulation for the PDE stated in (5.17) is well-posed both locally and globally if the assumptions (5.20) and (5.22) hold.*

Proof. Assumptions (A.1)-(A.3), (A4.b), and (A.5)-(A.6) hold true and can be easily verified. In addition, (A4.a) also holds if assumption (5.20) does. On the other hand, the eigendecomposition of A reads

$$(5.23) \quad |A| = \frac{1}{\sqrt{|\boldsymbol{\beta} \cdot \mathbf{n}|^2 + 4}} \begin{bmatrix} 2\mathbf{n}\mathbf{n}^T & (\boldsymbol{\beta} \cdot \mathbf{n}) \mathbf{n} \\ (\boldsymbol{\beta} \cdot \mathbf{n}) \mathbf{n}^T & |\boldsymbol{\beta} \cdot \mathbf{n}|^2 + 2 \end{bmatrix}.$$

Thus, by setting $\Phi = \frac{2}{\boldsymbol{\beta} \cdot \mathbf{n}}$ and $\Psi = \frac{\boldsymbol{\beta} \cdot \mathbf{n}}{\sqrt{|\boldsymbol{\beta} \cdot \mathbf{n}|^2 + 4}}$, hypothesis (F.1) holds since $\boldsymbol{\beta} \cdot \mathbf{n} \neq 0$ across all elements.

Since the stabilization parameter $T = \frac{1}{2} \left(\sqrt{|\boldsymbol{\beta} \cdot \mathbf{n}|^2 + 4} - \boldsymbol{\beta} \cdot \mathbf{n} \right)$ we have

- $\frac{1}{2}C + T = \frac{1}{2}\sqrt{|\boldsymbol{\beta} \cdot \mathbf{n}|^2 + 4} > 0$, and
- $\bigcap_{k=1}^d \operatorname{Range}(\mathbf{e}_k) = \{\mathbf{0}\}$ and $\mathcal{N}(\mathbf{e}_k) = \{\mathbf{0}\}$ for $\forall k = 1, \dots, d$.

By Lemma 3.9 and Theorem 3.10, we conclude that the hp-HDG formulation for (5.22) is well-posed both locally and globally. $\square$

5.4. Numerical settings and results. For the numerical experiments, we use the square domain $\Omega = (0, 1) \times (0, 1)$ for (E.1)-(E.2), (HP.1), and (HB.1), and the rectangular domain $\Omega = (0, 8.4) \times (0, 24)$ for (E.3). In addition, they are initially solved on the simple meshes as shown in Figure 4 with $p_K = 2$ for $\forall K \in \mathcal{T}_h$ at the 0-th cycle of adaptation.

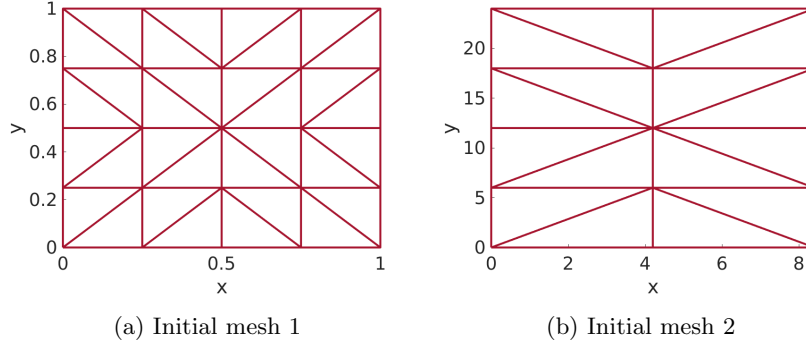


Fig. 4: (a) Initial mesh for (E.1)-(E.2),(HP.1), and (HB.1), and (b) initial mesh used for (E.3).

The solver developed in this work is built upon a **MATLAB** code discussed in [64]. For the numerical evaluation of integrals, cubature rule is used over elements and Gauss quadrature over the surfaces of elements. The adaptation is performed using Algorithm 4.1 with different error indicators stated in (4.11) for all problems. Convergence histories of $\mathcal{L}^2$ -error norm are also presented if the exact solutions are available.

(E.1) Poisson's problem with a corner singularity. Consider Poisson's problem stated in (5.3) where the diffusivity coefficient $\hat{\kappa}$ is set to be the identity matrix, the forcing term f is set to zero, and the exact solution is given below:

$$(5.24) \quad u(x_1, x_2) = 2(x_1^2 + x_2^2)^{-3/4} x_1 x_2 (1 - x_1)(1 - x_2).$$

Dirichlet boundary condition is applied to all the boundaries such that the solution can satisfy (5.24). It can be shown (see [6]) that the solution presented in (5.24) is singular at the origin $(x_1, x_2) = (0, 0)$, but is regular in the rest of the domain Ω . The problem is also studied in [43, 44].

We take $\omega = 0.01$ and Figure 5 shows the corresponding results at the adaptation cycle where the $\mathcal{L}^2$ -norm of $u - u_h$ have an order of magnitude $O(10^{-4})$ for both approaches. The p_K map shown in Figure 5 matches our expectations, that is, aggressive h -refinement takes place near the singularity while intensive p -refinement occurs in the other part of the domain. The p_K maps produced by the two different approaches are similar. Around the singularity, numerous small elements with low-order approximation are generated by the adaptation procedure, while away from it, a few large elements are generated with high-order approximation. However, overall, higher degrees of approximation are generated in the adjoint approach as opposed to Doleji's approach.

Despite the presence of intense oscillation near the singularity, hp -adaptation forces the oscillation zone to shrink. As demonstrated in Figure 5, the numerically polluted area is significantly reduced to a small region at the final cycle of the adaptation. This improvement can also be seen in Figure 6 which shows a convergence study of u_h with different tolerance values $\omega = 0.01$ and $\omega = 0.1$. For the lower tolerance $\omega = 0.01$, both approaches show good convergence behavior, but Doleji's method requires fewer degrees of freedom than the adjoint method at a given error level. For the higher tolerance $\omega = 0.1$, however, the convergence rate for Doleji's approach is flattened out near 10^3 degrees of freedom, whereas the adjoint counterpart still converges to the true solution.

(E.2) Anisotropic diffusion problem with the discontinuous Dirichlet boundary condition. In this example, we consider a strongly anisotropic diffusion problem stated in (5.3) where the diffusivity coefficient $\hat{\kappa}$ is set with $\theta_m = \pi/4$ and $\mathcal{A}_\kappa = 1000$, and the forcing term f is set to zero. In addition, the following piecewise constant Dirichlet boundary data is applied to (5.5c):

$$(5.25) \quad g_D = \begin{cases} 1 & \text{when } x_1 = 1 \text{ or } x_2 = 0, \\ 0 & \text{when } x_1 = 0 \text{ or } x_2 = 1. \end{cases}$$

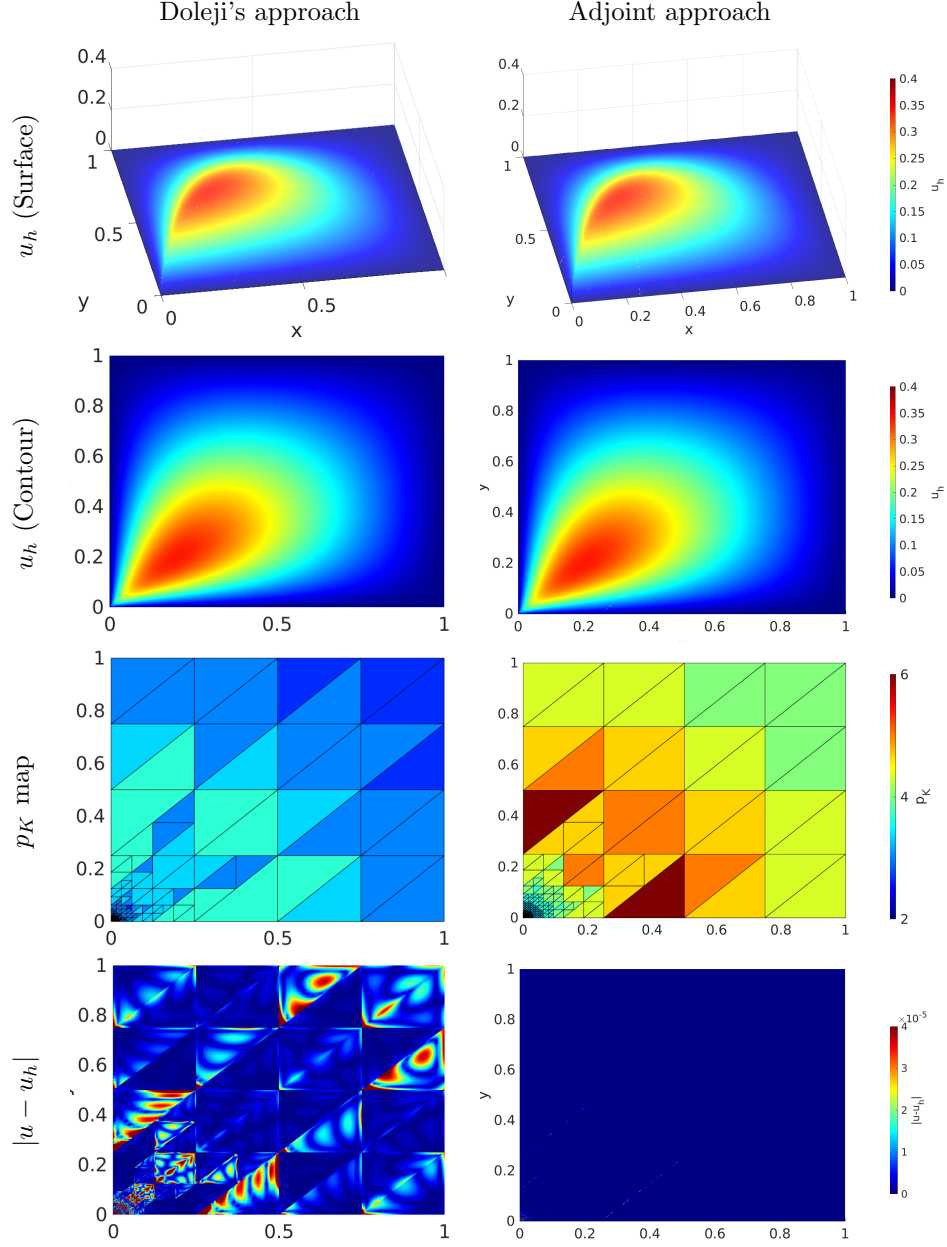


Fig. 5: Numerical results at the adaptation cycle where $\mathcal{L}^2$ -norm of $u - u_h$ is about $O(10^{-4})$ for the isotropic diffusion problem modeled by the elliptic problem (E.1) with the exact solution stated in (5.24). The tolerance is chosen as $\omega = 0.01$. The left column uses Doleji's approach (4.1) while the right column uses the adjoint approach (4.10). Surface and contour plots of the numerical solution are presented in the first two rows, the mesh configuration along with the arrangement of the degree of approximation p_K is presented in the third row, and the absolute error is presented in the fourth row.

903 The discontinuities at the corners $(0,0)$ and $(1,1)$ induce sharp gradients, making the problem difficult to
 904 solve. This problem is also investigated in [103, 80]. A semi-analytic solution for this test problem can be
 905 found by a sequence of geometric transformations which are numerically computed using MATLAB Schwarz-
 906 Christoffel toolbox [45]. Given that the accuracy of the mapping is sufficient, we treat this semi-analytic
 907 solution as “exact” to benchmark against our hp -HDG solution.

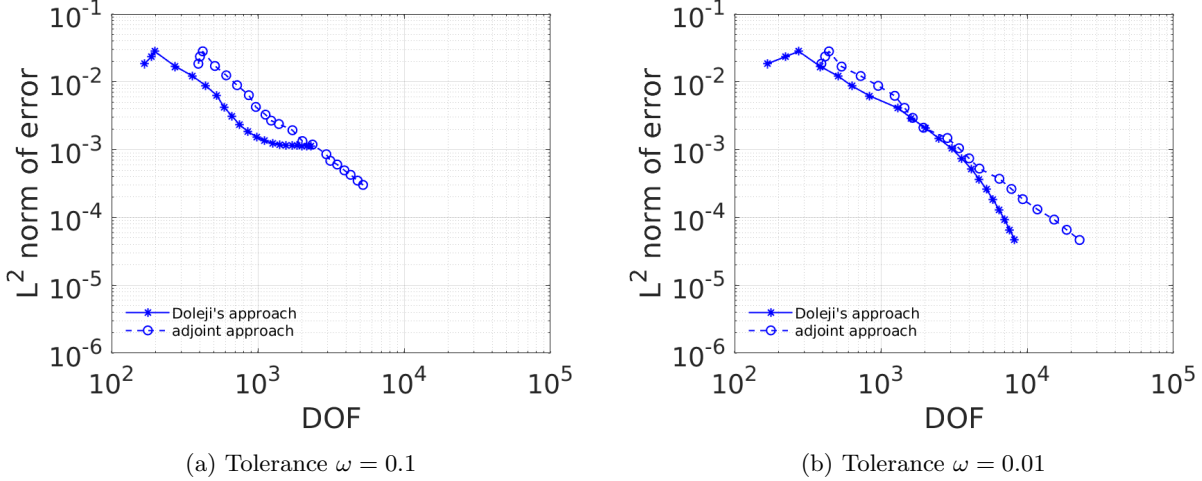


Fig. 6: Convergence histories of u_h measured in $\mathcal{L}^2$ -norm. The results are obtained by numerically solving the isotropic diffusion problem modeled by the elliptic problem (E.1) (with the exact solution in (5.24)) with different tolerance (a) $\omega = 0.1$ and (b) $\omega = 0.01$. In each plot, the results obtained by Doleji's and adjoint approaches are presented.

With the tolerance $\omega = 0.01$, we plot numerical results of both approaches in Figure 7 at the adaptation cycle where $\mathcal{L}^2$ -norm of $u - u_h$ are about $O(10^{-3})$. Due to the strong anisotropic feature, the solution behaves like convection where the amount of the flux transported in the specific direction is more than in the other direction. In this example, the dominant direction is 45 degrees from the x_1 -axis. As a result, “discontinuity”-like behavior occurs within the domain along the diagonal and becomes more substantial around the corners due to the presence of discontinuous Dirichlet boundary data. As can be seen in Figure 7, Gibbs phenomenon [60] occurs around the corners $(0, 0)$ and $(1, 1)$. Similar to the numerical result shown for (E.1), the numerically polluted area can be largely reduced though the oscillation cannot be completely removed via the adaptation process. This observation is also consistent with the convergence histories of u_h presented in Figure 8. On the other hand, Figure 8 demonstrates that a small tolerance value is required in this testing case to achieve acceptable convergence rates. Moreover, increasing the anisotropic ratio $\mathcal{A}_\kappa$ will make the problem more challenging to solve because the profile of the u_h will tend to be even steeper. A shock-like front may form, which causes more h -refinement and hence more DOFs. A closer look at Figure 8 reveals that the adaptation process for Doleji's approach stops first due to the satisfaction of criteria $\max_{K \in \mathcal{T}_h} \mathcal{E}_{h,K} \geq \omega \mathcal{E}_h$ while the adjoint approach proceeds further and stops owing to the maximum number of iterations. This observation suggests that more robust stopping criteria may be needed. However, the effort in designing robust stopping criteria may not be trivial and is beyond the scope of this paper. In summary, both approaches are comparable in this testing case using the adaptation algorithm outlined in Algorithm 4.1.

(E.3) Heterogeneous anisotropic diffusion problem with discontinuous field $\hat{\kappa}$. Here we consider the problem stated in (5.3), but with a piecewise constant diffusivity coefficient $\hat{\kappa}$ and Neumann/Robin mixed type boundary conditions

$$(5.26) \quad \hat{\kappa} \nabla u \cdot \mathbf{n} + \lambda u = g, \quad \text{on } \partial\Omega_N \cup \partial\Omega_R,$$

where $\partial\Omega = \partial\Omega_N \cup \partial\Omega_R$. Given that $\hat{\kappa}$ is now spatially varying the problem is not only anisotropic but also heterogeneous. The PDE can model the heat conduction in non-homogeneous materials, where u describes the temperature field. For example, the so-called “battery problem” [41], is of this type and is examined here. The domain is then modeled as a battery composed of five different materials which are indexed as numbers 1-5 in Table 1. The values of $\hat{\kappa}$ for different materials and the corresponding forcing term f are

summarized in Table 2. The boundary data is given in Table 3.

Figure 9 shows the numerical results of the two approaches at the final cycle of adaptation with $\omega = 0.01$. This problem is challenging in that the coefficient of the PDE is discontinuous across the entire domain. Without aligning the mesh skeleton with these discontinuities, a serious Gibb's phenomenon is easily induced. If we simply employ an isotropic h -refinement, then the numerically polluted area may still spread to some extent. One remedy for addressing this issue is to use anisotropic h -refinement [42, 80, 25, 9, 12]. However, such refinement requires a more delicate error estimator/indicator and needs to be equipped with a proper algorithm for generating a mesh. This task is left for future work.

Material	Region
1	$[0, 8.4] \times [0, 0.8), (8, 8.4] \times [0.8, 23.2], [0, 8.4] \times (23.2, 24]$
2	$[0, 6.1) \times [1.6, 3.6), [0, 6.1) \times [18.8, 21.2)$
3	$[0, 6.1) \times [3.6, 18.8)$
4	$[6.1, 6.5) \times [0.8, 21.2)$
5	$[0, 6.1) \times [0.8, 1.6), (6.5, 8) \times [0.8, 21.2), [0, 8) \times [21.2, 23.2)$

Table 1: The geometry of materials of the battery problem modeled with (E.3).

Material	κ_X	κ_Y	$\mathcal{A}_\kappa$	θ_m	f	BC data	λ	g
1	25.0	25.0	1.00	0.0	0.0	Left	0.0	0.0
2	7.0	0.8	8.75	0.0	0.0	Up	1.0	3.0
3	5.0	0.00001	5.00×10^5	0.0	1.0	Right	2.0	2.0
4	0.2	0.2	1.00	0.0	1.0	Bottom	3.0	1.0
5	0.05	0.05	1.00	0.0	0.0			

Table 2: Diffusivity coefficient $\hat{\kappa}$ and forcing term f for the battery problem.

Table 3: Boundary data used in (5.5c) for the battery problem.

(HP.1) Steady-state linear advection with variable speed and discontinuous inflow condition. In this experiment, we are going to solve the linear advection problem described in (5.9) along with the advection velocity $\beta = (1 + \sin(\pi x_2), 2)$ and the inflow data

$$(5.27) \quad g_D = \begin{cases} 1, & \text{for } x_1 = 0, 0 \leq x_2 \leq 1, \\ \sin^6(2\pi x_1), & \text{for } 0 \leq x_1 \leq 0.5, x_2 = 0, \\ 0, & \text{for } 0.5 \leq x_1 \leq 1, x_2 = 0, \end{cases}$$

where there is a discontinuity occurring right at the origin. The problem is also studied in [20, 84] and can be solved exactly by using the method of characteristics.

In Figure 10, we present the numerical results of both methods at the adaptation cycle where $\mathcal{L}^2$ -norm of $u - u_h$ is $O(10^{-1})$. Due to the discontinuous inflow boundary data and the nature of hyperbolic PDEs, we have a shock formed within the domain Ω . It is very challenging to remove the oscillation induced by Gibbs' phenomena unless all the discontinuities are well aligned with the skeleton of the mesh and the first order of approximation is used near the discontinuities. Given that we only consider isotropic h -refinement here, it is not possible to meet this condition. Nonetheless, we can still narrow down the region of shock-induced oscillation by the hp -adaptation process. As we expect, the aggressive h -refinement is performed around the shock, but less in near-outflow region even with strong discontinuities.

Figure 11 presents the convergence of u_h using both approaches along with two different tolerance values. It can be observed that the convergence rate can only be improved (not zero anymore) if the tolerance is set to be small enough. In addition, the two approaches are comparable in this example as well.

(HB.1) Steady-state convection-diffusion problem with discontinuous inflow condition. In this example, we focus on the steady-state convection-diffusion equation (5.17) and especially examine the problem first proposed by Eriksson and Johnson in [48]. This problem is also investigated in [26] using a discontinuous Petrov-Galerkin method. The diffusivity matrix is set to be $\hat{\kappa} := \varepsilon I_2$ where I_2 is the 2×2 identity matrix, and the velocity field is stated as $\beta := (0, 1)$. The boundaries are given as follows:

$$\begin{aligned}\partial\Omega^- &= \{(x_1, x_2) : x_1 = 0, 0 \leq x_2 \leq 1\}, \\ \partial\Omega^+ &= \{(x_1, x_2) : x_1 = 1, 0 \leq x_2 \leq 1\}, \\ \partial\Omega_0 &= \{(x_1, x_2) : 0 < x_1 < 1, x_2 = 0 \text{ or } 1\}.\end{aligned}$$

The boundary data in (5.18) read:

$$(5.28) \quad \begin{aligned}g_D &= 0, \text{ on } \partial\Omega^+ \\ g_{N,R} &= \begin{cases} (\beta u_0 + \sigma_0) \cdot \mathbf{n} & \text{on } \partial\Omega^-, \\ 0 & \text{on } \partial\Omega_0, \end{cases}\end{aligned}$$

and $u_0 := u(0, x_2)$, $\sigma_0 := \sigma(0, x_2)$. Further, the function u_0 is set to be a discontinuous function :

$$(5.29) \quad u_0(x_2) = \begin{cases} (x_2 - 1)^2, & x_2 > 0.5, \\ -x_2^2, & x_2 \leq 0.5. \end{cases}$$

The Eriksson-Johnson problem can be solved by the separation of variables and the solution is:

$$(5.30) \quad u(x_1, x_2) = C_0 + \sum_{i=1}^{\infty} C_i \frac{e^{s_2(x_1-1)} - e^{s_1(x_1-1)}}{e^{-s_2} - e^{-s_1}} \cos(i\pi x_2),$$

where

$$\begin{aligned}C_i &= \int_0^1 2u_0 \cos(i\pi x_2) dx_2, \\ s_{1,2} &= \frac{1 \pm \sqrt{1 + 4\varepsilon\sigma_i}}{2\varepsilon}, \\ \sigma_i &= \varepsilon i^2 \pi^2.\end{aligned}$$

In this testing case, we actually have $\partial\Omega^- = \partial\Omega_R$, $\partial\Omega^+ = \partial\Omega_D$, and $\partial\Omega_0 = \partial\Omega_N$. Note that we do not have a closed form of the exact solution. Therefore, for a convergence study, we approximate u_0 using the first 20 terms of the series in (5.30). Similarly, σ_0 can be approximated in the same way. The problem is tricky because there is not only discontinuous inflow data but also a boundary layer developed around the outflow boundary. In addition, the smaller the diffusivity coefficient ε is, the thinner the boundary layer which can only be well-resolved using a mesh with a proper resolution.

Figure 12 shows the numerical results of the two approaches with $\varepsilon = 10^{-3}$ and $\omega = 0.05$ at the adaptation cycle where the $\mathcal{L}^2$ -norm of $u - u_h$ is $O(10^{-3})$. As discussed previously, there is a boundary layer (sharp gradient in solution u) around the outflow boundary $x_1 = 1$. Furthermore, the resulting p_K maps for the two approaches are significantly different. Doleji's approach does not capture the sharp gradient induced by discontinuous Dirichlet boundary data at the inflow boundary $\partial\Omega^-$ and the smooth region near zero-flow boundaries $\partial\Omega_0$, but approximates the boundary layer near the outflow boundary $\partial\Omega^+$ well. Unless an even smaller tolerance value is provided, the value of local error indicator $\mathcal{E}_{h,K}^{\text{Doleji}}(u_h)$ surrounding the boundary layer is much higher than that of the rest of the domain, resulting in much less refinement in the area except for the region around $\partial\Omega^+$. On the other hand, the adjoint approach does capture almost every feature of the solution u . However, the local h -refinement is still not sufficient at the outflow boundary $\partial\Omega^+$, and the boundary layer is under-resolved due to flattened convergence rate of the adjoint approach in Figure 13.

Figure 13 shows the convergence histories of u_h together with three different diffusivities. To capture the thinner boundary layer caused by the smaller diffusivity, we employ a smaller tolerance value. As can be seen, Doleji's approach is slightly better than the adjoint approach in terms of accuracy and convergence rate.

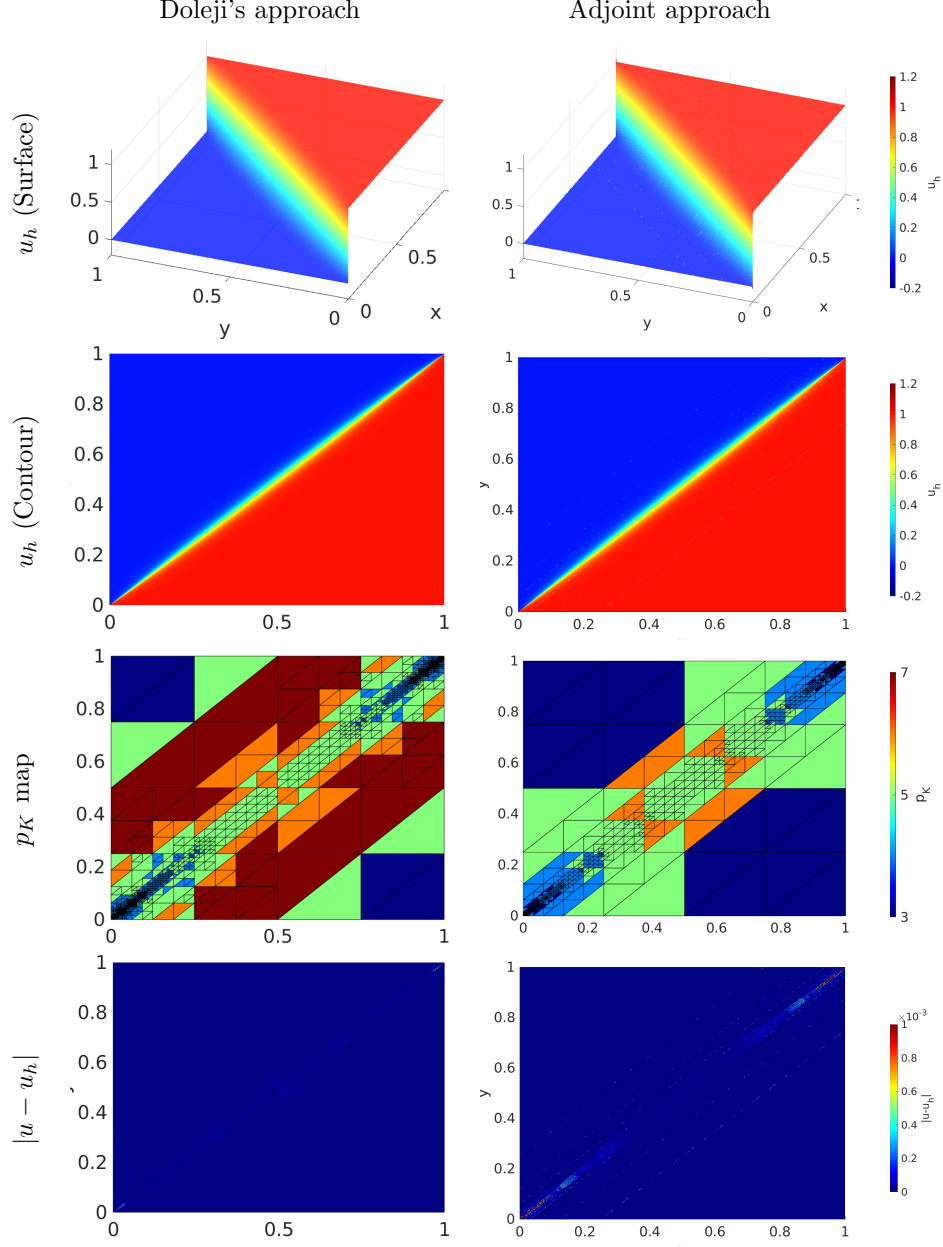


Fig. 7: Numerical results at the adaptation cycle where $\mathcal{L}^2$ -norm of $u - u_h$ is about $O(10^{-3})$. The results are collected by solving the anisotropic diffusion problem modeled by the elliptic problem (E.2) with anisotropic ratio $\mathcal{A}_\kappa = 1000$. Further, the semi-analytic solution can be obtained with the aid of accurate mappings. The tolerance is chosen as $\omega = 0.01$. The left column is for Doleji's approach (4.1) and the right column for the adjoint approach (4.10). Surface and contour plots of the numerical solution are presented in the first two rows, the mesh configuration along with the arrangement of the degree of approximation p_K is presented in the third row, and the absolute error is presented in the fourth row.

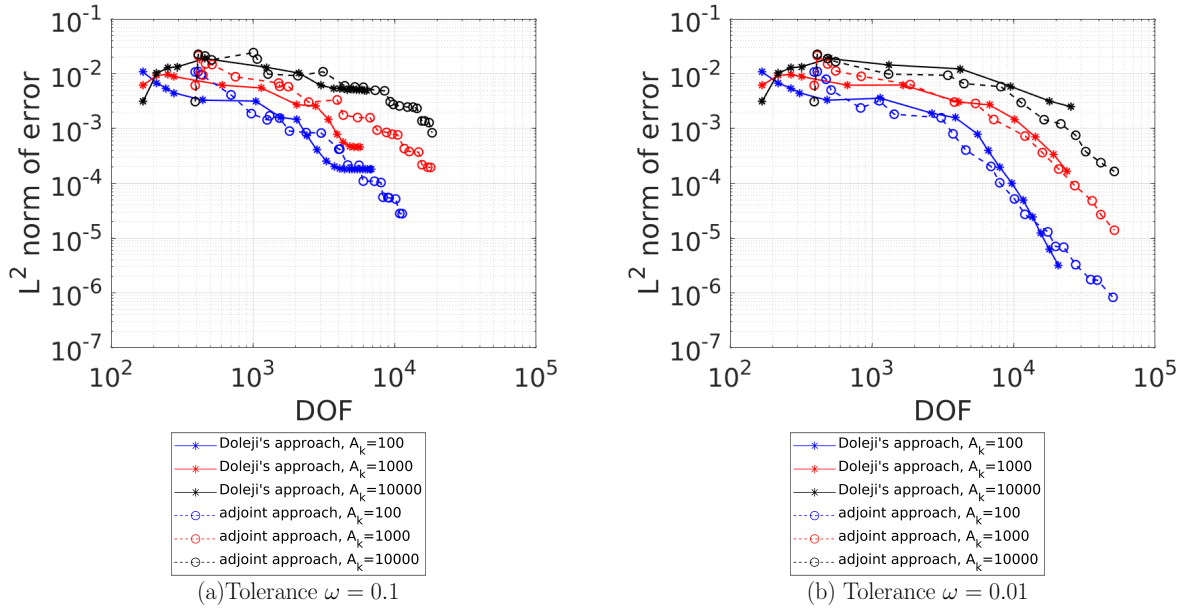


Fig. 8: Convergence histories of u_h measured in $\mathcal{L}^2$ -norm. The results are obtained by numerically solving the anisotropic diffusion problem modeled by the elliptic problem (E.2) where the semi-analytic solution can be obtained with the aid of accurate mappings. Different tolerance values of (a) $\omega = 0.1$ and (b) $\omega = 0.01$ are used. In each plot, the results with various anisotropy ratios A_k (denoted by different colors) obtained by different approaches (denoted by different marks) are presented.

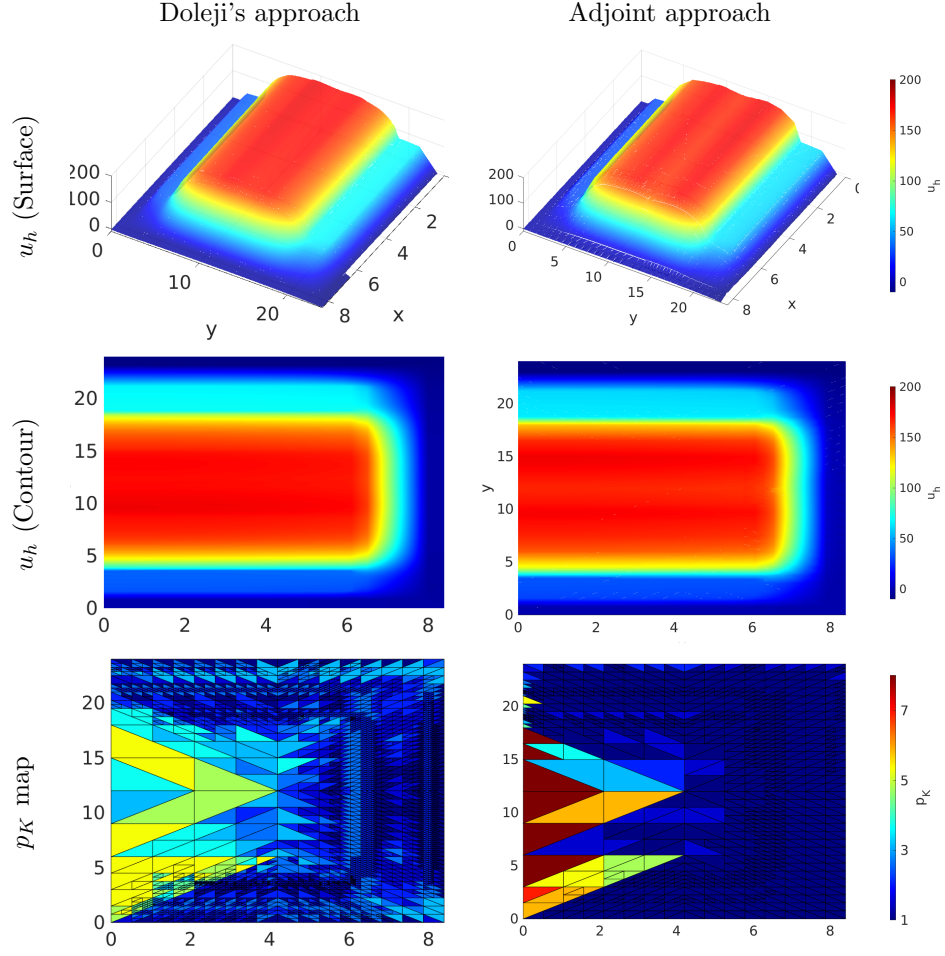


Fig. 9: Numerical results at the final cycle of adaptation for the battery problem modeled with the elliptic problem (E.3). The tolerance is chosen as $\omega = 0.01$. The left column is for Doleji's approach (4.1) and the right column for the adjoint approach (4.10). Surface and contour plots of the numerical solution are presented in the first two rows, and the mesh configuration along with the arrangement of the degree of approximation p_K is presented in the third row.

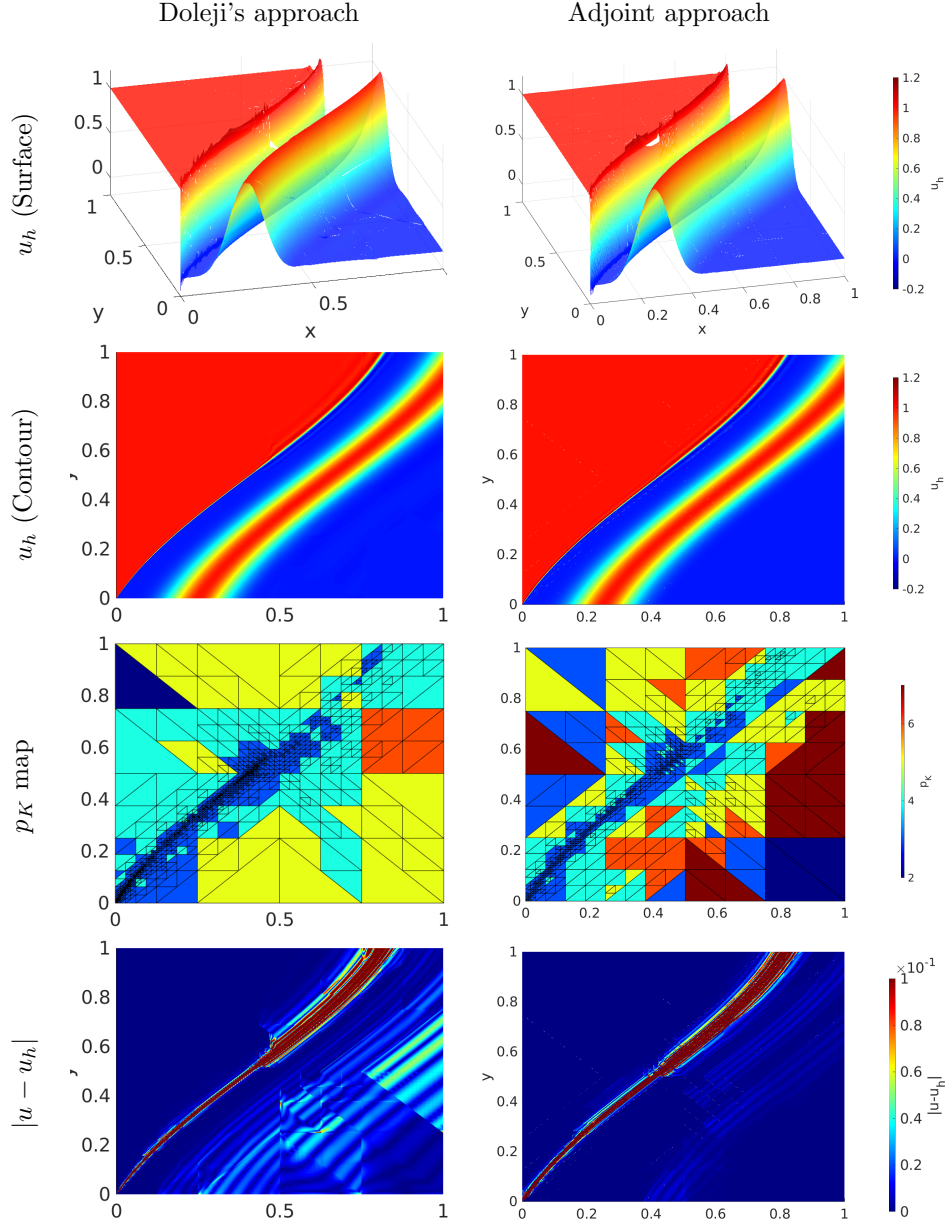


Fig. 10: Numerical results at the adaptation cycle where $\mathcal{L}^2$ -norm of $u - u_h$ is about $O(10^{-1})$ for the linear advection problem modeled with the hyperbolic problem (HP.1). Further, the exact solution can be obtained using the method of characteristics. The tolerance is chosen as $\omega = 0.05$. The left column is for Doleji's approach (4.1) and the right column for the adjoint approach (4.10). Surface and contour plots of the numerical solution are presented in the first two rows, the mesh configuration along with the arrangement of the degree of approximation p_K is presented in the third row, and the absolute error is presented in the fourth row.

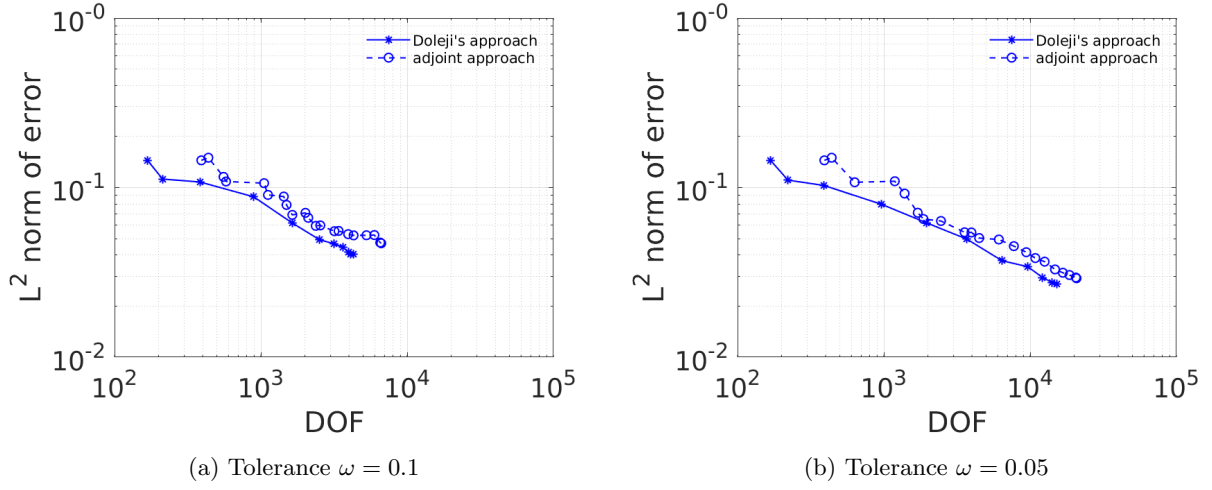


Fig. 11: Convergence histories of u_h measured in $\mathcal{L}^2$ -norm. The results are obtained by numerically solving the linear advection problem modeled with the hyperbolic problem (HP.1). The exact solution can be found by the method of characteristics. The results obtained with two different tolerances of (a) $\omega = 0.1$ and (b) $\omega = 0.05$ are presented.

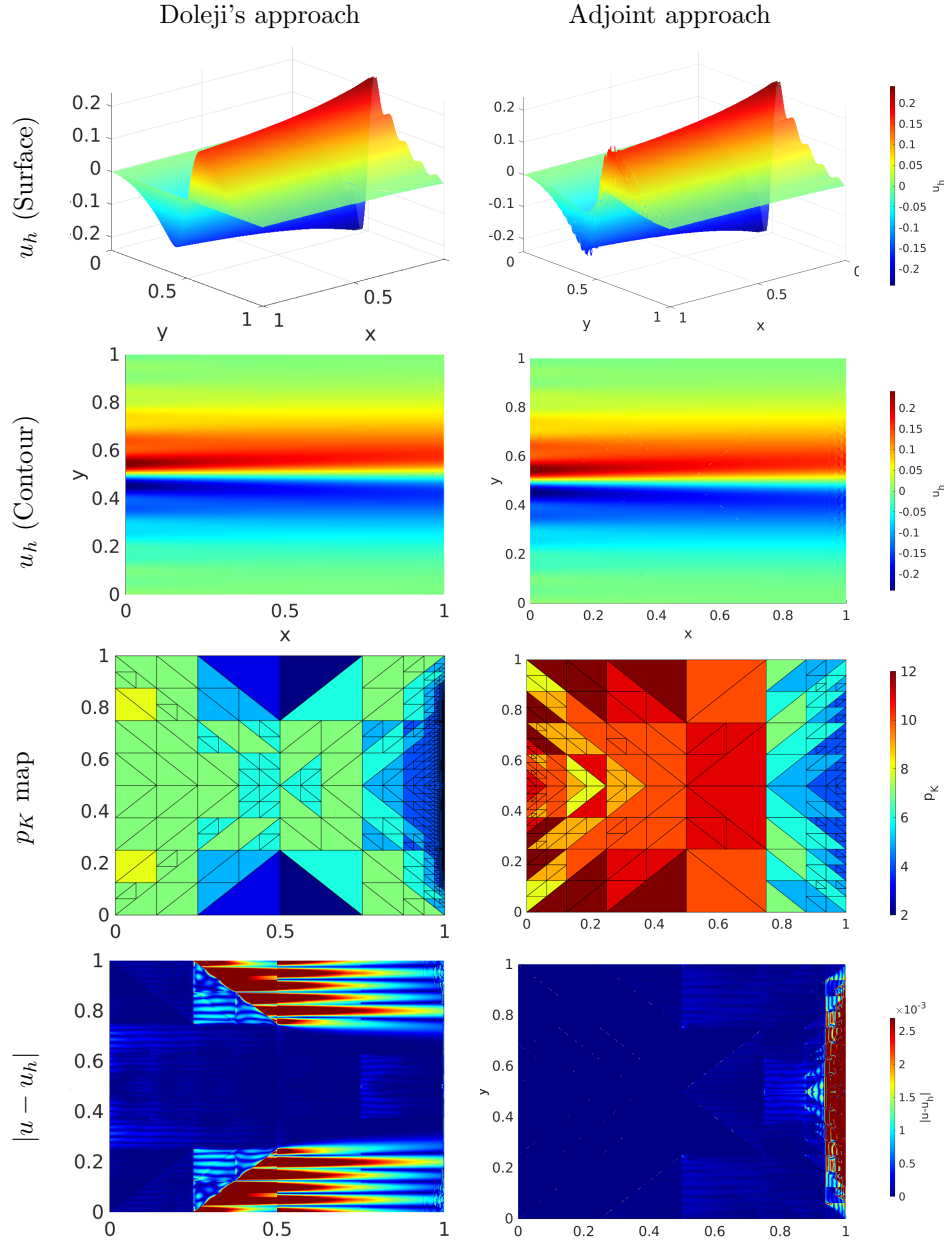


Fig. 12: Numerical results at the adaptation cycle where $\mathcal{L}^2$ -norm of $u - u_h$ for the convection-diffusion problem modeled by the mixed problem (HB.1) with the exact solution stated in (5.30) with the diffusivity $\varepsilon = 10^{-3}$. The tolerance is chosen as $\omega = 0.05$. The left column is for Doleji's approach (4.1) and the right column for the adjoint approach (4.10). Surface and contour plots of the numerical solution are presented in the first two rows, the mesh configuration along with the arrangement of the degree of approximation p_K is presented in the third row, and the absolute error is presented in the fourth row.

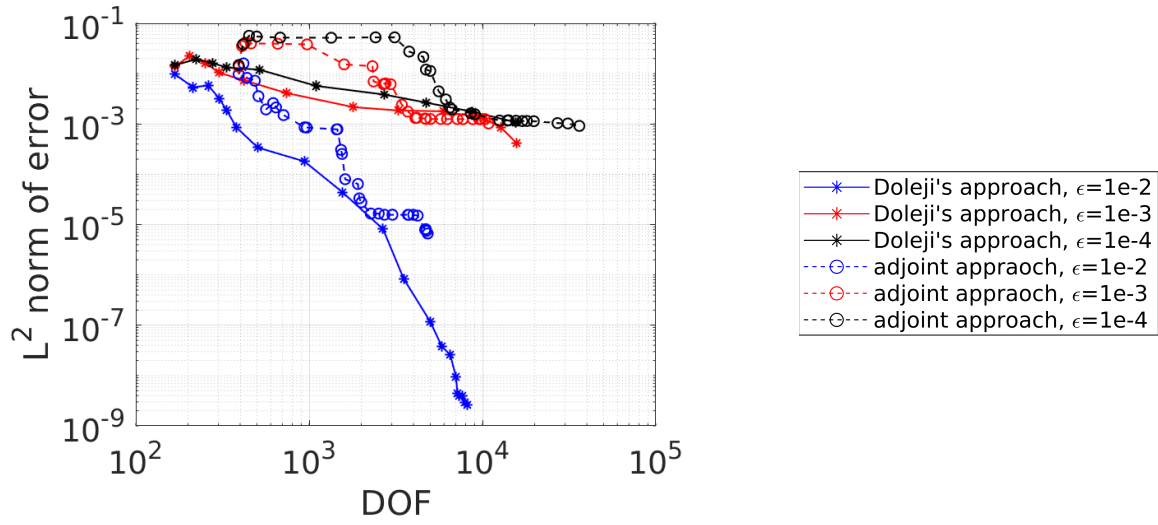


Fig. 13: Convergence histories of u_h measured in $\mathcal{L}^2$ -norm. The results are obtained by numerically solving the convection-diffusion problem modeled by the mixed problem (HB.1) that admits the exact solution stated in (5.30). The results with various diffusivity values ε (denoted by different colors) are presented. For different diffusivity values $\varepsilon = 10^{-2}, 10^{-3}$ and 10^{-4} , different tolerance $\omega = 0.1, 0.05$ and 0.01 are used respectively.

6. Conclusion. In this work, we proposed unified hp -HDG frameworks for Friedrichs' system that embraces large classes of PDEs. At the heart of our adaptivity strategy is to take advantage of natural built-in mortars in HDG methods. By choosing *split-sided* type of mortar, the degree $p_e = \max\{p_{K^+}, p_{K^-}\}$ for $\forall e \in \mathcal{E}_h^\circ$ and $p_e = p_K$ for $\forall e \subset \partial K \cap \mathcal{E}_h^\partial$, an upwind-based numerical flux can be naturally derived and be constructed over mortars. For one-field Friedrichs' systems, we have a general form of the numerical flux. For two-field Friedrichs' systems, we exploit the system's intrinsic structure to obtain the numerical flux in the reduced form. The existence of such a numerical flux is guaranteed as long as a few more additional assumptions are satisfied. We have shown that hp -HDG formulations are parameter-free and are well-posed for both one-field and two-field Friedrichs' systems. Leveraging the Friedrichs' framework we have systematically constructed and analyzed hp -HDG formulations for one-field and two-field systems. The unification opens an opportunity for us to develop a single universal code to solve various kinds of PDEs.

Besides the analysis, several numerical experiments are also carried out. In the experiments, three distinct types of PDEs are considered: elliptic, hyperbolic, and mixed-type. We showed that these PDEs fall into our framework and their hp -HDG formulations are hence well-posed. Additionally, a simple algorithm is proposed to drive the hp -adaptation and to verify its efficacy. Two different error indicators are used for our adaptivity strategy: Doleji's approach and the adjoint approach. The former relies on the estimation of the smoothness of the numerical solutions. The latter uses duality to derive an estimation of accuracy for a chosen output functional. Regarding performance, both approaches show improvement in the convergence rates and are comparable to each other for almost all the examples if an appropriate tolerance is picked. This is expected as the global objective function is used. (For a more localized objective, the adjoint is typically much more efficient, but it is not our focus here. We simply demonstrate that our hp -formulations and hp -adaptation can work with both of these popular approaches.) However, for the Poisson problem considered in (E.1) the exponential convergence rates are only observed at the first half of cycles of adaption in Doleji's approach with the tolerance $\omega = 0.1$. It is somehow stagnant for the rest half of the cycles of adaption. On the other hand, the convergence rate of the adjoint approach with the tolerance $\omega = 0.01$ is about 1.4-1.5, which is not appealing. Moreover, we also found that both approaches are quite sensitive to the user-defined tolerance value. A more robust hp -refinement strategy is required and it is the subject of future work.

Finally, we end the conclusion with a couple of remarks. In this work, we only considered one-field and two-field Friedrichs' systems. Another significant structure is the three-field, which includes the PDE governed by the linearized incompressible flow. However, such a system requires a significant amount of dedicated discussion and deserves another paper to cover it. Thus, it is left for future work. Although we only consider steady-state PDEs in this paper, it is not hard to extend our current work to time-dependent models where the temporal derivative is first discretized by some single-step time scheme. The discretized terms can be treated as reaction and forcing terms, and the resulting semi-discrete PDE can then be re-written as a general conservation form stated in (2.1). Thus, the analysis presented in this study is still applicable. However, the algorithm of hp -adaptation may need to be re-designed because different time steps can cause the solution to behave differently. Thus, h -coarsening operations may also be required in response to this change. Furthermore, a proper transfer of the solution between each adaptation needs to be carefully addressed.

Acknowledgments. Thanks to Geonyoung Lee for the fruitful discussion on the hp -adaptivity for elliptic equations. This research is partially funded by the National Science Foundation awards NSF-OAC-2212442, NSF-2108320, NSF-1808576 and NSF-CAREER-1845799; by the Department of Energy award DE-SC0018147 and DE-SC0022211.

Appendix A. Adjoint hp -HDG formulation. In this paper, we deployed a discrete weak adjoint approach that allows us to follow (almost identically) the framework outlined in [13] to derive adjoint-based error estimations. Instead of treating the governing equation (2.1) as a primal problem, here we consider hp -HDG formulation that is stated in (3.12) or (3.34) as a primal problem. The procedure of derivation of the adjoint problem is briefly described below:

1. Re-state the primal problem, either (3.12) or (3.34), as a bilinear form.
2. Define an output functional $\mathcal{J}(\cdot)$ (Recall that $\mathcal{J}(\cdot) = \mathcal{J}^{adjoint}(\cdot) + \mathcal{J}^{boundary}(\cdot)$) of the approximate solution $\mathcal{Z}_h$ given in (4.2a) and use it as an objective function of an optimization problem along with the constraint posed by the bilinear formulation given in step one.

3. Re-write the constrained optimization problem constructed in the previous step as the unconstrained optimization problem applying the Lagrangian approach in which the test function $\mathcal{W}_h$ presented in the bilinear formulation now becomes a Lagrange multiplier.
4. Solve this trivial optimization problem by taking advantage of the first optimality condition and the adjoint hp -HDG formulation can then be derived.

Note that the derivation of the adjoint PDEs can also be done by following a similar procedure. However, it should be noted that the adjoint hp -HDG formulations derived by following the above procedure are not necessarily the same as the ones discretized from the adjoint PDEs. In fact, they are different in this paper.

The adjoint hp -HDG formulation with regard to the one-field Friedrichs' system reads: seek $(\mathbf{w}_h, \widehat{\mathbf{w}}_h) \in \mathbf{W}_h \times \widehat{\mathbf{W}}_h$ such that:

$$(a.1a) \quad -\sum_{k=1}^d (\mathbf{w}_h, \partial_k (A_k \delta \mathbf{z}_h))_{\mathcal{T}_h} + \left(G^T \mathbf{w}_h, \delta \mathbf{z}_h \right)_{\mathcal{T}_h} + \langle -A \widehat{\mathbf{w}}_h + T(\mathbf{w}_h - \widehat{\mathbf{w}}_h), \delta \mathbf{z}_h \rangle_{\partial \mathcal{T}_h} = -\delta \mathcal{J}_{\mathbf{z}_h}^{adjoint}(\mathcal{Z}_h; \delta \mathbf{z}_h)$$

(a.1b)

$$\langle [-A \widehat{\mathbf{w}}_h + T(\mathbf{w}_h - \widehat{\mathbf{w}}_h)], \delta \widehat{\mathbf{z}}_h \rangle_{\mathcal{E}_h} = -\left\langle \frac{1}{2} (-A + M)^T \widehat{\mathbf{w}}_h, \delta \widehat{\mathbf{z}}_h \right\rangle_{\mathcal{E}_h^\partial} - \delta \mathcal{J}_{\widehat{\mathbf{z}}_h}^{adjoint}(\mathcal{Z}_h; \delta \widehat{\mathbf{z}}_h) - \delta \mathcal{J}_{\widehat{\mathbf{z}}_h}^{boundary}(\widehat{\mathbf{z}}_h; \delta \widehat{\mathbf{z}}_h),$$

for all $(\delta \mathbf{z}_h, \delta \widehat{\mathbf{z}}_h) \in \mathbf{W}_h \times \widehat{\mathbf{W}}_h$. The stabilization parameter is still set as $T := |A|$ and the boundary operator is set as $M := |A|$. In order to incorporate the boundary condition, the boundary-associated functional $\mathcal{J}^{boundary}(\widehat{\mathbf{z}}_h)$ is included in $\mathcal{J}(\mathcal{Z}_h)$ and has the following form:

$$(a.2) \quad \mathcal{J}^{boundary}(\widehat{\mathbf{z}}_h) = -\left\langle \frac{1}{2} (-A - M)^T \widehat{\mathbf{z}}_h, \mathbf{g}^{adjoint} \right\rangle_{\mathcal{E}_h^\partial},$$

where the function $\mathbf{g}^{adjoint} : \partial\Omega \rightarrow \mathbb{R}^m$ is defined as

$$(a.3) \quad \mathbf{g}^{adjoint} := \begin{cases} \mathbf{g}_D^{adjoint} & \text{if } (A + M) \neq 0, \\ \mathbf{0} & \text{if } (A + M) = 0, \end{cases}$$

and $\mathbf{g}_D^{adjoint}$ is the Dirichlet data for the adjoint system given in (a.1). Comparing (3.9) and (a.3), it can be observed that the inflow and outflow boundaries are switched between the primal formulation (3.12) and its adjoint formulation (a.1). It should be also noted that the adjoint problem (a.1) will automatically have homogeneous boundary conditions if the boundary-associated functional is zero (i.e., $\mathbf{g}_D^{adjoint} = 0$), which is the case in the numerical studies presented in this paper.

The well-posedness analysis is similar to Lemma 3.3 for the local equation (a.1a) and to Theorem 3.4 for the adjoint hp -HDG formulation (a.1). Thus, we simply outline the following lemma and theorem about well-posedness without any proof for the sake of brevity.

LEMMA A.1 (Well-posedness of the local equation). *Suppose that the assumptions (A.1)-(A.4) hold, the local solver (a.1a) is well-posed, that is, given $\widehat{\mathbf{w}}_h$ and $\delta \mathcal{J}_{\mathbf{z}_h}^{adjoint}(\mathcal{Z}_h; \delta \mathbf{z}_h)$, there exists a unique solution $\mathbf{w}_h$ of the local system.*

THEOREM A.2 (Well-posedness of the adjoint hp -HDG formulation). *Suppose that*

1. the assumptions (A.1)-(A.4) and (2.3b) hold,
2. $\mathcal{N}(A) = \{\mathbf{0}\}$,

the adjoint hp -HDG formulation stated in (a.1) is well-posed in the sense that given $\delta \mathcal{J}_{\mathbf{z}_h}^{adjoint}(\mathcal{Z}_h; \delta \mathbf{z}_h)$, $\delta \mathcal{J}_{\widehat{\mathbf{z}}_h}^{adjoint}(\mathcal{Z}_h; \delta \widehat{\mathbf{z}}_h)$ and $\delta \mathcal{J}_{\widehat{\mathbf{z}}_h}^{boundary}(\widehat{\mathbf{z}}_h; \delta \widehat{\mathbf{z}}_h)$, there exists a unique solution $(\mathbf{w}_h, \widehat{\mathbf{w}}_h)$.

On the other hand, the adjoint hp -HDG formulation with regard to the two-field Friedrichs' system

reads: seek $(\mathbf{s}_h, \mathbf{v}_h, \widehat{\mathbf{v}}_h) \in \boldsymbol{\Sigma}_h \times \mathbf{U}_h \times \widehat{\mathbf{U}}_h$ such that

$$(a.4a) \quad -\sum_{k=1}^d \left(\mathbf{v}_h, \partial_k (B_k^T \delta \boldsymbol{\sigma}_h) \right)_{\mathcal{T}_h} + \left((G^{\boldsymbol{\sigma}\boldsymbol{\sigma}})^T \mathbf{s}_h + ((G^{\mathbf{u}\boldsymbol{\sigma}})^T \mathbf{v}_h, \delta \boldsymbol{\sigma}_h)_{\mathcal{T}_h} - \langle B \widehat{\mathbf{v}}_h, \delta \boldsymbol{\sigma}_h \rangle_{\partial \mathcal{T}_h} \right. \\ \left. = -\delta \mathcal{J}_{\boldsymbol{\sigma}_h}^{adjoint}(\mathcal{Z}_h; \delta \boldsymbol{\sigma}_h) \right.$$

$$(a.4b) \quad -\sum_{k=1}^d (\mathbf{s}_h, \partial_k (B_k \delta \mathbf{u}_h))_{\mathcal{T}_h} - \sum_{k=1}^d (\mathbf{v}_h, \partial_k (C_k \delta \mathbf{u}_h))_{\mathcal{T}_h} + \left(G^{\boldsymbol{\sigma}\mathbf{u}}^T \mathbf{s}_h + (G^{\mathbf{u}\mathbf{u}})^T \mathbf{v}_h, \delta \mathbf{u}_h \right)_{\mathcal{T}_h} \\ + \left\langle -B^T \mathbf{s}_h - C^T \widehat{\mathbf{v}}_h + T(\mathbf{v}_h - \widehat{\mathbf{v}}_h), \delta \mathbf{u}_h \right\rangle_{\partial \mathcal{T}_h} = -\delta \mathcal{J}_{\mathbf{u}_h}^{adjoint}(\mathcal{Z}_h; \delta \mathbf{u}_h) \\ (a.4c)$$

$$(a.4d) \quad \left\langle \llbracket -B^T \mathbf{s}_h - C^T \widehat{\mathbf{v}}_h + T(\mathbf{v}_h - \widehat{\mathbf{v}}_h), \delta \mathbf{u}_h \rrbracket, \delta \widehat{\mathbf{u}}_h \right\rangle_{\mathcal{E}_h \setminus \partial \Omega_D} = \langle \varrho I_{m^{\mathbf{u}}} \widehat{\mathbf{v}}_h, \delta \widehat{\mathbf{u}}_h \rangle_{\mathcal{E}_h \cap (\partial \Omega_N \cup \partial \Omega_R)} - \delta \mathcal{J}_{\widehat{\mathbf{u}}_h}^{\partial \Omega_N \cup \partial \Omega_R}(\widehat{\mathbf{u}}_h; \delta \widehat{\mathbf{u}}_h) \\ \langle \widehat{\mathbf{v}}_h, \delta \widehat{\mathbf{u}}_h \rangle_{\mathcal{E}_h^{\partial} \cap \partial \Omega_D} = -\delta \mathcal{J}_{\widehat{\mathbf{u}}_h}^{adjoint}(\mathcal{Z}_h; \delta \widehat{\mathbf{u}}_h) - \delta \mathcal{J}_{\widehat{\mathbf{u}}_h}^{\partial \Omega_D}(\widehat{\mathbf{u}}_h; \delta \widehat{\mathbf{u}}_h),$$

for all $(\delta \boldsymbol{\sigma}_h, \delta \mathbf{u}_h, \delta \widehat{\mathbf{u}}_h) \in \boldsymbol{\Sigma}_h \times \mathbf{U}_h \times \widehat{\mathbf{U}}_h$. The following boundary-associated functional is contained in the output functional $\mathcal{J}(\mathcal{Z}_h)$ to account for the boundary conditions:

$$(a.5) \quad \mathcal{J}^{boundary}(\widehat{\mathbf{u}}_h) = \mathcal{J}^{\partial \Omega_N \cup \partial \Omega_R}(\widehat{\mathbf{u}}_h) + \mathcal{J}^{\partial \Omega_D}(\widehat{\mathbf{u}}_h),$$

where

$$(a.6) \quad \mathcal{J}^{\partial \Omega_N \cup \partial \Omega_R}(\widehat{\mathbf{u}}_h) = -\left\langle \widehat{\mathbf{u}}_h, \mathbf{g}_N^{adjoint} \right\rangle_{\mathcal{E}_h^{\partial} \cap \partial \Omega_N} - \left\langle \widehat{\mathbf{u}}_h, \mathbf{g}_R^{adjoint} \right\rangle_{\mathcal{E}_h^{\partial} \cap \partial \Omega_R}, \text{ and,} \\ \mathcal{J}^{\partial \Omega_D}(\widehat{\mathbf{u}}_h) = -\left\langle \widehat{\mathbf{u}}_h, \mathbf{g}_D^{adjoint} \right\rangle_{\mathcal{E}_h^{\partial} \cap \partial \Omega_D},$$

in which $\mathbf{g}_N^{adjoint} : \partial \Omega_N \mapsto \mathbb{R}$, $\mathbf{g}_R^{adjoint} : \partial \Omega_R \mapsto \mathbb{R}$, and $\mathbf{g}_D^{adjoint} : \partial \Omega_D \mapsto \mathbb{R}$. As indicated in Section 5, the homogeneous conditions are considered in this paper (i.e., $\mathbf{g}_D^{adjoint} = \mathbf{g}_N^{adjoint} = \mathbf{g}_R^{adjoint} = \mathbf{0}$). A similar analysis used in Lemma 3.7 and Theorem 3.8 can still be applied to the adjoint hp -HDG formulation (a.4) whose primal formulation (3.34) has full coercivity. Likewise, a similar argument presented in Lemma 3.9 and Theorem 3.10 can be used for the same adjoint formulation (a.4) whose primal formulation (3.34) only has partial coercivity.

LEMMA A.3 (Well-posedness of the local equation-with full coercivity). *The local solver composed by (a.4a) and (a.4b) is well-posed provided that:*

1. the assumptions (A.1)-(A.6) hold, and
2. $\frac{1}{2}C + T \geq 0$, and
3. B_k is a constant and is non-zero for $k = 1, \dots, d$.

By being well-posed, we mean that given $\widehat{\mathbf{v}}_h$, $\delta \mathcal{J}_{\boldsymbol{\sigma}_h}^{adjoint}(\mathcal{Z}_h; \delta \boldsymbol{\sigma}_h)$, and $\delta \mathcal{J}_{\mathbf{u}_h}^{adjoint}(\mathcal{Z}_h; \delta \mathbf{u}_h)$, there exists a unique solution $(\mathbf{s}_h, \mathbf{v}_h)$ of the local solver.

THEOREM A.4 (Well-posedness of the adjoint hp -HDG formulation-with full coercivity). *Suppose:*

1. the assumptions (A.1)-(A.6) and (2.3b) hold, and
2. $\frac{1}{2}C + T \geq 0$, and
3. B_k is constant and is nonzero for $k = 1, \dots, d$, and
4. $\bigcap_{k=1}^d \text{Range}(B_k) = \{\mathbf{0}\}$ and $\mathcal{N}(B_k) = \{\mathbf{0}\}$ for $\forall k = 1, \dots, d$.

Then, the adjoint hp -HDG formulation stated in (a.4) is well-posed in the sense that given

$\delta \mathcal{J}_{\boldsymbol{\sigma}_h}^{adjoint}(\mathcal{Z}_h; \delta \boldsymbol{\sigma}_h)$, $\delta \mathcal{J}_{\mathbf{u}_h}^{adjoint}(\mathcal{Z}_h; \delta \mathbf{u}_h)$, $\delta \mathcal{J}_{\widehat{\mathbf{u}}_h}^{adjoint}(\mathcal{Z}_h; \delta \widehat{\mathbf{u}}_h)$ and $\delta \mathcal{J}_{\widehat{\mathbf{u}}_h}^{boundary}(\widehat{\mathbf{u}}_h; \delta \widehat{\mathbf{u}}_h)$, there exists a unique solution $(\mathbf{s}_h, \mathbf{v}_h, \widehat{\mathbf{v}}_h)$.

LEMMA A.5 (Well-posedness of the local equation-with partial coercivity). *The local solver composed by (a.4a) and (a.4b) is well-posed provided that:*

1. the assumption (A.1)-(A.3), (A4.a)-(A4.b) and (A.5)-(A.6) hold, and
2. $\frac{1}{2}C + T > 0$, and
3. $\bigcap_{k=1}^d \text{Range}(B_k) = \{\mathbf{0}\}$ and $\mathcal{N}(B_k) = \{\mathbf{0}\}$ for $\forall k = 1, \dots, d$.

By being well-posed, we mean that given $\widehat{\mathbf{v}}_h$, $\delta \mathcal{J}_{\sigma_h}^{\text{adjoint}}(\mathcal{Z}_h; \delta \sigma_h)$, and $\delta \mathcal{J}_{\mathbf{u}_h}^{\text{adjoint}}(\mathcal{Z}_h; \delta \mathbf{u}_h)$, there exists a unique solution $(\mathbf{s}_h, \mathbf{v}_h)$ of the local solver.

THEOREM A.6 (Well-posedness of the adjoint hp -HDG formulation -with partial coercivity). *Suppose:*

1. the assumptions (A.1)-(A.3), (A4.a)-(A4.b), (A.5)-(A.6), and (2.3b) hold, and

2. $\frac{1}{2}C + T > 0$,

3. $\bigcap_{k=1}^d \text{Range}(B_k) = \{\mathbf{0}\}$ and $\mathcal{N}(B_k) = \{\mathbf{0}\}$ for $\forall k = 1, \dots, d$.

Then, the adjoint hp -HDG formulation stated in (a.4) is well-posed in the sense that given

$\delta \mathcal{J}_{\sigma_h}^{\text{adjoint}}(\mathcal{Z}_h; \delta \sigma_h)$, $\delta \mathcal{J}_{\mathbf{u}_h}^{\text{adjoint}}(\mathcal{Z}_h; \delta \mathbf{u}_h)$, $\delta \mathcal{J}_{\widehat{\mathbf{u}}_h}^{\text{adjoint}}(\mathcal{Z}_h; \delta \widehat{\mathbf{u}}_h)$ and $\delta \mathcal{J}_{\widehat{\mathbf{u}}_h}^{\text{boundary}}(\widehat{\mathbf{u}}_h; \delta \widehat{\mathbf{u}}_h)$, there exists a unique solution $(\mathbf{s}_h, \mathbf{v}_h, \widehat{\mathbf{v}}_h)$.

As can be seen, the assumptions needed for the well-posedness of primal formulations are the same as the corresponding adjoint formulation. In fact, this observation holds true for all HDG methods. To see it, we can express both volume and trace unknowns as discrete vectors (i.e., each element in the vectors represents a nodal value) instead of functions. The system of primal equations can then be rewritten in the matrix form and the transpose of the matrix is exactly the matrix in the corresponding system of adjoint equations [40, 105]. It is easy to see that the transpose of a square matrix is indeed invertible if the original square matrix is invertible.

Appendix B. Proof of the existence of the upwind flux in reduced form. We first look at the upwind flux stated in (3.23). Thanks to the assumption (A.5), the submatrix $A_k^{\sigma\sigma}$ contribute nothing and hence we have

$$(b.1) \quad \mathbf{F}(\sigma_h^*, \mathbf{u}_h^*) \mathbf{n} = \begin{bmatrix} B\mathbf{u}_h^* \\ B^T \sigma_h^* + C\mathbf{u}_h^* \end{bmatrix}.$$

By the equality (3.22), we then have⁹

$$(b.2) \quad \mathbf{F}^*(\sigma_h^-, \mathbf{u}_h^-, \sigma_h^*, \mathbf{u}_h^*) \mathbf{n} \underbrace{=}_{(3.22)} \mathbf{F}(\sigma_h^*, \mathbf{u}_h^*) \mathbf{n} \underbrace{=}_{(b.1)} \begin{bmatrix} B\mathbf{u}_h^* \\ B^T \sigma_h^* + C\mathbf{u}_h^* \end{bmatrix}.$$

Further, the numerical flux $\mathbf{F}^*(\sigma_h^-, \mathbf{u}_h^-, \sigma_h^*, \mathbf{u}_h^*) \mathbf{n}$ also satisfies Eq. (3.21) and thus we have flexibility in replacing one of components of the numerical flux. Given that the upwind state $B\mathbf{u}_h^*$ is desired to be kept in the numerical flux, we replace the second component in $\mathbf{F}^*$ using Eq. (3.21) and arrive at

$$(b.3) \quad \mathbf{F}^*(\sigma_h^-, \mathbf{u}_h^-, \sigma_h^*, \mathbf{u}_h^*) \mathbf{n} = \begin{bmatrix} B\mathbf{u}_h^* \\ B^T \sigma_h + C\mathbf{u}_h + \mathcal{A}^{\sigma\sigma}(\sigma_h - \sigma_h^*) + \mathcal{A}^{\mathbf{u}\mathbf{u}}(\mathbf{u}_h - \mathbf{u}_h^*) \end{bmatrix}.$$

The goal is to eliminate the state σ_h^* from the right-hand side of Eq. (b.3) via either the assumption (F.1) or (F.2).

We first consider that (F.1) holds true. Since the second row in the equality (3.22) is already used to rewrite (b.2) as (b.3) and the first row remains unused, we can take advantage of this observation to obtain

$$(b.4) \quad B\mathbf{u}_h^* = B\mathbf{u}_h + \mathcal{A}^{\sigma\sigma}(\sigma_h - \sigma_h^*) + \mathcal{A}^{\mathbf{u}\mathbf{u}}(\mathbf{u}_h - \mathbf{u}_h^*).$$

By invoking assumptions (F1.a) and (F1.b), Eq. (b.4) can be rearranged as

$$(b.5) \quad \mathcal{A}^{\sigma\sigma}(\sigma_h - \sigma_h^*) = -(\Phi^T B^T B \Phi)^{-1} \Phi^T B^T B (\Psi + I_{m^u}),$$

where the matrix $(\Phi^T B^T B \Phi)^{-1}$ is guaranteed to exist owing to the assumptions (F1.a) (Φ^{-1} exists) and (F1.c) (implies that $\mathcal{N}(B) = \{\mathbf{0}\}$). Now substitute (b.5) into (b.3) and define T in the way described in (3.25), we then can arrive at the formulation (3.24).

Now assume that (F.2) holds true, it is obvious that the state σ_h^* will vanish. By applying the definition (3.26), the formation (3.24) can then be obtained.

⁹As mentioned in Footnote 6, it should be noted that both of upwind states σ_h^* and $\mathbf{u}_h^*$ are the function of states from adjacent elements. That is, $\sigma_h^* = \sigma_h^*(\sigma_h^-, \mathbf{u}_h^-, \sigma_h^+, \mathbf{u}_h^+)$ and $\mathbf{u}_h^* = \mathbf{u}_h^*(\sigma_h^-, \mathbf{u}_h^-, \sigma_h^+, \mathbf{u}_h^+)$.

REFERENCES

- [1] G. ANAGNOSTOU, A. PATERA, Y. MADAY, AND C. MAVRIPLIS, *On the mortar element method- generalizations and implementation*, in International Symposium on Domain Decomposition Methods for Partial Differential Equations, 3 rd, Houston, TX, 1990, pp. 157–173.
- [2] R. ARAYA, M. SOLANO, AND P. VEGA, *Analysis of an adaptive HDG method for the Brinkman problem*, IMA Journal of Numerical Analysis, 39 (2019), pp. 1502–1528. ISBN: 0272-4979 Publisher: Oxford University Press.
- [3] T. ARBOGAST, L. C. COWSAR, M. F. WHEELER, AND I. YOTOV, *Mixed finite element methods on nonmatching multiblock grids*, SIAM Journal on Numerical Analysis, 37 (2000), pp. 1295–1315. ISBN: 0036-1429 Publisher: SIAM.
- [4] I. BABUŠKA AND M. R. DORR, *Error estimates for the combined h and p versions of the finite element method*, Numerische Mathematik, 37 (1981), pp. 257–277. ISBN: 0029-599X Publisher: Springer.
- [5] I. BABUŠKA AND M. SURI, *The optimal convergence rate of the p -version of the finite element method*, SIAM Journal on Numerical Analysis, 24 (1987), pp. 750–776. ISBN: 0036-1429 Publisher: SIAM.
- [6] I. BABUŠKA AND M. SURI, *The p - and hp versions of the finite element method, an overview*, Computer methods in applied mechanics and engineering, 80 (1990), pp. 5–26. ISBN: 0045-7825 Publisher: Elsevier.
- [7] Y. BAI AND K. J. FIDKOWSKI, *Continuous artificial-viscosity shock capturing for hybrid discontinuous Galerkin on adapted meshes*, AIAA Journal, 60 (2022), pp. 5678–5691. ISBN: 1533-385X Publisher: American Institute of Aeronautics and Astronautics.
- [8] A. BALAN, M. WOOPEN, AND G. MAY, *Adjoint-based hp -adaptation for a class of high-order hybridized finite element schemes for compressible flows*, in 21st AIAA Computational Fluid Dynamics Conference, 2013, p. 2938.
- [9] A. BALAN, M. WOOPEN, AND G. MAY, *Adjoint-based hp -adaptivity on anisotropic meshes for high-order compressible flow simulations*, Computers & Fluids, 139 (2016), pp. 47–67. ISBN: 0045-7930 Publisher: Elsevier.
- [10] R. E. BANK, A. H. SHERMAN, AND A. WEISER, *Some refinement algorithms and data structures for regular local mesh refinement*, Scientific Computing, Applications of Mathematics and Computing to the Physical Sciences, 1 (1983), pp. 3–17. Publisher: by R. Stepleman et al. IMACS. North-Holland.
- [11] T. J. BARTH, M. GRIEBEL, D. E. KEYES, R. M. NIEMINEN, D. ROOSE, AND T. SCHLICK, *Lecture notes in computational science and engineering*, vol. 49, Springer, 2005.
- [12] O. BARTOŠ, V. DOLEJŠÍ, G. MAY, A. RANGARAJAN, AND F. ROSKOVEC, *A goal-oriented anisotropic hp -mesh adaptation method for linear convection–diffusion–reaction problems*, Computers & Mathematics with Applications, 78 (2019), pp. 2973–2993. ISBN: 0898-1221 Publisher: Elsevier.
- [13] R. BECKER AND R. RANNACHER, *An optimal control approach to a posteriori error estimation in finite element methods*, Acta numerica, 10 (2001), pp. 1–102. ISBN: 1474-0508 Publisher: Cambridge University Press.
- [14] F. B. BELGACEM, *The mortar finite element method with Lagrange multipliers*, Numerische Mathematik, 84 (1999), pp. 173–197. ISBN: 0029-599X Publisher: Springer.
- [15] C. BERNARDI, *A new nonconforming approach to domain decomposition: the mortar element method*, Technical Report, Universite Pierre at Marie Curie, (1990).
- [16] C. BERNARDI, Y. MADAY, AND A. T. PATERA, *Domain decomposition by the mortar element method*, in Asymptotic and numerical methods for partial differential equations with critical parameters, Springer, 1993, pp. 269–286.
- [17] C. BERNARDI, Y. MADAY, AND F. RAPETTI, *Basics and some applications of the mortar element method*, GAMM-Mitteilungen, 28 (2005), pp. 97–123. ISBN: 0936-7195 Publisher: Wiley Online Library.
- [18] S. C. BRENNER, *Poincaré–Friedrichs inequalities for piecewise h^1 functions*, SIAM Journal on Numerical Analysis, 41 (2003), pp. 306–324. ISBN: 0036-1429 Publisher: SIAM.
- [19] S. C. BRENNER, *Korn's inequalities for piecewise h^1 vector fields*, Math. Comput., 73 (2004), pp. 1067–1087.
- [20] T. BUI-THANH, *From Godunov to a unified hybridized discontinuous Galerkin framework for partial differential equations*, Journal of Computational Physics, 295 (2015), pp. 114–146. ISBN: 0021-9991 Publisher: Elsevier.
- [21] T. BUI-THANH, L. DEMKOWICZ, AND O. GHATTAS, *A unified discontinuous Petrov–Galerkin method and its analysis for Friedrichs' systems*, SIAM Journal on Numerical Analysis, 51 (2013), pp. 1933–1958. ISBN: 0036-1429 Publisher: SIAM.
- [22] T. BUI-THANH AND O. GHATTAS, *Analysis of an hp -nonconforming discontinuous Galerkin spectral element method for wave propagation*, SIAM Journal on Numerical Analysis, 50 (2012), pp. 1801–1826. ISBN: 0036-1429 Publisher: SIAM.
- [23] C. CARSTENSEN, R. H. HOPPE, N. SHARMA, AND T. WARBURTON, *Adaptive hybridized interior penalty discontinuous Galerkin methods for $h(\text{curl})$ -elliptic problems*, Numerical Mathematics: Theory, Methods and Applications, 4 (2011), pp. 13–37. ISBN: 1004-8979 Publisher: Cambridge University Press.
- [24] A. CEMELIOGLU, B. COCKBURN, AND W. QIU, *Analysis of a hybridizable discontinuous Galerkin method for the steady-state incompressible Navier-Stokes equations*, Mathematics of Computation, 86 (2017), pp. 1643–1670. ISBN: 0025-5718.
- [25] M. CEZE AND K. J. FIDKOWSKI, *Anisotropic hp -adaptation framework for functional prediction*, AIAA journal, 51 (2013), pp. 492–509. ISBN: 0001-1452 Publisher: American Institute of Aeronautics and Astronautics.
- [26] J. CHAN, N. HEUER, T. BUI-THANH, AND L. DEMKOWICZ, *A robust DPG method for convection-dominated diffusion problems II: Adjoint boundary conditions and mesh-dependent test norms*, Computers & Mathematics with Applications, 67 (2014), pp. 771–795. ISBN: 0898-1221 Publisher: Elsevier.
- [27] H. CHEN, J. LI, AND W. QIU, *Robust a posteriori error estimates for HDG method for convection-diffusion equations*, IMA Journal of Numerical Analysis, 36 (2016), pp. 437–462. ISBN: 1464-3642 Publisher: Oxford University Press.
- [28] Y. CHEN AND B. COCKBURN, *Analysis of variable-degree HDG methods for convection–diffusion equations. part i: general nonconforming meshes*, IMA journal of numerical analysis, 32 (2012), pp. 1267–1293. ISBN: 1464-3642 Publisher: Oxford University Press.
- [29] Y. CHEN AND B. COCKBURN, *Analysis of variable-degree HDG methods for convection-diffusion equations. part II:*

- Semimatching nonconforming meshes*, Mathematics of Computation, 83 (2014), pp. 87–111. ISBN: 0025-5718.
- [30] B. COCKBURN, *Static condensation, hybridization, and the devising of the HDG methods*, Building bridges: connections and challenges in modern approaches to numerical partial differential equations, (2016), pp. 129–177. ISBN: 3319416383 Publisher: Springer.
- [31] B. COCKBURN AND J. GOPALAKRISHNAN, *The derivation of hybridizable discontinuous Galerkin methods for Stokes flow*, SIAM Journal on Numerical Analysis, 47 (2009), pp. 1092–1125. Publisher: Society for Industrial and Applied Mathematics.
- [32] 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), pp. 1319–1365. Publisher: Society for Industrial and Applied Mathematics.
- [33] B. COCKBURN, J. GOPALAKRISHNAN, N. NGUYEN, J. PERAIRE, AND F.-J. SAYAS, *Analysis of HDG methods for Stokes flow*, Mathematics of Computation, 80 (2011), pp. 723–760. ISBN: 0025-5718.
- [34] B. COCKBURN, J. GOPALAKRISHNAN, AND F.-J. SAYAS, *A projection-based error analysis of HDG methods*, Mathematics of Computation, 79 (2010), pp. 1351–1367. ISBN: 0025-5718.
- [35] B. COCKBURN, R. NOCHETTO, AND W. ZHANG, *Contraction property of adaptive hybridizable discontinuous Galerkin methods*, Mathematics of Computation, 85 (2016), pp. 1113–1141. ISBN: 0025-5718.
- [36] B. COCKBURN, W. QIU, AND K. SHI, *Conditions for superconvergence of HDG methods for second-order elliptic problems*, Mathematics of Computation, 81 (2012), pp. 1327–1353. ISBN: 0025-5718.
- [37] B. COCKBURN AND F.-J. SAYAS, *Divergence-conforming HDG methods for Stokes flows*, Mathematics of Computation, 83 (2014), pp. 1571–1598. ISBN: 0025-5718.
- [38] B. COCKBURN AND S. XIA, *An adjoint-based adaptive error approximation of functionals by the hybridizable discontinuous Galerkin method for second-order elliptic equations*, Journal of Computational Physics, 457 (2022), p. 111078. ISBN: 0021-9991 Publisher: Elsevier.
- [39] J. CUI AND W. ZHANG, *An analysis of HDG methods for the Helmholtz equation*, IMA Journal of Numerical Analysis, 34 (2014), pp. 279–295. ISBN: 1464-3642 Publisher: Oxford University Press.
- [40] J. P. DAHM AND K. FIDKOWSKI, *Error estimation and adaptation in hybridized discontinuous Galerkin methods*, in 52nd Aerospace Sciences Meeting, 2014, p. 0078.
- [41] L. DEMKOWICZ, *Computing with hp Finite Elements. I. One- and Two-Dimensional Elliptic and Maxwell Problems*, CRC press, 2006.
- [42] V. DOLEJŠÍ, *Anisotropic mesh adaptation for finite volume and finite element methods on triangular meshes*, Computing and Visualization in Science, 1 (1998), pp. 165–178. ISBN: 1432-9360 Publisher: Springer.
- [43] V. DOLEJŠÍ, *hp-DGFEM for nonlinear convection-diffusion problems*, Mathematics and Computers in Simulation, 87 (2013), pp. 87–118. ISBN: 0378-4754 Publisher: Elsevier.
- [44] V. DOLEJŠÍ AND M. FEISTAUER, *Discontinuous Galerkin Method: Analysis and Applications to Compressible Flow*, vol. 48, Springer, 2015.
- [45] T. A. DRISCOLL, *Algorithm 756: A MATLAB toolbox for Schwarz-Christoffel mapping*, ACM Transactions on Mathematical Software (TOMS), 22 (1996), pp. 168–186. ISBN: 0098-3500 Publisher: ACM New York, NY, USA.
- [46] H. EGGER AND C. WALUGA, *hp analysis of a hybrid DG method for Stokes flow*, IMA Journal of Numerical Analysis, 33 (2013), pp. 687–721.
- [47] K. ERIKSSON, D. ESTEP, P. HANSBO, AND C. JOHNSON, *Introduction to adaptive methods for differential equations*, Acta numerica, 4 (1995), pp. 105–158. ISBN: 1474-0508 Publisher: Cambridge University Press.
- [48] K. ERIKSSON AND C. JOHNSON, *Adaptive streamline diffusion finite element methods for stationary convection-diffusion problems*, mathematics of computation, 60 (1993), pp. 167–188. ISBN: 0025-5718.
- [49] A. ERN AND J.-L. GUERMOND, *Discontinuous Galerkin methods for Friedrichs’ systems. Part I. general theory*, SIAM journal on numerical analysis, 44 (2006), pp. 753–778. ISBN: 0036-1429 Publisher: SIAM.
- [50] A. ERN AND J.-L. GUERMOND, *Discontinuous Galerkin methods for Friedrichs’ systems. Part II. second-order elliptic PDEs*, SIAM journal on numerical analysis, 44 (2006), pp. 2363–2388. ISBN: 0036-1429 Publisher: SIAM.
- [51] A. ERN AND J.-L. GUERMOND, *Discontinuous Galerkin methods for Friedrichs’ systems. Part III. multifield theories with partial coercivity*, SIAM journal on numerical analysis, 46 (2008), pp. 776–804. ISBN: 0036-1429 Publisher: SIAM.
- [52] K. J. FIDKOWSKI AND G. CHEN, *Output-based mesh optimization for hybridized and embedded discontinuous Galerkin methods*, International Journal for Numerical Methods in Engineering, 121 (2020), pp. 867–887. ISBN: 0029-5981 Publisher: Wiley Online Library.
- [53] L. FRIEDRICH, A. R. WINTERS, D. C. DEL REY FERNÁNDEZ, G. J. GASSNER, M. PARSANI, AND M. H. CARPENTER, *An entropy stable h/p non-conforming discontinuous Galerkin method with the summation-by-parts property*, Journal of Scientific Computing, 77 (2018), pp. 689–725. ISBN: 0885-7474 Publisher: Springer.
- [54] K. O. FRIEDRICHS, *Symmetric positive linear differential equations*, Communications on Pure and Applied Mathematics, 11 (1958), pp. 333–418. ISBN: 0010-3640 Publisher: Wiley Online Library.
- [55] G. FU, W. QIU, AND W. ZHANG, *An analysis of HDG methods for convection-dominated diffusion problems*, ESAIM: Mathematical Modelling and Numerical Analysis, 49 (2015), pp. 225–256. ISBN: 0764-583X Publisher: EDP Sciences.
- [56] M. GIACOMINI AND R. SEVILLA, *Discontinuous Galerkin approximations in computational mechanics: hybridization, exact geometry and degree adaptivity*, SN Applied Sciences, 1 (2019), p. 1047. ISBN: 2523-3963 Publisher: Springer.
- [57] G. GIORGIANI, S. FERNÁNDEZ-MÉNDEZ, AND A. HUERTA, *Hybridizable discontinuous Galerkin with degree adaptivity for the incompressible Navier–Stokes equations*, Computers & Fluids, 98 (2014), pp. 196–208. ISBN: 0045-7930 Publisher: Elsevier.
- [58] G. GIORGIANI, S. FERNÁNDEZ-MÉNDEZ, AND A. HUERTA, *Hybridizable discontinuous Galerkin p-adaptivity for wave propagation problems*, International Journal for Numerical Methods in Fluids, 72 (2013), pp. 1244–1262. ISBN: 0271-2091 Publisher: Wiley Online Library.

- [59] S. K. GODUNOV AND I. BOHACHEVSKY, *Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics*, Matematičeskij sbornik, 47 (1959), pp. 271–306.
- [60] D. GOTTLIEB AND C.-W. SHU, *On the Gibbs phenomenon and its resolution*, SIAM review, 39 (1997), pp. 644–668. ISBN: 0036-1445 Publisher: SIAM.
- [61] R. GRIESMAIER AND P. MONK, *Error analysis for a hybridizable discontinuous Galerkin method for the Helmholtz equation*, Journal of Scientific Computing, 49 (2011), pp. 291–310. ISBN: 0885-7474 Publisher: Springer.
- [62] B. GUO AND I. BABUŠKA, *The hp version of the finite element method: Part 1: The basic approximation results*, Computational Mechanics, 1 (1986), pp. 21–41. ISBN: 0178-7675 Publisher: Springer.
- [63] B. GUO AND I. BABUŠKA, *The hp version of the finite element method: Part 2: General results and applications*, Computational Mechanics, 1 (1986), pp. 203–220. ISBN: 0178-7675 Publisher: Springer.
- [64] J. S. HESTHAVEN AND T. WARBURTON, *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*, Springer Science & Business Media, 2007.
- [65] J. M. HOERMANN, C. BERTOGLIO, M. KRONBICHLER, M. R. PFALLER, R. CHABINIOK, AND W. A. WALL, *An adaptive hybridizable discontinuous Galerkin approach for cardiac electrophysiology*, International Journal for Numerical Methods in Biomedical Engineering, 34 (2018), p. e2959.
- [66] A. HUERTA, A. RODRÍGUEZ-FERRAN, P. DíEZ, AND J. SARRATE, *Adaptive finite element strategies based on error assessment*, International Journal for Numerical Methods in Engineering, 46 (1999), pp. 1803–1818. ISBN: 0029-5981 Publisher: Wiley Online Library.
- [67] M. JENSEN, *Discontinuous Galerkin methods for Friedrichs' systems with irregular solutions*, PhD thesis, University of Oxford, 2005.
- [68] S. KANG, T. BUI-THANH, AND T. ARBOGAST, *A hybridized discontinuous Galerkin method for a linear degenerate elliptic equation arising from two-phase mixtures*, Computer Methods in Applied Mechanics and Engineering, 350 (2019), pp. 315–336. ISBN: 0045-7825 Publisher: Elsevier.
- [69] G. KARNIADAKIS AND S. J. SHERWIN, *Spectral/hp element methods for computational fluid dynamics*, Oxford University Press, USA, 2005.
- [70] R. M. KIRBY, S. J. SHERWIN, AND B. COCKBURN, *To CG or to HDG: A comparative study*, J Sci Comput, 51 (2012), pp. 183–212.
- [71] D. A. KOPRIVA AND J. H. KOLIAS, *A conservative staggered-grid Chebyshev multidomain method for compressible flows*, Journal of computational physics, 125 (1996), pp. 244–261. ISBN: 0021-9991 Publisher: Elsevier.
- [72] D. A. KOPRIVA, S. L. WOODRUFF, AND M. Y. HUSSAINI, *Computation of electromagnetic scattering with a non-conforming discontinuous spectral element method*, International journal for numerical methods in engineering, 53 (2002), pp. 105–122. ISBN: 0029-5981 Publisher: Wiley Online Library.
- [73] J. E. KOZDON AND L. C. WILCOX, *An energy stable approach for discretizing hyperbolic equations with nonconforming discontinuous Galerkin methods*, Journal of Scientific Computing, 76 (2018), pp. 1742–1784. ISBN: 0885-7474 Publisher: Springer.
- [74] J. J. LEE, S. J. SHANNON, T. BUI-THANH, AND J. N. SHADID, *Analysis of an HDG method for linearized incompressible resistive MHD equations*, SIAM Journal on Numerical Analysis, 57 (2019), pp. 1697–1722. ISBN: 0036-1429 Publisher: SIAM.
- [75] H. LENG, *Adaptive HDG methods for the steady-state incompressible Navier–Stokes equations*, Journal of Scientific Computing, 87 (2021), p. 37. ISBN: 0885-7474 Publisher: Springer.
- [76] H. LENG AND H. CHEN, *Adaptive HDG methods for the Brinkman equations with application to optimal control*, Journal of Scientific Computing, 87 (2021), p. 46. ISBN: 0885-7474 Publisher: Springer.
- [77] H. LENG AND Y. CHEN, *Adaptive hybridizable discontinuous Galerkin methods for nonstationary convection diffusion problems*, Advances in Computational Mathematics, 46 (2020), pp. 1–23. ISBN: 1019-7168 Publisher: Springer.
- [78] T. LEVÝ AND G. MAY, *Conservative solution transfer between anisotropic meshes for adaptive time-accurate hybridized discontinuous Galerkin methods*, in AIAA SCITECH 2023 Forum, 2023, p. 1794.
- [79] L. LI, S. LANTERI, AND R. PERRUSSEL, *A hybridizable discontinuous Galerkin method combined to a Schwarz algorithm for the solution of 3d time-harmonic Maxwell's equation*, Journal of Computational Physics, 256 (2014), pp. 563–581. ISBN: 0021-9991 Publisher: Elsevier.
- [80] X. LI AND W. HUANG, *An anisotropic mesh adaptation method for the finite element solution of heterogeneous anisotropic diffusion problems*, Journal of Computational Physics, 229 (2010), pp. 8072–8094. ISBN: 0021-9991 Publisher: Elsevier.
- [81] Y. MADAY, C. MAVRIPLIS, AND A. T. PATERA, *Nonconforming mortar element methods: Application to spectral discretizations*, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, 1988.
- [82] G. MAY, K. DEVESE, A. RANGARAJAN, AND T. MAGIN, *A hybridized discontinuous Galerkin solver for high-speed compressible flow*, Aerospace, 8 (2021), p. 322. ISBN: 2226-4310 Publisher: MDPI.
- [83] A. MUIXI, A. RODRÍGUEZ-FERRAN, AND S. FERNÁNDEZ-MÉNDEZ, *A hybridizable discontinuous Galerkin phase-field model for brittle fracture with adaptive refinement*, International Journal for Numerical Methods in Engineering, 121 (2020), pp. 1147–1169. ISBN: 0029-5981 Publisher: Wiley Online Library.
- [84] S. MURALIKRISHNAN, M.-B. TRAN, AND T. BUI-THANH, *iHDG: An iterative HDG framework for partial differential equations*, SIAM Journal on Scientific Computing, 39 (2017), pp. S782–S808. ISBN: 1064-8275 Publisher: SIAM.
- [85] N. C. NGUYEN AND J. PERAIRE, *An adaptive shock-capturing HDG method for compressible flows*, in 20th AIAA Computational Fluid Dynamics Conference, 2011, p. 3060.
- [86] N. C. NGUYEN, J. PERAIRE, AND B. COCKBURN, *An implicit high-order hybridizable discontinuous Galerkin method for linear convection-diffusion equations*, Journal of Computational Physics, 228 (2009), pp. 3232–3254.
- [87] N. C. NGUYEN, J. PERAIRE, AND B. COCKBURN, *An implicit high-order hybridizable discontinuous Galerkin method for nonlinear convection-diffusion equations*, Journal of Computational Physics, 228 (2009), pp. 8841–8855. ISBN: 0021-9991 Publisher: Elsevier.

- [88] N. C. NGUYEN, J. PERAIRE, AND B. COCKBURN, *A hybridizable discontinuous Galerkin method for Stokes flow*, Computer Methods in Applied Mechanics and Engineering, 199 (2010), pp. 582–597. ISBN: 0045-7825 Publisher: Elsevier.
- [89] N. C. NGUYEN, J. PERAIRE, AND B. COCKBURN, *High-order implicit hybridizable discontinuous Galerkin methods for acoustics and elastodynamics*, Journal of Computational Physics, 230 (2011), pp. 3695–3718. ISBN: 0021-9991 Publisher: Elsevier.
- [90] N. C. NGUYEN, J. PERAIRE, AND B. COCKBURN, *Hybridizable discontinuous Galerkin methods for the time-harmonic Maxwell's equations*, Journal of Computational Physics, 230 (2011), pp. 7151–7175. ISBN: 0021-9991 Publisher: Elsevier.
- [91] N. C. NGUYEN, J. PERAIRE, AND B. COCKBURN, *An implicit high-order hybridizable discontinuous Galerkin method for the incompressible Navier–Stokes equations*, Journal of Computational Physics, 230 (2011), pp. 1147–1170. ISBN: 0021-9991 Publisher: Elsevier.
- [92] G. PENCHEVA AND I. YOTOV, *Balancing domain decomposition for mortar mixed finite element methods*, Numerical linear algebra with applications, 10 (2003), pp. 159–180. ISBN: 1070-5325 Publisher: Wiley Online Library.
- [93] M. C. RIVARA, *Algorithms for refining triangular grids suitable for adaptive and multigrid techniques*, International journal for numerical methods in Engineering, 20 (1984), pp. 745–756. ISBN: 0029-5981 Publisher: Wiley Online Library.
- [94] M.-C. RIVARA, *Mesh refinement processes based on the generalized bisection of simplices*, SIAM Journal on Numerical Analysis, 21 (1984), pp. 604–613. ISBN: 0036-1429 Publisher: SIAM.
- [95] E. ROTHE, *Zweidimensionale parabolische Randwertaufgaben als Granzfall eindimensionaler randwertaufgaben*, Mathematische Annalen, 102 (1930), pp. 650–670. ISBN: 0025-5831 Publisher: Springer.
- [96] A. SAMII, C. MICHOSKI, AND C. DAWSON, *A parallel and adaptive hybridized discontinuous Galerkin method for anisotropic nonhomogeneous diffusion*, Computer Methods in Applied Mechanics and Engineering, 304 (2016), pp. 118–139. ISBN: 0045-7825 Publisher: Elsevier.
- [97] R. SEVILLA AND A. HUERTA, *HDG-NEFEM with degree adaptivity for Stokes flows*, Journal of Scientific Computing, 77 (2018), pp. 1953–1980. ISBN: 0885-7474 Publisher: Springer.
- [98] J. S. SHANNON, *Hybridized Discontinuous Galerkin Methods for Magnetohydrodynamics*, PhD thesis, The University of Texas at Austin, 2018.
- [99] D. SHIN, Y. HWANG, AND E.-J. PARK, *Novel adaptive hybrid discontinuous Galerkin algorithms for elliptic problems*, Computational Methods in Applied Mathematics, 21 (2021), pp. 929–951. ISBN: 1609-4840 Publisher: De Gruyter.
- [100] R. STEVENSON, *Optimality of a standard adaptive finite element method*, Foundations of Computational Mathematics, 7 (2007), pp. 245–269. ISBN: 1615-3375 Publisher: Springer.
- [101] T. SÁNCHEZ-VIZUET, M. E. SOLANO, AND A. J. CERFON, *Adaptive hybridizable discontinuous Galerkin discretization of the Grad–Shafranov equation by extension from polygonal subdomains*, Computer Physics Communications, 255 (2020), p. 107239. ISBN: 0010-4655 Publisher: Elsevier.
- [102] E. F. TORO, *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*, Springer, 2009.
- [103] M. V. UMANSKY, M. S. DAY, AND T. D. ROGNLIEN, *On numerical solution of strongly anisotropic diffusion equation on misaligned grids*, Numerical Heat Transfer, Part B: Fundamentals, 47 (2005), pp. 533–554. ISBN: 1040-7790 Publisher: Taylor & Francis.
- [104] M. F. WHEELER AND I. YOTOV, *Multigrid on the interface for mortar mixed finite element methods for elliptic problems*, Computer methods in applied mechanics and engineering, 184 (2000), pp. 287–302. ISBN: 0045-7825 Publisher: Elsevier.
- [105] M. WOOPEN, A. BALAN, G. MAY, AND J. SCHÜTZ, *A comparison of hybridized and standard DG methods for target-based hp-adaptive simulation of compressible flow*, Computers & Fluids, 98 (2014), pp. 3–16. ISBN: 0045-7930 Publisher: Elsevier.
- [106] M. WOOPEN AND G. MAY, *An anisotropic adjoint-based hp-adaptive HDG method for compressible turbulent flow*, in 53rd AIAA Aerospace Sciences Meeting, 2015, p. 2042.
- [107] M. WOOPEN, G. MAY, AND J. SCHÜTZ, *Adjoint-based error estimation and mesh adaptation for hybridized discontinuous Galerkin methods*, International journal for numerical methods in fluids, 76 (2014), pp. 811–834. ISBN: 0271-2091 Publisher: Wiley Online Library.