SIAM J. SCI. COMPUT.
Vol. 38, No. 6, pp. A3696–A3719

CONSTRUCTION AND ANALYSIS OF HDG METHODS FOR LINEARIZED SHALLOW WATER EQUATIONS*

TAN BUI-THANH†

Abstract. We present a systematic and constructive methodology to devise various hybridized discontinuous Galerkin (HDG) methods for linearized shallow water equations. It is shown that using the Rankine–Hugoniot condition to solve the Riemann problem is a natural approach to deriving HDG methods. At the heart of our development is an upwind HDG framework obtained by hybridizing the upwind flux in the standard discontinuous Galerkin (DG) approach. Essentially, the HDG framework is a redesign of the standard DG approach to reducing the number of coupled unknowns. An upwind and three other HDG methods are constructed and analyzed for linearized shallow water systems. Rigorous stability and convergence analysis for both semidiscrete and fully discrete systems are provided. We extend the upwind HDG method to a family of penalty HDG schemes and rigorously analyze their well-posedness, stability, and convergence rates. Numerical results for the linear standing wave and the Kelvin wave for oceanic shallow water systems are presented to verify our theoretical findings.

Key words. discontinuous Galerkin methods, hybridized discontinuous Galerkin methods, upwind hybridized discontinuous Galerkin methods, Godunov method, Riemann flux, Lax–Friedrichs flux, upwind, well-posedness, stability, convergence, shallow water equation

AMS subject classifications. 65N30, 65N12, 65N15, 65N22

DOI. 10.1137/16M1057243

1. Introduction.

The discontinuous Galerkin (DG) method was originally developed by Reed and Hill [44] for the neutron transport equation, first analyzed in [35, 31], and since has been extended to other problems governed by partial differential equations (PDEs) [15]. Roughly speaking, DG combines advantages of classical finite volume and finite element methods. In particular, it provides the flexibility to deal with complex geometries, and it is highly parallelizable due to its compact stencil. In the past decade, the DG method and its various extensions [50, 36, 1, 19, 45, 34, 25, 24, 23, 49, 52] have been widespread in the geophysical fluid dynamics community. However, for steady state problems or time-dependent ones that require implicit time-integrators, DG methods typically have many more (coupled) unknowns compared to the other existing numerical methods, and are more expensive in general.

Recently, Cockburn, coauthors, and others have introduced hybridizable (also known as hybridized) discontinuous Galerkin (HDG) methods for various types of PDEs including Poisson equation [12, 14, 32], convection-diffusion equation [39, 10, 20], Stokes equation [11, 40], Euler and Navier–Stokes equations [43, 38], Maxwell equation [42, 37], acoustics and elastodynamics equations [41], Helmholtz equation [26, 17], and eigenvalue problems [16], to name a few. The beauty of the HDG method is that it reduces the number of coupled unknowns substantially while retaining all other attractive properties of the DG method. The coupled unknowns are, in fact,

*Submitted to the journal's Methods and Algorithms for Scientific Computing section January 19, 2016; accepted for publication (in revised form) September 17, 2016; published electronically November 29, 2016.

http://www.siam.org/journals/sisc/38-6/M105724.html

Funding: This work was supported by DOE grants DE-SC0010518 and DE-SC0011118.

†Department of Aerospace Engineering and Engineering Mechanics, and Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX 78712 (tanbui@ices.utexas.edu).

new unknown traces introduced on the mesh skeleton (i.e., the faces) to hybridize the DG method. Once they are solved for, the usual DG unknowns can be recovered in an element-by-element fashion, completely independent of each other. Thus, the HDG methods are well suited for current and future supercomputing systems. The HDG framework is, however, less well-known in the geophysical fluid dynamics community. We have attempted to develop HDG schemes for shallow water equations and nonhydrostatic atmospheric flows, and the results are very promising [6].

This paper is a continuation of our previous work [6]. In particular, we will, for the first time, constructively develop and analyze various HDG methods for linearized shallow water equations. This class of PDEs is important for various branches of geophysical fluid dynamics (GFD), especially for hydrostatic flows, since it carries the fastest wave, i.e., the gravity wave. It is the gravity wave that limits the time step size in explicit time integration schemes (see, e.g., [25]). For long time integration, which is not uncommon in GFD, this can lead to an excessive number of time steps, and hence substantially taxes computing and storage resources. *Thus, there is a critical need for implicit time stepping methods.* However, implicit methods could be expensive, especially for nonlinear PDEs for which Newton-like methods are often required. Semi-implicit time-integrators have been designed to balance the time step size restriction due to fast waves and computational expense due to nonlinearities (see, e.g., [48, 18, 22, 46, 47]). In particular, they typically employ implicit time stepping schemes for the linear(ized) part of the PDE under consideration that contains the fastest waves, and explicit time-integrators for the (resulting) nonlinear part for which the fastest waves are removed.

The *objective* of this paper is to develop efficient implicit hybridized discontinuous Galerkin methods that can be used as stand-alone solvers for linear(ized) shallow water equations or as implicit solvers in semi-implicit schemes for nonlinear shallow water equations.[1] Clearly, our development can be easily extended to other linear(ized) PDEs [6].

In the following we summarize the contributions of this paper. In section 2, we start by using the Rankine–Hugoniot condition to introduce a constructive methodology to systematically devise an upwind HDG method for the linear shallow water system. *Unlike most of the existing works, we derive the HDG fluxes instead of proposing them.* At the heart of our development is the hybridization of the upwind flux. This is the key step in redesigning the standard upwind DG to substantially reduce the number of coupled unknowns. We rigorously show that the upwind HDG is equivalent to the standard upwind DG, and that the trace unknowns are in fact the upwind states. By exploiting the structure of the PDE under consideration, we can reduce the number of trace unknowns to one, the minimum possible, and construct an equivalent upwind HDG method with the least number of trace unknowns. We also hybridize the Lax–Friedrichs flux, which leads to two other HDG methods. Unlike the upwind case, the most economic HDG method in this case still has two trace unknowns. Next, we provide a detailed analysis for our proposed HDG schemes in section 3. Specifically, well-posedness, stability, and convergence rates are rigorously analyzed for both semidiscrete and fully discrete forms. As shall be shown, the HDG framework allows a straightforward analysis using an energy approach. In section 4, we extend the proposed upwind HDG method to a family of penalty HDG schemes and analyze the well-posedness, stability, and convergence rates for them as well. Nu-

[1] Clearly, one can develop similar HDG methods for each Newton step when solving nonlinear PDEs, but this is out of the scope of this paper.

merical results for the linear standing wave and the linear Kelvin wave to verify our theoretical developments will be presented in section 5. Section 6 concludes the paper with a discussion on future work.

2. HDG schemes for linearized shallow water system. In this section, we show that the Rankine–Hugoniot condition is a natural approach to systematically constructing upwind HDG methods. This approach has been presented in [7] to derive HDG methods for the convection-diffusion problem. In this paper, we apply it to the linearized shallow water system. In particular, we consider the following linearized shallow water system (see, e.g., [25, 24] for a derivation from oceanic shallow water model):

$$\frac{\partial}{\partial t}\begin{pmatrix}\phi\\ \Phi u\\ \Phi v\end{pmatrix}+\frac{\partial}{\partial x}\begin{pmatrix}\Phi u\\ \Phi\phi\\ 0\end{pmatrix}+\frac{\partial}{\partial y}\begin{pmatrix}\Phi v\\ 0\\ \Phi\phi\end{pmatrix}=\begin{pmatrix}0\\ f\Phi v-\gamma\Phi u+\frac{\tau_x}{\rho}\\ -f\Phi u-\gamma\Phi v+\frac{\tau_y}{\rho}\end{pmatrix}, \tag{2.1}$$

where $\phi = gH$ is the geopotential height, with g and H being the gravitational constant and the perturbation of the free surface height, $\Phi > 0$ is a constant mean flow geopotential height (see Appendix B for a general case), $\boldsymbol{\vartheta} := (u, v)$ is the perturbed velocity, $\gamma \geq 0$ is the bottom friction, $\boldsymbol{\tau} := (\tau_x, \tau_y)$ is the wind stress, and ρ is the density of the water. Here, $f = f_0 + \beta\,(y - y_m)$ is the Coriolis parameter, where f_0, β, and y_m are given constants. Note that we have excluded bathymetry and viscosity for simplicity.

If we define $\mathbf{u} := (\phi, \Phi\boldsymbol{\vartheta})^T := (\phi, \Phi u, \Phi v)^T$, we can cast (2.1) into the first order system of PDEs as

$$\frac{\partial \mathbf{u}}{\partial t}+\sum_{k=1}^{d}\partial_k \mathbf{F}_k\left(\mathbf{u}\right) := \frac{\partial \mathbf{u}}{\partial t}+\sum_{k=1}^{d}\partial_k\left(\mathbf{A}_k\mathbf{u}\right)=\mathbf{f} \quad \text{in } \Omega, \tag{2.2}$$

where $\mathbf{f}$ is the right-hand side of (2.1). Here, the flux tensor $\mathbf{F}$ is given by $\mathbf{F}\left(\mathbf{u}\right) := \mathbf{A}\mathbf{u}$, and $\mathbf{A}$ is a tensor with two components defined as

$$\mathbf{A}_1=\begin{bmatrix}0 & 1 & 0\\ \Phi & 0 & 0\\ 0 & 0 & 0\end{bmatrix}, \quad \mathbf{A}_2=\begin{bmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ \Phi & 0 & 0\end{bmatrix}.$$

It is easy to see that $\mathcal{A} := \mathbf{A}\cdot\mathbf{n} = \mathbf{A}_1\mathbf{n}_1 + \mathbf{A}_2\mathbf{n}_2$ has three real eigenvalues

$$[c_1, c_2, c_3] := \left[-\sqrt{\Phi}, 0, \sqrt{\Phi}\right], \tag{2.3}$$

and independent eigenvectors. Thus, (2.1) is a hyperbolic system (see, e.g., [51] for a definition of hyperbolicity). As such, it can be spatially discretized using DG methods [44, 35, 31]. In the following, we first briefly discuss an upwind DG discretization and then employ the Rankine–Hugoniot condition as a means for deriving an upwind HDG method.

In order to discretize (2.2) using the discontinuous Galerkin finite element method, let us partition the domain Ω into N_{el} nonoverlapping elements $K_j, j = 1, \ldots, N_{\text{el}}$ with Lipschitz boundaries such that $\Omega_h := \cup_{j=1}^{N_{\text{el}}} K_j$ and $\overline{\Omega} = \overline{\Omega}_h$. Here, h is defined as $h := \max_{j\in\{1,\ldots,N_{\text{el}}\}} diam\left(K_j\right)$. We denote the skeleton of the mesh by $\mathcal{E}_h := \cup_{j=1}^{N_{\text{el}}}\partial K_j$; it is the set of all (uniquely defined) faces e. 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^+). Furthermore, we use $\mathbf{n}$ to denote either $\mathbf{n}^-$ or $\mathbf{n}^+$ in an expression that is valid for both cases, and this convention is also used for other quantities (restricted) on a face $e \in \mathcal{E}_h$. For the sake of convenience, we denote by $\mathcal{E}_h^\partial$ the set of all boundary faces on $\partial\Omega$ and by $\mathcal{E}_h^o := \mathcal{E}_h \setminus \mathcal{E}_h^\partial$ the set of all interior faces, and $\partial\Omega_h := \{\partial K : K \in \Omega_h\}$.

For simplicity in writing we define $(\cdot,\cdot)_K$ as the L^2-inner product on a domain $K \in \mathbb{R}^d$ and $\langle \cdot,\cdot \rangle_K$ as the L^2-inner product on a domain K if $K \in \mathbb{R}^{d-1}$. We shall use $\|\cdot\|_K := \|\cdot\|_{L^2(K)}$ as the induced norm for both cases, and the particular value of K in a context will indicate which inner product the norm is coming from. We also denote the ε-weighted norm of a function u as $\|u\|_{K,\varepsilon} := \|\sqrt{\varepsilon}u\|_K$ for any positive ε. We shall use boldface lowercase letters for vector-valued functions, in which case the inner product is defined as $(\mathbf{u},\mathbf{v})_K := \sum_{i=1}^m (\mathbf{u}_i,\mathbf{v}_i)_K$, and similarly $\langle \mathbf{u},\mathbf{v} \rangle_K := \sum_{i=1}^m \langle \mathbf{u}_i,\mathbf{v}_i \rangle_K$, where m is the number of components $(\mathbf{u}_i, i = 1,\ldots,m)$ of $\mathbf{u}$. Moreover, we define $(\mathbf{u},\mathbf{v})_\Omega := \sum_{K\in\Omega_h} (\mathbf{u},\mathbf{v})_K$ and $\langle \mathbf{u},\mathbf{v} \rangle_{\mathcal{E}_h} := \sum_{e\in\mathcal{E}_h} \langle \mathbf{u},\mathbf{v} \rangle_e$ whose induced (weighted) norms are clear, and hence their definitions are omitted. We employ boldface uppercase letters, e.g., $\mathbf{L}$, to denote matrices and tensors. In addition, subscripts are used to denote the components of vectors, matrices, and tensors.

We define $\mathcal{P}^p(K)$ as the space of polynomials of degree at most p on a domain K. Next, we introduce two discontinuous piecewise polynomial spaces

$$\mathbf{V}_h(\Omega_h) := \left\{ \mathbf{v} \in \left[L^2(\Omega_h)\right]^m : \mathbf{v}|_K \in [\mathcal{P}^p(K)]^m \ \forall K \in \Omega_h \right\},$$
$$\mathbf{\Lambda}_h(\mathcal{E}_h) := \left\{ \boldsymbol{\lambda} \in \left[L^2(\mathcal{E}_h)\right]^m : \boldsymbol{\lambda}|_e \in [\mathcal{P}^p(e)]^m \ \forall e \in \mathcal{E}_h \right\},$$

and similarly,

$$\mathbf{V}_h(K) := \{\mathbf{v} \in [\mathcal{P}^p(K)]^m\}, \quad \mathbf{\Lambda}_h(e) := \{\boldsymbol{\lambda} \in [\mathcal{P}^p(e)]^m\},$$

where m is either 2 or 3 depending on the context. For scalar-valued functions, we denote the corresponding spaces as

$$V_h(\Omega_h) := \left\{ v \in L^2(\Omega_h) : v|_K \in \mathcal{P}^p(K) \ \forall K \in \Omega_h \right\},$$
$$\Lambda_h(\mathcal{E}_h) := \left\{ \lambda \in L^2(\mathcal{E}_h) : \lambda|_e \in \mathcal{P}^p(e) \ \forall e \in \mathcal{E}_h \right\}.$$

From now on *we conventionally use* $\mathbf{u}^e$ *for the exact solution.* We would like to find a local finite element solution $\mathbf{u} \in \mathbf{V}_h(K)$ on each element $K \in \Omega_h$. To that end, multiplying (2.2) by $\mathbf{v}$ and integrating by parts, we have

$$\left(\frac{\partial \mathbf{u}}{\partial t},\mathbf{v}\right)_K - (\mathbf{F}(\mathbf{u}),\nabla\mathbf{v})_K + \left\langle \mathbf{F}^*\left(\mathbf{u}^-,\mathbf{u}^+\right)\cdot\mathbf{n},\mathbf{v}\right\rangle_{\partial K} = (\mathbf{f},\mathbf{v})_K, \tag{2.4}$$

where $\mathbf{F}^*(\mathbf{u}^-,\mathbf{u}^+)$ is a numerical flux. For simplicity in writing, we have ignored the fact that (2.4) must hold for all test functions $\mathbf{v} \in \mathbf{V}_h(K)$; throughout this paper, this should be implicitly understood.

It is the numerical flux $\mathbf{F}^*$ that couples local unknowns on elements K^+ and K^- that share a face $e \in \partial K$. Consequently, the local unknowns on all elements are coupled, and they must be solved together. In the following, we break this coupling, and mitigate the computational expense associated with standard DG methods *(for steady state problems or time-dependent problem with implicit time-integrators)*, by *hybridizing the classical Godunov/Riemann flux* (see Appendix A for a reason why HDG can be much more advantageous than DG).

The key of the hybridization is the introduction of new trace unknowns that live on the mesh skeleton. Unlike our previous work [6], *we will show that the Rankine–Hugoniot condition (see, e.g.,* [51]*) provides all the necessary ingredients for this decoupling task*. To this end, let us sketch in Figure 1 the wave structure of the Riemann problem for the linearized shallow water system (2.1) along the normal direction of the interface between K^- and K^+.

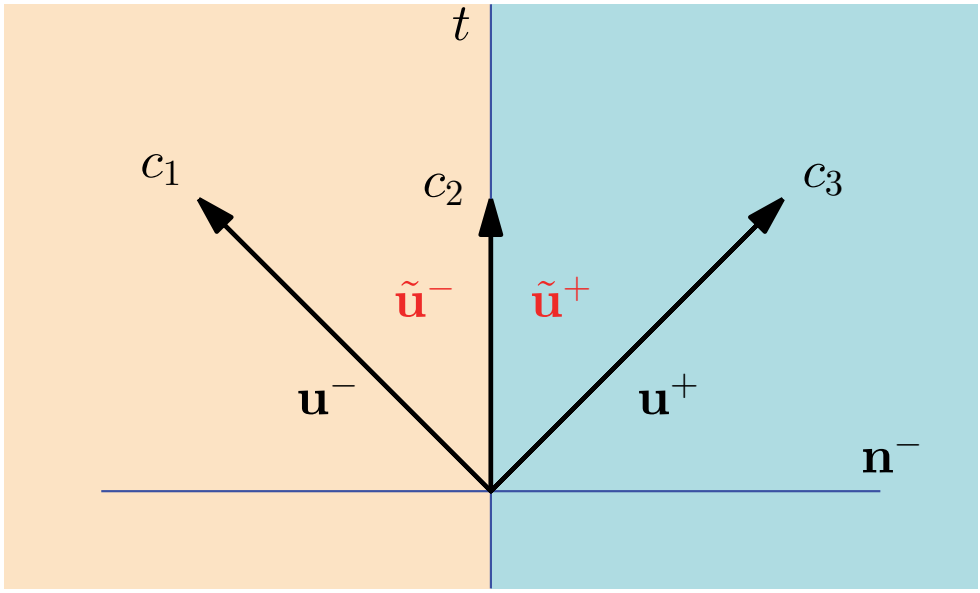


FIG. 1. *Wave structure in the Riemann problem for the linearized shallow water system* (2.1).

Now we apply the Rankine–Hugoniot condition across each wave to obtain

$$\left(\mathbf{A}^- \cdot \mathbf{n}^-\right) \tilde{\mathbf{u}}^- - \left(\mathbf{A}^- \cdot \mathbf{n}^-\right) \mathbf{u}^- = c_1 \left(\tilde{\mathbf{u}}^- - \mathbf{u}^-\right), \tag{2.5a}$$

$$\left(\mathbf{A}^+ \cdot \mathbf{n}^-\right) \tilde{\mathbf{u}}^+ - \left(\mathbf{A}^- \cdot \mathbf{n}^-\right) \tilde{\mathbf{u}}^- = 0, \tag{2.5b}$$

$$\left(\mathbf{A}^+ \cdot \mathbf{n}^-\right) \mathbf{u}^+ - \left(\mathbf{A}^+ \cdot \mathbf{n}^-\right) \tilde{\mathbf{u}}^+ = c_3 \left(\mathbf{u}^+ - \tilde{\mathbf{u}}^+\right), \tag{2.5c}$$

where $\tilde{\mathbf{u}}^\pm := \left(\tilde{\phi}^\pm, \Phi\tilde{\boldsymbol{\vartheta}}^\pm\right)$ are the states in the corresponding regions indicated in Figure 1. From the definition of $\mathbf{A}$ and the continuity of Φ we see that

$$\mathbf{A}^- = \mathbf{A}^+ = \mathbf{A} \quad \text{and} \quad (\mathbf{A} \cdot \mathbf{n})\, \mathbf{u} = [\Phi\boldsymbol{\vartheta} \cdot \mathbf{n}, \phi\Phi\mathbf{n}],$$

which, together with (2.5b), imply

$$\phi^* := \tilde{\phi}^- = \tilde{\phi}^+ \quad \text{and} \quad \Phi\boldsymbol{\vartheta}^* \cdot \mathbf{n} := \Phi\tilde{\boldsymbol{\vartheta}}^- \cdot \mathbf{n} = \Phi\tilde{\boldsymbol{\vartheta}}^+ \cdot \mathbf{n}, \tag{2.6}$$

where $\mathbf{u}^* := [\phi^*, \Phi\boldsymbol{\vartheta}^*]$ is the Riemann solution (see (2.18a) and (2.18b)). We define the "Rankine–Hugoniot" flux as

$$\mathbf{F}^{RH} \cdot \mathbf{n} := (\mathbf{A} \cdot \mathbf{n})\, \mathbf{u}^* = [\Phi\boldsymbol{\vartheta}^* \cdot \mathbf{n}, \phi^*\Phi\mathbf{n}]. \tag{2.7}$$

Using the definition of c_1, c_3 from (2.3), $\mathbf{A}$, (2.7), and (2.6), we can rewrite both (2.5a) and (2.5c) in a short form, referring to either K^- or K^+, as

$$\mathbf{F}^{RH} \cdot \mathbf{n} = \begin{bmatrix} \Phi\boldsymbol{\vartheta} \cdot \mathbf{n} + \sqrt{\Phi}\,(\phi - \phi^*) \\ \phi\Phi\mathbf{n} + \sqrt{\Phi}\,(\Phi\boldsymbol{\vartheta} - \Phi\boldsymbol{\vartheta}^*) \end{bmatrix}. \tag{2.8}$$

Since, from (2.7), the last two components of the left-hand side of (2.8) compose a vector parallel to $\mathbf{n}$, we define the upwind flux—containing the first component and the normal part of the last two components of the Rankine–Hugoniot flux (2.8)—as

$$\mathbf{F}^* \cdot \mathbf{n} = \begin{bmatrix} \Phi\boldsymbol{\vartheta} \cdot \mathbf{n} + \sqrt{\Phi}\,(\phi - \phi^*) \\ \phi\Phi\mathbf{n} + \sqrt{\Phi}\,(\Phi\boldsymbol{\vartheta} \cdot \mathbf{n} - \Phi\boldsymbol{\vartheta}^* \cdot \mathbf{n})\,\mathbf{n} \end{bmatrix}. \tag{2.9}$$

By the uniqueness of the upwind flux $\mathbf{F}^* \cdot \mathbf{n}$ across the interface, we have

$$[\![\mathbf{F}^* \cdot \mathbf{n}]\!] = \mathbf{0}, \tag{2.10}$$

where we have defined the "jump" operator as $[\![(\cdot)]\!] := (\cdot)^- + (\cdot)^+$. We also define the "average" operator as $2\,\{\!\{(\cdot)\}\!\} := [\![(\cdot)]\!]$.

We observe that the upwind flux (2.9) depends on the DG unknowns of only one side of a face $e \in \partial K$ and the single-valued solution $\mathbf{u}^*$ of the Riemann problem. This observation suggests that we can decouple the computation of $\mathbf{u}$ and $\mathbf{u}^*$ by treating the latter as the extra unknown and solve for it on the skeleton of the mesh. To signify this step, let us rename $\mathbf{u}^*$ to $\hat{\mathbf{u}}$ and $\mathbf{F}^*$ to $\hat{\mathbf{F}}$, i.e.,

$$\hat{\mathbf{F}} \cdot \mathbf{n} = \begin{bmatrix} \Phi\boldsymbol{\vartheta} \cdot \mathbf{n} + \sqrt{\Phi}\big(\phi - \hat{\phi}\big) \\ \Phi\phi\mathbf{n}_1 + \sqrt{\Phi}\big(\Phi\boldsymbol{\vartheta} - \Phi\hat{\boldsymbol{\vartheta}}\big) \cdot \mathbf{n}\mathbf{n}_1 \\ \Phi\phi\mathbf{n}_2 + \sqrt{\Phi}\big(\Phi\boldsymbol{\vartheta} - \Phi\hat{\boldsymbol{\vartheta}}\big) \cdot \mathbf{n}\mathbf{n}_2 \end{bmatrix}. \tag{2.11}$$

We see that there are three trace unknowns: two from $\hat{\boldsymbol{\vartheta}}$ and one from $\hat{\phi}$. *Let us term this HDG method with three trace unknowns as HDG-I.*

Now, as can be seen from (2.9), there are only two independent quantities, namely, the traces ϕ^* and $\boldsymbol{\vartheta}^* \cdot \mathbf{n}$, that are needed in order to compute the upwind flux. In fact, as will be shown, even these two quantities are not independent of each other, and thus either of them can be eliminated. This observation is the key leading to the most economic HDG methods with the least number of trace unknowns. In the following, we will first remove all of the dependencies, and then construct these economic HDG methods. To begin, let us present the following important result.

LEMMA 2.1. *The upwind/Riemann solution* $\mathbf{u}^* = (\phi^*, \Phi\boldsymbol{\vartheta}^*)^T$ *satisfies the following identity on any* $e \in \mathcal{E}_h^o$*:*

$$(\Phi\boldsymbol{\vartheta} - \Phi\boldsymbol{\vartheta}^*) \cdot \mathbf{n} = -\sqrt{\Phi}\,(\phi - \phi^*)\,, \tag{2.12}$$

where $\mathbf{u} = [\phi, \Phi\boldsymbol{\vartheta}]$ *is the trace of the DG solution from either side of* e*.*

Proof. The proof is simple, using the identity (2.10) for the upwind flux. Indeed, from (2.10) we have (2.18), which in turn leads to (2.12) as a direct consequence. ∎

Lemma 2.1 states that the flux of the first equation in (2.1) (mass) is opposite the one corresponding to the last two equations (momentum). As a direct consequence, the upwind flux (2.9) can be written as follows.

COROLLARY 2.2. *The upwind flux* (2.9) *is equivalent to*

$$\mathbf{F}^* \cdot \mathbf{n} = \begin{bmatrix} \Phi\boldsymbol{\vartheta} \cdot \mathbf{n} + \sqrt{\Phi}\,(\phi - \phi^*) \\ \Phi\phi^*\mathbf{n}_1 \\ \Phi\phi^*\mathbf{n}_2 \end{bmatrix}. \tag{2.13}$$

Proof. Substituting (2.12) into the last two components of upwind flux (2.9) yields the result. ∎

Unlike (2.9), the upwind flux (2.13) depends only on the first component of the upwind state, namely ϕ^*. At this point, we can follow the recipe discussed above to derive a new HDG method. To begin, we can replace "starred" quantities by "hatted" ones to define a new HDG flux, i.e.,

$$\hat{\mathbf{F}}\cdot\mathbf{n} = \begin{bmatrix} \Phi\boldsymbol{\vartheta}\cdot\mathbf{n} + \sqrt{\Phi}\left(\phi - \hat{\phi}\right) \\ \Phi\hat{\phi}\mathbf{n}_1 \\ \Phi\hat{\phi}\mathbf{n}_2 \end{bmatrix}, \tag{2.14}$$

where $\hat{\phi}$ is the single-valued unknown defined on the skeleton of the mesh.

Using either the HDG flux (2.11) or (2.14), we define semidiscretization in space for the linearized shallow water system (2.1) as follows: seek $\mathbf{u} = (\phi, \Phi\boldsymbol{\vartheta}) \in \mathbf{V}_h(K)$ and $\hat{\mathbf{u}} \in \boldsymbol{\Lambda}_h(\mathcal{E}_h^o)$ such that

$$\left(\frac{\partial \mathbf{u}}{\partial t}, \mathbf{v}\right)_K - \left(\mathbf{F}(\mathbf{u}), \nabla\mathbf{v}\right)_K + \left\langle \hat{\mathbf{F}}(\mathbf{u}, \hat{\mathbf{u}})\cdot\mathbf{n}, \mathbf{v}\right\rangle_{\partial K} = (\mathbf{f}, \mathbf{v})_K \quad \forall \mathbf{v} \in \mathbf{V}_h(K), \tag{2.15a}$$

$$\left\langle [\![\hat{\mathbf{F}}(\mathbf{u}, \hat{\mathbf{u}})\cdot\mathbf{n}]\!], \boldsymbol{\mu}\right\rangle_e = \mathbf{0} \quad \forall e \in \partial K \cap \mathcal{E}_h^o, \quad \forall \boldsymbol{\mu} \in \boldsymbol{\Lambda}_h(e), \tag{2.15b}$$

where we have enforced the uniqueness of the HDG flux weakly, which is sufficient for local conservation. *We have constructively devised an upwind HDG approach which consists of local solver* (2.15a) *and trace equation* (2.15b). For example, using the HDG flux (2.14) allows us to write the local solver (2.15a) as

$$\left(\frac{\partial \phi}{\partial t}, \varphi\right)_K - (\Phi\boldsymbol{\vartheta}, \nabla\varphi)_K + \left\langle \Phi\boldsymbol{\vartheta}\cdot\mathbf{n} + \sqrt{\Phi}\left(\phi - \hat{\phi}\right), \varphi\right\rangle_{\partial K} = 0, \tag{2.16a}$$

$$\left(\frac{\partial (\Phi u)}{\partial t}, \mathbf{w}_1\right)_K - \left(\Phi\phi, \frac{\partial \mathbf{w}_1}{\partial x}\right)_K + \left\langle \Phi\hat{\phi}\mathbf{n}_1, \mathbf{w}_1\right\rangle_{\partial K} = \left(f\Phi v - \gamma\Phi u + \frac{\tau_x}{\rho}, \mathbf{w}_1\right)_K, \tag{2.16b}$$

$$\left(\frac{\partial (\Phi v)}{\partial t}, \mathbf{w}_2\right)_K - \left(\Phi\phi, \frac{\partial \mathbf{w}_2}{\partial y}\right)_K + \left\langle \Phi\hat{\phi}\mathbf{n}_2, \mathbf{w}_2\right\rangle_{\partial K} = \left(-f\Phi u - \gamma\Phi v + \frac{\tau_y}{\rho}, \mathbf{w}_2\right)_K, \tag{2.16c}$$

and the conservation constraint (2.15b) as

$$\left\langle [\![\Phi\boldsymbol{\vartheta}\cdot\mathbf{n} + \sqrt{\Phi}(\phi - \hat{\phi})]\!], \mu\right\rangle_e = 0. \tag{2.17}$$

Here, $\varphi, \mathbf{w} = (\mathbf{w}_1, \mathbf{w}_2)$ and μ denote test functions in $V_h(\Omega_h)$ and $\Lambda_h(e)$, respectively. It is important to see that the conservation condition (2.17) is only enforced for the first component of the HDG flux (2.14) since the last two satisfy this condition automatically. *Let us term this new HDG method with a single trace unknown as HDG*-II; some of its properties are now discussed.

THEOREM 2.3. *Assuming that interior edges $e \in \mathcal{E}_h^o$ are straight lines, then the following hold for both HDG*-I *and HDG*-II *methods:*

(i) *The trace unknowns coincide with the upwind states.*

(ii) *The HDG fluxes* (2.11) *and* (2.14) *are identical to the upwind flux* (2.9).

(iii) *The upwind HDG scheme is equivalent to the standard upwind DG.*

Proof. We provide proofs for only the HDG-II method since those for the HDG-I method are similar, and hence are omitted. To show the first assertion, we first see from (2.10), namely the continuity of the Riemann flux, that

$$\begin{aligned}\Phi\boldsymbol{\vartheta}^- \cdot \mathbf{n}^- + \sqrt{\Phi}\left(\phi^- - \phi^*\right) + \Phi\boldsymbol{\vartheta}^+ \cdot \mathbf{n}^+ + \sqrt{\Phi}\left(\phi^+ - \phi^*\right) = 0,\\ \phi^-\Phi\mathbf{n}^- + \sqrt{\Phi}\left(\Phi\boldsymbol{\vartheta}^- \cdot \mathbf{n}^- - \Phi\boldsymbol{\vartheta}^* \cdot \mathbf{n}^-\right)\mathbf{n}^- + \phi^+\Phi\mathbf{n}^+\\ + \sqrt{\Phi}\left(\Phi\boldsymbol{\vartheta}^+ \cdot \mathbf{n}^+ - \Phi\boldsymbol{\vartheta}^* \cdot \mathbf{n}^+\right)\mathbf{n}^+ = 0,\end{aligned}$$

which, owing to fact that ϕ^* and $\boldsymbol{\vartheta}^*$ are single-valued, yields

$$\phi^* = \{\!\{\phi\}\!\} + \frac{\sqrt{\Phi}}{2}[\![\boldsymbol{\vartheta} \cdot \mathbf{n}]\!], \tag{2.18a}$$

$$\boldsymbol{\vartheta}^* \cdot \mathbf{n} = \{\!\{\boldsymbol{\vartheta}\}\!\} \cdot \mathbf{n} + \frac{1}{2\sqrt{\Phi}}[\![\phi]\!]. \tag{2.18b}$$

Next, we use the conservation condition (2.17) to obtain

$$\left\langle \hat{\phi}, \mu \right\rangle_e = \left\langle \{\!\{\phi\}\!\} + \frac{\sqrt{\Phi}}{2}[\![\boldsymbol{\vartheta} \cdot \mathbf{n}]\!], \mu \right\rangle_e = \langle \phi^*, \mu \rangle_e ,$$

where we have used (2.18a) in the second equality. Taking $\mu = \hat{\phi} - \phi^*$ yields

$$\left\| \hat{\phi} - \phi^* \right\|_e^2 = 0,$$

which shows that $\hat{\phi}$ coincides with the upwind state ϕ^*. The second assertion follows by substituting $\hat{\phi}$ from (2.18a) into (2.14) and using the identity (2.18b). The third assertion is a direct consequence of the first two. ∎

Let us briefly summarize what we have done up to this point. We have hybridized upwind/Riemann fluxes to construct the HDG-I scheme with three trace unknowns on the mesh skeleton. However, a closer look at the upwind flux shows that these trace unknowns are dependent. Indeed, we have shown that the number of trace unknowns can be reduced to one, which leads to the development of the most economic HDG-II method with a single scalar trace unknown. We have also shown that these two HDG schemes are equivalent to each other and to the upwind DG method. A natural question to be addressed is whether we can hybridize other DG methods. The development of upwind HDG in this section suggests that one can develop new HDG methods by simply hybridizing the DG numerical fluxes. In the following, we will carry out this idea for the Lax–Friedrichs flux and leave the hybridization of other fluxes for future work.

Following [6] we define the hybridized Lax–Friedrichs flux[2] as

$$\hat{\mathbf{F}} \cdot \mathbf{n} = \begin{bmatrix} \Phi\boldsymbol{\vartheta} \cdot \mathbf{n} + \sqrt{\Phi}\left(\phi - \hat{\phi}\right) \\ \Phi\phi\mathbf{n}_1 + \sqrt{\Phi}\left(\Phi u - \Phi\hat{u}\right) \\ \Phi\phi\mathbf{n}_2 + \sqrt{\Phi}\left(\Phi v - \Phi\hat{v}\right) \end{bmatrix}, \tag{2.19}$$

[2]The equivalence to the standard form involving information from both sides can be seen by enforcing the continuity equation as in (2.10), solving for $\hat{\phi}$ and $\hat{\boldsymbol{\vartheta}}$ as in (2.18), and then substituting the resulting $\hat{\phi}$ and $\hat{\boldsymbol{\vartheta}}$ into (2.19).

where we have used the fact that the largest eigenvalue of $\mathcal{A}$ is $\sqrt{\Phi}$ (see [6] and the references therein for the hybridization of the local Lax–Friedrichs flux for nonlinear problems). Note that, for the shallow water system (2.1), the Lax–Friedrichs flux coincides with the Rankine–Hugoniot flux (2.8) (if we replace ϕ^* and $\boldsymbol{\vartheta}^*$ by $\hat{\phi}$ and $\hat{\boldsymbol{\vartheta}}$, respectively). We can then define the corresponding HDG scheme, called *HDG*-III, similar to HDG-II, but now replace the HDG flux in (2.16) by (2.19). Since all three components of flux (2.19) in general do not satisfy the conservation condition, we have three conservation constraints instead of one compared to (2.17), i.e.,

$$\left\langle [\![\Phi\boldsymbol{\vartheta}\cdot\mathbf{n}+\sqrt{\Phi}(\phi-\hat{\phi})]\!], \mu\right\rangle_e = 0, \tag{2.20a}$$

$$\left\langle [\![\Phi\phi\mathbf{n}_1+\sqrt{\Phi}\,(\Phi u-\Phi\hat{u})]\!], \mu\right\rangle_e = 0, \tag{2.20b}$$

$$\left\langle [\![\Phi\phi\mathbf{n}_2+\sqrt{\Phi}\,(\Phi v-\Phi\hat{v})]\!], \mu\right\rangle_e = 0. \tag{2.20c}$$

As can be seen, the HDG-III method with hybridized Lax–Friedrichs flux (2.19) has three trace unknowns, and hence is more expensive than HDG-II. The question is whether we can eliminate some of the trace unknowns to reduce the cost of HDG-III. We will employ the conservation constraints (2.20) to perform the elimination. We start with a result similar to Lemma 2.1.

LEMMA 2.4. *Assuming that interior edges $e \in \mathcal{E}_h^o$ are straight lines, then on any $e \in \mathcal{E}_h^o$, the trace unknowns $\left(\hat{\phi}, \Phi\hat{\boldsymbol{\vartheta}}\right)$ of HDG-*III *scheme satisfy*

$$\left(\Phi\boldsymbol{\vartheta}-\Phi\hat{\boldsymbol{\vartheta}}\right)\cdot\mathbf{n} = -\sqrt{\Phi}\left(\phi-\hat{\phi}\right), \tag{2.21}$$

where $\mathbf{u}=[\phi,\Phi\boldsymbol{\vartheta}]$ is trace of the DG solution from either side of e.

Proof. We proceed with (2.20a). Taking $\mu = [\![\Phi\boldsymbol{\vartheta}\cdot\mathbf{n}+\sqrt{\Phi}(\phi-\hat{\phi})]\!]$ we conclude that $[\![\Phi\boldsymbol{\vartheta}\cdot\mathbf{n}+\sqrt{\Phi}(\phi-\hat{\phi})]\!] = 0$ since e is straight (and hence having constant Jacobian and normal vector). Now, arguing similarly to the proof of the first assertion of Theorem 2.3, we obtain

$$\sqrt{\Phi}\left(\phi-\hat{\phi}\right) = \frac{\sqrt{\Phi}}{2}[\![\phi\mathbf{n}]\!]\cdot\mathbf{n} - \frac{1}{2}[\![\Phi\boldsymbol{\vartheta}\cdot\mathbf{n}]\!].$$

Similarly, from (2.20b) and (2.20c) we have

$$\sqrt{\Phi}\left(\Phi\boldsymbol{\vartheta}-\Phi\hat{\boldsymbol{\vartheta}}\right)\cdot\mathbf{n} = \frac{\sqrt{\Phi}}{2}[\![\Phi\boldsymbol{\vartheta}\cdot\mathbf{n}]\!] - \frac{\Phi}{2}[\![\phi\mathbf{n}]\!]\cdot\mathbf{n}.$$

Combining these results ends the proof. ∎

The result in (2.21) allows us to eliminate $\hat{\phi}$ so that the hybridized Lax–Friedrichs flux becomes

$$\hat{\mathbf{F}}\cdot\mathbf{n} = \begin{bmatrix} \Phi\hat{\boldsymbol{\vartheta}}\cdot\mathbf{n} \\ \Phi\phi\mathbf{n}_1+\sqrt{\Phi}\,(\Phi u-\Phi\hat{u}) \\ \Phi\phi\mathbf{n}_2+\sqrt{\Phi}\,(\Phi v-\Phi\hat{v}) \end{bmatrix}. \tag{2.22}$$

Consequently, we need to enforce only two conservation conditions (2.20b) and (2.20c), since the first is automatically satisfied with the economic form (2.22). *Let us term*

this HDG method as HDG-IV. The local solver for HDG-IV reads

(2.23a)
$$\left(\frac{\partial \phi}{\partial t}, \varphi\right)_K - (\Phi \boldsymbol{\vartheta}, \nabla \varphi)_K + \left\langle \Phi \hat{\boldsymbol{\vartheta}} \cdot \mathbf{n}, \varphi \right\rangle_{\partial K} = 0,$$

$$\left(\frac{\partial (\Phi u)}{\partial t}, \mathbf{w}_1\right)_K - \left(\Phi\phi, \frac{\partial \mathbf{w}_1}{\partial x}\right)_K + \left\langle \Phi\phi \mathbf{n}_1 + \sqrt{\Phi}\,(\Phi u - \Phi \hat{u})\,, \mathbf{w}_2 \right\rangle_{\partial K}$$

(2.23b)
$$= \left(f\Phi v - \gamma \Phi u + \frac{\tau_x}{\rho}, \mathbf{w}_1\right)_K,$$

$$\left(\frac{\partial (\Phi v)}{\partial t}, \mathbf{w}_2\right)_K - \left(\Phi\phi, \frac{\partial \mathbf{w}_2}{\partial y}\right)_K + \left\langle \Phi\phi \mathbf{n}_2 + \sqrt{\Phi}\,(\Phi v - \Phi \hat{v})\,, \mathbf{w}_2 \right\rangle_{\partial K}$$

(2.23c)
$$= \left(-f\Phi u - \gamma \Phi v + \frac{\tau_y}{\rho}, \mathbf{w}_2\right)_K,$$

while the conservation constraints are

(2.24a)
$$\left\langle [\![\Phi\phi \mathbf{n}_1 + \sqrt{\Phi}\,(\Phi u - \Phi \hat{u})]\!], \mu \right\rangle_e = 0,$$

(2.24b)
$$\left\langle [\![\Phi\phi \mathbf{n}_2 + \sqrt{\Phi}\,(\Phi v - \Phi \hat{v})]\!], \mu \right\rangle_e = 0.$$

3. Analysis. In this section we study the well-posedness, stability, and convergence rates of our HDG schemes. Since HDG-II and HDG-IV are equivalent to HDG-I and HDG-III, it is sufficient to carry out the analysis for HDG-II and HDG-IV. To that end, we need to equip the linearized shallow water system with an appropriate boundary condition. For simplicity we consider only the wall boundary condition in this paper, and this is enforced using a reflection principle. In particular, for an element K^- that is adjacent to the domain boundary (i.e., $\partial K^- \cap \partial\Omega \neq \emptyset$), we assume that there is an imaginary neighbor element K^+ whose state $\mathbf{u}^+ = \left(\phi^+, \Phi\boldsymbol{\vartheta}^+\right)^T$ is determined as

(3.1a)
$$\phi^+ = \phi^-,$$

(3.1b)
$$\Phi u^+ = \Phi u^- - 2\left(\Phi \boldsymbol{\vartheta}^- \cdot \mathbf{n}\right)\mathbf{n}_1,$$

(3.1c)
$$\Phi v^+ = \Phi v^- - 2\left(\Phi \boldsymbol{\vartheta}^- \cdot \mathbf{n}\right)\mathbf{n}_2,$$

and for the exact solution we have $\boldsymbol{\vartheta}^e \cdot \mathbf{n} = 0$.

For the upwind HDG-II scheme, we also impose conservation condition (2.17) on boundary faces $e \in \partial K^- \cap \partial\Omega$, which, using (3.1), can be simplified to

(3.2)
$$\left\langle \Phi \boldsymbol{\vartheta} \cdot \mathbf{n} + \sqrt{\Phi}\left(\phi - \hat{\phi}\right), \mu \right\rangle_e = 0 \quad \forall e \in \partial K^- \cap \partial\Omega.$$

For HDG-IV, additional conservation constraints when applying (2.24) on boundary faces $e \in \partial K^- \cap \partial\Omega$ read

(3.3a)
$$\left\langle -\sqrt{\Phi}\Phi \boldsymbol{\vartheta} \cdot \mathbf{n}\mathbf{n}_1 + \sqrt{\Phi}\,(\Phi u - \Phi \hat{u}), \mu \right\rangle_e = 0 \quad \forall e \in \partial K^- \cap \partial\Omega,$$

(3.3b)
$$\left\langle -\sqrt{\Phi}\Phi \boldsymbol{\vartheta} \cdot \mathbf{n}\mathbf{n}_2 + \sqrt{\Phi}\,(\Phi v - \Phi \hat{v}), \mu \right\rangle_e = 0 \quad \forall e \in \partial K^- \cap \partial\Omega,$$

which implies

(3.3c)
$$\hat{\mathbf{u}} \cdot \mathbf{n} = 0 \quad \text{on} \quad e \in \partial K^- \cap \partial\Omega.$$

For simplicity in the analysis, we assume that all integrals are evaluated exactly, though quadrature is used in practice. Note that all of the following results still hold if quadrature error is taken into account (see, e.g., [8] and the references therein in the DG context). Thus, we omit the details. *Unlike most existing works (see, e.g.,* [14, 13]*) we do not introduce projection operators but develop a "direct" analysis. Furthermore, our analysis is for a pure hyperbolic PDE system.*

3.1. Semidiscrete HDG system: Well-posedness, stability, and convergence. In this section we will study the well-posedness, stability, and convergence of the proposed HDG schemes. Using an energy approach we will prove semidiscrete stability, from which we deduce well-posedness. We will also show that our HDG schemes have the same convergence rates as those of standard DG methods for linear hyperbolic systems. To that end, we define the energy for the linearized shallow water system (2.1) as

$$\mathcal{E} := \frac{1}{2}\left(\|\phi\|^2_{\Omega_h} + \|\boldsymbol{\vartheta}\|^2_{\Omega_h,\Phi}\right).$$

Here $\mathcal{E}$ is a function of time t, but we ignore this dependency when there is no ambiguity. For simplicity of exposition, this convention also applies to other quantities as well.

THEOREM 3.1 (semidiscrete stability). *Both HDG-* II *and HDG-*IV *are stable in the following sense:*

$$\frac{d\mathcal{E}}{dt} + \left\|\phi - \hat{\phi}\right\|^2_{\partial\Omega_h,c} \leq \begin{cases} \frac{1}{4}\|\boldsymbol{\tau}\|^2_{\Omega_h,(\gamma\rho^2\Phi)^{-1}} & \gamma > 0, \\ \frac{1}{2}\left(\|\boldsymbol{\tau}\|^2_{\Omega_h,\rho^2\Phi} + \mathcal{E}\right) & \gamma = 0, \end{cases} \tag{3.4}$$

where $c = \sqrt{\Phi}$ *for HDG-* II, $c = \sqrt{\Phi^3}$ *for HDG-*IV, *and sgn*$(\cdot)$ *the standard signum function.*

Proof. We proceed with HDG-II. We first take $\varphi = \phi$ in (2.16a), $\mathbf{w}_1 = u$ in (2.16b), and $\mathbf{w}_2 = v$ in (2.16c). We then integrate the second term of (2.16a) by parts, add the resulting equations together, and sum over all elements to arrive at

$$\frac{d\mathcal{E}}{dt} + \left\langle\sqrt{\Phi}\phi,\phi\right\rangle_{\partial\Omega_h} - \left\langle\sqrt{\Phi}\hat{\phi},\phi\right\rangle_{\partial\Omega_h} + \left\langle\Phi\boldsymbol{\vartheta}\cdot\mathbf{n},\hat{\phi}\right\rangle_{\partial\Omega_h} = -\|\boldsymbol{\vartheta}\|^2_{\Omega_h,\gamma\Phi} + \left(\frac{\boldsymbol{\tau}}{\rho},\boldsymbol{\vartheta}\right)_{\Omega_h},$$

where we have used the boundary conditions (3.2) on the boundary faces. Recall that $\partial\Omega_h := \{\partial K : K \in \Omega_h\}$. Next, taking $\mu = \hat{\phi}$ in (2.17) and (3.2), and then summing over all faces in the mesh skeleton (together with the boundary condition (3.2)) gives

$$\left\langle\Phi\boldsymbol{\vartheta}\cdot\mathbf{n},\hat{\phi}\right\rangle_{\partial\Omega_h} + \left\langle\sqrt{\Phi}\phi,\hat{\phi}\right\rangle_{\partial\Omega_h} - \left\langle\sqrt{\Phi}\hat{\phi},\hat{\phi}\right\rangle_{\partial\Omega_h} = 0.$$

Subtracting the above two equations yields

$$\frac{d\mathcal{E}}{dt} + \left\|\phi - \hat{\phi}\right\|^2_{\partial\Omega_h,\sqrt{\Phi}} = -\|\boldsymbol{\vartheta}\|^2_{\Omega_h,\gamma\Phi} + \left(\frac{\boldsymbol{\tau}}{\rho},\boldsymbol{\vartheta}\right)_{\Omega_h}, \tag{3.5}$$

which gives (3.4) after a simple application of Young's inequality for the last term on the right-hand side of (3.5). By the same token, one can show that the desired result holds for HDG-IV. ∎

Remark 1. It can be observed that when $\gamma > 0$ (with bottom friction), the energy grows at most linearly in time. On the other hand, for $\gamma = 0$, an application of the Gronwall lemma shows that the energy is bounded from above by $2\left\|\boldsymbol{\tau}\right\|^2_{\Omega_h,\rho^2\Phi}\exp\left(t/2\right)$.

COROLLARY 3.2. *At any point in time, there exists a unique solution for the HDG schemes HDG-*II *and HDG-*IV*. In particular, both local and global solvers are well-posed.*

Proof. Since the HDG solution $\mathbf{u} = (\phi, \Phi\boldsymbol{\vartheta})$ resides in a finite element space with finite dimensions, the well-posedness is equivalent to uniqueness. Furthermore, it is sufficient to show that HDG solutions vanish for zero initial conditions and $\boldsymbol{\tau} = \mathbf{0}$. Again, we omit the proof for HDG-IV for brevity. Integrating (3.5) from 0 to t we have

$$\mathcal{E}\left(t\right) = -\int_0^t \left\|\phi - \hat{\phi}\right\|^2_{\partial\Omega_h,\sqrt{\Phi}}\, dt + \int_0^t -\left\|\boldsymbol{\vartheta}\right\|^2_{\Omega_h,\gamma\Phi}\, dt,$$

whose left-hand side is nonnegative and right-hand side is nonpositive. This can only be true if both sides vanish, i.e., $\mathcal{E}\left(t\right) = \left\|\phi-\hat{\phi}\right\|_{\partial K,\sqrt{\Phi}} = \left\|\boldsymbol{\vartheta}\right\|_{\Omega_h,\gamma\Phi} = 0$, which implies that $\mathbf{u} = \mathbf{0}$ and $\hat{\phi} = 0$, and hence uniqueness. ∎

Next, we will obtain the convergence rate for our HDG schemes. Since the proof is relatively long and it is similar for both HDG-II and HDG-IV, we leave out the details of the latter. We start with some auxiliary results.

Recall that interpolation introduces truncation and aliasing errors [9, 33], and thus interpolation is generally different from projection, which has only truncation error. For sufficiently smooth functions, however, the aliasing error either is spectrally small [9, 27, 33] or can be made equal to zero [28]. Following [28], we shall make no distinction between interpolation and projection in what remains.

Again, we reserve superscript "e" for the exact solution. Let $\mathbb{P}$ be the L^2-projection (or interpolation) operator on $\mathcal{P}^p$. In addition, c denotes a generic constant that may have different values in different contexts. To begin, we recall the following fundamental hp approximation error bounds [3, 4, 5] for $q \in H^s\left(K\right)$:

$$\left\|q - \mathbb{P}q\right\|_{H^r(K)} \leq c\frac{h^{\sigma-r}}{p^{s-r}}\left\|q\right\|_{H^s(K)}, \quad 0 \leq r \leq s, \tag{3.6a}$$

$$\left\|q - \mathbb{P}q\right\|_{\partial K} \leq c\frac{h^{\sigma-1/2}}{p^{s-1/2}}\left\|q\right\|_{H^s(K)}, \quad s > \frac{1}{2}, \tag{3.6b}$$

with $\sigma = \min\left\{p+1, s\right\}$, and $\left\|\cdot\right\|_{H^r(K)}$ denoting the usual Sobolev norm.

Since the exact solution $\mathbf{u}^e$ is assumed to satisfy the linearized shallow water equations (2.1), the following equations hold:

$$\left(\frac{\partial\phi^e}{\partial t}, \varphi\right)_K - \left(\Phi\boldsymbol{\vartheta}^e, \nabla\varphi\right)_K + \left\langle\Phi\boldsymbol{\vartheta}^e\cdot\mathbf{n}, \varphi\right\rangle_{\partial K} = 0, \tag{3.7a}$$

$$\left(\frac{\partial\left(\Phi u^e\right)}{\partial t}, \mathbf{w}_1\right)_K - \left(\Phi\phi^e, \frac{\partial\mathbf{w}_1}{\partial x}\right)_K + \left\langle\Phi\phi^e\mathbf{n}_1, \mathbf{w}_1\right\rangle_{\partial K} = \left(f\Phi v^e - \gamma\Phi u^e + \frac{\tau_x}{\rho}, \mathbf{w}_1\right)_K, \tag{3.7b}$$

$$\left(\frac{\partial\left(\Phi v^e\right)}{\partial t}, \mathbf{w}_2\right)_K - \left(\Phi\phi^e, \frac{\partial\mathbf{w}_2}{\partial y}\right)_K + \left\langle\Phi\phi^e\mathbf{n}_2, \mathbf{w}_2\right\rangle_{\partial K} = \left(-f\Phi u^e - \gamma\Phi v^e + \frac{\tau_y}{\rho}, \mathbf{w}_2\right)_K. \tag{3.7c}$$

In particular, the exact solution also satisfies the conservation, i.e.,

$$[\![\Phi\boldsymbol{\vartheta}^e\cdot\mathbf{n}]\!]=0 \quad \text{on } e\in\mathcal{E}_h^o, \tag{3.7d}$$

and the boundary condition (3.2).

Let the error E between the projection of the exact solution $\mathbf{u}^e$ and the HDG counterpart $\mathbf{u}$ be defined as

$$\mathsf{E}:=\frac{1}{2}\left(\|\mathbb{P}\phi^e-\phi\|_{\Omega_h}^2+\|\mathbb{P}\boldsymbol{\vartheta}^e-\boldsymbol{\vartheta}\|_{\Omega_h,\Phi}^2\right),$$

and, for the rest of the paper, *we assume that the initial error is zero, i.e.,* $\mathsf{E}(0)=0$. Due to the triangle inequality and (3.6a), it is sufficient to analyze the convergence rate (to zero) of E since it is also the convergence of the error between the HDG solution and the exact one. We are now in position to prove the first convergence result with $\gamma=f=0$ and $\boldsymbol{\tau}=\mathbf{0}$.

THEOREM 3.3. *Assume* $\mathbf{u}^e|_K=(\phi^e,\Phi\boldsymbol{\vartheta}^e)|_K\in[H^s(K)]^3, s\geq 3/2$, $\gamma=f=0$, *and* $\boldsymbol{\tau}=\mathbf{0}$. *There exists a constant* c *that depends only on the angle condition of* K, s, *and on* Φ *such that*

$$\mathsf{E}(t)\leq c\frac{h^{2\sigma-1}}{p^{2s-1}}t\max_{\theta\in[0,t]}\mathcal{E}^e(\theta), \tag{3.8}$$

with $\sigma=\min\{p+1,s\}$ *and*

$$\mathcal{E}^e(t):=\sum_K\|\phi^e(t)\|_{H^s(K)}^2+\|\boldsymbol{\vartheta}^e(t)\|_{H^s(K)}^2.$$

Proof. We denote by $\mathbb{P}$ and Π the local projections on an element and an edge, respectively. We start the proof by defining the following errors:

$$\begin{aligned}
\varepsilon_\phi^h&:=\mathbb{P}\phi^e-\phi, \quad \varepsilon_\phi^I:=\phi^e-\mathbb{P}\phi^e,\\
\varepsilon_{\hat\phi}^h&:=\Pi\phi^e-\hat\phi, \quad \varepsilon_{\hat\phi}^I:=\phi^e-\Pi\phi^e,\\
\boldsymbol{\varepsilon}_{\boldsymbol{\vartheta}}^h&:=\mathbb{P}\boldsymbol{\vartheta}^e-\boldsymbol{\vartheta}, \quad \boldsymbol{\varepsilon}_{\boldsymbol{\vartheta}}^I:=\boldsymbol{\vartheta}^e-\mathbb{P}\boldsymbol{\vartheta}^e.
\end{aligned}$$

Subtracting (2.16) from the corresponding equations in (3.7) we obtain

$$\begin{aligned}
&\left(\frac{\partial\varepsilon_\phi^h}{\partial t},\varphi\right)_K-\left(\Phi\boldsymbol{\varepsilon}_{\boldsymbol{\vartheta}}^h,\nabla\varphi\right)_K+\left\langle\Phi\boldsymbol{\varepsilon}_{\boldsymbol{\vartheta}}^h\cdot\mathbf{n}+\sqrt{\Phi}\left(\varepsilon_\phi^h-\varepsilon_{\hat\phi}^h\right),\varphi\right\rangle_{\partial K}\\
&\qquad=-\left\langle\Phi\boldsymbol{\varepsilon}_{\boldsymbol{\vartheta}}^I\cdot\mathbf{n}+\sqrt{\Phi}\varepsilon_\phi^I,\varphi\right\rangle_{\partial K}, \qquad (3.9a)\\
&\left(\Phi\frac{\partial\boldsymbol{\varepsilon}_{\boldsymbol{\vartheta}}^h}{\partial t},\mathbf{w}\right)_K-\left(\Phi\varepsilon_\phi^h,\nabla\cdot\mathbf{w}\right)_K+\left\langle\Phi\varepsilon_{\hat\phi}^h,\mathbf{w}\cdot\mathbf{n}\right\rangle_{\partial K}=-\left\langle\Phi\varepsilon_{\hat\phi}^I,\mathbf{w}\cdot\mathbf{n}\right\rangle_{\partial K}, \qquad (3.9b)
\end{aligned}$$

where we have used the property of projection operators $\mathbb{P}$ and Π.

Subtracting (2.17) from (3.7d) yields

$$\left\langle[\![\Phi\boldsymbol{\varepsilon}_{\boldsymbol{\vartheta}}^h\cdot\mathbf{n}+\sqrt{\Phi}(\varepsilon_\phi^h-\varepsilon_{\hat\phi}^h)]\!],\mu\right\rangle_e=-\left\langle[\![\Phi\boldsymbol{\varepsilon}_{\boldsymbol{\vartheta}}^I\cdot\mathbf{n}+\sqrt{\Phi}\varepsilon_\phi^I]\!],\mu\right\rangle_e. \tag{3.10}$$

Now first taking $\varphi=\varepsilon_\phi^h$, $\mathbf{w}=\boldsymbol{\varepsilon}_{\boldsymbol{\vartheta}}^h$, and $\mu=\varepsilon_{\hat\phi}^h$, second subtracting (3.10) from (3.9), and then using an energy argument similar to the proof of Theorem 3.1, we

obtain

$$\frac{d\mathsf{E}}{dt} + \left\langle \sqrt{\Phi}\left(\varepsilon_\phi^h - \varepsilon_{\hat{\phi}}^h\right), \varepsilon_\phi^h - \varepsilon_{\hat{\phi}}^h \right\rangle_{\partial\Omega_h} + \left\langle \Phi\boldsymbol{\varepsilon}_{\boldsymbol{\vartheta}}^I \cdot \mathbf{n} + \sqrt{\Phi}\varepsilon_\phi^I, \varepsilon_\phi^h - \varepsilon_{\hat{\phi}}^h \right\rangle_{\partial\Omega_h} = 0.$$

Completing the square for the last two terms, we arrive at

$$\frac{d\mathsf{E}}{dt} + \left\| \Phi^{1/4}\left(\varepsilon_\phi^h - \varepsilon_{\hat{\phi}}^h\right) + \Phi^{3/4}\boldsymbol{\varepsilon}_{\boldsymbol{\vartheta}}^I \cdot \mathbf{n} + \Phi^{1/4}\varepsilon_\phi^I \right\|_{\partial\Omega_h}^2 = \frac{1}{4}\left\| \Phi^{1/4}\varepsilon_\phi^I + \Phi^{3/4}\boldsymbol{\varepsilon}_{\boldsymbol{\vartheta}}^I \right\|_{\partial\Omega_h}^2,$$

which, together with an application of the Cauchy–Schwarz inequality, implies

$$\frac{d\mathsf{E}}{dt} \leq \frac{1}{4}\sum_K \left\| \Phi^{3/4}\left(\boldsymbol{\vartheta}^e - \mathbb{P}\boldsymbol{\vartheta}^e\right)\cdot\mathbf{n} \right\|_{\partial K}^2 + \left\| \Phi^{1/4}\left(\phi^e - \mathbb{P}\phi^e\right) \right\|_{\partial K}^2. \tag{3.11}$$

Combining (3.6b) and (3.11) we obtain

$$\frac{d\mathsf{E}(t)}{dt} \leq c\frac{h^{2\sigma-1}}{p^{2s-1}}\mathcal{E}^e(t). \tag{3.12}$$

The desired result (3.8) is a direct consequence of (3.12) after integrating both sides from 0 to t. ∎

For general γ, f, and $\boldsymbol{\tau}$, we have the following convergence result.

THEOREM 3.4 (convergence). *Assume* $\mathbf{u}^e = (\phi^e, \Phi\boldsymbol{\vartheta}^e) \in [H^s(K)]^3, s \geq 3/2$. *There exists a constant* c *that depends only on the angle condition of* K, s, *and on* γ, f, Φ *such that*

$$\mathsf{E}(t) \leq c\left[sgn(\gamma)\, t + (1 - sgn(\gamma))\exp\left(\frac{t}{2}\right)\right]\frac{h^{2\sigma-1}}{p^{2s-1}}\max_{\theta\in[0,t]}\mathcal{E}^e(\theta), \tag{3.13}$$

where $\sigma = \min\{p+1, s\}$.

Proof. Following the same procedure as in the proof of Theorem 3.3 we obtain a similar result as in (3.12) but with an extra term due to the presence of γ and f, i.e.,

$$\frac{d\mathsf{E}(t)}{dt} \leq c\frac{h^{2\sigma-1}}{p^{2s-1}}\mathcal{E}^e(t) + \sum_K R_K, \tag{3.14}$$

where

$$\begin{aligned} R_K := {} & \left(f\Phi\left(v^e - \mathbb{P}v^e\right), \mathbb{P}u^e - u\right)_K - \left(f\Phi\left(u^e - \mathbb{P}u^e\right), \mathbb{P}v^e - v\right)_K \\ & - \left(\gamma\Phi\left(\boldsymbol{\vartheta}^e - \mathbb{P}\boldsymbol{\vartheta}^e\right), \mathbb{P}\boldsymbol{\vartheta}^e - \boldsymbol{\vartheta}\right)_K - \left\|\mathbb{P}\boldsymbol{\vartheta}^e - \boldsymbol{\vartheta}\right\|_{K,\gamma\Phi}^2. \end{aligned} \tag{3.15}$$

We now consider two cases: $\gamma > 0$ and $\gamma = 0$. For the first case, by a simple application of the Young inequality for the first three terms (in order to cancel out the last term) in (3.15) we obtain

$$R_K \leq \left\|\boldsymbol{\vartheta}^e - \mathbb{P}\boldsymbol{\vartheta}^e\right\|_{K, f^2\Phi/\gamma+\gamma\Phi}^2,$$

which, after applying the projection error bound (3.6a), becomes

$$\sum_K R_K \leq c\frac{h^{2\sigma}}{p^{2s}}\mathcal{E}^e(t) \leq c\frac{h^{2\sigma-1}}{p^{2s-1}}\mathcal{E}^e(t). \tag{3.16}$$

Combining (3.14) and (3.16) yields

$$\frac{d\mathsf{E}(t)}{dt} \leq c\frac{h^{2\sigma-1}}{p^{2s-1}}\mathcal{E}^e(t), \tag{3.17}$$

which is exactly (3.12) (clearly, with a different value of c). Consequently, the error bound (3.8) also holds for $\gamma > 0$.

For the second case, i.e., $\gamma = 0$, by the Cauchy–Schwarz inequality we have

$$\sum_K R_K \leq \sum_K \frac{1}{2}\|\boldsymbol{\vartheta}^e - \mathbb{P}\boldsymbol{\vartheta}^e\|^2_{K,f^2\Phi} + \frac{1}{2}\|\mathbb{P}\boldsymbol{\vartheta}^e - \boldsymbol{\vartheta}\|^2_{K,\Phi} \leq \frac{1}{2}\sum_K \|\boldsymbol{\vartheta}^e - \mathbb{P}\boldsymbol{\vartheta}^e\|^2_{K,f^2\Phi} + \frac{1}{2}\mathsf{E}(t),$$

which, together with (3.14) and (3.6a), gives

$$\frac{d\mathsf{E}(t)}{dt} \leq \frac{1}{2}\mathsf{E}(t) + c\frac{h^{2\sigma-1}}{p^{2s-1}}\mathcal{E}^e(t). \tag{3.18}$$

Using the Gronwall lemma we conclude

$$\mathsf{E}(t) \leq c\frac{h^{2\sigma-1}}{p^{2s-1}}\max_{\theta\in[0,t]}\mathcal{E}^e(\theta)\exp\left(\frac{t}{2}\right),$$

and this ends the proof. ∎

Remark 2. As can be seen in (3.13), when the bottom friction γ is taken into account the error grows at most linearly in time, though the rate may be exponential when there is no bottom friction. The reason is that the bottom friction provides a damping effect in the system, and reduces the growth rate. In both cases, the h-convergence rate is $p+\frac{1}{2}$ and the p-convergence rate is $s-\frac{1}{2}$. This is compatible with the DG convergence rates [31]. An extension to fully hp-convergence analysis, such as in [8], is straightforward, so we leave out the details here.

3.2. Fully discrete HDG system: Well-posedness, stability, and convergence. In this section we present a fully discrete HDG scheme using, for simplicity, the backward Euler approach for HDG-II. Extension to other implicit time discretization strategies such as Crank–Nicholson is straightforward. A similar scheme for HDG-IV is obvious, and hence is omitted. It is sufficient to present the results for one time step, thus we will ignore the time index for clarity. The backward Euler discretization of the local solver (2.16) reads

$$\left(\frac{\phi}{\Delta t},\varphi\right)_K - (\Phi\boldsymbol{\vartheta},\nabla\varphi)_K + \left\langle \Phi\boldsymbol{\vartheta}\cdot\mathbf{n} + \sqrt{\Phi}\left(\phi-\hat{\phi}\right),\varphi\right\rangle_{\partial K} = \left(\frac{\phi^0}{\Delta t},\varphi\right)_K, \tag{3.19a}$$

$$\left(\frac{\Phi u}{\Delta t},\varphi\right)_K - \left(\Phi\phi,\frac{\partial\varphi}{\partial x}\right)_K + \left\langle \Phi\hat{\phi}\mathbf{n}_1,\varphi\right\rangle_{\partial K} = \left(f\Phi v - \gamma\Phi u + \frac{\tau_x}{\rho},\varphi\right)_K + \left(\frac{\Phi u^0}{\Delta t},\varphi\right)_K, \tag{3.19b}$$

$$\left(\frac{\Phi v}{\Delta t},\varphi\right)_K - \left(\Phi\phi,\frac{\partial\varphi}{\partial y}\right)_K + \left\langle \Phi\hat{\phi}\mathbf{n}_2,\varphi\right\rangle_{\partial K} = \left(-f\Phi u - \gamma\Phi v + \frac{\tau_y}{\rho},\varphi\right)_K + \left(\frac{\Phi v^0}{\Delta t},\varphi\right)_K, \tag{3.19c}$$

where quantities with superscript "0" are the quantities from the previous time step while the others are those with the current time step.

PROPOSITION 3.5 (discrete stability). *The fully discrete HDG scheme with the local solver* (3.19) *and global solver* (2.17) *is stable in the following sense:*

$$\frac{\mathcal{E}}{\Delta t}+\left\|\phi-\hat{\phi}\right\|^2_{\partial\Omega_h,\sqrt{\Phi}}\leq\frac{\mathcal{E}^0}{\Delta t}+\begin{cases}\frac{1}{4}\left\|\boldsymbol{\tau}\right\|^2_{\Omega_h,(\gamma\rho^2\Phi)^{-1}} & \gamma>0,\\ \frac{1}{2}\left(\left\|\boldsymbol{\tau}\right\|^2_{\Omega_h,\rho^2\Phi}+\mathcal{E}\right) & \gamma=0,\end{cases}$$

where sgn$(\cdot)$ *is the standard signum function, and* $\mathcal{E}^0$ *is the energy from the previous time step.*

Proof. The proof is a direct consequence of Theorem 3.1. ∎

COROLLARY 3.6. *At any point in time, there is a unique solution for the proposed fully discrete HDG schemes. In particular, the discrete local and global solvers, namely* (3.19) *and* (2.17)*, are well-posed.*

Proof. It is enough to show that the only HDG solution for zero initial condition and $\boldsymbol{\tau}=\mathbf{0}$ must be zero for the first time step. Using a similar energy approach as in Theorem 3.1 we obtain

$$\frac{\mathcal{E}}{\Delta t}+\left\|\phi-\hat{\phi}\right\|^2_{\partial\Omega_h,\sqrt{\Phi}}+\left\|\boldsymbol{\vartheta}\right\|^2_{\Omega_h,\gamma\Phi}=0,$$

where we have again ignored the time dependency for simplicity in writing. Since the left-hand side of the preceding equation is nonnegative, we conclude that $\mathcal{E}=\left\|\phi-\hat{\phi}\right\|_{\partial K,\sqrt{\Phi}}=\left\|\boldsymbol{\vartheta}\right\|_{\Omega_h,\gamma\Phi}=0$. That is, $\mathbf{u}=\mathbf{0}$ and $\hat{\phi}=0$, which ends the proof. ∎

We now provide the convergence rate for the fully discrete HDG system for the most general case in which γ, f, and $\boldsymbol{\tau}$ could be nonzero.

PROPOSITION 3.7 (convergence of discrete system). *Assume* $\mathbf{u}^e=(\phi^e,\Phi\boldsymbol{\vartheta}^e)\in\left[H^s\left(K\right)\right]^3, s\geq 3/2$. *In addition, suppose that the exact solution also satisfies*[3] *the fully discrete variational equation* (3.19). *The errors* E^0 *and* E *at two consecutive time steps satisfy*

$$\left[\frac{1}{\Delta t}-\frac{(1-\ sgn(\gamma))}{2}\right]\mathsf{E}\leq c\frac{h^{2\sigma-1}}{p^{2s-1}}\mathcal{E}^e+\frac{\mathsf{E}^0}{\Delta t}.$$

Thus, if $\Delta t<2$*, then the solution of the fully discrete HDG scheme with the local solver* (3.19) *and global solver* (2.17) *converges to the exact solution with h-convergence rate of* $p+\frac{1}{2}$ *and p-convergence rate of* $s-\frac{1}{2}$.

Proof. The result is a direct consequence of (3.17) and (3.18) and the backward Euler discretization. The convergence rate is clear by induction since *the error* E *is zero initially.* ∎

4. A class of penalty HDG schemes. In this section, we generalize our upwind HDG approach to a class of penalty HDG schemes, which encompasses the upwind HDG. The objective is two-fold: first, to show that our upwind HDG construction is parameter free, and second, to provide a constructive reasoning to derive penalty HDG methods and their connection with upwind HDG schemes. To that end, we first observe that $\phi-\hat{\phi}$ is the mismatch between the volume unknown, restricted on the mesh skeleton, and the trace unknown. This mismatch vanishes for the exact

[3] An error of order $\mathcal{O}\left(\Delta t\right)$ occurs with backward Euler discretization, but can be made arbitrarily small relative to the spatial discretization error, and hence is negligible.

solutions, but converges to zero for the HDG solutions as the mesh (or solution order) is refined. This suggests that one can control the mismatch by introducing a penalty parameter λ to form a penalized family of HDG fluxes given by

$$\hat{\mathbf{F}}\cdot\mathbf{n} = \begin{bmatrix} \Phi\boldsymbol{\vartheta}\cdot\mathbf{n} + \lambda\left(\phi - \hat{\phi}\right) \\ \Phi\hat{\phi}\mathbf{n}_1 \\ \Phi\hat{\phi}\mathbf{n}_2 \end{bmatrix}. \tag{4.1}$$

Clearly, when $\lambda = \sqrt{\Phi}$ we recover the upwind HDG scheme. As shall be shown below, both theoretically and numerically, *HDG schemes are robust with respect to the penalty parameter λ, and provide flexibility. In particular, deviating from the fully upwind scheme does not deteriorate the convergence rate unless $\lambda = 0$, though there is some loss in accuracy.* Let us summarize stability and convergence results whose proofs are similar to those in section 3.

THEOREM 4.1 (stability). *Let $\lambda \geq 0$. The class of HDG schemes with HDG flux* (4.1) *is stable in the following sense:*

$$\frac{d\mathcal{E}}{dt} + \sum_K \left\|\phi - \hat{\phi}\right\|^2_{\partial K,\lambda} \leq \begin{cases} \frac{1}{4}\left\|\boldsymbol{\tau}\right\|^2_{\Omega_h,(\gamma\rho^2\Phi)^{-1}} & \gamma > 0, \\ \frac{1}{2}\left(\left\|\boldsymbol{\tau}\right\|^2_{\Omega_h,\rho^2\Phi} + \mathcal{E}\right) & \gamma = 0, \end{cases}$$

where $sgn(\cdot)$ is the standard signum function. Consequently, at any point in time there exists a unique HDG solution $\left(\phi, \Phi\boldsymbol{\vartheta}, \hat{\phi}\right)$.

THEOREM 4.2 (convergence). *Assume $\mathbf{u}^e = (\phi^e, \Phi\boldsymbol{\vartheta}^e) \in \left[H^s(K)\right]^3, s \geq 3/2$, and $\lambda > 0$. There exists a constant c that depends only on the angle condition of K, s, λ and on γ, f, Φ such that*

$$\mathsf{E}(t) \leq c\left[\, sgn(\gamma)\, t + (1 - \ sgn(\gamma))\exp\left(\frac{t}{2}\right)\right]\frac{h^{2\sigma-1}}{p^{2s-1}}\max_{\theta\in[0,t]}\mathcal{E}^e(\theta),$$

where $\sigma = \min\{p+1, s\}$.

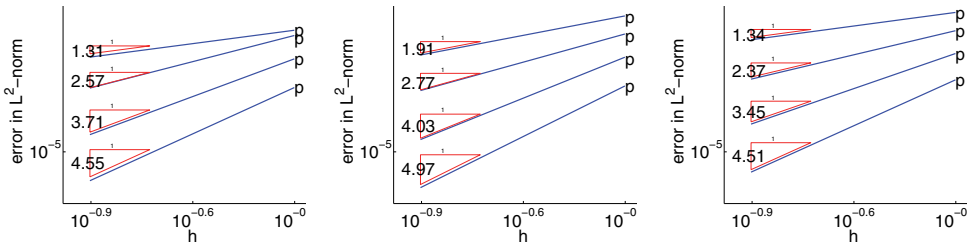
5. Numerical results. In this section, we present numerical results to verify our theoretical developments. To facilitate the computation of convergence rates, we shall consider two oceanic flows, the linear standing wave and the linear Kelvin wave, for which exact solutions are available. We adopt the nodal approach developed in [29] for our implementation on triangle meshes. The Crank–Nicholson discretization is used in the following computations to simply obtain second order accuracy in time.

We first consider the linear standing wave where we take $\Phi = g = 1$, $f = 0$ (zero Coriolis force), $\gamma = 0$ (zero bottom friction), and $\boldsymbol{\tau} = \mathbf{0}$ (zero wind stress). The computational domain is chosen as $\Omega = [0,1]^2$ with wall boundary conditions on $\partial\Omega$ and we take the following exact solution [30, 25, 24]:

$$\begin{aligned} \phi &= \cos(\pi x)\cos(\pi y)\cos\left(\sqrt{2}\pi t\right), \\ u &= \frac{1}{\sqrt{2}}\sin(\pi x)\cos(\pi y)\sin\left(\sqrt{2}\pi t\right), \\ v &= \frac{1}{\sqrt{2}}\cos(\pi x)\sin(\pi y)\sin\left(\sqrt{2}\pi t\right). \end{aligned}$$

To demonstrate the theoretical convergence rates, we take $\Delta t = 10^{-6}$ with 10^5 time steps.

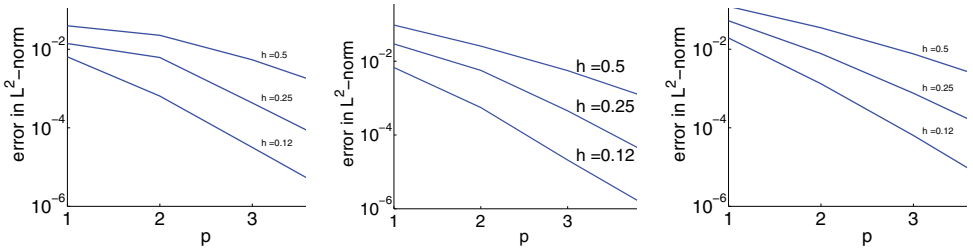
Figure 2 shows the h-convergence rate plots for three different values of λ, namely $\lambda = 0$ (this case is not covered by our analysis), $\lambda = 5$, and $\lambda = \sqrt{\Phi} = 1$ (upwind case). As can be observed, the convergence rate for the first two cases is $p + 1/2$, which is predicted by our proof for $\lambda > 0$. The upwind result is better than the theoretical prediction by half an order. We also note that the case with $\lambda = 5$ is less accurate than the others. Our numerical experiments (not included here) showed that the bigger λ is, the less accurate the result. This discrepancy may be due to the fact that the last scheme is more dissipative (larger λ); however, we leave the rigorous analysis for future work. This result also shows that the upwind HDG seems to be the most accurate among the penalty HDG family.



(a) HDG with $\lambda = 0$ (b) Upwind ($\lambda = 1$) HDG (c) HDG with $\lambda = 5$

FIG. 2. *h-convergence rate for linear standing wave problem: Log-log scale plot of the error of HDG methods in the L^2-norm, i.e.,* $\sqrt{\mathsf{E}} = \sqrt{0.5\sum_K \|\mathbb{P}\phi^e - \phi\|_K^2 + \|\mathbb{P}\boldsymbol{\vartheta}^e - \boldsymbol{\vartheta}\|_{K,\Phi}^2}$. *The mesh is refined in h for different polynomial orders from $p = 1$ to $p = 4$. The convergence is shown for three different mesh sizes $h = \{0.5, 0.25, 0.125\}$.*

We also show the p-convergence rates corresponding to $\lambda = 0$, $\lambda = \sqrt{\Phi} = 1$ (upwind case), and $\lambda = 5$ in Figure 3. As can be seen, the convergence curves in the logarithmic scale are almost straight lines (especially for $p = 2, 3, 4$), and hence the convergence rate is exponential in p. Again, the upwind HDG scheme is the most accurate in general.



(a) HDG with $\lambda = 0$ (b) Upwind ($\lambda = 1$) HDG (c) HDG with $\lambda = 5$

FIG. 3. *p-convergence rate for linear standing wave problem: log-log scale plot of the error of HDG methods in the L^2-norm, i.e.,* $\sqrt{\mathsf{E}} = \sqrt{0.5\sum_K \|\mathbb{P}\phi^e - \phi\|_K^2 + \|\mathbb{P}\boldsymbol{\vartheta}^e - \boldsymbol{\vartheta}\|_{K,\Phi}^2}$. *The mesh is refined in p for different mesh sizes, i.e.,* $\{0.5, 0.25, 0.125\}$. *The convergence is shown for four different solution orders $p = \{1, 2, 3, 4\}$.*

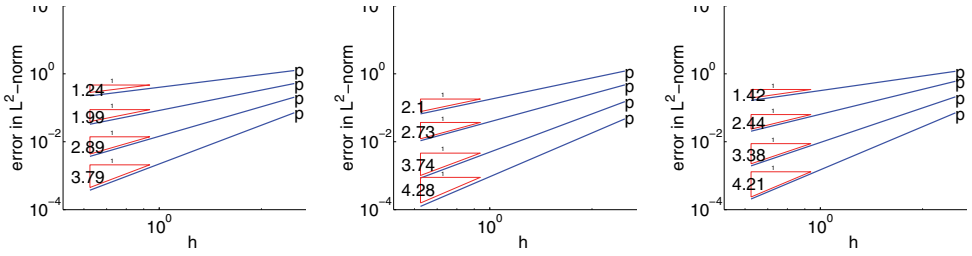
The next example considered in this paper is the linear Kelvin wave [21, 25, 24] where we take $\Phi = g = 1$, $f_0 = y_m = 0$, and $\beta = 1$ (nonzero Coriolis force), $\gamma = 0$ (zero

bottom friction), and $\boldsymbol{\tau} = \mathbf{0}$ (zero wind stress). The computational domain is chosen as $\Omega = [-10, 10] \times [-5, 5]$ (with periodic boundary condition in the x-direction and wall boundary condition in the y-direction) and we take the following exact solution:

$$\phi = 1 + \exp\left(-\frac{y^2}{2}\right) \exp\left[-\frac{(x+5-t)^2}{2}\right],$$
$$u = \exp\left(-\frac{y^2}{2}\right) \exp\left[-\frac{(x+5-t)^2}{2}\right],$$
$$v = 0,$$

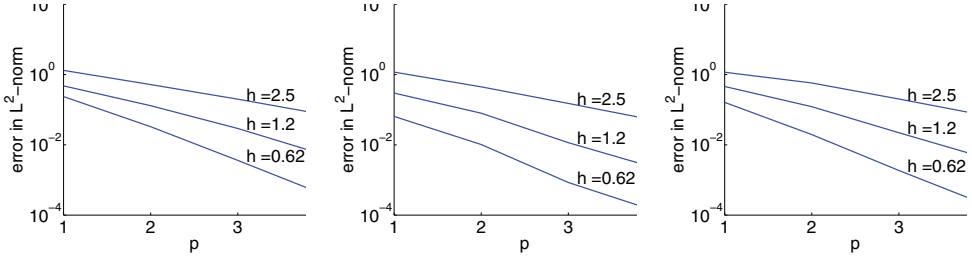
and $\Delta t = 10^{-4}$ with 10^4 time steps.

We again study the convergence rates and compare them with the theoretical rates. The results for h- and p-convergence are presented in Figures 4 and 5, respectively. For this example, the numerical convergence rate in h is compatible with the theoretical rates, i.e., $p + 1/2$, and the p-convergence rates are close to exponential. Note that for $\lambda = 0$ the rate is p, which is suboptimal. We again observe that the upwind HDG scheme is the most accurate, which is consistent with the results for the linear standing wave.



(a) HDG with $\lambda = 0$ (b) Upwind ($\lambda = 1$) HDG (c) HDG with $\lambda = 5$

FIG. 4. *h-convergence rate for the linear Kelvin wave problem: Log-log scale plot of the error of HDG methods in the L^2-norm, i.e., $\sqrt{\mathsf{E}} = \sqrt{0.5 \sum_K \|\mathbb{P}\phi^e - \phi\|_K^2 + \|\mathbb{P}\boldsymbol{\vartheta}^e - \boldsymbol{\vartheta}\|_{K,\Phi}^2}$. The mesh is refined in h for different polynomial orders from $p = 1$ to $p = 4$. The convergence is shown for three different mesh sizes $h = \{2.5, 1.25, 0.625\}$.*



(a) HDG with $\lambda = 0$ (b) Upwind ($\lambda = 1$) HDG (c) HDG with $\lambda = 5$

FIG. 5. *p-convergence rate for linear Kelvin wave problem: Log-log scale plot of the error of HDG methods in the L^2-norm, i.e., $\sqrt{\mathsf{E}} = \sqrt{0.5 \sum_K \|\mathbb{P}\phi^e - \phi\|_K^2 + \|\mathbb{P}\boldsymbol{\vartheta}^e - \boldsymbol{\vartheta}\|_{K,\Phi}^2}$. The mesh is refined in p for different mesh sizes in $h = \{2.5, 1.25, 0.625\}$. The convergence is shown for four different solution orders $p = \{1, 2, 3, 4\}$.*

6. Conclusions and future work. Using the Rankine–Hugoniot condition we have introduced a constructive methodology to systematically devise hybridized discontinuous Galerkin (HDG) methods for linear(ized) shallow water systems. The key step is the hybridization of upwind fluxes that is naturally available using the Rankine–Hugoniot condition for solving the Riemann problem. Using an energy approach we are able to rigorously analyze the stability and convergence of our proposed methods, in particular the upwind HDG scheme. We have also shown that the upwind HDG method can be naturally extended to a family of penalty HDG schemes that share the same stability and convergence rates. Numerical results for linear standing and Kelvin waves for oceanic shallow water systems are presented to verify our theoretical developments.

Ongoing work is to couple the proposed upwind HDG method with a standard DG approach in a semi-implicit, e.g., implicit-explicit (IMEX), framework for nonlinear shallow water systems. In particular, the former is expected to provide an efficient implicit solver for linear(ized) parts that contain fast waves, while the latter can employ large step sizes for the resulting nonlinear part for which fast waves are no longer present. This will be tested and demonstrated on regional and global hydrostatic oceanic and atmospheric flows. As part of the future work, we will study the dispersive and dissipative behaviors of our HDG schemes.

Appendix A. Complexity comparison between HDG and DG. In this section we attempt to derive conditions under which HDG can be more advantageous than DG. To that end, we assume that we are addressing either steady state problems or unsteady ones with implicit time-stepping schemes, for which the number of coupled unknowns can be used as the measure of complexity.[4] In order to have a definitive computation, we also restrict ourselves to structured meshes. In particular, we consider structured quadrilateral/hexahedral meshes in two dimensions/three dimensions with n elements along each dimension. To form triangular meshes, we split each quadrilateral into two triangles, while we construct tetrahedral meshes by partitioning each hexahedron into five tetrahedra (see [2, Fig. 3(d)] for a demonstration). We choose a nodal approach from [29], and it follows that there are $(p+1)^d$ degrees of freedom (dofs) for each quadrilateral/hexahedron and $\frac{(p+d)!}{p!\,d!}$ dofs for each triangle/tetrahedron, where, again, d is the spatial dimension.

Table 1 computes the number of coupled dofs on quadrilateral/hexahedral meshes for both HDG and DG methods in details. As can be seen, HDG becomes beneficial with $p \geq 2$ on quadrilateral meshes with $n \geq 3$; for $n = 1$, we need $p \geq 4$, and $p \geq 3$ for $n = 2$. On hexahedral meshes, for HDG to be more advantageous than DG, the solution order needs to be at least three, i.e., $p \geq 3$, and the mesh has at least $n = 4$ elements in each dimension; for $n < 4$, the saving from HDG starts with $p = 4$. Triangular/Tetrahedral meshes are more favorable to the HDG approach, as can be observed from Table 2. In particular, for triangular meshes with more than one element, i.e., $n > 1$, in each dimension, HDG starts to have a lesser number of coupled dofs even with a $p = 1$ solution; for $n = 1$, we need $p \geq 2$. For tetrahedral meshes, HDG is more beneficial when: (i) $p \geq 2$ and $n < 4$, or (ii) $p \geq 1$ and $n \geq 4$. Note that these conclusions are for HDG methods with the number of trace unknowns equal to the number of volume unknowns, i.e., HDG-I and HDG-III approaches. For HDG methods with a lesser number of trace unknowns relative to the number of volume unknowns, i.e., HDG-II and HDG-IV, the gain of HDG over DG is even greater.

[4] Of course, the sparsity structure plays an important role, but is ignored here for simplicity.

TABLE 1

The number of coupled dofs for HDG and DG on quadrilateral (quad) and hexahedral (hex) meshes. The last row is the condition under which HDG is more advantageous than DG. Here we assume that there is one volume (DG) unknown u and one trace (HDG) unknown $\hat{u}$.

	quad	hex
# elements	n^2	n^3
element dofs	$(p+1)^2$	$(p+1)^3$
DG dofs	$n^2 \times (p+1)^2$	$n^3 \times (p+1)^3$
# faces	$2n\,(n+1)$	$3n^2\,(n+1)$
face dofs	$(p+1)$	$(p+1)^2$
HDG dofs	$2n\,(n+1)\,(p+1)$	$3n^2\,(n+1)\,(p+1)^2$
$\frac{\text{HDG dofs}}{\text{DG dofs}} < 1$	$\boldsymbol{p > 1 + \frac{2}{n}}$	$\boldsymbol{p > 2 + \frac{3}{n}}$

TABLE 2

The number of coupled dofs for HDG and DG on triangular (tri) and tetrahedral (tet) meshes. The last row is the condition under which HDG is more advantageous than DG. Here we assume that there are one volume (DG) unknown u and one trace (HDG) unknown $\hat{u}$.

	tri	tet
# elements	$2n^2$	$5n^3$
element dofs	$\frac{(p+1)(p+2)}{2}$	$\frac{(p+1)(p+2)(p+3)}{6}$
DG dofs	$n^2\,(p+1)\,(p+2)$	$\frac{5n^3(p+1)(p+2)(p+3)}{6}$
# faces	$n\,(2n+1)$	$2n^2\,(3n+1)$
face dofs	$(p+1)$	$\frac{(p+1)(p+2)}{2}$
HDG dofs	$n\,(2n+1)\,(p+1)$	$2n^2\,(3n+1)\,\frac{(p+1)(p+2)}{2}$
$\frac{\text{HDG dofs}}{\text{DG dofs}} < 1$	$\boldsymbol{p > \frac{1}{n}}$	$\boldsymbol{p > \frac{3}{5} + \frac{6}{5n}}$

Appendix B. Extension to spatially varying mean flow. Let us first extend the HDG construction to a more general case in which the geopotential height Φ is spatially varying. In this case, the linearized shallow system (2.1) has $\frac{1}{2}\nabla\left(\Phi^2\right)$ as an additional term on the right-hand side of the last two equations. This is just a "forcing" (or source) term and hence does not change the differential operator. Thus, the upwind and Lax–Friedrichs fluxes remain unchanged. In other words, the four proposed HDG schemes are still the same. The theoretical results and proofs remain essentially the same with some minor changes. For example, the first assertion of Theorem 2.3 should now read "the trace unknowns are the L^2-projection of the upwind states." The reason is that Φ is no longer constant, and hence the upwind states are not polynomials. Note that for semi-implicit schemes, e.g., IMEX Runge–Kutta integrators, this extension is sufficient to separate the fast gravity wave.[5]

For implicit Newton-type methods, we need to develop HDG schemes for full linearized systems around points with spatially varying geopotential height and velocity. To this end, let us denote by $\boldsymbol{U} := (U, V)$ the mean flow velocity vector. After some algebraic calculations the linearized system around $(\Phi, \Phi\boldsymbol{U})$ reads

[5]We will report our work in this direction in the near future.

(B.1)
$$\frac{\partial}{\partial t}\begin{pmatrix}\phi\\ \Phi u\\ \Phi v\end{pmatrix}+\frac{\partial}{\partial x}\begin{pmatrix}\Phi u+U\phi\\ \Phi\phi+U\Phi u\\ U\Phi v\end{pmatrix}+\frac{\partial}{\partial y}\begin{pmatrix}\Phi v+V\phi\\ V\Phi u\\ \Phi\phi+V\Phi v\end{pmatrix}=\mathbf{g}\left(\Phi,\Phi\boldsymbol{U},\phi,\Phi\boldsymbol{\vartheta}\right),$$

where $\mathbf{g}(\cdot)$ is not a function of the derivatives of $\mathbf{u}:=(\phi,\Phi\boldsymbol{\vartheta})^T$, and hence does not contribute to the HDG flux construction. The tensor $\mathbf{A}$ now reads

$$\mathbf{A}_1=\begin{bmatrix}U&1&0\\ \Phi&U&0\\ 0&0&U\end{bmatrix},\quad \mathbf{A}_2=\begin{bmatrix}V&0&1\\ 0&V&0\\ \Phi&0&V\end{bmatrix},$$

which has three real eigenvalues

$$[c_1,c_2,c_3]:=\left[\boldsymbol{U}\cdot\mathbf{n}-\sqrt{\Phi},\boldsymbol{U}\cdot\mathbf{n},\boldsymbol{U}\cdot\mathbf{n}+\sqrt{\Phi}\right].$$

Without loss of generality, let us assume that $\sqrt{\Phi}>\boldsymbol{U}\cdot\mathbf{n}$ (the other case is treated similarly). The proposed upwind HDG framework then provides the following upwind HDG flux:

(B.2)
$$\hat{\mathbf{F}}\cdot\mathbf{n}=\begin{bmatrix}\boldsymbol{U}\cdot\mathbf{n}\phi+\Phi\boldsymbol{\vartheta}\cdot\mathbf{n}+\left(\frac{\boldsymbol{U}\cdot\mathbf{n}}{\sqrt{\Phi}}-1\right)\left(\Phi\boldsymbol{\vartheta}\cdot\mathbf{n}-\Phi\hat{\boldsymbol{\vartheta}}\cdot\mathbf{n}\right)\\ \phi\Phi\mathbf{n}+\boldsymbol{U}\cdot\mathbf{n}\Phi\boldsymbol{\vartheta}+\boldsymbol{U}\cdot\mathbf{n}\left(\Phi\boldsymbol{\vartheta}\cdot\mathbf{n}-\Phi\hat{\boldsymbol{\vartheta}}\cdot\mathbf{n}\right)\mathbf{n}-\mathbf{k}\cdot\left[\left(\Phi\boldsymbol{\vartheta}-\Phi\hat{\boldsymbol{\vartheta}}\right)\times\mathbf{n}\right](\mathbf{k}\times\mathbf{n})\end{bmatrix},$$

where $\mathbf{k}=(0,0,1)^T$. Note that the scalar and vector products in the second line of (B.2) are three-dimensional operators, and their operands are zero-extension into the third dimension if they are two-dimensional vectors. In this case, we can only reduce the number of trace unknowns down to two, namely $\Phi\hat{\boldsymbol{\vartheta}}$, which has one more trace unknown compared to the case with zero mean velocity studied above. For the Lax–Friedrichs flux, we are not be able to reduce any trace unknown, and hence have three trace unknowns $\left(\hat{\phi},\Phi\hat{\boldsymbol{\vartheta}}\right)$ instead of two $\Phi\hat{\boldsymbol{\vartheta}}$. Since the analysis of these schemes is beyond the scope of this paper, it will be presented elsewhere in our future work.

Acknowledgments. The author would like to thank Ellen B. Le and Stephen Shannon for careful proofreading of the paper. We are indebted to the anonymous referees for their critical and useful comments that improved this paper substantially.

REFERENCES

[1] V. Aizinger and C. Dawson, *A discontinuous Galerkin method for two-dimensional flow and transport in shallow water*, Adv. Water Resour., 25 (2002), pp. 67–84.
[2] T. Apel and N. Düvelmeyer, *Transformation of hexaedral finite element meshes into tetrahedral meshes according to quality criteria*, Computing, 71 (2003), pp. 293–304, https://doi.org/10.1007/s00607-003-0031-5.
[3] I. Babuvska and M. Suri, *The h-p version of the finite element method with quasi-uniform meshes*, RAIRO Modél. Math. Anal. Numér., 21 (1987), pp. 199–238.
[4] I. Babuvska and M. Suri, *The optimal convergence rate of the p-version of the finite element method*, SIAM J. Numer. Anal., 24 (1987), pp. 750–776, https://doi.org/10.1137/0724049.
[5] I. Babuvska and M. Suri, *The p and h-p versions of the finite element method, basic principles and properties*, SIAM Rev., 36 (1994), pp. 578–632, https://doi.org/10.1137/1036141.
[6] T. Bui-Thanh, *From Godunov to a unified hybridized discontinuous Galerkin framework for partial differential equations*, J. Comput. Phys., 295 (2015), pp. 114–146.

[7] T. Bui-Thanh, *From Rankine-Hugoniot condition to a constructive derivation of HDG methods*, Spectral and High Order Methods for Partial Differential Equations ICOSAHOM 2014, Lecture Notes in Comput. Sci. Eng. 106, Springer, Cham, 2015, pp. 483–491, https://doi.org/10.1007/978-3-319-19800-2_45.

[8] T. Bui-Thanh and O. Ghattas, *Analysis of an hp-nonconforming discontinuous Galerkin spectral element method for wave propagation*, SIAM J. Numer. Anal., 50 (2012), pp. 1801–1826, https://doi.org/10.1137/110828010.

[9] C. Canuto, M. Yousuff Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods*, Fundamentals in Single Domains, Sci. Comput., Springer-Verlag, Berlin, 2006.

[10] B. Cockburn, B. Dong, J. Guzmán, M. Restelli, and R. Sacco, *A hybridizable discontinuous Galerkin method for steady state convection-diffusion-reaction problems*, SIAM J. Sci. Comput., 31 (2009), pp. 3827–3846, https://doi.org/10.1137/080728810.

[11] B. Cockburn and J. Gopalakrishnan, *The derivation of hybridizable discontinuous Galerkin methods for Stokes flow*, SIAM J. Numer. Anal, 47 (2009), pp. 1092–1125, https://doi.org/10.1137/080726653.

[12] 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, https://doi.org/10.1137/070706616.

[13] B. Cockburn, J. Gopalakrishnan, N. C. Nguyen, J. Peraire, and F.-J. Sayas, *Analysis of HDG methods for Stokes flow*, Math. Comp., 80 (2011), pp. 723–760.

[14] B. Cockburn, J. Gopalakrishnan, and F.-J. Sayas, *A projection-based error analysis of HDG methods*, Math. Comp., 79 (2010), pp. 1351–1367.

[15] B. Cockburn, G. E. Karniadakis, and C.-W. Shu, *Discontinuous Galerkin Methods: Theory, Computation and Applications*, Lecture Notes in Comput. Sci. Eng. 11, Springer-Verlag, Berlin, Heidelberg, 2000, https://doi.org/10.1007/978-3-642-59721-3.

[16] B. Cockburn, F. Li, N.-C. Nguyen, and J. Peraire, *Hybridization and postprocessing techniques for mixed eigenfunctions*, SIAM J. Numer. Anal, 48 (2010), pp. 857–881, https://doi.org/10.1137/090765894.

[17] J. Cui and W. Zhang, *An analysis of HDG methods for the Helmholtz equation*, IMA J. Numer. Anal., 34 (2014), pp. 279–295.

[18] V. Dolejsi and M. Feistauer, *A semi-implicit discontinuous Galerkin finite element method for the numerical solution of inviscid compressible flow*, J. Comput. Phys., 198 (2004), pp. 727–746.

[19] F. Dupont and C. A. Lin, *The adaptive spectral element method and comparisons with more traditional formulations for ocean modeling*, J. Atmo. Oceanic Technol., 21 (2004), pp. 135–147, https://doi.org/10.1175/1520-0426(2004)021⟨0135:TASEMA⟩2.0.CO;2.

[20] H. Egger and J. Schöberl, *A hybrid mixed discontinuous Galerkin finite element method for convection-diffusion problems*, IMA J. Numer. Anal., 30 (2010), pp. 1206–1234.

[21] C. Eskilsson and S. J. Sherwin, *A triangular spectral/hp discontinuous Galerkin method for modelling* 2*d shallow water equations*, Internat. J. Numer. Methods Fluids, 45 (2004), pp. 605–623.

[22] M. Feistauer, V. Dolejsi, and V. Kucera, *On the discontinuous Galerkin method for the simulation of compressible flow with wide range of Mach numbers*, Comput. Vis. Sci., 10 (2007), pp. 17–27.

[23] R. Gandham, D. Medina, and T. Warburton, *GPU Accelerated Discontinuous Galerkin Methods for Shallow Water Equations*, https://arxiv.org/abs/1403.1661, 2014.

[24] F. X. Giraldo and M. Restelli, *High-order semi-implicit time-integrators for a triangular discontinous Galerkin oceanic shallow water model*, Int. J. Numer. Meth. Fluids, 63 (2010), pp. 1077–1102, https://doi.org/10.1002/fld.2118.

[25] F. X. Giraldo and T. Warburton, *A high-order triangular discontinuous Galerkin oceanic shallow water model*, Int. J. Numer. Meth. Fluids, 56 (2008), pp. 899–925, https://doi.org/10.1002/fld.1562.

[26] R. Griesmaier and P. Monk, *Error analysis for a hybridizable discontinuous Galerkin method for the Helmholtz equation*, J. Sci. Comput., 49 (2011), pp. 291–310.

[27] J. S. Hesthaven, S. Gottlieb, and D. Gottlieb, *Spectral Methods for Time-Dependent Problems*, Cambridge Monogr. Appl. Comput. Math. 21, Cambridge University Press, Cambridge, 2007.

[28] J. S. Hesthaven and T. Warburton, *Nodal high-order methods on unstructured grids. I. Time-domain solution of Maxwell's equations*, J. Comput. Phys., 181 (2002), pp. 186–221.

[29] J. S. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*, Texts Appl. Math. 54, Springer, Cham, 2008.

[30] M. Iskandarani, D. B. Haidvogel, and J. P. Boyd, *A staggered spectral element model with application to the oceanic shallow water equations*, Int. J. Numer. Meth. Fluids, 20 (1995), pp. 393–414, https://doi.org/10.1002/fld.1650200504.
[31] C. Johnson and J. Pitkäranta, *An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation*, Math. Comp., 46 (1986), pp. 1–26.
[32] 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.
[33] D. A. Kopriva, *Implementing Spectral Methods for Partial Differential Equations*, Springer, Berlin, 2009.
[34] E. J. Kubatko, J. J. Westerink, and C. Dawson, *Hp-discontinuous Galerkin methods for advection dominated problems in shallow water flow*, Comput. Method in Appl. Mech. Eng., 196 (2006), pp. 437–451, https://doi.org/10.1016/j.cma.2006.05.002.
[35] P. LeSaint and P. A. Raviart, *On a Finite Element Method for Solving the Neutron Transport Equation*, in Mathematical Aspects of Finite Element Methods in Partial Differential Equations, C. de Boor, ed., Academic Press, New York, 1974, pp. 89–123.
[36] H. Li and R. X. Liu, *The discontinuous Galerkin finite element method for the* $2D$ *shallow water equations*, Math. Comput. Simulation, 56 (2001), pp. 171–184.
[37] L. Li, S. Lanteri, and R. Perrrussel, *A hybridizable discontinuous Galerkin method for solving* $3D$ *time harmonic Maxwell's equations*, in Numerical Mathematics and Advanced Applications 2011, Springer, Heidelberg, 2013, pp. 119–128.
[38] D. Moro, N. C. Nguyen, and J. Peraire, *Navier-Stokes Solution using Hybridizable Discontinuous Galerkin Methods*, American Institute of Aeronautics and Astronautics, 2011-3407, Reston, VA, 2011.
[39] N. C. Nguyen, J. Peraire, and B. Cockburn, *An implicit high-order hybridizable discontinuous Galerkin method for linear convection-diffusion equations*, J. Comput. Phys., 228 (2009), pp. 3232–3254.
[40] N. C. Nguyen, J. Peraire, and B. Cockburn, *A hybridizable discontinous Galerkin method for Stokes flow*, Comput Method Appl. Mech. Eng., 199 (2010), pp. 582–597, https://doi.org/10.1016/j.cma.2009.10.007.
[41] N. C. Nguyen, J. Peraire, and B. Cockburn, *High-order implicit hybridizable discontinuous Galerkin method for acoustics and elastodynamics*, J. Comput. Phys., 230 (2011), pp. 3695–3718.
[42] N. C. Nguyen, J. Peraire, and B. Cockburn, *Hybridizable discontinuous Galerkin method for the time harmonic Maxwell's equations*, J. Comput. Phys., 230 (2011), pp. 7151–7175.
[43] N. C. Nguyen, J. Peraire, and B. Cockburn, *An implicit high-order hybridizable discontinuous Galerkin method for the incompressible Navier-Stokes equations*, J. Comput. Phys., 230 (2011), pp. 1147–1170.
[44] W. H. Reed and T. R. Hill, *Triangular Mesh Methods for the Neutron Transport Equation*, Tech. Report LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos, New Mexico, 1973.
[45] J. F. Remacle, S. S. Frazao, X. G. Li, and M. S. Shephard, *An adaptive discretization of shallow-water equations based on discontinuous Galerkin methods*, Internat. J. Numer. Methods Fluids, 52 (2006), pp. 903–923.
[46] M. Restelli, *Semi-Lagrangian and Semi-Implicit Discontinuous Galerkin Methods for Atmospheric Modeling Applications*, Ph.D. thesis, Politecnico di Milano, Milano, Italy, 2007.
[47] M. Restelli and F. X. Giraldo, *A conservative discontinuous Galerkin semi-implicit formulation for the Navier-Stokes equations in non-hydrostatic mesoscale modeling*, SIAM J. Sci. Comput., 31 (2009), pp. 2231–2257, https://doi.org/10.1137/070708470.
[48] A. Robert, J. Henderson, and C. Turnbull, *An implicit time integration scheme for baroclinic models of the atmosphere*, Mon. Weather Rev., 100 (1972), pp. 329–335, https://doi.org/10.1080/07055900.1981.9649098.
[49] O. San and K. Kara, *High-order accurate spectral difference method for shallow water equations*, Int. J. Res. Rev. Appl. Sci., 6 (2011), pp. 41–54.
[50] D. Schwanenberg and J. Köngeter, *A discontinuous Galerkin method for the shallow water equations with source terms*, in Discontinuous Galerkin Methods, Lect. Notes Comput. Sci. Eng. 11, B. Cockburn, G. E. Karniadakis, and C-W. Shu, eds., Springer, Berlin, 2000, pp. 419–424.
[51] E. F. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*, Springer, Berlin, Heidelberg, 2009.
[52] Y. Xing and X. Zhang, *Positivity-preserving well-balanced discontinuous Galerkin methods for the shallow water equations on unstructured triangular meshes*, J. Sci. Comput., 57 (2013), pp. 19–41.