

OPTIMAL ERROR ESTIMATES OF DISCONTINUOUS GALERKIN METHODS WITH GENERALIZED FLUXES FOR WAVE EQUATIONS ON UNSTRUCTURED MESHES

ZHENG SUN AND YULONG XING

ABSTRACT. L^2 stable discontinuous Galerkin method with a family of numerical fluxes was studied for the one-dimensional wave equation by Cheng, Chou, Li, and Xing in [Math. Comp. 86 (2017), pp. 121–155]. Although optimal convergence rates were numerically observed with wide choices of parameters in the numerical fluxes, their error estimates were only proved for a sub-family with the construction of a local projection. In this paper, we first complete the one-dimensional analysis by providing optimal error estimates that match all numerical observations in that paper. The key ingredient is to construct an optimal global projection with the characteristic decomposition. We then extend the analysis on optimal error estimate to multidimensions by constructing a global projection on unstructured meshes, which can be considered as a perturbation away from the local projection studied by Cockburn, Gopalakrishnan, and Sayas in [Math. Comp. 79 (2010), pp. 1351–1367] for hybridizable discontinuous Galerkin methods. As a main contribution, we use a novel energy argument to prove the optimal approximation property of the global projection. This technique does not require explicit assembly of the matrix for the perturbed terms and hence can be easily used for unstructured meshes in multidimensions. Finally, numerical tests in two dimensions are provided to validate our analysis is sharp and at least one of the unknowns will degenerate to suboptimal rates if the assumptions are not satisfied.

1. INTRODUCTION

In this paper, we study discontinuous Galerkin (DG) methods with generalized fluxes for the linear wave equation

$$(1.1) \quad u_{tt} = \Delta u,$$

in the form of first-order hyperbolic system,

$$(1.2) \quad u_t = \nabla \cdot \mathbf{q}, \quad \mathbf{q}_t = \nabla u.$$

Here $\mathbf{x} = (x_1, \dots, x_d) \in \Omega \subset \mathbb{R}^d$, $d \geq 1$ and $t \in [0, T]$. $u = u(\mathbf{x}, t)$ and $\mathbf{q} = \mathbf{q}(\mathbf{x}, t)$ are unknowns to be solved. One- and two-dimensional time-domain linear Maxwell's equations can be viewed as a special case of (1.2). The goal of the paper is to provide the DG approximation of (1.2) with a class of generalized numerical fluxes, and establish rigorous analysis to demonstrate that such DG methods have simultaneous

Received by the editor April 12, 2020, and, in revised form, September 20, 2020.

2020 *Mathematics Subject Classification*. Primary 65M12, 65M15, 65M60; Secondary 35L05, 35L45.

Key words and phrases. Discontinuous Galerkin methods, wave equations, numerical fluxes, optimal error estimates, global projections.

The work of the second author was partially supported by the NSF grant DMS-1753581.

optimal convergence rates for both u and \mathbf{q} in L^2 norm. For simplicity, we consider rectangular (cubic) domains with periodic boundaries, although similar analysis can be extended to more complicated geometries with boundary conditions of other types.

The wave equation (1.1) widely occurs in scientific and engineering applications, such as acoustics, electromagnetics and fluid dynamics, for modeling propagation of mechanical or light waves. A vast amount of research has been dedicated to numerical approximations of (1.1). These efforts include the finite difference methods, the finite volume methods, the integral equation methods, the continuous Galerkin finite element methods, the mixed finite element methods, the DG methods and so on. See [6, 9, 16, 20] and references therein. In this paper, we confine our attention to the DG methods, which belong to a class of finite element methods using discontinuous piecewise polynomial spaces. The methods were originally proposed by Reed and Hill in [34] for the transport equation, then received their major development in a series of work by Cockburn et al. for hyperbolic conservation laws [14, 15, 17–19]. The methods feature with advantages, such as preservation of local conservation, flexibility of fitting complex geometries, good $h-p$ adaptivity and high parallel efficiency, and have been favored in various of applications. For approximations of the wave equation, a stream of research concerns the first-order form (1.2), sometimes as a particular case of linear symmetric hyperbolic systems. These studies include the space-time DG methods [21, 32], the staggered DG method [9], the hybridizable DG methods [16, 33, 37], the DG methods in [6] and [22], and the central DG method [29]. Another stream of research is based on the second-order form (1.1), and numerical techniques for steady state elliptic problems can be adopted in the treatment of the discrete Laplacian. See the interior penalty DG methods [24], the hybridizable DG methods [11, 35], and the local DG methods [8, 43], for some of the related works.

It is known that the choice of numerical fluxes is one of the main ingredients for designing DG methods. It has crucial influences on the stability, convergence rate and dispersive behavior of the numerical schemes [1, 2, 21, 25, 30, 38, 45]. For linear equations, the common choices of the numerical fluxes include the upwind and alternating fluxes. To provide more flexibility of numerical dissipation with potential applications to complex systems, there is a growing attention on studying numerical fluxes with general patterns very recently, for example, the upwind-biased fluxes for linear hyperbolic equations [28, 31, 44], the generalized alternating fluxes for the convection diffusion equation [7] and the Burgers–Poisson equation in [27], the $\alpha\beta$ -fluxes for the linear two-way wave equation [6] and other equations with high-order derivatives [23], the generalized local Lax–Friedrichs fluxes for nonlinear scalar conservation laws [26] and the generalized fluxes for Hamiltonian partial differential equations [42]. The error analysis of these schemes may possibly require the construction of a global projection operator, based on the technique developed in [5] and is usually for structured meshes in one or two dimensions.

Our work focuses on the first-order form (1.2), as a continuation of [6] on analyzing subtle effects of different numerical fluxes on the convergence rates. In [6], Cheng et al. studied the DG method for (1.2) in one spatial dimension, and presented a systematic study on a family of numerical fluxes parametrized by three constants α , β_1 and β_2 . The same family of numerical fluxes has also been analyzed earlier by Ainsworth [2] et al. with emphasis on dispersive behaviors. Numerically,

optimal convergence rates were reported in [6] with very general choices of parameters, while at the time, optimal error estimates were only proved for a sub-family among them. This sub-family of numerical fluxes, up to a small relaxation, was identified in that paper as the $\alpha\beta$ -fluxes: $\beta_1 \geq 0, \beta_2 \geq 0$ and

$$(1.3) \quad \alpha^2 + \beta_1\beta_2 = \frac{1}{4}.$$

The assumption on nonnegativity of β_1 and β_2 was required for the stability of the scheme.

The first contribution of our paper is to provide optimal error estimates for unproved cases in [6]. Instead of assuming (1.3), we prove optimal convergence rates for all of the stable DG schemes under a relaxed assumption: $\beta_1 \geq 0, \beta_2 \geq 0$ and

$$(1.4) \quad \alpha^2 + \beta_1\beta_2 \neq 0,$$

which matches the numerical observations in [6]. The key ingredient is to construct a novel global projection pair of the unknowns u and q , which is coupled through the flux terms. We apply the characteristic decomposition on the definition of the projection pair, so that the conditions on the transformed unknowns are decoupled. We then apply the generalized Gauss–Radau projections [7, 31] on the transformed unknowns, and transform them back, to obtain the desired tailored operators for our problem. In the special case of $\alpha\beta$ -fluxes, namely $\alpha^2 + \beta_1\beta_2 = 1/4$, the projection is a linear combination of classical Gauss–Radau projections, and can be locally constructed as that in [6]. In general, the projection is a global operator coupled through all mesh cells.

Our second contribution is to extend the analysis on optimal error estimate to unstructured simplex meshes in multidimensions. We firstly note that the multi-dimensional DG scheme generalized from [6] has close connections with the hybridizable DG methods in [33] and [16]. Indeed, if one rewrites the hybridizable DG methods into the local DG form [12], it can be seen that the two schemes are very similar, except that the hybridizable DG methods particularly uses the $\alpha\beta$ -fluxes (1.3) under a different parametrization,¹ and we refer to Remark 3.3 for details. Optimal error estimates were obtained in [16] for the hybridizable DG methods with the $\alpha\beta$ -fluxes, by employing the local projection pair constructed in [13].

The optimal error estimate is provided in this paper for the multi-dimensional DG scheme with generalized numerical fluxes, including but not limited to the $\alpha\beta$ -fluxes. Compared with the one-dimensional case, the analysis of generalized fluxes is more involved in multidimensions, and the technique of characteristic decomposition can no longer be applied. The complication mainly comes from two facts. Firstly, since $\mathbf{q} = \nabla u$ is a vector, the symmetry between \mathbf{q} and u is broken. Secondly, there is no global projection, similar to the one-dimensional generalized Gauss–Radau projections, at hand on unstructured meshes, which can be served as building blocks to construct the required projection. As a result, we have to go through the entire procedure to construct the global projection pair and derive

¹The choice of $\alpha\beta$ -fluxes seems to be essential in the hybridizable DG methods to guarantee the locality of the solvers when inverting the spatial operator. While in particular applications of solving (1.2), the problem is hyperbolic in nature and explicit time stepping methods are widely used, therefore other types of numerical fluxes are also acceptable in these applications.

its approximation property. Previously, in [31] and [7], the unisolvency and approximation property of (global) generalized Gauss–Radau projections (2.14) were proved by explicitly assembling the linear system for the perturbation term with respect to the classical Gauss–Radau projection. Similar technique was also used in an earlier work [5] for analyzing the generalized KdV equation. However, this argument may not be easily generalized to unstructured meshes in multidimensions. Instead, we apply a novel energy argument to circumvent the steps for assembling matrices, and the proof is sketched as follows. We consider the difference between the global projection and the local projection constructed in [13]; then with an energy argument, it can be shown that the difference is a high-order term; finally, the unisolvency and approximation property follow as consequences of those of the local projection pair. With the global projection pair, we obtain optimal error estimate for both u and \mathbf{q} if

$$(1.5) \quad \alpha^2 + \beta_1\beta_2 \neq 0, \quad \beta_1 \geq 0 \quad \text{and} \quad \beta_2 > 0.$$

Compared with the one-dimensional case, we require the additional assumption on nondegeneracy of β_2 . The condition in (1.5) is also considered to be necessary among stable DG schemes: if the condition is violated, we observe numerically that at least one of u and \mathbf{q} will converge at a suboptimal rate on general unstructured meshes.

The main novelty of this paper is to establish the global projections and analyze their optimal convergence properties, with techniques of characteristic decomposition in one dimension and the energy argument in multidimensions. The constructed projections can also be used in other contexts for error estimates. For example, besides the DG schemes in the first-order form (1.2), the same analysis can be extended to DG schemes in [8, 43] based on the second-order mixed form

$$(1.6) \quad u_{tt} = \nabla \cdot \mathbf{q}, \quad \mathbf{q} = \nabla u.$$

A brief discussion is outlined in Section 2.4 for the one-dimensional case. These projections can also be used for analyzing DG schemes with generalized fluxes for other equations with high-order spatial derivatives, including the heat equation, time-dependent biharmonic equation, Schrödinger equation and dispersive equation studied in [23]. The one-dimensional and multi-dimensional analysis on these equations will be reported in a future work.

The rest of the paper is planned as follows. We present the one-dimensional analysis in Section 2 and multi-dimensional analysis in Section 3. Numerical tests in two dimensions are provided in Section 4 to validate our analysis. Conclusions are given in Section 5 at the end.

2. ONE-DIMENSIONAL CASE

In this section, we analyze the DG scheme for (1.2) in one dimension

$$(2.1) \quad u_t = q_x, \quad q_t = u_x.$$

We start with introducing notations and the DG scheme in Section 2.1, then state the optimal error estimates in Section 2.2, with the properties of the required projection postponed to Section 2.3. Finally, extension to the DG scheme based on the second-order in time formulation is discussed in Section 2.4.

2.1. Notations and the DG scheme. Let $\Omega = \cup_{j=1}^N I_j$, where $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, be a quasi-uniform mesh partition of the spatial domain with mesh size h . The finite element space is chosen as the discontinuous piecewise polynomial space

$$(2.2) \quad V_h = \{v : v|_{I_j} \in P_k(I_j), j = 1, \dots, N\}.$$

Here $P_k(I_j)$ is the linear space spanned by polynomials of degree less or equal to k on I_j . Note that functions in V_h can be double-valued across cell interfaces. We denote by $v_{j+\frac{1}{2}}^+$ and $v_{j+\frac{1}{2}}^-$ the right and left limit of v at $x_{j+\frac{1}{2}}$. $\{v\}_{j+\frac{1}{2}} = (v_{j+\frac{1}{2}}^+ + v_{j+\frac{1}{2}}^-)/2$ and $[[v]]_{j+\frac{1}{2}} = v_{j+\frac{1}{2}}^+ - v_{j+\frac{1}{2}}^-$ are used to represent averages and jumps at $x_{j+\frac{1}{2}}$. We introduce the shorthand notation $(v, w) = \sum_j (v, w)_{I_j}$, with $(v, w)_{I_j} = \int_{I_j} v w dx$, the L^2 inner product on Ω . The associated L^2 norm is denoted by $\|v\| = \sqrt{(v, v)}$.

The DG scheme in [6] for approximating (2.1) is given as follows: find $u_h, q_h \in V_h$, such that

$$(2.3a) \quad ((u_h)_t, w)_{I_j} + (q_h, w_x)_{I_j} - (\mathcal{F}_q(q_h, u_h)w^-)_{j+\frac{1}{2}} + (\mathcal{F}_q(q_h, u_h)w^+)_{j-\frac{1}{2}} = 0, \quad \forall w \in V_h,$$

$$(2.3b) \quad ((q_h)_t, v)_{I_j} + (u_h, v_x)_{I_j} - (\mathcal{F}_u(u_h, q_h)v^-)_{j+\frac{1}{2}} + (\mathcal{F}_u(u_h, q_h)v^+)_{j-\frac{1}{2}} = 0, \quad \forall v \in V_h.$$

Here the numerical flux is given by

$$(2.4a) \quad \mathcal{F}_q(q_h, u_h) = \{q_h\} + \alpha [[q_h]] + \beta_1 [[u_h]],$$

$$(2.4b) \quad \mathcal{F}_u(u_h, q_h) = \{u_h\} + \beta_2 [[q_h]] - \alpha [[u_h]].$$

The scheme can be equivalently written in the global form: find $q_h, u_h \in V_h$, such that

$$(2.5) \quad ((q_h)_t, v)_{I_j} + ((u_h)_t, w)_{I_j} + a_h(q_h, u_h; v, w) = 0, \quad \forall v, w \in V_h,$$

where

$$(2.6) \quad a_h(q_h, u_h; v, w) = (u_h, v_x) + (q_h, w_x) + \sum_j (\mathcal{F}_q(q_h, u_h) [[w]] + \mathcal{F}_u(u_h, q_h) [[v]])_{j+\frac{1}{2}}.$$

Proposition 2.1 (Stability). *The bilinear form a_h satisfies the following property*

$$(2.7) \quad a_h(q_h, u_h; q_h, u_h) = \sum_j \left(\beta_1 [[u_h]]^2 + \beta_2 [[q_h]]^2 \right)_{j+\frac{1}{2}}.$$

Suppose $\beta_1 \geq 0$ and $\beta_2 \geq 0$. Then the DG scheme (2.3) is stable, with the energy equality

$$(2.8) \quad \frac{1}{2} \frac{d}{dt} (\|q_h\|^2 + \|u_h\|^2) + a_h(q_h, u_h; q_h, u_h) = 0.$$

Remark 2.1 (Flux parameters). Note that in (2.4), $\alpha = 0, \beta_1 = \beta_2 = 1/2$ gives the upwind flux for the hyperbolic system; $\alpha = 1/2, \beta_1 = \beta_2 = 0$ retrieves the alternating flux widely used for local DG methods; $\alpha = \beta_1 = \beta_2 = 0$ corresponds to the central flux. Compared with these commonly used fluxes, one can see from Proposition 2.1 that the generalized fluxes allow the fine-tuning of numerical dissipation in the scheme, similar to upwind-biased fluxes investigated in [31]. In addition, without being restricted to specific choices, the generalized fluxes are more flexible and could be potentially useful in the design of numerical schemes for complex systems.

2.2. Error estimates. The key ingredient for error estimates is to construct a projection pair Π_q^* and Π_u^* , satisfying the following requirements

$$(2.9a) \quad (\Pi_q^* q, v)_{I_j} = (q, v)_{I_j}, \quad \forall v \in P_{k-1}(I_j), \quad \forall j,$$

$$(2.9b) \quad (\Pi_u^* u, w)_{I_j} = (u, w)_{I_j}, \quad \forall w \in P_{k-1}(I_j), \quad \forall j,$$

$$(2.9c) \quad \mathcal{F}_q(\Pi_q^* q, \Pi_u^* u) = \mathcal{F}_q(q, u), \quad \text{at } x = x_{j+\frac{1}{2}}, \quad \forall j,$$

$$(2.9d) \quad \mathcal{F}_u(\Pi_u^* u, \Pi_q^* q) = \mathcal{F}_u(u, q), \quad \text{at } x = x_{j+\frac{1}{2}}, \quad \forall j.$$

Lemma 2.1. *Suppose $\alpha^2 + \beta_1\beta_2 \neq 0$. Then the projection pair Π_q^* and Π_u^* in (2.9) is well-defined. Furthermore, we have²*

$$(2.10) \quad \|q - \Pi_q^* q\|^2 + \|u - \Pi_u^* u\|^2 \leq C_{\alpha, \beta} h^{2k+2} (|q|_{H^{k+1}}^2 + |u|_{H^{k+1}}^2).$$

The proof of Lemma 2.1 is postponed to Section 2.3.

With the definition of Π_q^* and Π_u^* in (2.9), it can be easily shown that

$$(2.11) \quad a_h(q - \Pi_q^* q, u - \Pi_u^* u; v, w) = 0, \quad \forall v, w \in V_h.$$

By utilizing the same argument as in [6, Theorem 2.5], the following error estimate of the semidiscrete DG scheme can be obtained. See also Appendix A for details.

Theorem 2.1 (Error estimate). *Suppose $\beta_1 \geq 0, \beta_2 \geq 0$, and $\alpha^2 + \beta_1\beta_2 \neq 0$. Then the scheme (2.3) has the error estimate*

$$(2.12) \quad (\|q - q_h\|^2 + \|u - u_h\|^2)^{\frac{1}{2}} \Big|_{t=T} \leq (\|\Pi_q^* q - q_h\|^2 + \|\Pi_u^* u - u_h\|^2)^{\frac{1}{2}} \Big|_{t=0} + C(T+1)h^{k+1}.$$

Here C is a constant that is independent of h , but may depend on $(k+1)$ th-order Sobolev norms of q, u, q_t and u_t , as well as values of α, β_1 and β_2 .

Remark 2.2. The condition $\alpha^2 + \beta_1\beta_2 \neq 0$ is considered to be necessary as well. In [6, Table 3.8], the suboptimal convergence rate was observed for $(\|q - q_h\|^2 + \|u - u_h\|^2)^{\frac{1}{2}}$ with $\alpha = \beta_1 = 0$ and $\beta_2 = 1/10$.

If a strongly stable explicit Runge–Kutta method is used for time integration [40], one can apply [41, Corollary 3.1] to obtain the fully discrete error estimate.

Theorem 2.2 (Fully discrete error estimate). *After applying a p th-order strongly stable Runge–Kutta method to discretize (2.3), the resulted fully discrete solution satisfies the error estimate*

$$(2.13) \quad (\|q^n - q_h^n\|^2 + \|u^n - u_h^n\|^2)^{\frac{1}{2}} \leq (\|\Pi_q^* q^0 - q_h^0\|^2 + \|\Pi_u^* u^0 - u_h^0\|^2)^{\frac{1}{2}} + C(T+1) (\tau^p + h^{k+1}).$$

Here $q^n(x) = q(x, n\tau)$ and $u^n(x) = u(x, n\tau)$ with τ being the time step size. For $i = 0$ or n , q_h^i and u_h^i are the fully discrete solutions at the i th time step.

Remark 2.3 (Temporal discretization methods). For the fully discrete scheme, the choice of the time integrator is not quite essential. One can also use other numerical schemes in time, such as implicit Runge–Kutta methods and multi-step methods. As long as the stability can be ensured, we expect a similar error estimate as that in (2.13). We also note that changing flux parameters would perturb the spectrum of the discrete DG operator, and may either increase or decrease the maximum

²Without further specification, we use C , possibly with subscript(s), for generic constants independent of mesh size h throughout the paper.

allowable time step associated with the stability region of the time integrator. For particular parameters and given temporal discretization methods, it is possible to use eigenvalue analysis to find the time step constraint, while having an analytical characterization in general settings could be challenging.

2.3. Proof of Lemma 2.1. Under general settings, the resulted projection pair may not be locally defined. We hence need to introduce generalized Gauss–Radau projections as building blocks.

Lemma 2.2 (Generalized Gauss–Radau projections, [7, 31]). *Suppose $\lambda \neq 0$. Then there exists a uniquely defined linear operator Π^λ satisfying*

$$(2.14a) \quad (\Pi^\lambda u, w)_{I_j} = (u, w)_{I_j}, \quad w \in P_{k-1}(I_j), \quad \forall j,$$

$$(2.14b) \quad \{\Pi^\lambda u\} + \lambda \llbracket \Pi^\lambda u \rrbracket = \{u\} + \lambda \llbracket u \rrbracket, \quad \text{at } x = x_{j+\frac{1}{2}}, \quad \forall j.$$

Furthermore, the following approximation property holds

$$(2.15) \quad \|u - \Pi^\lambda u\|^2 + h \sum_j \llbracket u - \Pi^\lambda u \rrbracket_{j+\frac{1}{2}}^2 \leq C_\lambda h^{2k+2} |u|_{H^{k+1}}^2.$$

Remark 2.4. When $\lambda = \pm 1/2$, the projection in Lemma 2.2 retrieves the classical Gauss–Radau projection, which is a local operator. We denote by $\Pi^+ = \Pi^{+\frac{1}{2}}$ and $\Pi^- = \Pi^{-\frac{1}{2}}$.

When $\beta_1 = \beta_2 = 0$, the definition of the projection pair Π_q^* and Π_u^* in (2.9) is decoupled, and one can easily observe that $\Pi_q^* = \Pi^\alpha$ and $\Pi_u^* = \Pi^{-\alpha}$. The complication for constructing the projection pair (2.9) in the general case, is that conditions on Π_u^* and Π_q^* are coupled through flux terms in (2.9c) and (2.9d). In [6], the authors point out that one can explicitly define

$$(2.16a) \quad \Pi_q^* q = \Pi^+ \left(\left(\frac{1}{2} + \alpha \right) q + \beta_1 u \right) + \Pi^- \left(\left(\frac{1}{2} - \alpha \right) q - \beta_1 u \right),$$

$$(2.16b) \quad \Pi_u^* u = \Pi^+ \left(\left(\frac{1}{2} - \alpha \right) u + \beta_2 q \right) + \Pi^- \left(\left(\frac{1}{2} + \alpha \right) u - \beta_2 q \right).$$

for (2.9) when $\alpha^2 + \beta_1\beta_2 = 1/4$. Here Π^+ and Π^- are classical Gauss–Radau projections specified in Remark 2.4. Note we have

$$\beta_2 \left(\left(\frac{1}{2} \pm \alpha \right) q \pm \beta_1 u \right) = \left(\frac{1}{2} \pm \alpha \right) \left(\left(\frac{1}{2} \mp \alpha \right) u \pm \beta_2 q \right)$$

under the condition $\alpha^2 + \beta_1\beta_2 = 1/4$. As a result, the Gauss–Radau projections are applied to $(1/2 + \alpha)q + \beta_1 u$ and $(1/2 - \alpha)q - \beta_1 u$ separately, which indicates that certain transformed unknowns may admit to decoupled conditions in (2.9). This motivates us to look into the characteristic decomposition of the coefficient matrix

$A = \begin{pmatrix} \alpha & \beta_1 \\ \beta_2 & -\alpha \end{pmatrix}$ for finding the proper transformation. With these in mind, we have the following proof of Lemma 2.1.

Proof. Let $A = \begin{pmatrix} \alpha & \beta_1 \\ \beta_2 & -\alpha \end{pmatrix}$. Note that the two eigenvalues of A are

$$(2.17) \quad \lambda_1 = \sqrt{\alpha^2 + \beta_1\beta_2} \quad \text{and} \quad \lambda_2 = -\sqrt{\alpha^2 + \beta_1\beta_2}.$$

Under the assumption $\alpha^2 + \beta_1\beta_2 \neq 0$, the matrix A is nonsingular and diagonalizable.

We now switch to vector notations. Assume

$$(2.18) \quad \Lambda = S^{-1}AS = \text{diag}(\lambda_1, \lambda_2), \quad \lambda_1\lambda_2 \neq 0 \quad \text{and} \quad \begin{pmatrix} q \\ u \end{pmatrix} = S \begin{pmatrix} \sigma \\ \rho \end{pmatrix}.$$

Let us denote by

$$(2.19) \quad \Pi^* \begin{pmatrix} q \\ u \end{pmatrix} = \begin{pmatrix} \Pi_q^* q \\ \Pi_u^* u \end{pmatrix},$$

and we will show that

$$(2.20) \quad \Pi^* \begin{pmatrix} q \\ u \end{pmatrix} = S \begin{pmatrix} \Pi^{\lambda_1} \sigma \\ \Pi^{\lambda_2} \rho \end{pmatrix}$$

satisfies the definition (2.9), where Π^{λ_1} and Π^{λ_2} are the generalized Gauss–Radau projections defined in Lemma 2.2.

Similar to the scalar case, let us use $(\cdot, \cdot)_{I_j}$ for L^2 inner product on I_j . $\{\cdot\}$ and $\llbracket \cdot \rrbracket$ are used to represent the component-wise averages and jumps. Note (2.9a) and (2.9b) can be equivalently written as

$$(2.21) \quad \left(\Pi^* \begin{pmatrix} q \\ u \end{pmatrix}, \begin{pmatrix} v \\ w \end{pmatrix} \right)_{I_j} = \left(\begin{pmatrix} q \\ u \end{pmatrix}, \begin{pmatrix} v \\ w \end{pmatrix} \right)_{I_j}, \quad \forall w, v \in P_{k-1}(I_j), \quad \forall j.$$

By taking $\begin{pmatrix} v \\ w \end{pmatrix} = (S^{-1})^T \begin{pmatrix} \tilde{v} \\ \tilde{w} \end{pmatrix}$ and using the linearity of Π^* , one can see that (2.21) holds if and only if

$$(2.22) \quad \left(\Pi^* \begin{pmatrix} \sigma \\ \rho \end{pmatrix}, \begin{pmatrix} \tilde{v} \\ \tilde{w} \end{pmatrix} \right)_{I_j} = \left(\begin{pmatrix} \sigma \\ \rho \end{pmatrix}, \begin{pmatrix} \tilde{v} \\ \tilde{w} \end{pmatrix} \right)_{I_j}, \quad \forall \tilde{v}, \tilde{w} \in P_{k-1}(I_j), \quad \forall j.$$

Similarly, (2.9c) and (2.9d) can be rephrased in the vector form

$$(2.23) \quad \left\{ \Pi^* \begin{pmatrix} q \\ u \end{pmatrix} \right\} + A \llbracket \Pi^* \begin{pmatrix} q \\ u \end{pmatrix} \rrbracket = \left\{ \begin{pmatrix} q \\ u \end{pmatrix} \right\} + A \llbracket \begin{pmatrix} q \\ u \end{pmatrix} \rrbracket, \quad \text{at } x = x_{j+\frac{1}{2}}, \quad \forall j.$$

Left multiplying S^{-1} on both sides of (2.23) leads to

$$(2.24) \quad \left\{ \Pi^* \begin{pmatrix} \sigma \\ \rho \end{pmatrix} \right\} + \Lambda \llbracket \Pi^* \begin{pmatrix} \sigma \\ \rho \end{pmatrix} \rrbracket = \left\{ \begin{pmatrix} \sigma \\ \rho \end{pmatrix} \right\} + \Lambda \llbracket \begin{pmatrix} \sigma \\ \rho \end{pmatrix} \rrbracket, \quad \text{at } x = x_{j+\frac{1}{2}}, \quad \forall j.$$

Since Λ is a diagonal matrix, one can then rewrite (2.22) and (2.24) into the scalar form,

$$(2.25a) \quad (\Pi_q^* \sigma, \tilde{v})_{I_j} = (\sigma, \tilde{v})_{I_j}, \quad \forall \tilde{v} \in P_{k-1}(I_j), \quad \forall j,$$

$$(2.25b) \quad \{\Pi_q^* \sigma\} + \lambda_1 \llbracket \Pi_q^* \sigma \rrbracket = \{\sigma\} + \lambda_1 \llbracket \sigma \rrbracket, \quad \text{at } x = x_{j+\frac{1}{2}}, \quad \forall j;$$

$$(2.25c) \quad (\Pi_u^* \rho, \tilde{w})_{I_j} = (\rho, \tilde{w})_{I_j}, \quad \forall \tilde{w} \in P_{k-1}(I_j), \quad \forall j,$$

$$(2.25d) \quad \{\Pi_u^* \rho\} + \lambda_2 \llbracket \Pi_u^* \rho \rrbracket = \{\rho\} + \lambda_2 \llbracket \rho \rrbracket, \quad \text{at } x = x_{j+\frac{1}{2}}, \quad \forall j;$$

which is equivalent to (2.9). According to Lemma 2.2, $\Pi_q^* \sigma$ and $\Pi_u^* \rho$ are uniquely defined as

$$(2.26) \quad \Pi_q^* \sigma = \Pi^{\lambda_1} \sigma, \quad \Pi_u^* \rho = \Pi^{\lambda_2} \rho.$$

Therefore,

$$(2.27) \quad \Pi^* \begin{pmatrix} q \\ u \end{pmatrix} = S \Pi^* \begin{pmatrix} \sigma \\ \rho \end{pmatrix} = S \begin{pmatrix} \Pi^{\lambda_1} \sigma \\ \Pi^{\lambda_2} \rho \end{pmatrix}$$

is also uniquely defined, and satisfies the definition (2.9). The approximation property (2.10) of Π^* follows from those of Π^{λ_1} and Π^{λ_2} in Lemma 2.2. \square

Remark 2.5 (Rewriting with generalized Gauss–Radau projections). With $\lambda = \lambda_1 = \sqrt{\alpha^2 + \beta_1\beta_2}$, the transformation matrix may take the form $S = \begin{pmatrix} \alpha + \lambda & \beta_1 \\ \beta_2 & -\alpha - \lambda \end{pmatrix}$. Then the projection in Lemma 2.1 can be explicitly written as

$$(2.28a) \quad \Pi_q^* q = \frac{1}{2\lambda} (\Pi^\lambda ((\lambda + \alpha) q + \beta_1 u) + \Pi^{-\lambda} ((\lambda - \alpha) q - \beta_1 u)),$$

$$(2.28b) \quad \Pi_u^* u = \frac{1}{2\lambda} (\Pi^\lambda ((\lambda - \alpha) u + \beta_2 q) + \Pi^{-\lambda} ((\lambda + \alpha) u - \beta_2 q)),$$

which retrieves the local projection (2.16) in [6] in the special case $\alpha^2 + \beta_1\beta_2 = \lambda^2 = 1/4$. Furthermore, if $\beta_1 = \beta_2 = 0$, then we have $\Pi_q^* = \Pi^\alpha$ and $\Pi_u^* = \Pi^{-\alpha}$.

Remark 2.6 (The critical case $\alpha^2 + \beta_1\beta_2 = 0$). The optimal error estimate in Theorem 2.1 excludes the case of $\alpha^2 + \beta_1\beta_2 = 0$. Recall that, in the proof of Lemma 2.2 in [7, Lemma 3.2], the following estimate for the approximation constant has been obtained when λ is bounded

$$(2.29) \quad C_\lambda \leq C \left(1 + (|1 - 2\lambda|^2 - |1 + 2\lambda|^2)^{-1} \right).$$

Note that the right side blows up when λ shrinks to zero, which indicates the failure of the approximation property (2.15) in the critical case. Our constructed projection $(\Pi_q^* q, \Pi_u^* u)$ is essentially a linear combination of generalized Gauss–Radau projections, see Remark 2.5. Hence it suffers a similar singularity with the critical parameter $\alpha^2 + \beta_1\beta_2 = 0$. In this situation, the optimal error estimate in Theorem 2.1 would also break up, since C in (2.13) depends on C_λ .

Remark 2.7 (The four-parameter family of fluxes). The analysis can be easily extended to a larger class of numerical fluxes with $A = \begin{pmatrix} \alpha_1 & \beta_1 \\ \beta_2 & \alpha_2 \end{pmatrix}$, where $\alpha_2 \neq -\alpha_1$ in general. The scheme (2.3) is stable under the assumption $\beta_1 \geq 0, \beta_2 \geq 0$, and $(\alpha_1 + \alpha_2)^2 \leq 4\beta_1\beta_2$. On the top of it, if we further assume $\alpha_1\alpha_2 + \beta_1\beta_2 \neq 0$, it can be shown that A is diagonalizable and nonsingular, and the projection in Lemma 2.1 is still well-defined with the same approximation estimate.

2.4. Extension to the second-order in time DG scheme. In this paper, we mainly focus on the first-order system of the wave equation (2.1). Directly solving the second-order in time wave equation (1.1) may be preferred under some circumstances, as illustrated in [8, 43]. The same optimal error estimate can also be obtained for the local DG scheme in [43], which is based on the second-order mixed form

$$(2.30) \quad u_{tt} = q_x, \quad q = u_x.$$

The local DG scheme takes the form of

$$(2.31a) \quad ((u_h)_{tt}, w)_{I_j} + (q_h, w_x)_{I_j} - (\mathcal{F}_q(q_h, u_h)w^-)_{j+\frac{1}{2}} + (\mathcal{F}_q(q_h, u_h)w^+)_{j-\frac{1}{2}} = 0, \quad \forall w \in V_h,$$

$$(2.31b) \quad (q_h, v)_{I_j} + (u_h, v_x)_{I_j} - (\mathcal{F}_u(u_h, q_h)v^-)_{j+\frac{1}{2}} + (\mathcal{F}_u(u_h, q_h)v^+)_{j-\frac{1}{2}} = 0, \quad \forall v \in V_h,$$

where the numerical flux is given by (2.4).

Differentiate (2.31b) with respect to t . After combining the two equations, it yields that

$$(2.32) \quad ((q_h)_t, v) + ((u_h)_{tt}, w) + \tilde{a}_h(q_h, u_h; v, w) = 0.$$

where

$$(2.33) \quad \begin{aligned} &\tilde{a}_h(q_h, u_h; v, w) \\ &= ((u_h)_t, v_x) + (q_h, w_x) + \sum_j (\mathcal{F}_u((u_h)_t, (q_h)_t) \llbracket v \rrbracket + \mathcal{F}_q(q_h, u_h) \llbracket w \rrbracket)_{j+\frac{1}{2}}. \end{aligned}$$

Proposition 2.2 (Stability). *The bilinear form \tilde{a}_h satisfies the following property*

$$(2.34) \quad \tilde{a}_h(q_h, u_h; q_h, (u_h)_t) = \frac{1}{2} \frac{d}{dt} \sum_j \left(\beta_1 \llbracket u_h \rrbracket^2 + \beta_2 \llbracket q_h \rrbracket^2 \right)_{j+\frac{1}{2}}.$$

The DG method (2.31) is energy-conserving, in the sense that $\frac{d}{dt} \mathcal{E}(q_h, u_h) = 0$ with

$$(2.35) \quad \mathcal{E}(q_h, u_h) = \left(\|q_h\|^2 + \|(u_h)_t\|^2 + \sum_j \left(\beta_1 \llbracket u_h \rrbracket^2 + \beta_2 \llbracket q_h \rrbracket^2 \right)_{j+\frac{1}{2}} \right)^{\frac{1}{2}}.$$

Theorem 2.3. *With $\mathcal{E}(\cdot, \cdot)$ defined in (2.35), the DG scheme (2.31) has the error estimate*

$$(2.36) \quad \mathcal{E}(q - q_h, u - u_h)|_{t=T} \leq \mathcal{E}(\Pi_q^* q - q_h, \Pi_u^* u - u_h)|_{t=0} + C(T + 1)h^{k+1}.$$

In particular, with the initial condition $q_h(x, 0) = \Pi_q^* q(x, 0)$ and $u_h(x) = \Pi_u^* u(x, 0)$, we have the error estimate in L^2 norm

$$(2.37) \quad (\|q - q_h\|^2 + \|u_t - (u_h)_t\|^2)^{\frac{1}{2}} \leq C(T + 1)h^{k+1}.$$

Here, both constants C are independent of h , but may depend on $(k + 1)$ th-order Sobolev norms of q , q_t , u_t and u_{tt} , as well as values of α , β_1 and β_2 .

Theorem 2.4. *With $q_h(x, 0) = \Pi_q^* q(x, 0)$ and $u_h(x, 0) = \Pi_u^* u(x, 0)$, we have*

$$(2.38) \quad \sup_{t \in [0, T]} \|u - u_h\| \leq C(T + 1)h^{k+1},$$

where the constant C is independent of h , but may depend on $(k + 1)$ th-order Sobolev norms of q , q_t , u_t and u_{tt} , as well as values of α , β_1 and β_2 .

The proof of Theorem 2.3 is very similar to that of Theorem 2.1 in Appendix A, and is skipped. The proof of Theorems 2.4 is provided in Appendix B.

3. MULTI-DIMENSIONAL CASE

In this section, we extend the previous analysis to multidimensions with unstructured simplex meshes. Notations and the DG scheme are given in Section 3.1 and the optimal error estimates are stated in Section 3.2. The detailed proof on properties of the required projection pair is given in Section 3.3.

3.1. Notations and the DG scheme. Let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, be a rectangular (cubic) domain with periodic boundaries. Consider a quasi-uniform triangulation of the domain $\Omega = \cup_{K \in \mathcal{T}_h} K$ with simplex mesh cells. The collection of faces (mesh skeleton) is denoted by Γ . Given a face $F \in \partial K$, we denote by \mathbf{n} the unit outer normal along F with respect to K . With the mesh cell K unspecified, we use \mathbf{n}_F for the unit normal across F , whose direction is not essential. We also assume there exists a fixed constant vector $\mathbf{r} \in \mathbb{R}^2$, $|\mathbf{r}| = 1$, such that $|\mathbf{r} \cdot \mathbf{n}_F| \geq \kappa > 0$.

For finite element spaces, let

$$(3.1) \quad V_h = \{v_h : v_h|_K \in P_k(K), \forall K \in \mathcal{T}_h\}$$

be the space of discontinuous piecewise polynomials, where $P_k(K)$ is the linear space spanned by polynomials on K of degree less or equal to k . We use the notations $\mathbf{V}_h = [V_h]^d$ and $\mathbf{P}_k(K) = [P_k(K)]^d$ for corresponding product spaces for vector functions. Given a face F , we denote by K^+ and K^- the two neighboring cells. w^\pm and \mathbf{v}^\pm are the traces of w and \mathbf{v} on $K^+ \cap K^-$ taken within K^\pm . \mathbf{n}^\pm is the unit outer normal of F from K^\pm . The following notations will be used for averages and jumps across a face F

$$(3.2) \quad \{w\} = \frac{1}{2} (w^+ + w^-), \quad \{\mathbf{v}\} = \frac{1}{2} (\mathbf{v}^+ + \mathbf{v}^-),$$

$$(3.3) \quad \llbracket w\mathbf{n} \rrbracket = w^+ \mathbf{n}^+ + w^- \mathbf{n}^-, \quad \llbracket \mathbf{v} \cdot \mathbf{n} \rrbracket = \mathbf{v}^+ \cdot \mathbf{n}^+ + \mathbf{v}^- \cdot \mathbf{n}^-,$$

which are all single-valued on F . For ease of presentation, we denote by

$$(3.4) \quad (\cdot, \cdot) = \sum_{K \in \mathcal{T}_h} (\cdot, \cdot)_K \quad \text{and} \quad \langle \cdot, \cdot \rangle = \sum_{F \in \Gamma} \langle \cdot, \cdot \rangle_F,$$

where

$$(3.5) \quad (u, w)_K = \int_K u w dx, \quad (\mathbf{q}, \mathbf{v})_K = \int_K \mathbf{q} \cdot \mathbf{v} dx,$$

$$(3.6) \quad \langle \mu, \nu \rangle_F = \int_F \mu \nu ds, \quad \langle \boldsymbol{\zeta}, \boldsymbol{\xi} \rangle_F = \int_F \boldsymbol{\zeta} \cdot \boldsymbol{\xi} ds.$$

$\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ is used for the L^2 norm over the domain Ω and $\|\cdot\|_\Gamma = \sqrt{\langle \cdot, \cdot \rangle}$ is used for the L^2 norm over the mesh skeleton Γ . Note the following identities hold for jump terms

$$(3.7) \quad \llbracket w\mathbf{n} \rrbracket = \mathbf{n}_F (\llbracket w\mathbf{n} \rrbracket \cdot \mathbf{n}_F), \quad |\llbracket w\mathbf{n} \rrbracket|^2 = (\llbracket w\mathbf{n} \rrbracket \cdot \mathbf{n}_F)^2, \quad \mathbf{r} \cdot \llbracket w\mathbf{n} \rrbracket = \llbracket \mathbf{r}w \cdot \mathbf{n} \rrbracket,$$

and integration over faces [3]

$$(3.8) \quad \sum_K \langle w, \mathbf{v} \cdot \mathbf{n} \rangle_{\partial K} = \langle \{w\mathbf{n}\}, \llbracket \mathbf{v} \rrbracket \rangle + \langle \llbracket w\mathbf{n} \rrbracket, \{\mathbf{v}\} \rangle.$$

Now we are ready to state the numerical scheme for (1.2): find $u_h \in V_h$ and $\mathbf{q}_h \in \mathbf{V}_h$, such that

$$(3.9a) \quad ((u_h)_t, w)_K + (\mathbf{q}_h, \nabla w)_K - \langle \mathcal{F}_q(\mathbf{q}_h, u_h), \mathbf{n}w \rangle_{\partial K} = 0, \quad \forall w \in V_h,$$

$$(3.9b) \quad ((\mathbf{q}_h)_t, \mathbf{v})_K + (u_h, \nabla \cdot \mathbf{v})_K - \langle \mathcal{F}_u(u_h, \mathbf{q}_h), \mathbf{n}, \mathbf{v} \rangle_{\partial K} = 0, \quad \forall \mathbf{v} \in \mathbf{V}_h,$$

where the numerical flux is chosen as

$$(3.10a) \quad \mathcal{F}_q(\mathbf{q}_h, u_h) = \{\mathbf{q}_h\} - \alpha \llbracket \mathbf{q}_h \cdot \mathbf{n} \rrbracket - \beta_1 \llbracket u_h \mathbf{n} \rrbracket,$$

$$(3.10b) \quad \mathcal{F}_u(u_h, \mathbf{q}_h) = \{u_h\} - \beta_2 \llbracket \mathbf{q}_h \cdot \mathbf{n} \rrbracket + \alpha \cdot \llbracket u_h \mathbf{n} \rrbracket.$$

Here

$$(3.11) \quad \boldsymbol{\alpha} = \alpha \text{sign}(\mathbf{r} \cdot \mathbf{n}_F) \mathbf{n}_F$$

and α , β_1 and β_2 are parameters.

After summing over all mesh cells K , the scheme (3.9) can be written in the following global form: find $\mathbf{q}_h \in \mathbf{V}_h$ and $u_h \in V_h$, such that

$$(3.12) \quad ((\mathbf{q}_h)_t, \mathbf{v}) + ((u_h)_t, w) + a_h(\mathbf{q}_h, u_h; \mathbf{v}, w) = 0, \quad \forall \mathbf{v} \in \mathbf{V}_h, \quad \forall w \in V_h,$$

where

$$(3.13) \quad \begin{aligned} a_h(\mathbf{q}_h, u_h; \mathbf{v}, w) &= (\mathbf{q}_h, \nabla w) + (u_h, \nabla \cdot \mathbf{v}) - \sum_K (\langle \mathcal{F}_q(\mathbf{q}_h, u_h), \mathbf{n}w \rangle_{\partial K} + \langle \mathcal{F}_u(u_h, \mathbf{q}_h) \mathbf{n}, \mathbf{v} \rangle_{\partial K}) \\ &= (\mathbf{q}_h, \nabla w) + (u_h, \nabla \cdot \mathbf{v}) - \langle \mathcal{F}_q(\mathbf{q}_h, u_h), \llbracket w \mathbf{n} \rrbracket \rangle - \langle \mathcal{F}_u(u_h, \mathbf{q}_h), \llbracket \mathbf{v} \cdot \mathbf{n} \rrbracket \rangle. \end{aligned}$$

We start by presenting the stability of the proposed method.

Proposition 3.1 (Stability). *The bilinear form a_h satisfies the following property*

$$(3.14) \quad a_h(\mathbf{q}_h, u_h; \mathbf{q}_h, u_h) = \beta_1 \|\llbracket u_h \mathbf{n} \rrbracket\|_\Gamma^2 + \beta_2 \|\llbracket \mathbf{q}_h \cdot \mathbf{n} \rrbracket\|_\Gamma^2.$$

Suppose $\beta_1 \geq 0$ and $\beta_2 \geq 0$. Then the DG scheme (3.12) is stable, with the energy equality

$$(3.15) \quad \frac{1}{2} \frac{d}{dt} (\|\mathbf{q}_h\|^2 + \|u_h\|^2) + a_h(\mathbf{q}_h, u_h; \mathbf{q}_h, u_h) = 0.$$

Proof. With integration by parts, it can be seen that

$$(3.16) \quad \begin{aligned} a_h(\mathbf{q}_h, u_h; \mathbf{q}_h, u_h) &= (\mathbf{q}_h, \nabla u_h) + (u_h, \nabla \cdot \mathbf{q}_h) - \langle \mathcal{F}_q(\mathbf{q}_h, u_h), \llbracket u_h \mathbf{n} \rrbracket \rangle - \langle \mathcal{F}_u(u_h, \mathbf{q}_h), \llbracket \mathbf{q}_h \cdot \mathbf{n} \rrbracket \rangle \\ &= (\nabla \cdot (u_h \mathbf{q}_h), 1) - \langle \mathcal{F}_q(\mathbf{q}_h, u_h), \llbracket u_h \mathbf{n} \rrbracket \rangle - \langle \mathcal{F}_u(u_h, \mathbf{q}_h), \llbracket \mathbf{q}_h \cdot \mathbf{n} \rrbracket \rangle \\ &= \sum_K \langle u_h, \mathbf{q}_h \cdot \mathbf{n} \rangle_{\partial K} - \langle \mathcal{F}_q(\mathbf{q}_h, u_h), \llbracket u_h \mathbf{n} \rrbracket \rangle - \langle \mathcal{F}_u(u_h, \mathbf{q}_h), \llbracket \mathbf{q}_h \cdot \mathbf{n} \rrbracket \rangle \\ &= \sum_K \langle u_h, \mathbf{q}_h \cdot \mathbf{n} \rangle_{\partial K} - \langle \{\mathbf{q}_h\}, \llbracket u_h \mathbf{n} \rrbracket \rangle - \langle \{u_h\}, \llbracket \mathbf{q}_h \cdot \mathbf{n} \rrbracket \rangle \\ &\quad + \langle \alpha \llbracket \mathbf{q}_h \cdot \mathbf{n} \rrbracket + \beta_1 \llbracket u_h \mathbf{n} \rrbracket, \llbracket u_h \mathbf{n} \rrbracket \rangle + \langle \beta_2 \llbracket \mathbf{q}_h \cdot \mathbf{n} \rrbracket - \alpha \cdot \llbracket u_h \mathbf{n} \rrbracket, \llbracket \mathbf{q}_h \cdot \mathbf{n} \rrbracket \rangle \\ &= \beta_1 \|\llbracket u_h \mathbf{n} \rrbracket\|_\Gamma^2 + \beta_2 \|\llbracket \mathbf{q}_h \cdot \mathbf{n} \rrbracket\|_\Gamma^2, \end{aligned}$$

where we have used the definition of numerical flux (3.10) in the second last equality, and the identity (3.8) in the last equality. The stability (3.15) of the DG scheme follows from choosing $\mathbf{v} = \mathbf{q}_h$ and $w = u_h$ in (3.12). \square

3.2. Error estimates. Similar to the one-dimensional case, the key ingredient is to construct the projection pair satisfying

$$(3.17a) \quad (\Pi_{\mathbf{q}}^* \mathbf{q}, \mathbf{v})_K = (\mathbf{q}, \mathbf{v})_K, \quad \forall \mathbf{v} \in \mathbf{P}_{k-1}(K), \quad \forall K \in \mathcal{T}_h,$$

$$(3.17b) \quad (\Pi_u^* u, w)_K = (u, w)_K, \quad \forall w \in P_{k-1}(K), \quad \forall K \in \mathcal{T}_h,$$

$$(3.17c) \quad \langle \mathcal{F}_q(\Pi_{\mathbf{q}}^* \mathbf{q}, \Pi_u^* u), \mathbf{n}_F \mu \rangle_F = \langle \mathcal{F}_q(\mathbf{q}, u), \mathbf{n}_F \mu \rangle_F, \quad \forall \mu \in P_k(F), \quad \forall F \in \Gamma,$$

$$(3.17d) \quad \langle \mathcal{F}_u(\Pi_u^* u, \Pi_{\mathbf{q}}^* \mathbf{q}), \nu \rangle_F = \langle \mathcal{F}_u(u, \mathbf{q}), \nu \rangle_F, \quad \forall \nu \in P_k(F), \quad \forall F \in \Gamma.$$

Lemma 3.1. *Suppose $\beta_1 \geq 0$, $\beta_2 > 0$ and $\alpha^2 + \beta_1\beta_2 \neq 0$. Then the projection pair in (3.17) is well-defined. Furthermore, we have*

$$(3.18) \quad \|\mathbf{q} - \Pi_{\mathbf{q}}^* \mathbf{q}\|^2 + \|u - \Pi_u^* u\|^2 + h \left(\|\llbracket (\mathbf{q} - \Pi_{\mathbf{q}}^* \mathbf{q}) \cdot \mathbf{n}_F \rrbracket\|_{\Gamma}^2 + \|\llbracket (u - \Pi_u^* u) \mathbf{n}_F \rrbracket\|_{\Gamma}^2 \right) \leq C_* h^{2k+2} (|\mathbf{q}|_{\mathbf{H}^{k+1}}^2 + |u|_{H^{k+1}}^2).$$

Here C_* is a constant independent of h , but may depend on values of α , β_1 and β_2 . It may also depend on κ if $\beta_1 = 0$.

The proof of Lemma 3.1 is postponed to the next subsection. Next, we can establish the optimal L^2 error estimate of the proposed methods.

Theorem 3.1 (Error estimate). *With α , β_1 and β_2 prescribed in Lemma 3.1, the scheme (3.9) admits the following error estimate.*

$$(3.19) \quad (\|\mathbf{q} - \mathbf{q}_h\|^2 + \|u - u_h\|^2)^{\frac{1}{2}} \Big|_{t=T} \leq (\|\Pi_{\mathbf{q}}^* \mathbf{q} - \mathbf{q}_h\|^2 + \|\Pi_u^* u - u_h\|^2)^{\frac{1}{2}} \Big|_{t=0} + C(T+1)h^{k+1},$$

where C is a constant independent of h , but depends on C_* in Lemma 3.1 and $(k + 1)$ th-order Sobolev norms of \mathbf{q} , u , \mathbf{q}_t and u_t .

Proof. We denote by

$$(3.20) \quad e_{\mathbf{q}} = \mathbf{q} - \mathbf{q}_h \quad \text{and} \quad e_u = u - u_h.$$

It can be seen that the following error equation holds

$$(3.21) \quad ((e_{\mathbf{q}})_t, \mathbf{v}) + ((e_u)_t, w) + a_h(e_{\mathbf{q}}, e_u; \mathbf{v}, w) = 0, \quad \forall \mathbf{v} \in \mathbf{V}_h, \quad \forall w \in V_h.$$

Let us then decompose the error terms as the projection error and the projected error

$$(3.22) \quad e_{\mathbf{q}} = \eta_{\mathbf{q}} + \Pi_{\mathbf{q}}^* e_{\mathbf{q}} \quad \text{and} \quad e_u = \eta_u + \Pi_u^* e_u,$$

where $\eta_{\mathbf{q}} = \mathbf{q} - \Pi_{\mathbf{q}}^* \mathbf{q}$ and $\eta_u = u - \Pi_u^* u$. According to the construction of $\Pi_{\mathbf{q}}^*$ and Π_u^* in (3.17), we have

$$(3.23) \quad a_h(\eta_{\mathbf{q}}, \eta_u; \mathbf{v}, w) = 0, \quad \forall \mathbf{v} \in \mathbf{V}_h, \quad \forall w \in V_h.$$

Hence (3.21) can be rewritten as

$$(3.24) \quad ((\Pi_{\mathbf{q}}^* e_{\mathbf{q}})_t, \mathbf{v}) + ((\Pi_u^* e_u)_t, w) + a_h(\Pi_{\mathbf{q}}^* e_{\mathbf{q}}, \Pi_u^* e_u; \mathbf{v}, w) = -((\eta_{\mathbf{q}})_t, \mathbf{v}) - ((\eta_u)_t, w).$$

Then by taking $\mathbf{v} = \Pi_{\mathbf{q}}^* \mathbf{q}$ and $w = \Pi_u^* u$, it can be deduced that

$$(3.25) \quad \frac{1}{2} \frac{d}{dt} (\|\Pi_{\mathbf{q}}^* e_{\mathbf{q}}\|^2 + \|\Pi_u^* e_u\|^2) \leq -((\eta_{\mathbf{q}})_t, \Pi_{\mathbf{q}}^* e_{\mathbf{q}}) - ((\eta_u)_t, \Pi_u^* e_u).$$

Here we have used the fact that $a_h(\Pi_{\mathbf{q}}^* e_{\mathbf{q}}, \Pi_u^* e_u; \Pi_{\mathbf{q}}^* e_{\mathbf{q}}, \Pi_u^* e_u) \geq 0$ from Proposition 3.1 under the assumptions $\beta_1 \geq 0$ and $\beta_2 \geq 0$. One can then apply the Cauchy–Schwarz inequality and the approximation property in Lemma 3.1 to obtain

$$(3.26) \quad \begin{aligned} \frac{d}{dt} (\|\Pi_{\mathbf{q}}^* e_{\mathbf{q}}\|^2 + \|\Pi_u^* e_u\|^2)^{\frac{1}{2}} &\leq (\|(\eta_{\mathbf{q}})_t\|^2 + \|(\eta_u)_t\|^2)^{\frac{1}{2}} = (\|\eta_{\mathbf{q}t}\|^2 + \|\eta_{ut}\|^2)^{\frac{1}{2}} \\ &\leq Ch^{k+1} (|\mathbf{q}_t(\cdot, t)|_{\mathbf{H}^{k+1}} + |u_t(\cdot, t)|_{H^{k+1}}), \end{aligned}$$

which implies

(3.27)

$$\begin{aligned} & (\|\Pi_{\mathbf{q}}^* e_{\mathbf{q}}\|^2 + \|\Pi_u^* e_u\|^2)^{\frac{1}{2}} \Big|_{t=T} \\ & \leq (\|\Pi_{\mathbf{q}}^* e_{\mathbf{q}}\|^2 + \|\Pi_u^* e_u\|^2)^{\frac{1}{2}} \Big|_{t=0} + CT h^{k+1} \sup_{t \in [0, T]} (|\mathbf{q}_t|_{\mathbf{H}^{k+1}} + |u_t|_{H^{k+1}}). \end{aligned}$$

The proof can be completed after invoking triangle inequality with respect to the product norm $\sqrt{\|\cdot\|^2 + \|\cdot\|^2}$ and the approximation property of $\Pi_{\mathbf{q}}^*$ and Π_u^* in Lemma 3.1. □

Remark 3.1 (Fully discrete error estimate). Similar to that in one dimension, if explicit Runge–Kutta methods are used for time marching, one can obtain fully discrete error estimates using [41, Corollary 3.1].

Remark 3.2. If one uses the standard L^2 projections, instead of $\Pi_{\mathbf{q}}^*$ and Π_u^* in (3.17) for error estimates, only suboptimal convergence can be obtained. This suboptimal rate is $(k + 1/2)$ th-order if $\beta_1 > 0$ and $\beta_2 > 0$, and is k th-order in general.

Remark 3.3 (Connections with hybridizable DG methods). We remark the case $\alpha^2 + \beta_1\beta_2 = 1/4$ and $\beta_2 > 0$ has been studied in [16] in the context of hybridizable DG method (in particular, the LDG-hybridizable method). The coefficients of the numerical flux are given under a different parametrization [12]

$$(3.28a) \quad \mathcal{F}_{\mathbf{q}}(\mathbf{q}_h, u_h) = \left(\frac{\tau^-}{\tau^- + \tau^+}\right) \mathbf{q}_h^+ + \left(\frac{\tau^+}{\tau^- + \tau^+}\right) \mathbf{q}_h^- - \left(\frac{\tau^+ \tau^-}{\tau^- + \tau^+}\right) \llbracket u_h \mathbf{n} \rrbracket,$$

$$(3.28b) \quad \mathcal{F}_u(u_h, \mathbf{q}_h) = \left(\frac{\tau^+}{\tau^- + \tau^+}\right) u_h^+ + \left(\frac{\tau^-}{\tau^- + \tau^+}\right) u_h^- - \left(\frac{1}{\tau^- + \tau^+}\right) \llbracket \mathbf{q}_h \cdot \mathbf{n} \rrbracket,$$

which can be expressed in the form of (3.10) with

$$\alpha = \frac{1}{2} - \frac{\tau^-}{\tau^- + \tau^+}, \quad \beta_1 = \frac{\tau^+ \tau^-}{\tau^- + \tau^+}, \quad \beta_2 = \frac{1}{\tau^- + \tau^+}.$$

Remark 3.4 (Numerical fluxes and energy conservation). From the energy estimate in Proposition 3.1 and the prescribed family of flux parameters in Lemma 3.1, it can be seen that we need the numerical scheme to be energy-dissipative for optimal convergence in Theorem 3.1. This has also been numerically validated in Tables 4.9, 4.10 and 4.11: if we take $\beta_1 = \beta_2 = 0$ for an energy-conserving scheme, the numerical scheme may lose order on unstructured meshes. This is in contrast to the one-dimensional case, where the same choices of numerical fluxes will lead to optimal energy-conserving schemes. However, we point out it is possible to use the second-order in time formulation to develop an energy-conserving DG method with optimal convergence rate. This has been studied in the context of hybridizable DG method in [11].

Remark 3.5 (With other boundary conditions). Although for simplicity we only consider periodic boundaries, the analysis can also be extended to boundary conditions of other types. For example, if homogeneous Dirichlet boundaries $u = 0$ are assumed, one can keep the numerical flux as (3.10) on interior faces Γ^0 , and replace its definition with

$$(3.29) \quad \mathcal{F}_{\mathbf{q}}(\mathbf{q}_h, u_h) = (\mathbf{q}_h - u_h \mathbf{n}) \Big|_K, \quad \mathcal{F}_u(u, \mathbf{q}_h) = 0, \quad \forall F \in K \cap \Gamma^\partial$$

along boundary faces Γ^∂ , as studied in the hybridizable DG method [16]. The projection in Lemma 3.1 can still be constructed for the optimal error analysis.

Remark 3.6 (In heterogeneous media). The analysis can also be extended to problems in heterogeneous media. The treatment of variable coefficients follows similar lines along [16] and [11].

3.3. Proof of Lemma 3.1. In this subsection, we provide the proof of Lemma 3.1. We will start by presenting a special case when the projections $\Pi_{\mathbf{q}}^*$ and Π_u^* defined in (3.17) become the local projections. The general case will be studied as a perturbation away from these local projections, and the main analytical tool is based on an energy argument.

3.3.1. Local projections. When $\alpha^2 + \beta_1\beta_2 = 1/4$ and $\beta_2 > 0$, the projection pair $\Pi_{\mathbf{q}}^*$ and Π_u^* becomes the local projection studied in [13]. In particular, let us denote by $\bar{\Pi}_{\mathbf{q}}$ and $\bar{\Pi}_u$ the projection pair corresponding to $\alpha = 0$ and $\beta_1 = \beta_2 = 1/2$, which satisfies the following equations

$$(3.30a) \quad (\bar{\Pi}_{\mathbf{q}}\mathbf{q}, \mathbf{v})_K = (\mathbf{q}, \mathbf{v})_K, \quad \forall \mathbf{v} \in \mathbf{P}_{k-1}(K), \forall K \in \mathcal{T}_h,$$

$$(3.30b) \quad (\bar{\Pi}_u u, w)_K = (u, w)_K, \quad \forall w \in P_{k-1}(K), \forall K \in \mathcal{T}_h,$$

$$(3.30c)$$

$$\left\langle \{\bar{\Pi}_{\mathbf{q}}\mathbf{q}\} - \frac{1}{2} \llbracket \bar{\Pi}_u u \mathbf{n} \rrbracket, \mathbf{n}_F \mu \right\rangle_F = \left\langle \{\mathbf{q}\} - \frac{1}{2} \llbracket u \mathbf{n} \rrbracket, \mathbf{n}_F \mu \right\rangle_F, \quad \forall \mu \in P_k(F), \quad \forall F \in \Gamma,$$

$$(3.30d)$$

$$\left\langle \{\bar{\Pi}_u u\} - \frac{1}{2} \llbracket \bar{\Pi}_{\mathbf{q}}\mathbf{q} \cdot \mathbf{n} \rrbracket, \nu \right\rangle_F = \left\langle \{u\} - \frac{1}{2} \llbracket \mathbf{q} \cdot \mathbf{n} \rrbracket, \nu \right\rangle_F, \quad \forall \nu \in P_k(F), \quad \forall F \in \Gamma.$$

We claim that (3.30c) and (3.30d) can be equivalently written as

$$(3.31)$$

$$\langle \bar{\Pi}_{\mathbf{q}}\mathbf{q} \cdot \mathbf{n} - \bar{\Pi}_u u, \mu \rangle_F = \langle \mathbf{q} \cdot \mathbf{n} - u, \mu \rangle_F, \quad \forall \mu \in P_k(F), \quad \forall F \in \partial K, \quad \forall K \in \mathcal{T}_h.$$

To show (3.30c) and (3.30d) imply (3.31), we consider a face $F \in \partial K$ and take \mathbf{n}_F as the unit outer normal of F with respect to K . (3.31) can be obtained by subtracting (3.30d) from (3.30c). On the other hand, adding (3.31) along F in two neighboring cells will lead to (3.30d) and taking the subtraction will lead to (3.30c).

Note that the projection pair $\Pi_{\mathbf{q}}^*$ and Π_u^* , defined by (3.30a), (3.30b) and (3.31), retrieves the same form as that in [13, Theorem 2.1], therefore we have the following approximation result.

Lemma 3.2. *The projection pair $\bar{\Pi}_{\mathbf{q}}$ and $\bar{\Pi}_u$ in (3.30) is well-defined, with the approximation property*

$$(3.32) \quad \|\mathbf{q} - \bar{\Pi}_{\mathbf{q}}\mathbf{q}\|^2 + \|u - \bar{\Pi}_u u\|^2 \leq Ch^{2k+2} (|\mathbf{q}|_{\mathbf{H}^{k+1}}^2 + |u|_{H^{k+1}}^2),$$

$$(3.33) \quad \|\llbracket (\mathbf{q} - \bar{\Pi}_{\mathbf{q}}\mathbf{q}) \cdot \mathbf{n} \rrbracket\|_\Gamma^2 + \|\llbracket (u - \bar{\Pi}_u u) \mathbf{n} \rrbracket\|_\Gamma^2 \leq Ch^{2k+1} (|\mathbf{q}|_{\mathbf{H}^{k+1}}^2 + |u|_{H^{k+1}}^2).$$

The approximation (3.32) follows from the result in [13, Theorem 2.1], and the approximation (3.33) can be observed by adding and subtracting the L^2 projection of u and \mathbf{q} , and then applying Cauchy–Schwartz inequality, inverse inequality, the result (3.32) and the optimal projection error of the L^2 projection.

3.3.2. *The general case.* Next, we will analyze the general case as a perturbation away from the above local projection pair $\bar{\Pi}_{\mathbf{q}}$ and $\bar{\Pi}_u$. With a dimension count (see Appendix C), it can be seen that (3.17) forms a square linear system, therefore the approximation estimate in (3.18) implies that the projection pair (3.17) must be unique, hence unisolvent. It suffices to prove (3.18) holds for any $\Pi_{\mathbf{q}}^*$ and Π_u^* satisfying (3.17).

To this end, we denote the projection error of $\bar{\Pi}_{\mathbf{q}}$ and $\bar{\Pi}_u$ by

$$(3.34) \quad \bar{\eta}_{\mathbf{q}} = \mathbf{q} - \bar{\Pi}_{\mathbf{q}}\mathbf{q} \quad \text{and} \quad \bar{\eta}_u = u - \bar{\Pi}_u u,$$

and define perturbation terms

$$(3.35) \quad \delta_{\mathbf{q}} = (\Pi_{\mathbf{q}}^* - \bar{\Pi}_{\mathbf{q}})\mathbf{q} \quad \text{and} \quad \delta_u = (\Pi_u^* - \bar{\Pi}_u)u.$$

By subtracting (3.30) from (3.17), it can be seen that $\delta_{\mathbf{q}}$ and δ_u satisfy the following conditions

$$(3.36a) \quad \langle \delta_{\mathbf{q}}, \mathbf{v} \rangle_K = 0, \quad \forall \mathbf{v} \in \mathbf{P}_{k-1}(K), \quad \forall K \in \mathcal{T}_h,$$

$$(3.36b) \quad \langle \delta_u, w \rangle_K = 0, \quad \forall w \in P_{k-1}(K), \quad \forall K \in \mathcal{T}_h,$$

$$(3.36c)$$

$$\langle \mathcal{F}_{\mathbf{q}}(\delta_{\mathbf{q}}, \delta_u), \mathbf{n}_F \mu \rangle_F = - \left\langle \boldsymbol{\alpha} \llbracket \bar{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket + (\beta_1 - \frac{1}{2}) \llbracket \bar{\eta}_u \mathbf{n} \rrbracket, \mathbf{n}_F \mu \right\rangle_F, \quad \forall \mu \in P_k(F), \quad \forall F \in \Gamma,$$

$$(3.36d)$$

$$\langle \mathcal{F}_u(\delta_u, \delta_{\mathbf{q}}), \nu \rangle_F = - \left\langle (\beta_2 - \frac{1}{2}) \llbracket \bar{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket - \boldsymbol{\alpha} \cdot \llbracket \bar{\eta}_u \mathbf{n} \rrbracket, \nu \right\rangle_F, \quad \forall \nu \in P_k(F), \quad \forall F \in \Gamma.$$

It can be equivalently written as

$$(3.37a) \quad \langle \delta_{\mathbf{q}}, \mathbf{v} \rangle_K = 0, \quad \forall \mathbf{v} \in \mathbf{P}_{k-1}(K), \quad \forall K \in \mathcal{T}_h,$$

$$(3.37b) \quad \langle \delta_u, w \rangle_K = 0, \quad \forall w \in P_{k-1}(K), \quad \forall K \in \mathcal{T}_h,$$

$$(3.37c) \quad \left\langle \{ \delta_{\mathbf{q}} \} - \frac{1}{2} \llbracket \delta_u \mathbf{n} \rrbracket, \mathbf{n}_F \mu \right\rangle_F = \langle \boldsymbol{\zeta} \cdot \mathbf{n}_F, \mu \rangle_F, \quad \forall \mu \in P_k(F), \quad \forall F \in \Gamma,$$

$$(3.37d) \quad \left\langle \{ \delta_u \} - \frac{1}{2} \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket, \nu \right\rangle_F = \langle \boldsymbol{\xi}, \nu \rangle_F, \quad \forall \nu \in P_k(F), \quad \forall F \in \Gamma,$$

where

$$(3.38a) \quad \boldsymbol{\zeta} = -\boldsymbol{\alpha}(\llbracket \bar{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket - \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket) - (\beta_1 - \frac{1}{2})(\llbracket \bar{\eta}_u \mathbf{n} \rrbracket - \llbracket \delta_u \mathbf{n} \rrbracket),$$

$$(3.38b) \quad \boldsymbol{\xi} = -(\beta_2 - \frac{1}{2})(\llbracket \bar{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket - \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket) + \boldsymbol{\alpha} \cdot (\llbracket \bar{\eta}_u \mathbf{n} \rrbracket - \llbracket \delta_u \mathbf{n} \rrbracket).$$

Lemma 3.3.

$$(3.39) \quad \|\delta_{\mathbf{q}}\|^2 + \|\delta_u\|^2 \leq C_{\alpha, \beta} h \left(\|\boldsymbol{\zeta} \cdot \mathbf{n}\|_{\Gamma}^2 + \|\boldsymbol{\xi}\|_{\Gamma}^2 \right).$$

Lemma 3.3 essentially states a property of the local projection pair $\bar{\Pi}_{\mathbf{q}}$ and $\bar{\Pi}_u$. Its proof follows similar arguments as that in [13, Appendix A], and is sketched in Appendix D.

Lemma 3.4. *The approximation estimate (3.18) holds if*

$$(3.40) \quad \|\llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket\|_{\Gamma}^2 + \|\llbracket \delta_u \mathbf{n} \rrbracket\|_{\Gamma}^2 \leq C_{\star} \left(\|\llbracket \bar{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket\|_{\Gamma}^2 + \|\llbracket \bar{\eta}_u \mathbf{n} \rrbracket\|_{\Gamma}^2 \right).$$

Proof. Combining the definition of ζ and ξ in (3.38) with the assumption (3.40), and applying Cauchy–Schwarz inequality, it yields that

$$(3.41) \quad \|\zeta \cdot \mathbf{n}\|_{\Gamma}^2 + \|\xi\|_{\Gamma}^2 \leq C_{\star} \left(\|\llbracket \bar{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket\|_{\Gamma}^2 + \|\llbracket \bar{\eta}_u \mathbf{n} \rrbracket\|_{\Gamma}^2 \right).$$

With Lemma 3.3, one can obtain that

$$(3.42) \quad \|\delta_{\mathbf{q}}\|^2 + \|\delta_u\|^2 \leq C_{\star} h \left(\|\llbracket \bar{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket\|_{\Gamma}^2 + \|\llbracket \bar{\eta}_u \mathbf{n} \rrbracket\|_{\Gamma}^2 \right).$$

Recall the approximation estimate for $\bar{\eta}_{\mathbf{q}}$ and $\bar{\eta}_u$ in Lemma 3.2, we have

$$(3.43) \quad \|\delta_{\mathbf{q}}\|^2 + \|\delta_u\|^2 \leq C_{\star} h^{2k+2} \left(|\mathbf{q}|_{\mathbf{H}^{k+1}}^2 + |u|_{H^{k+1}}^2 \right).$$

The estimates for $\|\mathbf{q} - \Pi_{\mathbf{q}}^{\star} \mathbf{q}\|$ and $\|u - \Pi_u^{\star} u\|$ can be obtained after applying triangle inequality. By adding and subtracting $\bar{\Pi}_u u$ and $\bar{\Pi}_{\mathbf{q}} \mathbf{q}$, and then applying Cauchy–Schwarz inequality, inverse inequality, (3.40) and the optimal approximation property of $\bar{\Pi}_{\mathbf{q}}$ and $\bar{\Pi}_u$, one can also obtain estimates of $\|\llbracket (\mathbf{q} - \Pi_{\mathbf{q}}^{\star} \mathbf{q}) \cdot \mathbf{n} \rrbracket\|_{\Gamma}$ and $\|\llbracket (u - \Pi_u^{\star} u) \mathbf{n} \rrbracket\|_{\Gamma}$. Details are omitted. \square

It remains to verify (3.40) for the listed cases. The main idea is to use the semi-positivity of the bilinear form $a_h(\cdot, \cdot; \cdot, \cdot)$ to estimate the jump terms.

Take $\mathbf{v} = \nabla \delta_u$ in (3.36a), $w = \nabla \cdot \delta_{\mathbf{q}}$ in (3.36b), $\mu = \llbracket \delta_u \mathbf{n} \rrbracket \cdot \mathbf{n}_F$ in (3.36c) and $\nu = \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket$ in (3.36d). Summing these equations over all the elements yields that

$$(3.44) \quad \begin{aligned} & (\delta_{\mathbf{q}}, \nabla \delta_u) + (\delta_u, \nabla \cdot \delta_{\mathbf{q}}) - \langle \mathcal{F}_{\mathbf{q}}(\delta_u, \delta_{\mathbf{q}}), \llbracket \delta_u \mathbf{n} \rrbracket \rangle - \langle \mathcal{F}_u(\delta_u, \delta_{\mathbf{q}}), \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket \rangle \\ &= \left\langle \alpha \llbracket \bar{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket + \left(\beta_1 - \frac{1}{2}\right) \llbracket \bar{\eta}_u \mathbf{n} \rrbracket, \llbracket \delta_u \mathbf{n} \rrbracket \right\rangle \\ & \quad + \left\langle \left(\beta_2 - \frac{1}{2}\right) \llbracket \bar{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket - \alpha \cdot \llbracket \bar{\eta}_u \mathbf{n} \rrbracket, \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket \right\rangle. \end{aligned}$$

Note the left side of the equation is simply $a_h(\delta_{\mathbf{q}}, \delta_u; \delta_{\mathbf{q}}, \delta_u)$. Recall the identity (3.14), therefore the left hand side term becomes $\beta_1 \|\llbracket \delta_u \mathbf{n} \rrbracket\|_{\Gamma}^2 + \beta_2 \|\llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket\|_{\Gamma}^2$. Applying Cauchy–Schwarz inequality to right hand side of (3.44), it yields that

$$(3.45) \quad \begin{aligned} & \beta_1 \|\llbracket \delta_u \mathbf{n} \rrbracket\|_{\Gamma}^2 + \beta_2 \|\llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket\|_{\Gamma}^2 \\ & \leq \frac{C_{\alpha, \beta}}{\varepsilon} \left(\|\llbracket \bar{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket\|_{\Gamma}^2 + \|\llbracket \bar{\eta}_u \mathbf{n} \rrbracket\|_{\Gamma}^2 \right) + \varepsilon \left(\|\llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket\|_{\Gamma}^2 + \|\llbracket \delta_u \mathbf{n} \rrbracket\|_{\Gamma}^2 \right), \end{aligned}$$

for any positive constant ε .

Case 1 ($\beta_1 > 0$ and $\beta_2 > 0$). In this case, we can simply take $\varepsilon = \min(\beta_1, \beta_2)/2$ and rearrange terms to obtain (3.40), with the constant $C_{\star} = C_{\alpha, \beta}$ independent of κ .

Case 2 ($\beta_1 = 0, \beta_2 > 0$ and $\alpha \neq 0$). In this case, we have

$$(3.46) \quad (\beta_2 - \varepsilon) \|\llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket\|_{\Gamma}^2 \leq \frac{C_{\alpha, \beta}}{\varepsilon} \left(\|\llbracket \bar{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket\|_{\Gamma}^2 + \|\llbracket \bar{\eta}_u \mathbf{n} \rrbracket\|_{\Gamma}^2 \right) + \varepsilon \|\llbracket \delta_u \mathbf{n} \rrbracket\|_{\Gamma}^2.$$

Now it suffices to estimate $\|\llbracket \delta_u \mathbf{n} \rrbracket\|_{\Gamma}^2$ to verify (3.40). The main idea is to construct the bilinear form of the advection operator from (3.36) and use the associated semi-positivity for estimating $\|\llbracket \delta_u \mathbf{n} \rrbracket\|_{\Gamma}^2$.

We take $w = \mathbf{r} \cdot \nabla \delta_u$ in (3.36b) and $\nu = \llbracket \mathbf{r} \delta_u \cdot \mathbf{n} \rrbracket$ in (3.36d). Then it can be shown that

$$(3.47) \quad \begin{aligned} L_{\delta_{\mathbf{q}}}(\delta_u; \delta_u) &:= \sum_K (\delta_u \mathbf{r}, \nabla \delta_u)_K - \langle \{\delta_u\} - \beta_2 \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket + \boldsymbol{\alpha} \cdot \llbracket \delta_u \mathbf{n} \rrbracket, \llbracket \mathbf{r} \delta_u \cdot \mathbf{n} \rrbracket \rangle \\ &= \left\langle \left(\beta_2 - \frac{1}{2} \right) \llbracket \overline{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket - \boldsymbol{\alpha} \cdot \llbracket \overline{\eta}_u \mathbf{n} \rrbracket, \llbracket \mathbf{r} \delta_u \cdot \mathbf{n} \rrbracket \right\rangle. \end{aligned}$$

The left hand side of (3.47) corresponds to the DG discretization of the advection operator $\mathbf{r} \cdot \nabla$, which can be simplified as follows

$$(3.48) \quad \begin{aligned} L_{\delta_{\mathbf{q}}}(\delta_u; \delta_u) &= \sum_K \left(\nabla \cdot (\delta_u (\mathbf{r} \delta_u)), \frac{1}{2} \right)_K - \langle \{\delta_u\} - \beta_2 \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket + \boldsymbol{\alpha} \cdot \llbracket \delta_u \mathbf{n} \rrbracket, \llbracket \mathbf{r} \delta_u \cdot \mathbf{n} \rrbracket \rangle \\ &= \frac{1}{2} \sum_K \langle \delta_u, (\mathbf{r} \delta_u) \cdot \mathbf{n} \rangle_{\partial K} - \langle \{\delta_u\}, \llbracket \mathbf{r} \delta_u \cdot \mathbf{n} \rrbracket \rangle + \langle \beta_2 \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket - \boldsymbol{\alpha} \cdot \llbracket \delta_u \mathbf{n} \rrbracket, \llbracket \mathbf{r} \delta_u \cdot \mathbf{n} \rrbracket \rangle \\ &= \frac{1}{2} \left(\sum_K \langle \delta_u, (\mathbf{r} \delta_u) \cdot \mathbf{n} \rangle_{\partial K} - \langle \{\delta_u\}, \llbracket \mathbf{r} \delta_u \cdot \mathbf{n} \rrbracket \rangle - \langle \{\mathbf{r} \delta_u\}, \llbracket \delta_u \mathbf{n} \rrbracket \rangle \right) \\ &\quad + \langle \beta_2 \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket - \boldsymbol{\alpha} \cdot \llbracket \delta_u \mathbf{n} \rrbracket, \mathbf{r} \cdot \llbracket \delta_u \mathbf{n} \rrbracket \rangle \\ &= \langle \beta_2 \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket - \boldsymbol{\alpha} \cdot \llbracket \delta_u \mathbf{n} \rrbracket, \mathbf{r} \cdot \llbracket \delta_u \mathbf{n} \rrbracket \rangle. \end{aligned}$$

Here we have used the fact that $\llbracket \mathbf{r} \delta_u \cdot \mathbf{n} \rrbracket = \mathbf{r} \cdot \llbracket \delta_u \mathbf{n} \rrbracket$, hence $\langle \{\delta_u\}, \llbracket \mathbf{r} \delta_u \cdot \mathbf{n} \rrbracket \rangle = \langle \{\mathbf{r} \delta_u\}, \llbracket \delta_u \mathbf{n} \rrbracket \rangle$, and then the identity (3.8). One can further apply the relationship $\llbracket \delta_u \mathbf{n} \rrbracket = \mathbf{n}_F (\llbracket \delta_u \mathbf{n} \rrbracket \cdot \mathbf{n}_F)$ and $\boldsymbol{\alpha} = \alpha \text{sign}(\mathbf{r} \cdot \mathbf{n}_F) \mathbf{n}_F$ to obtain

$$(3.49) \quad \begin{aligned} L_{\delta_{\mathbf{q}}}(\delta_u; \delta_u) &= \langle \beta_2 \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket - \boldsymbol{\alpha} \cdot \llbracket \delta_u \mathbf{n} \rrbracket, \mathbf{r} \cdot \llbracket \delta_u \mathbf{n} \rrbracket \rangle \\ &= - \langle (\boldsymbol{\alpha} \cdot \mathbf{n}_F) (\mathbf{r} \cdot \mathbf{n}_F), (\llbracket \delta_u \mathbf{n} \rrbracket \cdot \mathbf{n}_F)^2 \rangle + \beta_2 (\llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket, (\mathbf{r} \cdot \mathbf{n}_F) \llbracket \delta_u \mathbf{n} \rrbracket \cdot \mathbf{n}_F) \\ &= - \alpha \langle |\mathbf{r} \cdot \mathbf{n}_F|, (\llbracket \delta_u \mathbf{n} \rrbracket \cdot \mathbf{n}_F)^2 \rangle + \beta_2 \langle (\mathbf{r} \cdot \mathbf{n}_F) \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket, \llbracket \delta_u \mathbf{n} \rrbracket \cdot \mathbf{n}_F \rangle. \end{aligned}$$

We then simplify the right hand side of (3.47) along similar lines to obtain

$$(3.50) \quad \begin{aligned} &\left\langle \left(\beta_2 - \frac{1}{2} \right) \llbracket \overline{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket - \boldsymbol{\alpha} \cdot \llbracket \overline{\eta}_u \mathbf{n} \rrbracket, \llbracket \mathbf{r} \delta_u \cdot \mathbf{n} \rrbracket \right\rangle \\ &= \left\langle -\alpha |\mathbf{r} \cdot \mathbf{n}_F| \llbracket \overline{\eta}_u \mathbf{n} \rrbracket \cdot \mathbf{n}_F + \left(\beta_2 - \frac{1}{2} \right) (\mathbf{r} \cdot \mathbf{n}_F) \llbracket \overline{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket, \llbracket \delta_u \mathbf{n} \rrbracket \cdot \mathbf{n}_F \right\rangle. \end{aligned}$$

Combining (3.49) with (3.50), one can deduce

$$(3.51) \quad \begin{aligned} \alpha \left\| |\mathbf{r} \cdot \mathbf{n}_F|^{\frac{1}{2}} \llbracket \delta_u \mathbf{n} \rrbracket \cdot \mathbf{n}_F \right\|_{\Gamma}^2 &= \alpha \langle |\mathbf{r} \cdot \mathbf{n}_F| \llbracket \overline{\eta}_u \mathbf{n} \rrbracket \cdot \mathbf{n}_F, \llbracket \delta_u \mathbf{n} \rrbracket \cdot \mathbf{n}_F \rangle \\ &\quad + \left\langle \beta_2 \llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket - \left(\beta_2 - \frac{1}{2} \right) \llbracket \overline{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket, (\mathbf{r} \cdot \mathbf{n}_F) (\llbracket \delta_u \mathbf{n} \rrbracket \cdot \mathbf{n}_F) \right\rangle. \end{aligned}$$

Applying Cauchy–Schwarz inequality and rearranging terms yield that

$$(3.52) \quad (|\alpha| - \tilde{\varepsilon}) \left\| |\mathbf{r} \cdot \mathbf{n}_F|^{\frac{1}{2}} \llbracket \delta_u \mathbf{n} \rrbracket \cdot \mathbf{n}_F \right\|_{\Gamma}^2 \leq \frac{C_{\alpha, \beta}}{\tilde{\varepsilon}} \left(\|\llbracket \overline{\eta}_u \mathbf{n} \rrbracket\|_{\Gamma}^2 + \|\llbracket \overline{\eta}_{\mathbf{q}} \cdot \mathbf{n} \rrbracket\|_{\Gamma}^2 + \|\llbracket \delta_{\mathbf{q}} \cdot \mathbf{n} \rrbracket\|_{\Gamma}^2 \right).$$

Here we have used the fact that $|\mathbf{r} \cdot \mathbf{n}_F| \leq 1$ and $\|[\overline{\eta}_u \mathbf{n}] \cdot \mathbf{n}_F\|_\Gamma^2 = \|[\overline{\eta}_u \mathbf{n}]\|_\Gamma^2$. Then we take $\tilde{\varepsilon} = |\alpha|/2$ and recall the assumption $0 < \kappa \leq |\mathbf{r} \cdot \mathbf{n}_F|$ to get

$$(3.53) \quad \|[\delta_u \mathbf{n}]\|_\Gamma^2 \leq C_{\alpha,\beta,\kappa} \left(\|[\overline{\eta}_u \mathbf{n}]\|_\Gamma^2 + \|[\overline{\eta}_q \cdot \mathbf{n}]\|_\Gamma^2 + \|[\delta_q \cdot \mathbf{n}]\|_\Gamma^2 \right).$$

We combine (3.46) and (3.53) to obtain

$$(3.54) \quad (\beta_2 - (1 + C_{\alpha,\beta,\kappa})\varepsilon) \|[\delta_q \cdot \mathbf{n}]\|_\Gamma^2 \leq \frac{C_{\alpha,\beta,\kappa}}{\varepsilon} \left(\|[\overline{\eta}_q \cdot \mathbf{n}]\|_\Gamma^2 + \|[\overline{\eta}_u \mathbf{n}]\|_\Gamma^2 \right).$$

The bound for $\|[\delta_q \cdot \mathbf{n}]\|_\Gamma^2$ can be obtained after taking ε to be sufficiently small. In the end, we obtain (3.40) with $C_* = C_{\alpha,\beta,\kappa}$ dependent on κ , by combining (3.54) and (3.53). This completes the proof of Lemma 3.1.

4. NUMERICAL TESTS IN TWO DIMENSIONS

Extensive numerical studies in one dimension can be found in [6], which match all of our theoretical findings. In this section, we only provide two-dimensional accuracy tests using the initial value problem in [22, Example 4.9].

We apply the DG scheme (3.9) to (1.2) on $\Omega = [0, 1] \times [0, 1]$ with the initial condition

$$(4.1) \quad u(x, y, 0) = \sin(2\pi(x + y)) \quad \text{and} \quad \mathbf{q}(x, y, 0) = \mathbf{0}.$$

The exact solution to the stated initial value problem is given by the following plane wave

$$(4.2) \quad u(x, y, t) = \frac{1}{2} \sin(2\pi(x + y - (\sqrt{2})t)) + \frac{1}{2} \sin(2\pi(x + y + (\sqrt{2})t)),$$

$$(4.3) \quad \mathbf{q}(x, y, t) = \left(\begin{array}{l} \frac{\sqrt{2}}{4} \sin(2\pi(x + y - (\sqrt{2})t)) - \frac{\sqrt{2}}{4} \sin(2\pi(x + y + (\sqrt{2})t)) \\ \frac{\sqrt{2}}{4} \sin(2\pi(x + y - (\sqrt{2})t)) - \frac{\sqrt{2}}{4} \sin(2\pi(x + y + (\sqrt{2})t)) \end{array} \right).$$

We use the four-stage fourth-order explicit Runge–Kutta method³ [39] with the time step $\tau = 1/(50N^{\max(1,k/4)})$ for time discretization and compute upto $T = 1/10$. The convergence rate is examined both on the structured triangular mesh and the unstructured triangular mesh, as illustrated in Figure 4.1. The structured mesh is created by splitting a uniform square mesh into triangles, and the unstructured mesh is created with Netgen [36] with the maximum mesh size specified as $1/N$. The numerical tests and observed convergence rates are summarized in Table 4.1. See also Tables 4.2–4.14 for details.

From the numerical tests, we see that for the proved cases: 1. $\beta_1 > 0$ and $\beta_2 > 0$; 2. $\beta_1 = 0, \beta_2 > 0$ and $\alpha \neq 0$, we do obtain optimal convergence rate for \mathbf{q}_h and u_h simultaneously on both the structured meshes and the unstructured meshes. The cases that are not covered in the proof may fall into the following categories.

- (1) $\beta_2 = 0$ and
 - (a) $\beta_1 \neq 0, \alpha \neq 0$; (b) $\beta_1 = 0, \alpha \neq 0$;
 - (c) $\beta_1 \neq 0, \alpha = 0$; (d) $\beta_1 = \alpha = 0$.
- (2) $\beta_2 > 0$ and $\beta_1 = \alpha = 0$.

³We choose the fourth-order method because our numerical tests are primarily performed for the cases with $k \leq 3$. The purpose of using the fourth-order method is to avoid the temporal error affect the exhibited convergence rate. Other high-order methods can also be used and we expect similar results.

We perform numerical tests for each of the listed cases. Note that the special case $\beta_1 = \beta_2 = 0$ and $\alpha = \pm 1/2$ in 1(b) is usually referred to as the alternating flux, and has been studied in [10] in the context of the minimal dissipation local DG method for convection-diffusion problems. The case $\alpha = \beta_1 = \beta_2 = 0$ in 1(d) is also known as the central flux proposed in Bassi and Rebay [4].

For all these listed choices, it can be seen that at least one of \mathbf{q}_h and u_h converge at a suboptimal rate on unstructured meshes. Hence the conditions in Lemma 3.1 should provide a complete characterization for parameters to guarantee simultaneous optimal convergence for both \mathbf{q}_h and u_h . At the same time, we do observe interesting convergence patterns for suboptimal cases. For example, if one uses alternating flux, optimal convergence rates are still observed on structured triangular meshes; the central flux have the even-odd patterns, comparable to the recent analysis on advection equation on Cartesian meshes in [30]; and if one adds jump penalties for one of the unknowns, this unknown will exhibit optimal converging rate, for which similar observation has been reported in [42] for Hamiltonian partial differential equations. We postpone the analysis of these suboptimal behaviors to our future work.

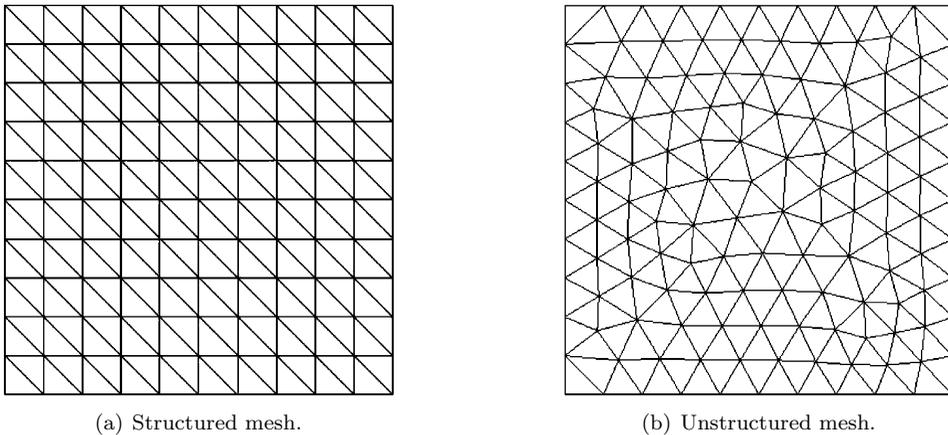


FIGURE 4.1. Structured and unstructured meshes for accuracy test with $N = 10$.

5. CONCLUSIONS

In this paper, we present optimal error estimates of the DG methods for the wave equation (1.2) with generalized numerical fluxes. For the one-dimensional case, this work completes the analysis in [6] by proving the optimal convergence rates which match all previous numerical observations. The analysis is mainly based on characteristic decomposition and uses generalized Gauss–Radau projections as building blocks to construct the required projection pair. For the multi-dimensional case, we introduce the technique with energy arguments to construct a global projection on unstructured meshes, and obtain optimal error estimates for a family of numerical fluxes. Two-dimensional numerical examples are provided to validate at least one of the unknowns will degenerate to suboptimal rates if the assumed assumptions are not satisfied.

TABLE 4.1. Summary of the observed convergence rates in various numerical tests. $\iota = k + \text{mod}(k + 1, 2)$. The sign * indicates unclean observed convergence rates. The DG scheme with central, \mathbf{q} -dissipative and u -dissipative fluxes are tested for $k = 1, 2, 3, 4, 5$. All other schemes are tested for $k = 1, 2, 3$.

Flux	Table	Parameters			Structured Mesh		Unstructured Mesh	
		α	β_1	β_2	u_h	\mathbf{q}_h	u_h	\mathbf{q}_h
Optimal schemes: $\beta_1 > 0$ and $\beta_2 > 0$								
Upwind	4.2	0	1/2	1/2	$k + 1$	$k + 1$	$k + 1$	$k + 1$
Underupwinding	4.3	0	1/10	1/10	$k + 1$	$k + 1$	$k + 1$	$k + 1$
Overupwinding	4.4	0	1	1	$k + 1$	$k + 1$	$k + 1$	$k + 1$
Nondegenerate	4.5	-3/10	1/20	1/5	$k + 1$	$k + 1$	$k + 1$	$k + 1$
Optimal schemes: $\beta_1 = 0, \beta_2 > 0$ and $\alpha \neq 0$								
$\alpha\beta$	4.6	1/2	0	1	$k + 1$	$k + 1$	$k + 1$	$k + 1$
non- $\alpha\beta$	4.7	-1/10	0	1/10	$k + 1$	$k + 1$	$k + 1$	$k + 1$
A few suboptimal schemes								
β_2 -degenerate	4.8	1/2	1/2	0	$k + 1$	$k + 1$	$k + 1$	k
Alternating	4.9	1/2	0	0	$k + 1$	$k + 1$	$k + 1$	k
Underalternating*	4.10	-1/10	0	0	$k + 1$	k	$k + 1$	k
Overalternating*	4.11	1	0	0	$k + 1$	k	$k + 1$	k
Central	4.12	0	0	0	$\max(2, \iota)$	ι	$\max(2, \iota)$	k
\mathbf{q} -dissipative	4.13	0	0	1/2	$\max(2, \iota)$	$k + 1$	$\max(2, \iota)$	$k + 1$
u -dissipative	4.14	0	1/2	0	$k + 1$	ι	$k + 1$	k

TABLE 4.2. Upwind flux (also an $\alpha\beta$ -flux), $\alpha = 0$ and $\beta_1 = \beta_2 = 1/2$.

k	N	Structured Mesh				Unstructured Mesh			
		u_h		\mathbf{q}_h		u_h		\mathbf{q}_h	
		L^2 error	order	L^2 error	order	L^2 error	order	L^2 error	order
1	10	6.8691e-03	-	9.1693e-03	-	9.7037e-03	-	3.4192e-02	-
	20	1.7834e-03	1.95	2.2355e-03	2.04	2.3273e-03	2.06	5.9651e-03	2.52
	40	4.5272e-04	1.98	5.5103e-04	2.02	6.0673e-04	1.94	1.2284e-03	2.28
	80	1.1393e-04	1.99	1.3676e-04	2.01	1.5447e-04	1.97	2.7184e-04	2.18
2	10	3.6510e-04	-	4.7225e-04	-	6.9628e-04	-	1.1838e-03	-
	20	4.6503e-05	2.97	5.7377e-05	3.04	7.5990e-05	3.20	1.1477e-04	3.37
	40	5.8584e-06	2.99	7.0725e-06	3.02	9.8094e-06	2.95	1.4542e-05	2.98
	80	7.3542e-07	2.99	8.7844e-07	3.01	1.2394e-06	2.98	1.8168e-06	3.00
3	10	1.1092e-05	-	1.5070e-05	-	2.7779e-05	-	5.0280e-05	-
	20	7.1733e-07	3.95	9.0186e-07	4.06	1.4771e-06	4.23	2.4288e-06	4.37
	40	4.5467e-08	3.98	5.5438e-08	4.02	9.5696e-08	3.95	1.5924e-07	3.93
	80	2.8618e-09	3.99	3.4355e-09	4.01	6.1118e-09	3.97	1.0017e-08	3.99

APPENDIX A. PROOF OF THEOREM 2.1

Proof. Let $e_q = q - q_h$ and $e_u = u - u_h$. The consistency of the DG scheme (2.5) gives

$$(A.1) \quad ((e_q)_t, v) + ((e_u)_t, w) + a_h(e_q, e_u; v, w) = 0.$$

With the projection error denoted by $\eta_q = q - \Pi_q^* q$ and $\eta_u = u - \Pi_u^* u$, it can be shown that

$$(A.2) \quad e_q = \Pi_q^* e_q + \eta_q \quad \text{and} \quad e_u = \Pi_u^* e_u + \eta_u.$$

TABLE 4.3. Upwind-biased flux (under-upwinding), $\alpha = 0$ and $\beta_1 = \beta_2 = 1/10$.

		Structured Mesh				Unstructured Mesh			
		u_h		\mathbf{q}_h		u_h		\mathbf{q}_h	
k	N	L^2 error	order	L^2 error	order	L^2 error	order	L^2 error	order
1	10	7.6539e-03	-	3.4827e-02	-	1.0118e-02	-	6.0910e-02	-
	20	1.6772e-03	2.19	8.3258e-03	2.06	2.1923e-03	2.21	1.2468e-02	2.29
	40	4.1207e-04	2.03	1.9435e-03	2.10	5.5475e-04	1.98	2.9344e-03	2.09
	80	1.0192e-04	2.02	4.5869e-04	2.08	1.3903e-04	2.00	6.9001e-04	2.09
2	10	3.4584e-04	-	5.6675e-04	-	7.6958e-04	-	2.0052e-03	-
	20	4.0524e-05	3.09	5.7198e-05	3.31	7.5713e-05	3.35	1.6821e-04	3.58
	40	4.9288e-06	3.04	6.5321e-06	3.13	9.1708e-06	3.05	2.0072e-05	3.07
	80	6.0775e-07	3.02	7.9036e-07	3.05	1.1338e-06	3.02	2.4245e-06	3.05
3	10	8.9671e-06	-	6.2266e-05	-	2.4349e-05	-	1.5225e-04	-
	20	8.2103e-07	3.45	3.7806e-06	4.04	1.3224e-06	4.20	7.5647e-06	4.33
	40	7.0739e-08	3.54	2.2390e-07	4.08	1.0008e-07	3.72	4.9201e-07	3.94
	80	5.0686e-09	3.80	1.3419e-08	4.06	7.1837e-09	3.80	3.0420e-08	4.02

TABLE 4.4. Upwind-biased flux (over-upwinding), $\alpha = 0$, $\beta_1 = \beta_2 = 1$.

		Structured Mesh				Unstructured Mesh			
		u_h		\mathbf{q}_h		u_h		\mathbf{q}_h	
k	N	L^2 error	order	L^2 error	order	L^2 error	order	L^2 error	order
1	10	7.2307e-03	-	8.2040e-03	-	1.0953e-02	-	3.3382e-02	-
	20	1.7641e-03	2.04	2.0067e-03	2.03	2.4806e-03	2.14	5.6655e-03	2.56
	40	4.3410e-04	2.02	5.0060e-04	2.00	6.3209e-04	1.97	1.1265e-03	2.33
	80	1.0762e-04	2.01	1.2531e-04	2.00	1.5847e-04	2.00	2.4270e-04	2.21
2	10	3.9554e-04	-	6.3507e-04	-	7.1351e-04	-	1.4281e-03	-
	20	5.6902e-05	2.80	7.7403e-05	3.04	8.5794e-05	3.06	1.4408e-04	3.31
	40	7.5662e-06	2.91	9.4806e-06	3.03	1.1743e-05	2.87	1.8097e-05	2.99
	80	9.7325e-07	2.96	1.1709e-06	3.02	1.5252e-06	2.94	2.2494e-06	3.01
3	10	1.2171e-05	-	1.2873e-05	-	3.2823e-05	-	4.9117e-05	-
	20	7.0635e-07	4.11	7.4530e-07	4.11	1.6527e-06	4.31	2.2420e-06	4.45
	40	4.2454e-08	4.06	4.5988e-08	4.02	1.0630e-07	3.96	1.4767e-07	3.92
	80	2.6016e-09	4.03	2.8726e-09	4.00	6.7309e-09	3.98	9.2682e-09	3.99

Recall the property (2.11). One can then rewrite (A.1) as

$$(A.3) \quad ((\Pi_q^* e_q)_t, v) + ((\Pi_u^* e_u)_t, w) + a_h(\Pi_q^* e_q, \Pi_u^* e_u; v, w) = -((\eta_q)_t, v) - ((\eta_u)_t, w).$$

Taking $v = \Pi_q^* e_q$ and $w = \Pi_u^* e_u$ in (A.3) leads to

$$(A.4) \quad \frac{1}{2} \frac{d}{dt} (\|\Pi_q^* e_q\|^2 + \|\Pi_u^* e_u\|^2) + a_h(\Pi_q^* e_q, \Pi_u^* e_u; \Pi_q^* e_q, \Pi_u^* e_u) = -((\eta_q)_t, \Pi_q^* e_q) - ((\eta_u)_t, \Pi_u^* e_u).$$

Note that $a_h(\Pi_q^* e_q, \Pi_u^* e_u; \Pi_q^* e_q, \Pi_u^* e_u) \geq 0$ under the stability assumption. After using Cauchy–Schwarz inequality, it yields that

$$(A.5) \quad \frac{d}{dt} (\|\Pi_q^* e_q\|^2 + \|\Pi_u^* e_u\|^2)^{\frac{1}{2}} \leq (\|(\eta_q)_t\|^2 + \|(\eta_u)_t\|^2)^{\frac{1}{2}} \leq Ch^{k+1} (|q_t|_{H^{k+1}} + |u_t|_{H^{k+1}}).$$

Here we have used the fact $(\eta_q)_t = \eta_{qt}$ and $(\eta_u)_t = \eta_{ut}$, as well as the approximation estimate (2.10), in the last inequality. The estimate for $(\|\Pi_q^* e_q\|^2 + \|\Pi_u^* e_u\|^2)^{\frac{1}{2}}$ can

TABLE 4.5. Nondegenerate flux, $\alpha = -3/10$, $\beta_1 = 1/20$ and $\beta_2 = 1/5$.

		Structured Mesh				Unstructured Mesh			
		u_h		\mathbf{q}_h		u_h		\mathbf{q}_h	
k	N	L^2 error	order	L^2 error	order	L^2 error	order	L^2 error	order
1	10	8.3802e-03	-	1.3567e-02	-	1.1764e-02	-	3.8089e-02	-
	20	2.0548e-03	2.03	3.2707e-03	2.05	2.7154e-03	2.12	6.8594e-03	2.47
	40	5.0956e-04	2.01	7.8743e-04	2.05	6.9580e-04	1.96	1.4946e-03	2.20
	80	1.2720e-04	2.00	1.9189e-04	2.04	1.7528e-04	1.99	3.4404e-04	2.12
2	10	3.7079e-04	-	4.5639e-04	-	7.8061e-04	-	1.2925e-03	-
	20	4.4681e-05	3.05	5.5149e-05	3.05	8.0687e-05	3.27	1.2152e-04	3.41
	40	5.5117e-06	3.02	6.7612e-06	3.03	1.0128e-05	2.99	1.5549e-05	2.97
	80	6.8493e-07	3.01	8.3824e-07	3.01	1.2653e-06	3.00	1.9372e-06	3.00
3	10	1.4030e-05	-	2.2236e-05	-	3.1754e-05	-	6.9300e-05	-
	20	9.0716e-07	3.95	1.3665e-06	4.02	1.7441e-06	4.19	3.2556e-06	4.41
	40	5.6901e-08	3.99	8.3301e-08	4.04	1.1507e-07	3.92	2.1349e-07	3.93
	80	3.5681e-09	4.00	5.1398e-09	4.02	7.3423e-09	3.97	1.3573e-08	3.98

TABLE 4.6. $\alpha\beta$ -flux, $\alpha = 1/2$, $\beta_1 = 0$ and $\beta_2 = 1/2$.

		Structured Mesh				Unstructured Mesh			
		u_h		\mathbf{q}_h		u_h		\mathbf{q}_h	
k	N	L^2 error	order	L^2 error	order	L^2 error	order	L^2 error	order
1	10	9.2272e-03	-	8.3926e-03	-	1.3941e-02	-	3.3658e-02	-
	20	2.3118e-03	2.00	2.1224e-03	1.98	3.1460e-03	2.15	5.8153e-03	2.53
	40	5.7642e-04	2.00	5.3535e-04	1.99	8.1850e-04	1.94	1.1819e-03	2.30
	80	1.4380e-04	2.00	1.3444e-04	1.99	2.0615e-04	1.99	2.6044e-04	2.18
2	10	4.7203e-04	-	4.3965e-04	-	9.4417e-04	-	1.0413e-03	-
	20	5.9159e-05	3.00	5.5626e-05	2.98	1.0193e-04	3.21	1.0284e-04	3.34
	40	7.3770e-06	3.00	6.9560e-06	3.00	1.2890e-05	2.98	1.3517e-05	2.93
	80	9.2110e-07	3.00	8.7157e-07	3.00	1.6188e-06	2.99	1.7063e-06	2.99
3	10	1.5993e-05	-	1.3628e-05	-	4.2163e-05	-	4.8535e-05	-
	20	1.0039e-06	3.99	8.7343e-07	3.96	2.1147e-06	4.32	2.3227e-06	4.39
	40	6.2776e-08	4.00	5.4328e-08	4.01	1.4311e-07	3.89	1.4521e-07	4.00
	80	3.9040e-09	4.01	3.4155e-09	3.99	9.0260e-09	3.99	9.1382e-09	3.99

TABLE 4.7. Non- $\alpha\beta$ -flux, $\alpha = -1/10$, $\beta_1 = 0$ and $\beta_2 = 1/10$.

		Structured Mesh				Unstructured Mesh			
		u_h		\mathbf{q}_h		u_h		\mathbf{q}_h	
k	N	L^2 error	order	L^2 error	order	L^2 error	order	L^2 error	order
1	10	8.5013e-03	-	3.4512e-02	-	1.0745e-02	-	6.0297e-02	-
	20	1.8383e-03	2.21	8.1529e-03	2.08	2.3561e-03	2.19	1.2223e-02	2.30
	40	4.4447e-04	2.05	1.9165e-03	2.09	5.9555e-04	1.98	2.8660e-03	2.09
	80	1.1007e-04	2.01	4.4986e-04	2.09	1.4887e-04	2.00	6.7383e-04	2.09
2	10	3.8131e-04	-	6.0458e-04	-	8.2656e-04	-	2.0093e-03	-
	20	4.2426e-05	3.17	6.1937e-05	3.29	8.0893e-05	3.35	1.7145e-04	3.55
	40	5.0922e-06	3.06	7.1245e-06	3.12	9.6360e-06	3.07	2.0571e-05	3.06
	80	6.2543e-07	3.03	8.5722e-07	3.06	1.1828e-06	3.03	2.4963e-06	3.04
3	10	1.8385e-05	-	6.0208e-05	-	3.2836e-05	-	1.4697e-04	-
	20	1.6217e-06	3.50	3.6273e-06	4.05	2.2783e-06	3.85	7.1396e-06	4.36
	40	1.1465e-07	3.82	2.1849e-07	4.05	1.6332e-07	3.80	4.6616e-07	3.94
	80	7.5798e-09	3.92	1.3338e-08	4.03	1.1098e-08	3.88	2.9112e-08	4.00

TABLE 4.8. β_2 -degenerate flux, $\alpha = 1/2$, $\beta_1 = 1/2$ and $\beta_2 = 0$.

		Structured Mesh				Unstructured Mesh			
		u_h		\mathbf{q}_h		u_h		\mathbf{q}_h	
k	N	L^2 error	order	L^2 error	order	L^2 error	order	L^2 error	order
1	10	7.7938e-03	-	1.0573e-02	-	1.2369e-02	-	6.1832e-02	-
	20	1.8935e-03	2.04	2.6601e-03	1.99	2.7261e-03	2.18	1.6980e-02	1.86
	40	4.6578e-04	2.02	6.6814e-04	1.99	7.1042e-04	1.94	8.0946e-03	1.07
	80	1.1551e-04	2.01	1.6750e-04	2.00	1.7861e-04	1.99	3.7740e-03	1.10
2	10	3.9345e-04	-	5.3209e-04	-	8.6998e-04	-	5.4702e-03	-
	20	4.8139e-05	3.03	6.6892e-05	2.99	8.9484e-05	3.28	8.0084e-04	2.77
	40	5.9444e-06	3.02	8.3302e-06	3.01	1.1126e-05	3.01	1.8873e-04	2.09
	80	7.3863e-07	3.01	1.0409e-06	3.00	1.3886e-06	3.00	4.3113e-05	2.13
3	10	1.2197e-05	-	1.9543e-05	-	3.2702e-05	-	2.8946e-04	-
	20	7.5333e-07	4.02	1.2214e-06	4.00	1.6560e-06	4.30	2.1453e-05	3.75
	40	4.6334e-08	4.02	7.6398e-08	4.00	1.1472e-07	3.85	3.1700e-06	2.76
	80	2.8546e-09	4.02	4.7653e-09	4.00	7.3417e-09	3.97	3.9948e-07	2.99

TABLE 4.9. Alternating flux, $\alpha = 1/2$ and $\beta_1 = \beta_2 = 0$.

		Structured Mesh				Unstructured Mesh			
		u_h		\mathbf{q}_h		u_h		\mathbf{q}_h	
k	N	L^2 error	order	L^2 error	order	L^2 error	order	L^2 error	order
1	10	7.4932e-03	-	8.5840e-03	-	1.2034e-02	-	3.1136e-01	-
	20	1.8871e-03	1.99	2.0795e-03	2.05	2.4554e-03	2.29	1.3124e-01	1.25
	40	4.6821e-04	2.01	5.3759e-04	1.95	7.2169e-04	1.77	8.0050e-02	0.71
	80	1.1438e-04	2.03	1.3395e-04	2.00	1.7451e-04	2.05	4.1493e-02	0.95
2	10	3.9653e-04	-	4.5186e-04	-	8.2522e-04	-	3.7836e-02	-
	20	4.5740e-05	3.12	5.6208e-05	3.01	8.9414e-05	3.21	7.1534e-03	2.40
	40	6.3587e-06	2.85	6.8942e-06	3.03	1.1309e-05	2.98	1.9966e-03	1.84
	80	6.8097e-07	3.22	8.6548e-07	2.99	1.4579e-06	2.96	4.9252e-04	2.02
3	10	1.1002e-05	-	1.4215e-05	-	3.1286e-05	-	2.2610e-03	-
	20	6.6831e-07	4.04	8.8831e-07	4.00	1.7160e-06	4.19	2.1120e-04	3.42
	40	5.0516e-08	3.73	5.4327e-08	4.03	1.1233e-07	3.93	3.5116e-05	2.59
	80	3.1901e-09	3.99	3.4331e-09	3.98	7.2164e-09	3.96	4.6718e-06	2.91

TABLE 4.10. Under-alternating flux, $\alpha = -1/10$ and $\beta_1 = \beta_2 = 0$.

		Structured Mesh				Unstructured Mesh			
		u_h		\mathbf{q}_h		u_h		\mathbf{q}_h	
k	N	L^2 error	order	L^2 error	order	L^2 error	order	L^2 error	order
1	10	9.0098e-03	-	7.4649e-02	-	1.0694e-02	-	1.2251e-01	-
	20	2.3119e-03	1.96	2.8799e-02	1.37	2.6285e-03	2.02	4.3635e-02	1.49
	40	5.6043e-04	2.04	1.2191e-02	1.24	6.8009e-04	1.95	1.9692e-02	1.15
	80	1.4049e-04	2.00	5.5256e-03	1.14	1.7988e-04	1.92	9.1492e-03	1.11
2	10	3.8036e-04	-	1.1562e-03	-	8.3339e-04	-	6.2218e-03	-
	20	4.4149e-05	3.11	1.6267e-04	2.83	8.3654e-05	3.32	1.0372e-03	2.58
	40	5.6599e-06	2.96	3.0007e-05	2.44	9.9132e-06	3.08	2.6203e-04	1.98
	80	6.2220e-07	3.19	6.5691e-06	2.19	1.2533e-06	2.98	6.2988e-05	2.06
3	10	1.6588e-05	-	2.0653e-04	-	4.0395e-05	-	5.4798e-04	-
	20	2.1448e-06	2.95	2.2071e-05	3.23	3.5180e-06	3.52	4.9863e-05	3.46
	40	4.6719e-08	5.52	2.5061e-06	3.14	1.5914e-07	4.47	6.4415e-06	2.95
	80	3.5514e-09	3.72	2.9723e-07	3.08	1.5932e-08	3.32	7.8985e-07	3.03

TABLE 4.11. Over-alternating flux, $\alpha = 1$ and $\beta_1 = \beta_2 = 0$.

k	N	Structured Mesh				Unstructured Mesh			
		u_h		q_h		u_h		q_h	
		L^2 error	order	L^2 error	order	L^2 error	order	L^2 error	order
1	10	7.5445e-03	-	5.3818e-02	-	1.1570e-02	-	8.6852e-02	-
	20	1.6139e-03	2.22	1.7616e-02	1.61	2.3281e-03	2.31	2.4799e-02	1.81
	40	3.8453e-04	2.07	6.0410e-03	1.54	6.4049e-04	1.86	9.2793e-03	1.42
	80	1.0052e-04	1.94	2.2353e-03	1.43	1.5789e-04	2.02	3.7932e-03	1.29
2	10	4.9145e-04	-	3.9924e-03	-	9.5467e-04	-	8.0941e-03	-
	20	6.8313e-05	2.85	7.1245e-04	2.49	1.0700e-04	3.16	1.3089e-03	2.63
	40	8.0090e-06	3.09	1.3552e-04	2.39	1.5096e-05	2.83	2.7658e-04	2.24
	80	9.9967e-07	3.00	2.8106e-05	2.27	1.8251e-06	3.05	6.0775e-05	2.19
3	10	1.0466e-05	-	1.4046e-04	-	3.3397e-05	-	4.2220e-04	-
	20	6.1685e-07	4.08	1.1292e-05	3.64	1.5874e-06	4.40	2.8180e-05	3.91
	40	4.0218e-08	3.94	1.0319e-06	3.45	1.1596e-07	3.77	3.2837e-06	3.10
	80	2.4527e-09	4.04	1.0490e-07	3.30	7.3146e-09	3.99	3.7745e-07	3.12

TABLE 4.12. Central flux, $\alpha = \beta_1 = \beta_2 = 0$.

k	N	Structured Mesh				Unstructured Mesh			
		u_h		q_h		u_h		q_h	
		L^2 error	order	L^2 error	order	L^2 error	order	L^2 error	order
1	10	8.0547e-03	-	8.1528e-02	-	1.1561e-02	-	1.2912e-01	-
	20	2.0318e-03	1.99	3.1726e-02	1.36	2.6735e-03	2.11	4.6672e-02	1.47
	40	3.6300e-04	2.48	1.3536e-02	1.23	6.5505e-04	2.03	2.1016e-02	1.15
	80	8.8736e-05	2.03	6.1701e-03	1.13	1.4980e-04	2.13	9.7653e-03	1.11
2	10	5.2336e-04	-	1.0297e-03	-	1.0457e-03	-	6.3339e-03	-
	20	5.5574e-05	3.24	1.0768e-04	3.26	1.0053e-04	3.38	1.0470e-03	2.60
	40	7.1120e-06	2.97	1.2208e-05	3.14	1.1529e-05	3.12	2.6773e-04	1.97
	80	8.9970e-07	2.98	1.4562e-06	3.07	1.3755e-06	3.07	6.4992e-05	2.04
3	10	3.5774e-05	-	2.2620e-04	-	6.0198e-05	-	5.8612e-04	-
	20	5.2642e-06	2.76	2.4305e-05	3.22	6.4845e-06	3.21	5.4978e-05	3.41
	40	7.6318e-07	2.79	2.7756e-06	3.13	8.7760e-07	2.89	7.0664e-06	2.96
	80	1.0198e-07	2.90	3.3033e-07	3.07	1.1370e-07	2.95	8.6774e-07	3.03
4	5	3.4217e-07	-	1.9074e-06	-	2.5687e-06	-	2.6427e-05	-
	10	1.1691e-08	4.87	5.0967e-08	5.23	6.1250e-08	5.39	1.0179e-06	4.70
	20	4.1002e-10	4.83	1.4759e-09	5.11	1.5999e-09	5.26	6.3606e-08	4.00
	40	9.5362e-12	5.43	4.2700e-11	5.11	4.7353e-11	5.08	3.8146e-09	4.06
5	5	1.4612e-08	-	2.3973e-07	-	7.5295e-08	-	1.1371e-06	-
	10	4.1897e-10	5.12	6.7241e-09	5.16	1.8766e-09	5.33	2.4269e-08	5.55
	20	1.3005e-11	5.01	1.9870e-10	5.08	5.2950e-11	5.15	8.6451e-10	4.81
	40	4.0247e-13	5.01	9.0715e-12	4.45	1.7536e-12	4.92	3.0559e-11	4.82

be obtained after integration from $t = 0$ to $t = T$. Then we apply the triangle inequality

$$(A.6) \quad (\|e_q\|^2 + \|e_u\|)^{\frac{1}{2}} \leq (\|\Pi_q^* q\|^2 + \|\Pi_u^* u\|)^{\frac{1}{2}} + (\|\eta_q\|^2 + \|\eta_u\|)^{\frac{1}{2}},$$

and the approximation estimate (2.10) to complete the proof. \square

TABLE 4.13. \mathbf{q} -dissipative flux, $\alpha = 0$, $\beta_1 = 0$ and $\beta_2 = 1/2$.

		Structured Mesh				Unstructured Mesh			
		u_h		\mathbf{q}_h		u_h		\mathbf{q}_h	
k	N	L^2 error	order	L^2 error	order	L^2 error	order	L^2 error	order
1	10	1.4881e-02	-	1.1676e-02	-	1.7837e-02	-	3.5105e-02	-
	20	3.4331e-03	2.12	3.1640e-03	1.88	3.9628e-03	2.17	6.4256e-03	2.45
	40	8.5600e-04	2.00	8.1205e-04	1.96	9.7125e-04	2.03	1.3775e-03	2.22
	80	2.1257e-04	2.01	2.0537e-04	1.98	2.4227e-04	2.00	3.1647e-04	2.12
2	10	6.0834e-04	-	4.3773e-04	-	1.1920e-03	-	1.0406e-03	-
	20	6.8614e-05	3.15	5.1948e-05	3.07	1.1414e-04	3.38	1.0235e-04	3.35
	40	8.4038e-06	3.03	6.3438e-06	3.03	1.3905e-05	3.04	1.3380e-05	2.94
	80	1.0388e-06	3.02	7.8970e-07	3.01	1.7222e-06	3.01	1.6796e-06	2.99
3	10	3.9384e-05	-	2.2997e-05	-	6.5884e-05	-	5.8191e-05	-
	20	5.8770e-06	2.74	1.5047e-06	3.93	6.5714e-06	3.33	3.1109e-06	4.23
	40	8.4586e-07	2.80	9.8013e-08	3.94	8.7841e-07	2.90	2.0501e-07	3.92
	80	1.1400e-07	2.89	6.2875e-09	3.96	1.2493e-07	2.81	1.3664e-08	3.91
4	5	7.5218e-07	-	4.3217e-07	-	2.8923e-06	-	1.9451e-06	-
	10	2.1901e-08	5.10	1.1739e-08	5.20	6.1400e-08	5.56	4.2881e-08	5.50
	20	6.7750e-10	5.01	3.6661e-10	5.00	1.7827e-09	5.11	1.4807e-09	4.86
	40	2.1541e-11	4.98	1.0975e-11	5.06	5.5384e-11	5.01	4.7525e-11	4.96
5	5	4.6857e-08	-	1.8869e-08	-	1.4166e-07	-	1.0427e-07	-
	10	1.9530e-09	4.58	3.3150e-10	5.83	3.1429e-09	5.49	1.2317e-09	6.40
	20	7.2504e-11	4.75	5.6570e-12	5.87	1.0517e-10	4.90	1.9082e-11	6.01
	40	2.5106e-12	4.85	2.0747e-13	4.77	3.8834e-12	4.76	4.0028e-13	5.58

TABLE 4.14. u -dissipative flux, $\alpha = 0$, $\beta_1 = 1/2$ and $\beta_2 = 0$.

		Structured Mesh				Unstructured Mesh			
		u_h		\mathbf{q}_h		u_h		\mathbf{q}_h	
k	N	L^2 error	order	L^2 error	order	L^2 error	order	L^2 error	order
1	10	7.6423e-03	-	8.0092e-02	-	1.1034e-02	-	1.2915e-01	-
	20	1.7911e-03	2.09	3.1504e-02	1.35	2.4093e-03	2.20	4.6686e-02	1.47
	40	4.3343e-04	2.05	1.3478e-02	1.22	6.1943e-04	1.96	2.1030e-02	1.15
	80	1.0686e-04	2.02	6.1556e-03	1.13	1.5486e-04	2.00	9.7687e-03	1.11
2	10	3.7076e-04	-	1.3085e-03	-	8.3589e-04	-	6.6912e-03	-
	20	4.0230e-05	3.20	1.4661e-04	3.16	8.0815e-05	3.37	1.0723e-03	2.64
	40	4.8534e-06	3.05	1.7309e-05	3.08	9.4463e-06	3.10	2.6957e-04	1.99
	80	5.9924e-07	3.02	2.0996e-06	3.04	1.1502e-06	3.04	6.5135e-05	2.05
3	10	1.8558e-05	-	2.6631e-04	-	4.0826e-05	-	6.0453e-04	-
	20	1.2137e-06	3.93	3.0866e-05	3.11	2.1479e-06	4.25	5.7666e-05	3.39
	40	8.1576e-08	3.90	3.6862e-06	3.07	1.4224e-07	3.92	7.3768e-06	2.97
	80	4.8797e-09	4.06	4.4932e-07	3.04	9.6926e-09	3.88	9.0796e-07	3.02
4	5	3.6692e-07	-	2.2798e-06	-	2.0421e-06	-	2.7528e-05	-
	10	9.5674e-09	5.26	6.7639e-08	5.07	4.3293e-08	5.56	1.0388e-06	4.73
	20	2.8146e-10	5.09	2.0648e-09	5.03	1.1560e-09	5.23	6.3943e-08	4.02
	40	8.5796e-12	5.04	6.2591e-11	5.04	3.6656e-11	4.98	3.8220e-09	4.06
5	5	2.2108e-08	-	2.9250e-07	-	5.4020e-08	-	1.1670e-06	-
	10	3.9557e-10	5.80	9.1281e-09	5.00	7.3914e-10	6.19	2.5465e-08	5.52
	20	6.7129e-12	5.88	2.8716e-10	4.99	1.3607e-11	5.76	9.0183e-10	4.82
	40	1.4282e-13	5.55	1.1507e-11	4.64	2.6472e-13	5.68	3.1748e-11	4.83

APPENDIX B. PROOF OF THEOREM 2.4

Proof. For any fixed $\tau \leq T$, we use the superscript t to indicate the integration in time from t to τ , for example,

$$(B.1) \quad q_h^t(x, t) = \int_t^\tau q_h(x, s) ds \quad \text{and} \quad u_h^t(x, t) = \int_t^\tau u_h(x, s) ds.$$

We define the bilinear form

$$(B.2) \quad \check{a}_h(q_h, u_h; v, w) = (u_h^t, v_x) + (q_h, w_x^t) + \sum_j (\mathcal{F}_u(u_h^t, q_h^t) \llbracket v \rrbracket + \mathcal{F}_q(q_h, u_h) \llbracket w^t \rrbracket)_{j+\frac{1}{2}}.$$

It can be shown that

$$(B.3) \quad \check{a}_h(q_h, u_h; q_h, u_h) = \frac{1}{2} \frac{d}{dt} \sum_j \left(\beta_1 \llbracket u_h^t \rrbracket^2 + \beta_2 \llbracket q_h^t \rrbracket^2 \right).$$

Since the time integration commutes with the projection operator Π_q^* and Π_u^* , we have $(\eta_q)^t = \eta_{q^t}$ and $(\eta_u)^t = \eta_{u^t}$. It can be shown that

$$(B.4) \quad \check{a}_h(\eta_q, \eta_u; v, w) = 0.$$

Note we have the error equations

$$(B.5) \quad ((e_u)_{tt}, w) + (e_q, w_x) + \sum_j (\mathcal{F}_q(e_q, e_u) \llbracket w \rrbracket)_{j+\frac{1}{2}} = 0,$$

$$(B.6) \quad (e_q, v) + (e_u, v_x) + \sum_j (\mathcal{F}_u(e_u, e_q) \llbracket v \rrbracket)_{j+\frac{1}{2}} = 0.$$

Using the chain rule and replacing w with w^t , (B.5) implies

$$(B.7) \quad ((\Pi_u^* e_u)_t, w) + (e_q, w_x^t) + \sum_j (\mathcal{F}_q(e_q, e_u) \llbracket w^t \rrbracket)_{j+\frac{1}{2}} = -((\eta_u)_{tt}, w^t) - \frac{d}{dt} ((\Pi_u^* e_u)_t, w^t),$$

where we have used the fact $(w^t)_t = -w$. Integrating (B.6) in time from t to τ gives

$$(B.8) \quad ((\Pi_q^* e_q)^t, v) + ((e_u)^t, v_x) + \sum_j (\mathcal{F}_u((e_u)^t, (e_q)^t) \llbracket v \rrbracket)_{j+\frac{1}{2}} = -((\eta_q)^t, v).$$

Combining the two equations and noting $(\eta_u)_{tt} = \eta_{u_{tt}}$, $(\eta_q)^t = \eta_{q^t}$, it then yields that

$$(B.9) \quad \begin{aligned} & ((\Pi_q^* e_q)^t, v) + ((\Pi_u^* e_u)_t, w) + \check{a}_h(e_q, e_u; v, w) \\ & = -(\eta_{q^t}, v) - ((\eta_{u_{tt}}), w^t) - \frac{d}{dt} ((\Pi_u^* e_u)_t, w^t), \end{aligned}$$

Hence by taking $v = \Pi_q^* e_q$ and $w = \Pi_u^* e_u$, one can apply (B.4) to obtain that

$$(B.10) \quad \begin{aligned} & \frac{1}{2} \frac{d}{dt} (\|\Pi_u^* e_u\|^2 - \|(\Pi_q^* e_q)^t\|^2) + \check{a}_h(\Pi_q^* e_q, \Pi_u^* e_u; \Pi_q^* e_q, \Pi_u^* e_u) \\ & = -(\eta_{q^t}, \Pi_q^* e_q) - ((\eta_{u_{tt}}), (\Pi_u^* e_u)^t) - \frac{d}{dt} ((\Pi_u^* e_u)_t, (\Pi_u^* e_u)^t). \end{aligned}$$

Combining it with (B.3), we have

$$\begin{aligned}
 \text{(B.11)} \quad & \frac{1}{2} \frac{d}{dt} \left(\|\Pi_u^* e_u\|^2 - \|(\Pi_q^* e_q)^t\|^2 + \sum_j \left(\beta_1 \llbracket (\Pi_u^* e_u)^t \rrbracket^2 + \beta_2 \llbracket (\Pi_q^* e_q)^t \rrbracket^2 \right)_{j+\frac{1}{2}} \right) \\
 & = -(\eta_{q^t}, \Pi_q^* e_q) - ((\eta_{u_{tt}}), (\Pi_u^* e_u)^t) - \frac{d}{dt} ((\Pi_u^* e_u)_t, (\Pi_u^* e_u)^t).
 \end{aligned}$$

We then integrate t from 0 to τ and note $w^t|_{t=\tau} = 0$ to obtain

$$\begin{aligned}
 \text{(B.12)} \quad & \frac{1}{2} \|\Pi_u^* e_u\|^2|_{t=\tau} - \frac{1}{2} (\|\Pi_u^* e_u\|^2 - \|(\Pi_q^* e_q)^t\|^2)|_{t=0} \\
 & = \frac{1}{2} \sum_j \left(\beta_1 \llbracket (\Pi_u^* e_u)^t \rrbracket^2 + \beta_2 \llbracket (\Pi_q^* e_q)^t \rrbracket^2 \right)_{j+\frac{1}{2}} \Big|_{t=0} \\
 & \quad - \int_0^\tau (\eta_{q^t}, \Pi_q^* e_q) + ((\eta_{u_{tt}}), (\Pi_u^* e_u)^t) dt + ((\Pi_u^* e_u)_t, (\Pi_u^* e_u)^t)|_{t=0} \\
 & := I + II + III.
 \end{aligned}$$

Using our previous estimates, we have

$$\text{(B.13)} \quad I \leq \frac{T^2}{2} \sup_{t \in [0, \tau]} \sum_j \left(\beta_1 \llbracket \Pi_u^* e_u \rrbracket^2 + \beta_2 \llbracket \Pi_q^* e_q \rrbracket^2 \right)_{j+\frac{1}{2}} \leq CT^2 h^{2k+2}.$$

By applying the Cauchy–Schwarz inequality, it can be shown that

$$\begin{aligned}
 \text{(B.14)} \quad II & \leq C \sup_{t \in [0, T]} (\|\eta_{q^t}\|^2 + \|\eta_{u_{tt}}\|^2 + T^2 \|\Pi_u^* e_u\|^2 + \|\Pi_q^* e_q\|^2) \\
 & \leq CT^2 h^{2k+2} + \frac{1}{4} \sup_{t \in [0, T]} \|\Pi_u^* e_u\|^2.
 \end{aligned}$$

With the special choice of initial condition, we have $III = 0$.

Substitute the estimates above into (B.12) and take supreme with respect to $t \in [0, T]$. It yields that

$$\text{(B.15)} \quad \sup_{t \in [0, T]} \|\Pi_u^* e_u\| \leq C(T + 1)h^{k+1}.$$

The estimate of $\|u - u_h\|$ then follows from the approximation property of Π_u^* . \square

APPENDIX C. DIMENSION COUNT

Note that total number of equations in (3.17) is

$$\text{(C.1)} \quad N_{\text{equation}} = \sum_{K \in \mathcal{T}_h} (\dim(P_{k-1}(K)) + \dim(\mathbf{P}_{k-1}(K))) + 2 \sum_{F \in \Gamma} \dim(P_k(F)).$$

The total degree of freedom is

$$\text{(C.2)} \quad N_{\text{unknown}} = \sum_{K \in \mathcal{T}_h} (\dim(P_k(K)) + \dim(\mathbf{P}_k(K))).$$

Since $2|\Gamma| = (d + 1)|\mathcal{T}_h|$ and

$$(C.3) \quad \dim(P_{k-1}(K)) + \dim(\mathbf{P}_{k-1}(K)) = \binom{k-1+d}{d} \times (d+1),$$

$$(C.4) \quad \dim(P_k(F)) = \binom{k+d-1}{d-1},$$

we have

$$(C.5) \quad N_{\text{equation}} = \left(\binom{k-1+d}{d} + \binom{k+d-1}{d-1} \right) \times (d+1) \times |\mathcal{T}_h|.$$

On the other hand, it can be seen that

$$(C.6) \quad N_{\text{unknown}} = \binom{k+d}{d} \times (d+1) \times |\mathcal{T}_h|.$$

We can deduce $N_{\text{equation}} = N_{\text{unknown}}$ by recalling the combinatorial identity

$$(C.7) \quad \binom{k-1+d}{d} + \binom{k+d-1}{d-1} = \binom{k+d}{d}.$$

APPENDIX D. PROOF OF LEMMA 3.3

Proof. We denote by $P_k^\perp(K)$ the orthogonal complement of $P_{k-1}(K)$ with respect to $P_k(K)$. Similar to the equivalence of (3.30c)-(3.30d) and (3.31), one can show that (3.36c)-(3.36d) lead to

$$(D.1) \quad \langle \delta_u - \delta_{\mathbf{q}} \cdot \mathbf{n}, \mu \rangle_F = \langle \xi - \zeta \cdot \mathbf{n}, \mu \rangle_F, \quad \forall \mu \in P_k(F), \quad \forall F \in \partial K, \quad \forall K \in \mathcal{T}_h.$$

(3.37a)-(3.37b) implies that $\delta_{\mathbf{q}}|_K \in \mathbf{P}_k^\perp(K)$ and $\delta_u|_K \in P_k^\perp(K)$. Note that

$$(D.2) \quad \langle \delta_{\mathbf{q}} \cdot \mathbf{n}, w \rangle_{\partial K} = \int_K \nabla \cdot (\delta_{\mathbf{q}} w) dx = \int_K (\nabla \cdot \delta_{\mathbf{q}}) w dx + \int_K \delta_{\mathbf{q}} \cdot \nabla w dx = 0, \quad \forall w \in P_k^\perp(K).$$

Taking $\mu = \delta_u|_K \in P_k^\perp(K)$ in (D.1) and summing over all $F \in \partial K$, it can be shown that

$$(D.3) \quad \|\delta_u\|_F \leq \|\delta_u\|_{\partial K} \leq \|\zeta \cdot \mathbf{n}\|_{\partial K} + \|\xi\|_{\partial K},$$

where $F \in \partial K$ be any face of K . One can also use a scaling argument to show [13, Lemma A.1]

$$(D.4) \quad \|\delta_u\|_K \leq Ch^{\frac{1}{2}} \|\delta_u\|_F.$$

Combine (D.3) with (D.4), take the square, and then sum over all elements K . It gives

$$(D.5) \quad \|\delta_u\|^2 \leq Ch \left(\|\zeta \cdot \mathbf{n}\|_\Gamma^2 + \|\xi\|_\Gamma^2 \right).$$

We now estimate $\|\delta_{\mathbf{q}}\|$. Let us first choose an arbitrary face $F^* \in \partial K$ and consider a fixed face $F \in \partial K \setminus \{F^*\}$. Rearranging terms in (D.1), it gives

$$(D.6) \quad \langle \delta_{\mathbf{q}} \cdot \mathbf{n}, \mu \rangle_F = \langle \zeta \cdot \mathbf{n} - \xi + \delta_u, \mu \rangle_F, \quad \forall \mu \in P_k(F).$$

We can then take $\mu = \delta_{\mathbf{q}} \cdot \mathbf{n}_F$ and apply (D.3) to obtain

$$(D.7) \quad \|\delta_{\mathbf{q}} \cdot \mathbf{n}_F\|_F \leq \|\zeta \cdot \mathbf{n}_F\|_F + \|\xi\|_F + \|\delta_u\|_F \leq 2(\|\zeta \cdot \mathbf{n}\|_{\partial K} + \|\xi\|_{\partial K}).$$

Note we have $\delta_{\mathbf{q}} \cdot \mathbf{n}_F \in P_k^\perp(K)$ for a fixed \mathbf{n}_F . Using the scaling argument as that for (D.4), one can obtain $\|\delta_{\mathbf{q}} \cdot \mathbf{n}_F\|_K \leq Ch^{\frac{1}{2}} \|\delta_{\mathbf{q}} \cdot \mathbf{n}\|_F$. With the estimate in (D.7), it can be seen that

$$(D.8) \quad \|\delta_{\mathbf{q}} \cdot \mathbf{n}_F\|_K \leq Ch^{\frac{1}{2}} (\|\zeta \cdot \mathbf{n}\|_{\partial K} + \|\xi\|_{\partial K}).$$

Note that $\delta_{\mathbf{q}} = \sum_{F \in \partial K \setminus \{F^*\}} (\delta_{\mathbf{q}} \cdot \mathbf{n}_F) \tilde{\mathbf{n}}_F$, where $\tilde{\mathbf{n}}_F$ is the dual of \mathbf{n}_F such that $\tilde{\mathbf{n}}_F \cdot \mathbf{n}_F = \delta_{FF^*}$. Therefore, we have

$$(D.9) \quad \|\delta_{\mathbf{q}}\|_K \leq \sum_{F \in \partial K \setminus \{F^*\}} \|\delta_{\mathbf{q}} \cdot \mathbf{n}_F\|_K \leq Ch^{\frac{1}{2}} (\|\zeta \cdot \mathbf{n}\|_{\partial K} + \|\xi\|_{\partial K}).$$

Finally, we take the square of (D.9) and then sum over all elements K to obtain

$$(D.10) \quad \|\delta_{\mathbf{q}}\|^2 \leq Ch \left(\|\zeta \cdot \mathbf{n}\|_\Gamma^2 + \|\xi\|_\Gamma^2 \right).$$

The combination of (D.5) and (D.10) leads to (3.39), which finishes the proof. \square

ACKNOWLEDGMENT

The first author would like to thank Professor Guosheng Fu at the University of Notre Dame for helpful comments on the implementation of the DG solver with NGSolve.

REFERENCES

- [1] Mark Ainsworth and Guosheng Fu, *Dispersive behavior of an energy-conserving discontinuous Galerkin method for the one-way wave equation*, J. Sci. Comput. **79** (2019), no. 1, 209–226, DOI 10.1007/s10915-018-0846-z. MR3936245
- [2] M. Ainsworth, P. Monk, and W. Muniz, *Dispersive and dissipative properties of discontinuous Galerkin finite element methods for the second-order wave equation*, J. Sci. Comput. **27** (2006), no. 1-3, 5–40, DOI 10.1007/s10915-005-9044-x. MR2285764
- [3] Douglas N. Arnold, Franco Brezzi, Bernardo Cockburn, and L. Donatella Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal. **39** (2001/02), no. 5, 1749–1779, DOI 10.1137/S0036142901384162. MR1885715
- [4] F. Bassi and S. Rebay, *A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations*, J. Comput. Phys. **131** (1997), no. 2, 267–279, DOI 10.1006/jcph.1996.5572. MR1433934
- [5] J. L. Bona, H. Chen, O. Karakashian, and Y. Xing, *Conservative, discontinuous Galerkin-methods for the generalized Korteweg-de Vries equation*, Math. Comp. **82** (2013), no. 283, 1401–1432, DOI 10.1090/S0025-5718-2013-02661-0. MR3042569
- [6] Yingda Cheng, Ching-Shan Chou, Fengyan Li, and Yulong Xing, *L^2 stable discontinuous Galerkin methods for one-dimensional two-way wave equations*, Math. Comp. **86** (2017), no. 303, 121–155, DOI 10.1090/mcom/3090. MR3557796
- [7] Yao Cheng, Xiong Meng, and Qiang Zhang, *Application of generalized Gauss-Radau projections for the local discontinuous Galerkin method for linear convection-diffusion equations*, Math. Comp. **86** (2017), no. 305, 1233–1267, DOI 10.1090/mcom/3141. MR3614017
- [8] Ching-Shan Chou, Chi-Wang Shu, and Yulong Xing, *Optimal energy conserving local discontinuous Galerkin methods for second-order wave equation in heterogeneous media*, J. Comput. Phys. **272** (2014), 88–107, DOI 10.1016/j.jcp.2014.04.009. MR3212263
- [9] Eric T. Chung and Björn Engquist, *Optimal discontinuous Galerkin methods for wave propagation*, SIAM J. Numer. Anal. **44** (2006), no. 5, 2131–2158, DOI 10.1137/050641193. MR2263043
- [10] Bernardo Cockburn and Bo Dong, *An analysis of the minimal dissipation local discontinuous Galerkin method for convection-diffusion problems*, J. Sci. Comput. **32** (2007), no. 2, 233–262, DOI 10.1007/s10915-007-9130-3. MR2320571
- [11] Bernardo Cockburn, Zhixing Fu, Allan Hungria, Liangyue Ji, Manuel A. Sánchez, and Francisco-Javier Sayas, *Stormer-Numerov HDG methods for acoustic waves*, J. Sci. Comput. **75** (2018), no. 2, 597–624, DOI 10.1007/s10915-017-0547-z. MR3780780

- [12] Bernardo Cockburn, Jayadeep Gopalakrishnan, and Raytcho Lazarov, *Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal. **47** (2009), no. 2, 1319–1365, DOI 10.1137/070706616. MR2485455
- [13] Bernardo Cockburn, Jayadeep Gopalakrishnan, and Francisco-Javier Sayas, *A projection-based error analysis of HDG methods*, Math. Comp. **79** (2010), no. 271, 1351–1367, DOI 10.1090/S0025-5718-10-02334-3. MR2629996
- [14] Bernardo Cockburn, Suchung Hou, and Chi-Wang Shu, *The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case*, Math. Comp. **54** (1990), no. 190, 545–581, DOI 10.2307/2008501. MR1010597
- [15] Bernardo Cockburn, San Yih Lin, and Chi-Wang Shu, *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. III. One-dimensional systems*, J. Comput. Phys. **84** (1989), no. 1, 90–113, DOI 10.1016/0021-9991(89)90183-6. MR1015355
- [16] Bernardo Cockburn and Vincent Quenneville-Bélaïr, *Uniform-in-time superconvergence of the HDG methods for the acoustic wave equation*, Math. Comp. **83** (2014), no. 285, 65–85, DOI 10.1090/S0025-5718-2013-02743-3. MR3120582
- [17] Bernardo Cockburn and Chi-Wang Shu, *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework*, Math. Comp. **52** (1989), no. 186, 411–435, DOI 10.2307/2008474. MR983311
- [18] Bernardo Cockburn and Chi-Wang Shu, *The Runge-Kutta local projection P^1 -discontinuous-Galerkin finite element method for scalar conservation laws* (English, with French summary), RAIRO Modél. Math. Anal. Numér. **25** (1991), no. 3, 337–361, DOI 10.1051/m2an/1991250303371. MR1103092
- [19] Bernardo Cockburn and Chi-Wang Shu, *The Runge-Kutta discontinuous Galerkin method for conservation laws. V. Multidimensional systems*, J. Comput. Phys. **141** (1998), no. 2, 199–224, DOI 10.1006/jcph.1998.5892. MR1619652
- [20] Gary C. Cohen, *Higher-order Numerical Methods for Transient Wave Equations: With a foreword by R. Glowinski*, Scientific Computation, Springer-Verlag, Berlin, 2002. MR1870851
- [21] Richard S. Falk and Gerard R. Richter, *Explicit finite element methods for symmetric hyperbolic equations*, SIAM J. Numer. Anal. **36** (1999), no. 3, 935–952, DOI 10.1137/S0036142997329463. MR1688992
- [22] Guosheng Fu and Chi-Wang Shu, *Optimal energy-conserving discontinuous Galerkin methods for linear symmetric hyperbolic systems*, J. Comput. Phys. **394** (2019), 329–363, DOI 10.1016/j.jcp.2019.05.050. MR3960714
- [23] Pei Fu, Yingda Cheng, Fengyan Li, and Yan Xu, *Discontinuous Galerkin methods with optimal L^2 accuracy for one dimensional linear PDEs with high order spatial derivatives*, J. Sci. Comput. **78** (2019), no. 2, 816–863, DOI 10.1007/s10915-018-0788-5. MR3918671
- [24] Marcus J. Grote, Anna Schneebeli, and Dominik Schötzau, *Discontinuous Galerkin finite element method for the wave equation*, SIAM J. Numer. Anal. **44** (2006), no. 6, 2408–2431, DOI 10.1137/05063194X. MR2272600
- [25] Guang Shan Jiang and Chi-Wang Shu, *On a cell entropy inequality for discontinuous Galerkin methods*, Math. Comp. **62** (1994), no. 206, 531–538, DOI 10.2307/2153521. MR1223232
- [26] Jia Li, Dazhi Zhang, Xiong Meng, Boying Wu, and Qiang Zhang, *Discontinuous Galerkin methods for nonlinear scalar conservation laws: generalized local Lax-Friedrichs numerical fluxes*, SIAM J. Numer. Anal. **58** (2020), no. 1, 1–20, DOI 10.1137/19M1243798. MR4046779
- [27] Hailiang Liu and Nattapol Ploymaklam, *A local discontinuous Galerkin method for the Burgers-Poisson equation*, Numer. Math. **129** (2015), no. 2, 321–351, DOI 10.1007/s00211-014-0641-1. MR3300422
- [28] Minghui Liu, Boying Wu, and Xiong Meng, *Optimal error estimates of the discontinuous Galerkin method with upwind-biased fluxes for 2D linear variable coefficients hyperbolic equations*, J. Sci. Comput. **83** (2020), no. 1, Paper No. 9, 19, DOI 10.1007/s10915-020-01197-x. MR4079266
- [29] Yong Liu, Jianfang Lu, Chi-Wang Shu, and Mengping Zhang, *Central discontinuous Galerkin methods on overlapping meshes for wave equations*, ESAIM Math. Model. Numer. Anal. **55** (2021), no. 1, 329–356, DOI 10.1051/m2an/2020069. MR4216835

- [30] Y. Liu, C.-W. Shu, and M. Zhang, *Sub-optimal convergence of discontinuous Galerkin methods with central fluxes for even degree polynomial approximations*, Journal of Computational Mathematics, to appear.
- [31] Xiong Meng, Chi-Wang Shu, and Boying Wu, *Optimal error estimates for discontinuous Galerkin methods based on upwind-biased fluxes for linear hyperbolic equations*, Math. Comp. **85** (2016), no. 299, 1225–1261, DOI 10.1090/mcom/3022. MR3454363
- [32] Peter Monk and Gerard R. Richter, *A discontinuous Galerkin method for linear symmetric hyperbolic systems in inhomogeneous media*, J. Sci. Comput. **22/23** (2005), 443–477, DOI 10.1007/s10915-004-4132-5. MR2142205
- [33] N. C. Nguyen, J. Peraire, and B. Cockburn, *High-order implicit hybridizable discontinuous Galerkin methods for acoustics and elastodynamics*, J. Comput. Phys. **230** (2011), no. 10, 3695–3718, DOI 10.1016/j.jcp.2011.01.035. MR2783813
- [34] W. H. Reed and T. Hill, *Triangular mesh methods for the neutron transport equation*, Technical report, Los Alamos Scientific Lab., N. Mex.(USA), 1973.
- [35] M. A. Sánchez, C. Ciuca, N. C. Nguyen, J. Peraire, and B. Cockburn, *Symplectic Hamiltonian HDG methods for wave propagation phenomena*, J. Comput. Phys. **350** (2017), 951–973, DOI 10.1016/j.jcp.2017.09.010. MR3707190
- [36] J. Schöberl, *NETGEN an advancing front 2D/3D-mesh generator based on abstract rules*, Computing and visualization in science, 1(1):41–52, 1997.
- [37] M. Stanglmeier, N. C. Nguyen, J. Peraire, and B. Cockburn, *An explicit hybridizable discontinuous Galerkin method for the acoustic wave equation*, Comput. Methods Appl. Mech. Engrg. **300** (2016), 748–769, DOI 10.1016/j.cma.2015.12.003. MR3452794
- [38] Zheng Sun and Chi-Wang Shu, *Stability analysis and error estimates of Lax-Wendroff discontinuous Galerkin methods for linear conservation laws*, ESAIM Math. Model. Numer. Anal. **51** (2017), no. 3, 1063–1087, DOI 10.1051/m2an/2016049. MR3666657
- [39] Zheng Sun and Chi-Wang Shu, *Stability of the fourth order Runge-Kutta method for time-dependent partial differential equations*, Ann. Math. Sci. Appl. **2** (2017), no. 2, 255–284, DOI 10.4310/AMSA.2017.v2.n2.a3. MR3690156
- [40] Zheng Sun and Chi-wang Shu, *Strong stability of explicit Runge-Kutta time discretizations*, SIAM J. Numer. Anal. **57** (2019), no. 3, 1158–1182, DOI 10.1137/18M122892X. MR3953465
- [41] Z. Sun and C.-W. Shu. Error analysis of Runge–Kutta discontinuous Galerkin methods for linear time-dependent partial differential equations. arXiv preprint arXiv:2001.00971, 2020.
- [42] Zheng Sun and Yulong Xing, *On structure-preserving discontinuous Galerkin methods for Hamiltonian partial differential equations: energy conservation and multi-symplecticity*, J. Comput. Phys. **419** (2020), 109662, 25, DOI 10.1016/j.jcp.2020.109662. MR4116802
- [43] Yulong Xing, Ching-Shan Chou, and Chi-Wang Shu, *Energy conserving local discontinuous Galerkin methods for wave propagation problems*, Inverse Probl. Imaging **7** (2013), no. 3, 967–986, DOI 10.3934/ipi.2013.7.967. MR3105364
- [44] Yuan Xu, Chi-Wang Shu, and Qiang Zhang, *Error Estimate of the Fourth-Order Runge–Kutta Discontinuous Galerkin Methods for Linear Hyperbolic Equations*, SIAM J. Numer. Anal. **58** (2020), no. 5, 2885–2914, DOI 10.1137/19M1280077. MR4161755
- [45] Qiang Zhang and Chi-Wang Shu, *Stability analysis and a priori error estimates of the third order explicit Runge-Kutta discontinuous Galerkin method for scalar conservation laws*, SIAM J. Numer. Anal. **48** (2010), no. 3, 1038–1063, DOI 10.1137/090771363. MR2669400

DEPARTMENT OF MATHEMATICS, THE OHIO STATE UNIVERSITY, COLUMBUS, OHIO 43210
 Email address: sun.2516@osu.edu

DEPARTMENT OF MATHEMATICS, THE OHIO STATE UNIVERSITY, COLUMBUS, OHIO 43210
 Email address: xing.205@osu.edu