ICES REPORT 11-45

December 2011

A Relation between the Discontinuous Petrov--Galerkin Method and the Discontinuous Galerkin Method

by

Tan Bui-Thanh, Omar Ghattas, and Leszek Demkowicz

The Institute for Computational Engineering and Sciences
The University of Texas at Austin
Austin, Texas 78712

Reference: Tan Bui-Thanh, Omar Ghattas, and Leszek Demkowicz, A Relation between the Discontinuous Petrov--Galerkin Method and the Discontinuous Galerkin Method, ICES REPORT 11-45, The Institute for Computational Engineering and Sciences, The University of Texas at Austin, December 2011.

A RELATION BETWEEN THE DISCONTINUOUS PETROV–GALERKIN METHOD AND THE DISCONTINUOUS GALERKIN METHOD

TAN BUI-THANH [†], OMAR GHATTAS [†‡§], AND LESZEK DEMKOWICZ [†]

Abstract. This paper is an attempt in seeking a connection between the discontinuous Petrov–Galerkin method of Demkowicz and Gopalakrishnan [13,15] and the popular discontinuous Galerkin method. Starting from a discontinuous Petrov–Galerkin (DPG) method with zero enriched order we re-derive a large class of discontinuous Galerkin (DG) methods for first order hyperbolic and elliptic equations. The first implication of this result is that the DG method can be considered as the least accurate DPG method. The second implication is that the DPG method can be viewed as a systematic way to improve the accuracy of the DG method when nonzero enriched orders are employed. A detailed derivation of the upwind DG, the local DG, and the hybridized DG from a DPG method with optimal test norms will be presented.

Key words. discontinuous Petrov–Galerkin methods; discontinuous Galerkin methods; well-posedness; partial differential equations; inf–sup condition;

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

1. Introduction.

The discontinuous Petrov–Galerkin (DPG) framework introduced by Demkowicz and Gopalakrishnan [13,15] has been evolving as a new numerical methodology for partial differential equations (PDEs). The method has been successfully applied to a wide range of PDEs including scalar transport [5,13,15], Laplace [14], convection-diffusion [14,15], Helmholtz [16,17,27], Burgers and Navier-Stokes [8], and linear elasticity [4] equations. The DPG framework starts by partitioning the domain of interest into non-overlapping elements. Variational formulations are posed for each element separately and then summed up to form a global variational statement. Elemental solutions are connected by introducing hybrid variables (also known as fluxes or traces) that live on the skeleton of the mesh. This is therefore a mesh-dependent variational approach in which both bilinear and linear forms depend on the mesh under consideration.

In general, the trial and test spaces are not related to each other. In the standard Bubnov–Galerkin (also known as Galerkin) approach, the trial and test spaces are identical, while they differ in a Petrov–Galerkin scheme. Traditionally, one chooses either Galerkin or Petrov–Galerkin approaches, then proves the consistency and stability in both infinite and finite dimensional settings (if possible). The DPG method introduces a new paradigm in which one selects both trial and test spaces at the same time to satisfy well-posedness. In particular, one can select trial and test function spaces for which the continuity and inf–sup constants are unity. Given a finite dimensional trial subspace, the finite dimensional test space is constructed in such a way that the well-posedness of the finite dimensional setting is automatically inherited from the infinite dimensional counterpart.

For example, the DPG method in [15] starts with a given norm in the trial space and then seeks a norm in the test space in order to achieve unity continuity and inf–sup constants. Another DPG method in [16] achieves the same goal but reverses the

[†]Institute for Computational Engineering & Sciences, The University of Texas at Austin, Austin, TX 78712, USA.

[‡]Jackson School of Geosciences, The University of Texas at Austin, Austin, TX 78712, USA.

[§]Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX 78712, USA.

process, i.e., it looks for a norm in the trial space corresponding to a given norm in the test space. Clearly, this is one of the advantages of the DPG methodology, since it allows one to choose a norm of interest to work with, while rendering the error optimal, i.e., smallest in that norm. Furthermore, the DPG methodology provides a natural framework for constructing robust versions of the method for singular perturbation problems, enabling automatic adaptivity. We shall not discuss the advantages of the DPG methods any further here, and the readers are referred to the original DPG papers [13–16] for more details.

The discontinuous Galerkin (DG) method, on the other hand, was originally developed by Reed and Hill [25] for the neutron transport equation, and has been extended to other partial differential equations (PDEs) [1, 9–11]. Roughly speaking, DG combines advantages of classical finite volume and finite element methods. In particular, it has the ability to treat solutions with large gradients including shocks, it provides the flexibility to deal with complex geometries, and it is highly parallelizable due to its compact stencil. The DG literature is vast and an extensive review is out of the scope of the paper.

It has been observed that the DPG method seems to be more accurate than the DG method for the transport equation, given the same polynomial order for the finite dimensional trial subspace [5, 13, 15]. In particular, the numerical results in [5, 13, 15] show that the DPG method is more accurate, more stable, and does not seem to have any noticeable dispersion compared to the DG method for several one- and two-dimensional examples. Moreover, the DG method has sub-optimal h convergence rate on Peterson's meshes [24] for $p < 3$ [5], while h convergence rate of the DPG method is uniformly optimal for all p. This paper is the first attempt in understanding why the DPG methodology could provide a numerical solution that is more accurate than that of the DG method.

Section 2 gives a short introduction to the mathematical theory behind the discontinuous Petrov–Galerkin framework. Section 3 shows in which sense the DG method corresponds the least accurate DPG method for the first order hyperbolic equations. In particular, we re-derive a large class of DG methods from a DPG method with zero enriched order (to be defined). In the similar manner, we re-derive the local DG method and the hybridized DG method from a DPG method with zero enriched order for elliptic equations in Sections 4 and 5, respectively. Finally, Section 6 concludes the paper.

2. A brief review of the discontinuous Petrov–Galerkin theory. In this section, we review the discontinuous Petrov–Galerkin framework presented in [5]. Let U and V be Hilbert spaces over the real line (generalization of our theory to the complex field is straightforward). Consider the following variational problem,

$$\begin{cases} \text{Seek } u \in U \text{ such that} \\ a(u,v) = \ell(v), \quad \forall v \in V, \end{cases} \tag{2.1}$$

where $\ell(\cdot)$ is a linear form on V, $a(\cdot,\cdot)$ is a bilinear form satisfying the continuity condition with continuity constant M,

$$|a(u,v)| \leq M \|u\|_U \|v\|_V , \tag{2.2}$$

the inf–sup condition with the inf–sup constant γ,

$$\exists \gamma > 0 : \inf_{u \in U} \sup_{v \in V} \frac{a(u,v)}{\|u\|_U \|v\|_V} \geq \gamma, \tag{2.3a}$$

and the injectivity of the adjoint operator (to be defined),

$$(a(u,v)=0, \quad \forall u \in U) \Rightarrow (v=0)\,. \tag{2.3b}$$

If (2.2) and (2.3) hold, then by the generalized Lax-Milgram theorem [3,23] (also known as the Banach-Nečas-Babuška theorem [19]), (2.1) has a unique solution and the solution is stable in the following sense,

$$\|u\|_U \leq \frac{1}{\gamma}\|\ell\|_{V'}\,,$$

where V' is the topological dual of V. Note that for convenience in writing, we have abused the notation $\sup_{v\in V}$ instead of $\sup_{v\in V, v\neq 0}$ (and similarly for inf).

Now let $U_h \subset U$ and $V_h \subset V$ be two finite dimensional trial and test spaces, and consider the following finite dimensional approximation problem,

$$\begin{cases} \text{Seek } u_h \in U_h \text{ such that} \\ a(u_h, v_h) = \ell(v_h), \quad \forall v_h \in V_h. \end{cases} \tag{2.4}$$

If $\dim U_h = \dim V_h = n$, and the following discrete inf–sup condition

$$\exists \gamma_h > 0 : \inf_{u_h\in U_h}\sup_{v_h\in V_h}\frac{a(u_h,v_h)}{\|u_h\|_U\|v_h\|_V} \geq \gamma_h \tag{2.5}$$

holds, then the finite dimensional problem (2.4) is well-posed by an application of the generalized Lax-Milgram theorem for finite dimensional problems (also known as the Babuška's Theorem [2,3]). In general, however, the finite dimensional problem (2.4) does not inherit the well-posedness of the infinite dimensional counterpart (2.1) except for some special circumstances. One, therefore, has to prove the non-trivial discrete inf–sup condition [19].

Our goal is to construct finite dimensional approximations that are guaranteed to be trivially well-posed with unity continuity and inf–sup constants. Here, trivial well-posedness means that the well-posedness of the finite dimensional problems is trivially inherited from their infinite dimensional counterparts. We begin by a result on the error between the exact and the finite dimensional approximate solutions.

THEOREM 2.1 (Babuška [2]). *Suppose that both the continuous problem* (2.1) *and discrete problem* (2.4) *are well-posed, then*

$$\|u-u_h\|_U \leq \frac{M}{\gamma_h}\inf_{w_h\in U_h}\|u-w_h\|_U\,.$$

Proof. A standard proof can be found in [12,26]. □

The following best approximation error result immediately follows from Theorem 2.1.

COROLLARY 2.2. *If* $M=\gamma_h$, *then*

$$\|u-u_h\|_U = \inf_{w_h\in U_h}\|u-w_h\|_U\,.$$

In particular, $M=\gamma_h=1$ satisfies Corollary 2.2. That is, if the continuity constant and the discrete inf–sup constant are unity, then the error incurred from the discrete approximation (2.4) is the best, i.e., it is smallest. As can be seen, there are two

spaces to work with, namely the trial and test spaces, respectively. The first DPG method [15] starts with a given norm in the trial space U, and then seeks a norm in the test space V so that $M = \gamma = 1$. In the second DPG method [16], on the other hand, one defines a norm in U from a given norm in V such that $M = \gamma = 1$. Clearly, this is one of the advantages of the DPG methodology since it allows one to choose a norm of interest to work with while making the error optimal, i.e. smallest, in that norm.

On the other hand, it is not necessary to prescribe either norms in the spaces U and V. Indeed, we let the problem speak out its "natural" energy norms, thus which norms to be chosen to work with is out of the question in our new approach. Nevertheless, care must be taken since our idea may not be applicable for cases in which one prefers to work with particular norms.

The following useful result will be used as guidelines to construct the "natural" norms in U and V spaces such that $M = \gamma = 1$.

THEOREM 2.3. *Suppose the continuity condition holds with unity continuity constant, i.e.,*

$$a(u,v) \leq \|u\|_U \|v\|_V .$$

Then there holds $M = \gamma = 1$ if either of the following conditions holds

i) For each $u \in U \setminus \{0\}$, there exists $v_u \in V \setminus \{0\}$ such that

$$a(u,v_u) = \|u\|_U \|v_u\|_V .$$

ii) For each $v \in V \setminus \{0\}$, there exists $u_v \in U \setminus \{0\}$ such that

$$a(u_v,v) = \|u_v\|_U \|v\|_V .$$

Proof. A proof can be found in [5]. ∎

REMARK 2.4. *In general, the continuity and the inf–sup conditions are not related to each other, and it is typically more difficult to establish the later. However, Theorem 2.3 shows that if the continuity constant is unity and the equality is attainable, then the continuity condition actually implies the inf–sup condition and the inf–sup constant is unity as well. To the end of the paper, we call the norms in U and V spaces* optimal norms *if both continuity and inf–sup constants are unity in these norms. Moreover, we also call the pair u and v_u (and hence for u_v and v) as the optimal trial and test functions, respectively. Here, optimality is in the sense of Corollary 2.2.*

We are now in position to construct the approximation subspaces U_h and V_h such that the discrete continuity and inf–sup constants are unity.

THEOREM 2.5. *Define the map $T : U \ni u \mapsto Tu \in V'$ as $\langle Tu, v\rangle_{V'\times V} = a(u,v)$. Denote v_{Tu} as the Riesz representation of Tu in V. Suppose $a(\cdot,\cdot)$ is continuous with unity constant and assumption i) of Theorem 2.3 holds. Take $U_h \subset U$ and define*

$$V_h = span\{v_{Tu_h} \in V : u_h \in U_h\} .$$

Then, the following hold,

(i) $M_h = \gamma_h = 1$.

(ii) Let $U_h = span\{\varphi_i\}_{i=1}$, where $\varphi_i \in U, i = 1,\ldots,n$. Then $\{v_{T\varphi_i}\}_{i=1}^n$ is a basis of V_h.

(iii) The discrete problem (2.4) is well-posed.

Proof. A proof can be found in [5]. ∎

At this point, it is important to point out that while the finite dimensional trial space $U_h = span\{\varphi_i\}_{i=1}^n$ is designed for good approximability, the test basis $V_h = span\{v_{T\varphi_i}\}_{i=1}^n$ is constructed for well-posedness of the discrete problem (2.4) via the Banach-Nečas-Babuška theorem (which is a re-statement of the closed range and the open mapping theorems [23]).

To the rest of the paper, we shall not distinguish v_u and v_{Tu} since we shall work exclusively with the Riesz representations.

3. A relation between DPG and DG for first order hyperbolic PDEs. The model problem in this section is the first order scalar linear hyperbolic equation of the form

$$\boldsymbol{\beta}\cdot\nabla u + \mu u = f, \quad \text{in } \Omega, \tag{3.1a}$$
$$u = g, \quad \text{on } \Gamma, \tag{3.1b}$$

where $\Gamma = \{\mathbf{x}\in\partial\Omega : \mathbf{n}(\mathbf{x})\cdot\boldsymbol{\beta} < 0\}$ is the inflow boundary; $\mathbf{n}(\mathbf{x})$ denotes the outward normal vector at $\mathbf{x}$ on the boundary $\partial\Omega$. Assume $\boldsymbol{\beta}\in\left[W^{1,\infty}(\Omega)\right]^d$ with $d\in\{1,2,3\}$ denoting the dimension of the problem, $\mu\in L^\infty(\Omega)$, $f\in L^2(\Omega)$, and $g\in L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(\Gamma)$ with

$$L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(\Gamma) = \left\{w : \|w\|^2_{L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(\Gamma)} = \int_\Gamma |\boldsymbol{\beta}\cdot\mathbf{n}|\,|w|^2\,d\Gamma < \infty\right\},$$

and $u\in H^1_{\boldsymbol{\beta}}$ with

$$H^1_{\boldsymbol{\beta}}(\Omega) = \left\{u\in L^2(\Omega) : \boldsymbol{\beta}\cdot\nabla u\in L^2(\Omega)\right\}.$$

We partition the polygonal domain Ω into N^{el} non-overlapping elements $K_j, j = 1,\dots,N^{\text{el}}$ such that $\Omega_h = \cup_{j=1}^{N^{\text{el}}} K_j$ and $\overline{\Omega} = \overline{\Omega}_h$ and that the mesh is assumed to be affine. Here, h is defined as $h = \max_{j\in\{1,\dots,N^{\text{el}}\}} diam(K_j)$. In addition, we denote by $\mathcal{E}_h$ the mesh skeleton, with cardinal number N^{ed}, which consists of all unique faces in the mesh, each of which comes with a normal vector $\mathbf{n}_e$. In this paper, the term "faces" is used to indicate boundary points of 1D elements, edges of 2D elements, or faces of 3D elements. Finally, we require $\boldsymbol{\beta}\cdot\mathbf{n}_e\in L^\infty(e)$ for $e = 1,\dots,N^{\text{ed}}$. Multiplying (3.1a) by a test function v, integrating by parts, and introducing the single-valued flux $q\in L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(\mathcal{E}_h)$ at the element interfaces, we have,

$$\sum_{j=1}^{N^{\text{el}}}\int_{K_j}\left[-u\nabla\cdot(\boldsymbol{\beta}v) + \mu uv\right]d\mathbf{x} + \int_{\partial K_j}\mathbf{1}_{\partial K_j\setminus\Gamma}\boldsymbol{\beta}\cdot\mathbf{n}qv\,ds = \sum_{j=1}^{N^{\text{el}}}\int_{K_j} fv\,d\mathbf{x} - \int_{\partial K_j\cap\Gamma}\boldsymbol{\beta}\cdot\mathbf{n}gv\,ds, \tag{3.2}$$

with $\mathbf{1}_{\partial K_j\setminus\Gamma}$ denoting the indicator function (also known as the characteristic function) of $\partial K_j\setminus\Gamma$. Clearly, for elements with characteristic faces, i.e., $\boldsymbol{\beta}\cdot\mathbf{n} = 0$ on ∂K_j, the boundary integrals corresponding to these faces simply drop out and q is allowed

to be undefined on these boundaries. Next, integrating by parts one more time gives

$$\sum_{j=1}^{N^{\text{el}}} \int_{K_j} (\boldsymbol{\beta}\cdot\nabla u + \mu u)\, v\, d\mathbf{x} + \int_{\partial K_j} \boldsymbol{\beta}\cdot\mathbf{n}\left(\mathbf{1}_{\partial K_j\setminus\Gamma} q - u\right) v\, ds = \sum_{j=1}^{N^{\text{el}}} \int_{K_j} f v\, d\mathbf{x} - \int_{\partial K_j\cap\Gamma} \boldsymbol{\beta}\cdot\mathbf{n} g v\, ds. \quad (3.3)$$

If we choose $v|_{K_j} \in L^2(K_j)$, then the trace $v|_{\partial K_j}$ is not defined. Therefore, we introduce a new hybrid variable r that lives in the space $\Pi_{j=1}^{N^{\text{el}}} L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(\partial K_j)$. Unlike q, which is single-valued on a face of the skeleton, r is allowed to have double values depending on the side of that face. With the introduction of r, (3.3) can be rewritten as

$$a(\mathbf{u},\mathbf{v}) = \sum_{j=1}^{N^{\text{el}}} \int_{K_j} (\boldsymbol{\beta}\cdot\nabla u + \mu u)\, v\, d\mathbf{x} + \int_{\partial K_j} \boldsymbol{\beta}\cdot\mathbf{n}\left(\mathbf{1}_{\partial K_j\setminus\Gamma} q - u\right) r\, ds = \ell(\mathbf{v}) = \sum_{j=1}^{N^{\text{el}}} \int_{K_j} f v\, d\mathbf{x} - \int_{\partial K_j\cap\Gamma} \boldsymbol{\beta}\cdot\mathbf{n} g r\, ds, \quad (3.4)$$

As shown in [5], applying the DPG framework in Section 2 to the weak form (3.4), obtained by integrating by parts twice, recovers an existing *hp* least-squares DG (LSDG) method presented in [21]. Thus, the DPG framework can be considered as a different, but constructive, methodology to derive *hp* LSDG methods. In particular, starting from the requirement of having unity continuity and inf–sup constants, and choosing the weak formulation (3.4) to apply our theory developed in Section 2, we constructively (and accidentally) derive a LSDG method from a well-posed infinite dimensional setting. The distinct feature of our method is that the flux is introduced as a new unknown, and then found by fulfilling stability. It should be pointed out that our theory is general in the following sense. If it is applied to different weak formulations, one will constructively obtain different numerical methods, again, with trivial well-posedness for the finite dimensional approximation problem, as we now show.

The starting point is the weak formulation (3.2) obtained by integrating (3.1) by parts once. This formulation can be written in a more compact form as

$$\sum_{j=1}^{N^{\text{el}}} \int_{K_j} u\left[-\nabla\cdot(\boldsymbol{\beta} v) + \mu v\right] d\mathbf{x} + \frac{1}{2}\int_{\partial K_j} q\boldsymbol{\beta}\cdot[\![v]\!]\, ds = \sum_{j=1}^{N^{\text{el}}} \int_{K_j} f v\, d\mathbf{x} - \int_{\partial K_j\cap\Gamma} \boldsymbol{\beta}\cdot\mathbf{n} g v\, ds, \quad (3.5)$$

with $[\![v]\!] = v^-\mathbf{n}^- + v^+\mathbf{n}^+$. In this section, we conventionally define $v^+\mathbf{n}^+ = v^-\mathbf{n}^-$ on the outflow, and on the inflow $v^+\mathbf{n}^+ = -v^-\mathbf{n}^-$. We allow the arbitrariness in assigning "-" and "+" quantities on a common face of two adjacent elements K_i and

K_j. We define the trial and test spaces as

$$U = \left\{ \mathbf{u} : \mathbf{u}|_{K_j} \in L^2(K_j) \times L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(\partial K_j), j = 1, \ldots, N^{\mathrm{el}} \right\} = L^2(\Omega_h) \times L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(\mathcal{E}_h),$$
$$V = \left\{ v : v|_{K_j} \in H^1_{\boldsymbol{\beta}}(K_j), j = 1, \ldots, N^{\mathrm{el}} \right\} = H^1_{\boldsymbol{\beta}}(\Omega_h).$$

Guided by Theorem 2.3, we are looking for natural norms in spaces U and V such that the continuity constant is unity and the equality in the continuity condition is achievable. An approach to obtain this goal is to apply the Cauchy–Schwarz inequality, i.e.,

$$\begin{aligned} a(\mathbf{u}, v) &\le \sum_{j=1}^{N^{\mathrm{el}}} \|-\nabla\cdot(\boldsymbol{\beta} v) + \mu v\|_{L^2(K_j)} \|u\|_{L^2(K_j)} \\ &\quad + \frac{1}{2} \|q\|_{L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(\partial K_j)} \|[\![v]\!]\|_{L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(\partial K_j)} \\ &\le \underbrace{\left\{ \sum_{j=1}^{N^{\mathrm{el}}} \|-\nabla\cdot(\boldsymbol{\beta} v) + \mu v\|^2_{L^2(K_j)} + \left\| \frac{1}{\sqrt{2}} [\![v]\!] \right\|^2_{L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(\partial K_j)} \right\}^{\frac{1}{2}}}_{\|\mathbf{v}\|_V} \\ &\quad \times \underbrace{\left\{ \sum_{j=1}^{N^{\mathrm{el}}} \|u\|^2_{L^2(K_j)} + \left\| \frac{1}{\sqrt{2}} q \right\|^2_{L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(\partial K_j)} \right\}^{\frac{1}{2}}}_{\|\mathbf{u}\|_U}. \end{aligned} \tag{3.6}$$

Equalities in (3.6) are achievable if we use the Riesz representations, i.e., given $\mathbf{u} = (u, q) \in U$,

$$u = -\nabla\cdot(\boldsymbol{\beta} v_{\mathbf{u}}) + \mu v_{\mathbf{u}}, \quad \text{in } K_j, \tag{3.7a}$$
$$q = \operatorname{sgn}(\boldsymbol{\beta}\cdot\mathbf{n})\, \mathbf{n}\cdot[\![v_{\mathbf{u}}]\!], \quad \text{on } \partial K_j. \tag{3.7b}$$

Once we know that the equality is obtainable under condition (3.7), the consistency and well-posedness of (3.5) with respect to the norms defined in (3.6) are readily available by Theorem 2.5 [5] or by a general DPG theory [7].

We next use (3.7) to find optimal pairs of trial and (corresponding) test basis functions. For basis functions of the form $\boldsymbol{\phi} = (0, \phi) \in U$, where ϕ is a function in $L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(\mathcal{E}_h)$, the corresponding basis functions in V are given by, for $j = 1, \ldots, N^{\mathrm{el}}$,

$$-\nabla\cdot(\boldsymbol{\beta} v_\phi) + \mu v_\phi = 0, \quad \text{in } K_j, \tag{3.8a}$$
$$\mathbf{n}\cdot[\![v_\phi]\!] = \operatorname{sgn}(\boldsymbol{\beta}\cdot\mathbf{n})\, \phi, \quad \text{on } \partial K_j. \tag{3.8b}$$

Similarly, for basis functions of the form $\boldsymbol{\varphi} = (\varphi, 0) \in U$, where $\varphi \in L^2(\Omega_h)$, the corresponding basis functions in V are given by, for $j = 1, \ldots, N^{\mathrm{el}}$,

$$-\nabla\cdot(\boldsymbol{\beta} v_\varphi) + \mu v_\varphi = \varphi, \quad \text{in } K_j, \tag{3.9a}$$
$$\mathbf{n}\cdot[\![v_\varphi]\!] = 0, \quad \text{on } \partial K_j. \tag{3.9b}$$

Once the test functions are found using (3.8) or (3.9), we can substitute them into (3.5) to establish equations to solve for the unknowns $\mathbf{u} = (u, q)$. Let us proceed with

the generic test basis in (3.8) first. If ϕ is different from zero on $e \in \mathcal{E}_h$ and zero elsewhere on the skeleton, then testing (3.5) with $\mathbf{v} = (0, v_\phi)$ yields,

$$\int_e |\boldsymbol{\beta}\cdot\mathbf{n}|\ q\ \phi\ ds = \int_{\Omega_h} f\ v_\phi\ d\mathbf{x} - \int_\Gamma \boldsymbol{\beta}\cdot\mathbf{n}\ g\ v_\phi\ ds. \tag{3.10}$$

As can be seen in (3.10), for each $\phi \in L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(e), e \in \mathcal{E}_h$, $q \in L^2_{\boldsymbol{\beta}\cdot\mathbf{n}}(e)$ can be locally solved independently of u.

Now if $\varphi|_{Kj}$ is a nonzero function in K_j but zero elsewhere, then testing (3.5) with $\mathbf{v} = (0, v_\varphi)$ gives,

$$\int_{K_j} u\varphi\ d\mathbf{x} = \int_{\Omega_h} f\ v_\varphi\ d\mathbf{x} - \int_\Gamma \boldsymbol{\beta}\cdot\mathbf{n}\ g\ v_\varphi\ ds, \tag{3.11}$$

which shows that the unknown u can also be computed locally element-by-element and independently of q. The detailed discussion on this DPG method can be found in [5].

As shown in the existing literature, the DPG methodology can provide different DPG methods depending the norm in the test space (and hence the corresponding norm in the trial space) [5, 6, 13–18, 27]. Typically, the chosen norms are equivalent with small (e.g., order one) constants. Most of DPG methods therefore share several attractive features including stability and accuracy due to the (quasi-)minimization of the approximation error.

In this paper, we choose to work with the DPG framework with optimal norms of the type defined in (3.6) to demonstrate the DPG stability and accuracy. More importantly, it allows us to recover several existing DG methods as its least accurate versions. To begin, we first need to address the question on how to solve for the optimal test functions in (3.8) and (3.9). Since we are only interested in the original variable u, we simply ignore (3.8), and hence the computation of the hybrid variable q. Denote $\mathcal{P}^p$ as the space of polynomials of order at most p, we seek $u_h|_{K_j}$ in the following finite dimensional piecewise polynomial subspace

$$U_h(K_j) = \left\{\varphi \in L^2(K_j) : \varphi \in \mathcal{P}^{p_j}\right\},$$

such that

$$\int_{K_j} u_h\varphi\ d\mathbf{x} = \int_{\Omega_h} f\ v_\varphi\ d\mathbf{x} - \int_\Gamma \boldsymbol{\beta}\cdot\mathbf{n}\ g\ v_\varphi\ ds, \quad \forall\varphi \in U_h(K_j).$$

Clearly, given $\varphi \in U_h(K_j)$, it is not possible to solve for the optimal test function in (3.9) exactly, and one has to resort to numerical approximations. This in fact provides the flexibility in choosing the accuracy of the computed test functions, and hence the accuracy of the approximation solution u_h, as we shall show. In particular, we choose to approximate the optimal test functions v_φ, the solution of (3.8), by v^h_φ in the following finite dimensional test subspace

$$V_h^{\Delta p} = \left\{v \in V : v|_{K_j} \in \mathcal{P}^{p_j+\Delta p_j}, j = 1, \ldots, N^{\text{el}}\right\} \subset V,$$

with Δp_j, $j = 1, \ldots, N^{\text{el}}$, as the enriched orders. As can be observed, $V_h^{\Delta p}$ asymptotically approaches V as $\Delta p_j \to \infty$, $j = 1, \ldots, N^{\text{el}}$, owing to the density of the space

of polynomials. That is, $v_\varphi^h \in V_h^{\Delta p}$ is increasingly accurate as Δp_j, $j = 1, \ldots, N^{\mathrm{el}}$, increase. The discrete equation for u_h reads

$$\int_{K_j} u_h \varphi \, d\mathbf{x} = \int_{\Omega_h} f \; v_\varphi^h \, d\mathbf{x} - \int_\Gamma \boldsymbol{\beta} \cdot \mathbf{n} \; g \; v_\varphi^h \, ds, \quad \forall \varphi \in \mathcal{P}^{p_j}. \tag{3.12}$$

Next, any reasonable numerical method for hyperbolic equations can be used to solve (3.9) for v_φ^h. Here, we choose a class of DG methods with the following numerical flux

$$\mathbf{n} \cdot \left(\boldsymbol{\beta} v_\varphi^h\right)^* = \boldsymbol{\beta} \cdot \mathbf{n} \{\!\{ v_\varphi^h \}\!\} - \frac{(1-\alpha)}{2} \mathrm{sgn}\left(\boldsymbol{\beta} \cdot \mathbf{n}\right) \boldsymbol{\beta} \cdot [\![v_\varphi^h]\!], \tag{3.13}$$

with $\{\!\{ v \}\!\}$ denoting the average $\frac{1}{2}\left(v^- + v^+\right)$, and $\alpha \in [0, 1]$. Note that $\alpha = 0$ corresponds to the downwind flux and $\alpha = 1$ the central flux. The DG scheme to approximate (3.9) reads, for each $j \in \left\{1, \ldots, N^{\mathrm{el}}\right\}$ and for each $\varphi \in \mathcal{P}^{p_j}$,

$$\sum_{K_i \in \Omega_j} \int_{K_i} v_\varphi^h \boldsymbol{\beta} \cdot \nabla \phi \, d\mathbf{x} - \int_{\partial K_i \setminus \partial \Omega_j^+} \mathbf{n} \cdot \left(\boldsymbol{\beta} v_\varphi^h\right)^* \phi \, ds = \int_{K_j} \varphi \phi \, d\mathbf{x}, \quad \forall \phi \in \mathcal{P}^{p_j + \Delta p_j}, \tag{3.14}$$

where Ω_j involves only elements downstream of the adjoint flow starting from element K_j, and $\partial \Omega_j^+ = \{\mathbf{x} \in \partial \Omega_j : -\boldsymbol{\beta} \cdot \mathbf{x} < 0\}$ is the adjoint inflow of the boundary of Ω_j. Moreover, the approximate solution on K_j satisfies the discrete equation (3.12). Since $\mathcal{P}^{p_j} \subseteq \mathcal{P}^{p_j + \Delta p_j}$, we can take $\phi = u_h$ so that (3.14) becomes

$$\sum_{K_i \in \Omega_j} \int_{K_i} v_\varphi^h \boldsymbol{\beta} \cdot \nabla u_h \, d\mathbf{x} - \int_{\partial K_i \setminus \partial \Omega_j^+} \mathbf{n} \cdot \left(\boldsymbol{\beta} v_\varphi^h\right)^* u_h \, ds = \int_{K_j} \varphi u_h \, d\mathbf{x}. \tag{3.15}$$

The following useful identity is easy to inspect for any vector $\boldsymbol{\tau}$ and scalar w

$$\sum_{K_i \in \Omega_j} \int_{\partial K_i \setminus \partial \Omega_j} \mathbf{n} \cdot \boldsymbol{\tau} w \, ds = \sum_{e \in \mathcal{E}_h^j} \int_{e \setminus \partial \Omega_j} \{\!\{ \boldsymbol{\tau} \}\!\} \cdot [\![w]\!] \, ds + \sum_{e \in \mathcal{E}_h^j} \int_{e \setminus \partial \Omega_j} [\![\boldsymbol{\tau}]\!] \{\!\{ w \}\!\} \, ds, \tag{3.16}$$

where $\mathcal{E}_h^j$ is the skeleton of Ω_j, and the jump for vector-valued quantity is defined as $[\![\boldsymbol{\tau}]\!] = \boldsymbol{\tau}^- \cdot \mathbf{n}^- + \boldsymbol{\tau}^+ \cdot \mathbf{n}^+$.

Now integrating the first term on the left side of (3.15) by parts, applying identity (3.16) to the second term on the left side of (3.15), substituting the adjoint numerical flux (3.13) into the left side of (3.15), and substituting (3.12) into the right side of (3.15) we can rewrite (3.15) as

$$\begin{aligned} -\sum_{K_i \in \Omega_j} \int_{K_i} u_h \nabla \cdot \left(\boldsymbol{\beta} v_\varphi^h\right) \, d\mathbf{x} + \int_{\partial K_i \setminus \Gamma} \mathbf{n} \cdot \left(\boldsymbol{\beta} u_h\right)^* v_\varphi^h \, ds = \\ \sum_{K_i \in \Omega_j} \int_{K_i} f \; v_\varphi^h \, d\mathbf{x} - \int_\Gamma \boldsymbol{\beta} \cdot \mathbf{n} \; g \; v_\varphi^h \, ds, \quad j = 1, \ldots, N^{\mathrm{el}}, \end{aligned} \tag{3.17}$$

where

$$\mathbf{n} \cdot \left(\boldsymbol{\beta} u_h\right)^* = \boldsymbol{\beta} \cdot \mathbf{n} \{\!\{ u_h \}\!\} + \frac{(1-\alpha)}{2} \mathrm{sgn}\left(\boldsymbol{\beta} \cdot \mathbf{n}\right) \boldsymbol{\beta} \cdot [\![u_h]\!]. \tag{3.18}$$

Here comes the important point. If $\Delta p_j = 0$, $j = 1, \ldots, N^{\text{el}}$, then (3.17) is equivalent to the following single variational equation

$$-\sum_{K_j \in \Omega} \int_{K_i} u_h \nabla \cdot (\boldsymbol{\beta} v_h)\, d\mathbf{x} + \int_{\partial K_i \setminus \Gamma} \mathbf{n} \cdot (\boldsymbol{\beta} u_h)^* v_h\, ds = \int_{\Omega_h} f\, v_h\, d\mathbf{x} - \int_{\Gamma} \boldsymbol{\beta} \cdot \mathbf{n}\, g\, v_h\, ds, \quad \forall v_h \in V_h^{\Delta p=0} = U_h(\Omega_h). \tag{3.19}$$

At this point, a closer look tells us that (3.19) and (3.18) embrace a class of DG discretizations for the original equation (3.1). In particular, $\alpha = 0$ corresponds to the DG method with upwind numerical flux [22, 25] while $\alpha = 1$ corresponds to the DG method with central numerical flux. On the other hand, as shown in [5], $\Delta p_j = 0$, $j = 1, \ldots, N^{\text{el}}$ correspond to the least accurate DPG method. In practice, one typically takes $\Delta p_j \geq 1$. The above analysis is therefore the first attempt in explaining why the DPG method can be more accurate than the DG method using the same solution order p_j, $j = 1, \ldots, N^{\text{el}}$, though this has already been observed numerically in [5, 13, 15]. Indeed, these papers show that the DPG is more accurate, more stable, and does not seems to have any noticeable artificial dispersion compared to the DG for several one- and two-dimensional examples. Moreover, the DG method has sub-optimal h convergence rate on Peterson's meshes [24] for $p < 3$, while the DPG h convergence rate is uniformly optimal for all p.

Let us now provide more numerical results to support the above analysis. For the first example, we take $\Omega = (0, 1)$, $\beta = 1, \mu = 0$, $g = 0.5 \exp\left(-0.8^2/\sigma^2\right)$, and the forcing function f is chosen such that the exact solution is given by

$$u(x) = \frac{x}{\left(1 + \sqrt{1/\exp\left(1/(\lambda a)\right)}\right) \exp\left(x^2/(4\lambda)\right)} + 0.5 \exp\left(-\frac{(x-0.8)^2}{\lambda a}\right),$$

where $a = 6, \lambda = 0.00045$, and $\sigma = 0.05$. The optimal test functions can be computed exactly as in [5]. The solution order is chosen to be 4, and the mesh is uniform with 50 elements, i.e., $h = 1/50$. Figure 3.1 shows that both upwind DG and DPG solutions are comparable in the linear part of the solution while the DG solution has more over-shootings in the sharp gradient region. In the smooth region with $x \geq 0.6$, the DPG solution is on top of the exact one whereas the DG solution is inaccurate. The better accuracy is due to the fact that the DPG solutions minimize the error in the energy norm (3.6), a component of which is the error in the L^2 norm. Therefore, oscillations with big amplitude that have significant L^2 norm are not allowed in the DPG solutions. The DG method, however, does not have such a property. The better stability reflects the fact that while most of numerical methods for hyperbolic equations introduce numerical dissipation either explicitly or implicitly (e.g., through the numerical fluxes as in many DG methods) to gain stability, and hence may not be enough (e.g., large over-shootings), the DPG stability comes directly from the functional setting through the Banach-Nečas-Babuška theorem on the infinite dimensional level.

In the second example, we study the dispersion error. To this end, we take $\Omega = (0, 1)$, $\beta = 1, \mu = 0$, $g = 1$, and the forcing function f is chosen such that the exact solution is given by

$$u(x) = \cos(4\pi x).$$

For the numerical solution, we use eight linear elements. Figure 3.2 plots the upwind DG, the DPG, and the exact solutions. Compared to the DPG, the DG is less

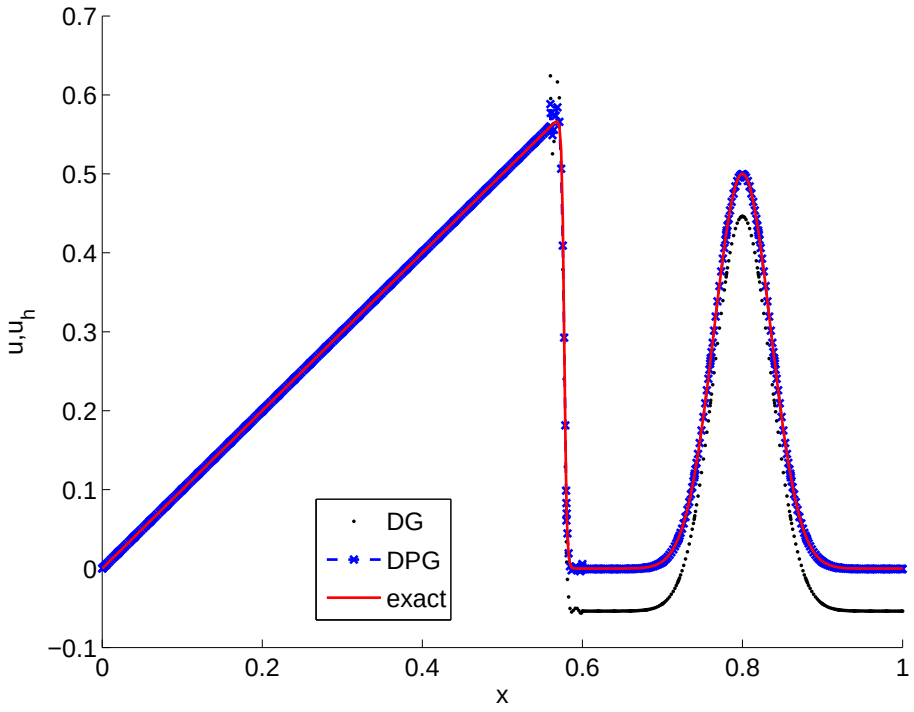

FIG. 3.1. *DPG and upwind DG solutions for the first example.*

accurate not only in magnitude but also in phase. On the other hand, the DPG solution does not seems to have any noticeable dispersion error. More importantly, the DPG approximability is uniform across the whole domain while the DG solution undershoots and overshoots at the negative and the positive peaks, respectively. This indicates that the upwind numerical flux introduces not only non-uniform dissipation but also unphysical dispersion.

For the first two examples, we have computed the optimal test functions exactly. A more general and practical approach is to solve for them approximately, as described above. Several two-dimensional examples using the upwind DG method to compute the optimal test functions with enriched exponent $\Delta p_j = 1$ can be found in [5]. The observations are the same as in the case with exact test functions including better accuracy, better stability, no noticeable dispersion, and uniform optimal convergence rates. The reader is referred to [5] for several two dimensional examples and the corresponding detailed comparison between DPG and DG methods.

We have shown that the DPG methodology provides flexibility in approximating the optimal test functions. The more accurately the test functions are computed, the more accurate the DPG approximation is, at least asymptotically. It is interesting to see that the least accurate DPG methods, i.e., with $\Delta p_j = 0$, $j = 1, \ldots, N^{\mathrm{el}}$, recover a class of DG discretizations. Note that negative Δp_j is not allowable, since otherwise the dimension of the test space is less than that of the trial space, and hence the approximate solution would not be unique.

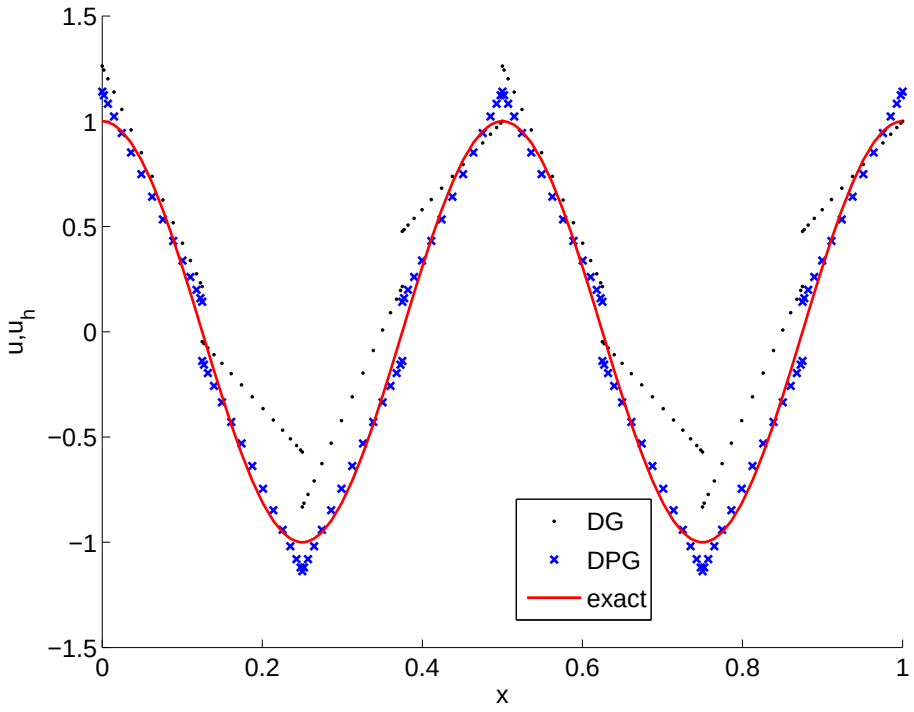

FIG. 3.2. *DPG and DG solutions for the second example.*

4. A relation between DPG and DG for elliptic PDEs. In this section, the model PDE we consider is the Poisson equation

$$-\Delta u = f, \quad \text{in } \Omega, \tag{4.1a}$$
$$u = 0, \quad \text{on } \Gamma = \partial\Omega, \tag{4.1b}$$

where $f \in L^2(\Omega)$. We first rewrite the equation in the first order form as

$$\boldsymbol{\sigma} + \nabla u = 0 \quad \text{in } \Omega, \tag{4.2a}$$
$$\nabla \cdot \boldsymbol{\sigma} = f \quad \text{in } \Omega, \tag{4.2b}$$
$$u = 0 \quad \text{on } \Gamma = \partial\Omega. \tag{4.2c}$$

Multiply (4.2) by the pair of test function $(\boldsymbol{\tau}, v)$ and integrate by parts we obtain

$$(\boldsymbol{\sigma}, \boldsymbol{\tau})_{\Omega_h} - (u, \nabla \cdot \boldsymbol{\tau})_{\Omega_h} + \langle u, \boldsymbol{\tau} \cdot \mathbf{n} \rangle_{\partial\Omega_h} - (\boldsymbol{\sigma}, \nabla v)_{\Omega_h} + \langle v, \boldsymbol{\sigma} \cdot \mathbf{n} \rangle_{\partial\Omega_h} = (f, v)_{\Omega_h},$$

where $(\cdot,\cdot)_{\Omega_h}$ is the broken L^2 inner product on Ω_h, and $\langle \cdot, \cdot \rangle_{\partial\Omega_h}$ the broken duality pairing $H^{\frac{1}{2}}(\partial\Omega_h) \times H^{-\frac{1}{2}}(\partial\Omega_h)$ with $\partial\Omega_h = \Pi_{j=1}^{N^{\text{el}}} \partial K_j$. Next, introduce single-valued unknown trace $\hat{u}$ and flux $\hat{\sigma}$ on the skeleton, the weak formulation can be rewritten as

$$a\left((\boldsymbol{\sigma}, u, \hat{u}, \hat{\sigma}), (\boldsymbol{\tau}, v)\right) = (\boldsymbol{\sigma}, \boldsymbol{\tau})_{\Omega_h} - (u, \nabla \cdot \boldsymbol{\tau})_{\Omega_h} + \langle \hat{u}, [\![\boldsymbol{\tau}]\!] \rangle_{\mathcal{E}_h} - (\boldsymbol{\sigma}, \nabla v)_{\Omega_h} + \langle [\![v]\!], \hat{\sigma} \rangle_{\mathcal{E}_h}$$
$$= \ell(\boldsymbol{\tau}, v) = (f, v)_{\Omega_h}, \tag{4.3}$$

where

$$[\![v]\!] = v^- \operatorname{sgn}\left(\mathbf{n}^-\right) + v^+ \operatorname{sgn}\left(\mathbf{n}^+\right),$$

with

$$\operatorname{sgn}\left(\mathbf{n}^\pm\right) = \begin{cases} 1 & \text{if } \mathbf{n}^\pm = \mathbf{n}_e \\ -1 & \text{if } \mathbf{n}^\pm = -\mathbf{n}_e \end{cases}.$$

Here, we conventionally define $[\![\boldsymbol{\tau}]\!] = \boldsymbol{\tau}\cdot\mathbf{n}$ and $[\![v]\!] = v\operatorname{sgn}(\mathbf{n})$ on the domain boundary. Following [7,14], we choose the trial space as

$$U = \left[L^2(\Omega_h)\right]^d \times L^2(\Omega_h) \times H_0^{\frac{1}{2}}(\mathcal{E}_h) \times H^{-\frac{1}{2}}(\mathcal{E}_h),$$

and the test space as

$$V = H(\operatorname{div}, \Omega_h) \times H^1(\Omega_h).$$

For each face $e \in \mathcal{E}_h$, we use the Riesz representation theorem to express a duality pairing as an equivalent inner product, i.e.,

$$\langle \hat{u}, [\![\boldsymbol{\tau}]\!] \rangle_e = \left(\hat{u}, \mathcal{R}[\![\boldsymbol{\tau}]\!]\right)_{H^{\frac{1}{2}}(e)}, \quad \langle [\![v]\!], \hat{\sigma} \rangle_e = \left([\![v]\!], \mathcal{R}\hat{\sigma}\right)_{H^{\frac{1}{2}}(e)},$$

where $\mathcal{R} : H^{-\frac{1}{2}}(e) \to H^{\frac{1}{2}}(e)$ is the Riesz map, and $(\cdot,\cdot)_{H^{\frac{1}{2}}(e)}$ denotes the inner product in $H^{\frac{1}{2}}(e)$. The variational problem (4.3) can be equivalently rewritten as

$$\begin{aligned} a\left((\boldsymbol{\sigma}, u, \hat{u}, \hat{\sigma}), (\boldsymbol{\tau}, v)\right) = (\boldsymbol{\sigma}, \boldsymbol{\tau} - \nabla v)_{\Omega_h} - (u, \nabla\cdot\boldsymbol{\tau})_{\Omega_h} + \left(\hat{u}, \mathcal{R}[\![\boldsymbol{\tau}]\!]\right)_{H^{\frac{1}{2}}(\mathcal{E}_h)} \\ + \left([\![v]\!], \mathcal{R}\hat{\sigma}\right)_{H^{\frac{1}{2}}(\mathcal{E}_h)} = \ell(\boldsymbol{\tau}, v) = (f, v)_{\Omega_h}. \end{aligned} \tag{4.4}$$

Similar to Section 3 we apply the Cauchy–Schwarz inequality to obtain

$$a\left((\boldsymbol{\sigma}, u, \hat{u}, \hat{\sigma}), (\boldsymbol{\tau}, v)\right) \le \|\mathbf{u}\|_U \|\mathbf{v}\|_V,$$

with the optimal norms given by

$$\|\mathbf{u}\|_U^2 = \sum_{j=1}^{N^{\mathrm{el}}} \|\boldsymbol{\sigma}\|_{L^2(K_j)}^2 + \|u\|_{L^2(K_j)}^2 + \frac{1}{2}\|\hat{u}\|_{H^{\frac{1}{2}}(\partial K_j)}^2 + \frac{1}{2}\|\mathcal{R}\hat{\sigma}\|_{H^{\frac{1}{2}}(\partial K_j)}^2,$$

$$\|\mathbf{v}\|_V^2 = \sum_{j=1}^{N^{\mathrm{el}}} \|\boldsymbol{\tau} - \nabla v\|_{L^2(K_j)}^2 + \|\nabla\cdot\boldsymbol{\tau}\|_{L^2(K_j)}^2 + \frac{1}{2}\|\mathcal{R}[\![\boldsymbol{\tau}]\!]\|_{H^{\frac{1}{2}}(\partial K_j)}^2 + \frac{1}{2}\|[\![v]\!]\|_{H^{\frac{1}{2}}(\partial K_j)}^2,$$

where we have defined $\mathbf{u} = (\boldsymbol{\sigma}, u, \hat{u}, \hat{\sigma})$ and $\mathbf{v} = (\boldsymbol{\tau}, v)$. Given $\mathbf{u} \in U$, by virtue of Theorem 2.5 the optimal test functions can be found by solving the following systems of PDEs

$$\boldsymbol{\tau} - \nabla v = \boldsymbol{\sigma}, \qquad \text{in } K_j, \tag{4.5a}$$

$$\nabla\cdot\boldsymbol{\tau} = -u, \qquad \text{in } K_j, \tag{4.5b}$$

$$\mathcal{R}[\![\boldsymbol{\tau}]\!] = \hat{u}, \qquad \text{on } \partial K_j, \tag{4.5c}$$

$$[\![v]\!] = \mathcal{R}\hat{\sigma}, \qquad \text{on } \partial K_j, \tag{4.5d}$$

for $j = 1, \ldots, N^{\text{el}}$. Since we are interested in the original unknown u, let us focus on trial basis functions of the form $\boldsymbol{\varphi} = (0, \varphi, 0, 0)$, where $\varphi \in L^2(\Omega_h)$. The corresponding optimal basis functions in V are given by, for $j = 1, \ldots, N^{\text{el}}$,

$$\begin{aligned} \boldsymbol{\tau}_\varphi - \nabla v_\varphi &= 0, && \text{in } K_j, && (4.6a)\\ \nabla \cdot \boldsymbol{\tau}_\varphi &= -\varphi, && \text{in } K_j, && (4.6b)\\ \mathcal{R}[\![\boldsymbol{\tau}_\varphi]\!] &= 0, && \text{on } \partial K_j, && (4.6c)\\ [\![v_\varphi]\!] &= 0, && \text{on } \partial K_j. && (4.6d) \end{aligned}$$

Once the test functions are found using (4.6), we can substitute them into (4.4) to establish equations to solve for the unknowns u. If φ is a nonzero function in K_j but zero elsewhere, then testing (4.4) with $\mathbf{v} = \mathbf{v}_\varphi = (\boldsymbol{\tau}_\varphi, v_\varphi)$ gives

$$\int_{K_j} u\varphi \, d\mathbf{x} = \int_{\Omega_h} f \, v_\varphi \, d\mathbf{x}, \tag{4.7}$$

which shows that the unknown u can be computed locally element-by-element independently by simply inverting the local mass matrix.

We seek $u_h|_{K_j}$ in the following finite dimensional piecewise polynomial subspace

$$U_h(K_j) = \left\{ u_h \in L^2(K_j) : u_h \in \mathcal{P}^{p_j} \right\},$$

such that

$$\int_{K_j} u_h \varphi \, d\mathbf{x} = \int_{\Omega_h} f \, v_\varphi \, d\mathbf{x}, \quad \forall \varphi \in U_h(K_j), j = 1, \ldots, N^{\text{el}}.$$

Given $\varphi \in U_h(K_j)$, we choose to approximate the corresponding optimal test function in (4.6) using the following finite dimensional test subspace

$$V_h^{\Delta p} = \left\{ (\boldsymbol{\tau}, v) \in V : (\boldsymbol{\tau}, v)|_{K_j} \in \left[\mathcal{P}^{p_j + \Delta p_j}\right]^d \times \mathcal{P}^{p_j + \Delta p_j}, j = 1, \ldots, N^{\text{el}} \right\} \subset V.$$

At this point, one can choose a method of interest to solve the adjoint problem (4.6) for an approximation of the optimal test functions $(\boldsymbol{\tau}_\varphi, v_\varphi)$ in $V_h^{\Delta p}$. Here we choose the local discontinuous Galerkin method (LDG) [11] to look for a finite dimensional approximation $\left(\boldsymbol{\tau}_\varphi^h, v_\varphi^h\right) \in V_h^{\Delta p}$ of $(\boldsymbol{\tau}_\varphi, v_\varphi)$. Testing (4.6) with $(\boldsymbol{\lambda}, \phi) \in V_h^{\Delta p}$, integrating by parts, and introducing the following adjoint LDG numerical flux

$$\begin{aligned} \mathbf{n} \cdot \left(\boldsymbol{\tau}_\varphi^h\right)^* &= \mathbf{n} \cdot \{\!\{\boldsymbol{\tau}_\varphi^h\}\!\} + \mathbf{n} \cdot \boldsymbol{\xi} [\![\boldsymbol{\tau}_\varphi^h]\!] + \alpha \mathbf{n} \cdot [\![v_\varphi^h]\!],\\ \left(v_\varphi^h\right)^* &= \{\!\{v_\varphi^h\}\!\} - \boldsymbol{\xi} \cdot [\![v_\varphi^h]\!], \end{aligned}$$

with some constants $\boldsymbol{\xi}$ and α, we obtain

$$\begin{aligned} \left(\boldsymbol{\tau}_\varphi^h, \boldsymbol{\lambda}\right)_{\Omega_h} + \left(\boldsymbol{\tau}_\varphi^h, \nabla\phi\right)_{\Omega_h} + \left(v_\varphi^h, \nabla \cdot \boldsymbol{\lambda}\right)_{\Omega_h} - \sum_{i=1}^{N^{\text{el}}} \left(\mathbf{n} \cdot \left(\boldsymbol{\tau}_\varphi^h\right)^*, \phi\right)_{\partial K_i} - \sum_{i=1}^{N^{\text{el}}} \left(\mathbf{n} \cdot \boldsymbol{\lambda}, \left(v_\varphi^h\right)^*\right)_{\partial K_i} \\ = (\varphi, \phi)_{K_j}, \quad \forall (\boldsymbol{\lambda}, \phi) \in V_h^{\Delta p}. \end{aligned} \tag{4.8}$$

Since $\mathcal{P}^{p_j} \subseteq \mathcal{P}^{p_j+\Delta p_j}$, we can take $(\boldsymbol{\lambda}, \phi) = (\boldsymbol{\sigma}_h, u_h)$ in (4.8) to have

$$\left(\boldsymbol{\tau}_\varphi^h, \boldsymbol{\sigma}_h\right)_{\Omega_h} + \left(\boldsymbol{\tau}_\varphi^h, \nabla u_h\right)_{\Omega_h} + \left(v_\varphi^h, \nabla \cdot \boldsymbol{\sigma}_h\right)_{\Omega_h} - \sum_{i=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot \left(\boldsymbol{\tau}_\varphi^h\right)^*, u_h\right)_{\partial K_i} - \sum_{i=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot \boldsymbol{\sigma}_h, \left(v_\varphi^h\right)^*\right)_{\partial K_i} = \left(\varphi, u_h\right)_{K_j}, \quad (4.9)$$

Now integrating the second and third terms on the left side of (4.9) by parts, applying identity (3.16) to the fourth and fifth terms on the left side of (4.9), using the adjoint LDG numerical flux, and substituting (4.7) to the right side of (4.9) we can rewrite (4.9) as

$$\left(\boldsymbol{\tau}_\varphi^h, \boldsymbol{\sigma}_h\right)_{\Omega_h} - \left(\nabla \cdot \boldsymbol{\tau}_\varphi^h, u_h\right)_{\Omega_h} + \sum_{i=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot \boldsymbol{\tau}_\varphi^h, (u_h)^*\right)_{\partial K_i} - \left(\nabla v_\varphi^h, \boldsymbol{\sigma}_h\right)_{\Omega_h} + \sum_{i=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot (\boldsymbol{\sigma}_h)^*, v_\varphi^h\right)_{\partial K_i} = \left(\varphi, f\right)_{\Omega_h}, \quad (4.10)$$

where

$$\mathbf{n} \cdot (\boldsymbol{\sigma}_h)^* = \mathbf{n} \cdot \{\!\{\boldsymbol{\sigma}_h\}\!\} - \mathbf{n} \cdot \boldsymbol{\xi} [\![\boldsymbol{\sigma}_h]\!] - \alpha \mathbf{n} \cdot [\![u_h]\!],$$
$$(u_h)^* = \{\!\{u_h\}\!\} + \boldsymbol{\xi} \cdot [\![u_h]\!].$$

It should be pointed out that (4.10) holds for each $\varphi \in U_h(K_j)$, and each $j \in \{1, \ldots, N^{\mathrm{el}}\}$. Now if we take $\Delta p_j = 0$, $j = 1, \ldots, N^{\mathrm{el}}$, then (4.10) can be written as a single variational formulation,

$$\left(\boldsymbol{\tau}, \boldsymbol{\sigma}_h\right)_{\Omega_h} - \left(\nabla \cdot \boldsymbol{\tau}, u_h\right)_{\Omega_h} + \sum_{j=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot \boldsymbol{\tau}, (u_h)^*\right)_{\partial K_j}$$
$$-\left(\nabla v, \boldsymbol{\sigma}_h\right)_{\Omega_h} + \sum_{j=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot (\boldsymbol{\sigma}_h)^*, v\right)_{\partial K_j} = \left(\varphi, f\right)_{\Omega_h}, \quad \forall (\boldsymbol{\tau}, v) \in V_h^{\Delta p=0} = \left[U_h(\Omega_h)\right]^{d+1},$$

which is exactly the LDG discretization [11] of the original problem (4.2).

We have shown that the DPG method with zero enriched exponent, i.e., $\Delta p_j = 0$, $j = 1, \ldots, N^{\mathrm{el}}$ coincides with the LDG method for elliptic PDEs. Generally, following the same reasoning as above, one can show that if one of the DG discretizations for elliptic problems summarized in [1] is used to solve (4.6) for the optimal test functions, the DPG method with zero enriched exponent is exactly the corresponding DG method for the original problem (4.2). In practice, one chooses $\Delta p_j \geq 1$ [14, 20] to obtain more accurate optimal test functions $\left(\boldsymbol{\tau}_\varphi^h, v_\varphi^h\right)$. As a result, the approximate solution u_h from (4.7) is more accurate as Δp_j, $j = 1, \ldots, N^{\mathrm{el}}$, increase.

5. A relation between DPG and the hybridized DG for elliptic PDEs. In this section, we will derive a relation between the hybridized DG method [9] and our DPG method. For simplicity in the exposition, let us assume that the polynomial order, p_j, is the same on all elements, i.e., $p_j = p$, $j = 1, \ldots, N^{\mathrm{el}}$. Unlike Section 4, we

use the hybridized DG to solve for an approximation $\left(\boldsymbol{\tau}_{\varphi}^{h}, v_{\varphi}^{h}\right) \in V_h^{\Delta p}$ of the optimal test functions. Given $\varphi \in U_h\left(K_j\right)$, we choose to approximate the corresponding optimal test function in (4.6) using the following finite dimensional test subspace

$$V_h^{\Delta p} = \left\{(\boldsymbol{\tau}, v) \in V : (\boldsymbol{\tau}, v)|_{K_j} \in \left[\mathcal{P}^{p+\Delta p}\right]^d \times \mathcal{P}^{p+\Delta p}, j = 1, \ldots, N^{\mathrm{el}}\right\} \subset V.$$

We also need the following piecewise polynomial space on the skeleton

$$M_h^{\Delta p} = \left\{q \in L^2\left(\mathcal{E}_h\right) : q|_e \in \mathcal{P}^{p+\Delta p}, e \in \mathcal{E}_h\right\}.$$

Testing (4.6) with $(\boldsymbol{\lambda}, \phi) \in V_h^{\Delta p}$, integrating by parts, and introducing the unknown trace $\hat{v}_{\varphi}^{h} \in M_h^{\Delta p}$ and the unknown flux $\hat{\boldsymbol{\tau}}_{\varphi}^{h} = \boldsymbol{\tau}_{\varphi}^{h} - \alpha\left(v_{\varphi}^{h} - \hat{v}_{\varphi}^{h}\right)\mathbf{n}$, give the adjoint version of the hybridized DG discretization [9] of (4.6) as follows,

$$\begin{aligned}\left(\boldsymbol{\tau}_{\varphi}^{h}, \boldsymbol{\lambda}\right)_{\Omega_h} + \left(v_{\varphi}^{h}, \nabla \cdot \boldsymbol{\lambda}\right)_{\Omega_h} - \sum_{i=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot \boldsymbol{\lambda}, \hat{v}_{\varphi}^{h}\right)_{\partial K_i} + \left(\boldsymbol{\tau}_{\varphi}^{h}, \nabla \phi\right)_{\Omega_h} - \sum_{i=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot \hat{\boldsymbol{\tau}}_{\varphi}^{h}, \phi\right)_{\partial K_i} \\ + \sum_{i=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot \hat{\boldsymbol{\tau}}_{\varphi}^{h}, \hat{u}_h\right)_{\partial K_i} = (\varphi, \phi)_{K_j}, \quad \forall\, (\boldsymbol{\lambda}, \phi) \in V_h^{\Delta p}, \hat{u}_h \in M_h^{\Delta p}. \quad (5.1)\end{aligned}$$

Since $\mathcal{P}^p \subseteq \mathcal{P}^{p+\Delta p}$, we can take $(\boldsymbol{\lambda}, \phi) = (\boldsymbol{\sigma}_h, u_h)$ in (5.1) and then substitute (4.7) into the right side of (5.1) to obtain

$$\begin{aligned}&\left(\boldsymbol{\tau}_{\varphi}^{h}, \boldsymbol{\sigma}_h\right)_{\Omega_h} + \left(v_{\varphi}^{h}, \nabla \cdot \boldsymbol{\sigma}_h\right)_{\Omega_h} - \sum_{i=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot \boldsymbol{\sigma}_h, \hat{v}_{\varphi}^{h}\right)_{\partial K_i} + \left(\boldsymbol{\tau}_{\varphi}^{h}, \nabla u_h\right)_{\Omega_h} - \sum_{i=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot \hat{\boldsymbol{\tau}}_{\varphi}^{h}, u_h\right)_{\partial K_i} \\ &+ \sum_{i=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot \hat{\boldsymbol{\tau}}_{\varphi}^{h}, \hat{u}_h\right)_{\partial K_i} = \int_{\Omega_h} f\, v_{\varphi}\, d\mathbf{x}, \quad \forall \hat{u}_h \in M_h^{\Delta p}, \forall \varphi \in U_h\left(K_j\right), j = 1, \ldots, N^{\mathrm{el}}. \quad (5.2)\end{aligned}$$

Now integrating by parts the second and fourth terms on the left side of (5.2), we can rewrite (5.2) as

$$\begin{aligned}&\left(\boldsymbol{\tau}_{\varphi}^{h}, \boldsymbol{\sigma}_h\right)_{\Omega_h} - \left(\nabla \cdot \boldsymbol{\tau}_{\varphi}^{h}, u_h\right)_{\Omega_h} + \sum_{i=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot \boldsymbol{\tau}_{\varphi}^{h}, \hat{u}_h\right)_{\partial K_i} \\ &\quad - \left(\nabla v_{\varphi}^{h}, \boldsymbol{\sigma}_h\right)_{\Omega_h} + \sum_{i=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot \hat{\boldsymbol{\sigma}}_h, v_{\varphi}^{h}\right)_{\partial K_i} - \sum_{i=1}^{N^{\mathrm{el}}} \left(\mathbf{n} \cdot \hat{\boldsymbol{\sigma}}_h, \hat{v}_{\varphi}^{h}\right)_{\partial K_i} \\ &\quad = \int_{\Omega_h} f\, v_{\varphi}\, d\mathbf{x}, \quad \forall \hat{u}_h \in M_h^{\Delta p}, \forall \varphi \in U_h\left(K_j\right), j = 1, \ldots, N^{\mathrm{el}}, \quad (5.3)\end{aligned}$$

where we have introduced $\hat{\boldsymbol{\sigma}}_h = \boldsymbol{\sigma}_h + \alpha\left(u_h - \hat{u}_h\right)\mathbf{n}$.

Next, if we choose $\Delta p_j = 0$, $j = 1, \ldots, N^{\mathrm{el}}$, then (5.3) can be written as a single

variational formulation,

$$
\begin{aligned}
&(\boldsymbol{\tau}, \boldsymbol{\sigma}_h)_{\Omega_h} - (\nabla \cdot \boldsymbol{\tau}, u_h)_{\Omega_h} + \sum_{i=1}^{N^{\mathrm{el}}} (\mathbf{n} \cdot \boldsymbol{\tau}, \hat{u}_h)_{\partial K_i} \\
&\qquad - (\nabla v, \boldsymbol{\sigma}_h)_{\Omega_h} + \sum_{i=1}^{N^{\mathrm{el}}} (\mathbf{n} \cdot \hat{\boldsymbol{\sigma}}_h, v)_{\partial K_i} - \sum_{i=1}^{N^{\mathrm{el}}} (\mathbf{n} \cdot \hat{\boldsymbol{\sigma}}_h, \mu)_{\partial K_i} \\
&\qquad = \int_{\Omega_h} f \; v_\varphi \; d\mathbf{x}, \quad \forall \mu \in M_h^{\Delta p=0}, \forall\, (\boldsymbol{\tau}, v) \in V_h^{\Delta p=0} = \left[U_h\left(\Omega_h\right)\right]^{d+1},
\end{aligned}
$$

which is exactly the hybridized DG discretization [9] of the original problem (4.2).

We have shown that the DPG method with zero enriched exponent, i.e., $\Delta p_j = 0$, $j = 1, \ldots, N^{\mathrm{el}}$ coincides with the hybridized DG method for elliptic PDEs. In practice, one chooses $\Delta p_j \geq 1$ to obtain more accurate optimal test functions $\left(\boldsymbol{\tau}_\varphi^h, v_\varphi^h\right)$. As a result, the DPG approximate solution u_h from (4.7) is more accurate than that of the hybridized DG method using the the same polynomial order for U_h.

6. Conclusions. We have shown that starting from a discontinuous Petrov–Galerkin (DPG) method with zero enriched order one can re-derive a large class of discontinuous Galerkin (DG) methods for first order hyperbolic and elliptic equations. The first implication of this result is that the DG method can be considered as the least accurate DPG method. The second implication is that the DPG method can be viewed as a systematic way to improve the accuracy of the DG method when nonzero enriched orders are employed. A detailed derivation of the upwind DG, the local DG, and the hybridized DG from a DPG method with optimal test norms is presented, and one can obtain similar results for other existing DG methods.

REFERENCES

[1] Douglas N. Arnold, Franco Brezzi, Bernardo Cockburn, and L. Donatella Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM Journal on Numerical Analysis, 39 (2001/02), pp. 1749–1779 (electronic).

[2] Ivo Babuška, *Error bounds for finite element method*, Numerische Mathematik, 16 (1971), pp. 322–333.

[3] I. Babuška and A. Aziz, *Survey lectures on the mathematical foundations of the finite element method*, in The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, A. Aziz, ed., New York, 1972, Academic Press, pp. 3–359.

[4] Jamie Bramwell, Leszek Demkowicz, Jayadeep Gopalakrishnan, and Weifeng Qiu, *A locking-free hp DPG method for linear elasticity with symmetric stresses*, (2011). Submitted.Also ICES report 11-18.

[5] Tan Bui-Thanh, Leszek Demkowicz, and Omar Ghattas, *Constructively well-posed approximation method with unity inf-sup and continuity constants for partial differential equations*, Submitted, (2011).

[6] ———, *A fast algorithm for inverse transport equation using a discontinuous Petrov-Galerkin method*, In preparation, (2011).

[7] ———, *A unified discontinuous Petrov-Galerkin method and its analysis for Friedrichs' systems*, Submitted to SIAM J. Numer. Anal., (2011).

[8] Jesse Chan, Leszek Demkowicz, Robert Mose, and Nate Roberts, *A new discontinuous Petrov-Galerkin method with optimal test functions. Part V: Solutions of 1D Burger and Navier-Stokes equations*, Tech. Report 10-25, ICES, UT Austin, June 2010.

[9] 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), pp. 1319–1365.

[10] Bernardo Cockburn, George E. Karniadakis, and Chi-Wang Shu, *Discontinuous Galerkin Methods: Theory, Computation and Applications*, Lecture Notes in Computa-

tional Science and Engineering, Vol. 11, Springer Verlag, Berlin, Heidelberg, New York, 2000.

[11] B. Cockburn and C.-W. Shu, *The local discontinuous Galerkin finite element method for convection-diffusion systems*, SIAM Journal on Numerical Analysis, 35 (1998), pp. 2440–2463.

[12] Leszek Demkowicz, *"Babuška ↔ Brezzi?"*, Tech. Report 06-08, Institute for Computational Engineering and Sciences, the University of Texas at Austin, Appril 2006.

[13] Leszek Demkowicz and Jayadeep Gopalakrishnan, *A class of discontinuous Petrov–Galerkin methods. Part I: The transport equation*, Computer Methods in Applied Mechanics and Engineering, 199 (2010), pp. 1558–1572.

[14] ———, *Analysis of the DPG method for the Poisson equation*, (2011). To appear in SIAM Journal of Numerical Analysis.

[15] ———, *A class of discontinuous Petrov–Galerkin methods. Part II: Optimal test functions*, Numerical methods for Partial Differential Equations, 27 (2011), pp. 70–105.

[16] ———, *A class of discontinuous Petrov–Galerkin methods. Part IV: The optimal test norm and time-harmonic wave propagation in 1D*, Journal Computational Physics, 230 (2011), pp. 2406–2432.

[17] Leszek Demkowicz, Jayadeep Gopalakrishnan, Ignacio Muga, and Jeff Zitelli, *Wave number explicit analysis for a DPG method for the multi-dimensional Helmholtz equation*, Tech. Report 11-24, ICES, UT Austin, July 2011.

[18] Leszek Demkowicz and Norbert Heuer, *Robust DPG method for convection-dominated diffusion problems*, Submitted, (2011). Also ICES 11-33 report.

[19] Alexandre Ern and Jean-Luc Guermond, *Theory and Practice of Finite Elements*, vol. 159 of Applied Mathematical Sciences, Spinger-Verlag, 2004.

[20] Jayadeep Gopalakrishnan and Weifeng Qiu, *An analysis of the practical DPG method*, Submitted, (2011).

[21] Paul Houston, Max Jensen, and Endre Süli, *hp-Discontinuous Galerkin finite element methods with least-squares stabilization*, Journal of Scientific Computing, 17 (2002), pp. 3–25.

[22] P. LeSaint and P. A. Raviart, *On a finite element method for solving the neutron transport equation*, in Mathematical Aspects of Finite Element Methods in Partial Differential Equations, C. de Boor, ed., Academic Press, 1974, pp. 89–145.

[23] J. Tinsley Oden and Leszek F. Demkowicz, *Applied functional analysis*, CRC Press, 2010.

[24] Todd E. Peterson, *A note on the convergence of the discontinuous Galerkin method for a scalar hyperbolic equation*, SIAM J. Numer. Anal., 28 (1991), pp. 133–140.

[25] W. H. Reed and T. R. Hill, *Triangular mesh methods for the neutron transport equation*, Tech. Report LA-UR-73-479, Los Alamos Scientific Laboratory, 1973.

[26] Jinchao Xu and Ludmil Zikatanov, *Some observations on Babuška and Brezzi theories*, Tech. Report AM222, Penn State University, September 2000. http://www.math.psu.edu/ccma/reports.html.

[27] Jeff Zitelli, Ignacio Muga, Leszek Demkowicz, Jayadeep Gopalakrishnan, David Pardo, and Victor M. Calo, *A class of discontinuous Petrov-Galerkin methods. Part IV: Wave propagation*, Tech. Report 10-17, ICES, UT Austin, May 2010.