A DUAL PETROV-GALERKIN FINITE ELEMENT METHOD FOR THE CONVECTION-DIFFUSION EQUATION

JESSE CHAN, JOHN A. EVANS, WEIFENG QIU

Abstract. We present a minimum-residual finite element method for convection-diffusion problems in a higher order, adaptive, continuous Galerkin setting. The method borrows concepts from both the Discontinuous Petrov-Galerkin (DPG) method by Demkowicz and Gopalakrishnan [1] and the method of variational stabilization by Cohen, Dahmen, and Welper [2], and it can also be interpreted as a variational multiscale method in which the fine-scales are defined through a dual-orthogonality condition. A key ingredient in the method is the proper choice of norm used to measure the residual, and we present two choices which are observed to be robust in both convection and diffusion-dominated regimes, as well as a proof of stability for quasi-uniform meshes and a method for the weak imposition of boundary conditions. Numerically obtained convergence rates in 2D are reported, and benchmark numerical examples are given to illustrate the behavior of the method.

1. Introduction. It is well known that the standard Galerkin finite element method performs very poorly for the convection-diffusion equation – in the convection-dominated case, it experiences spurious oscillations in the solution that grow as $\epsilon \to 0$. The problem is connected back to a loss of discrete coercivity with respect to the standard H^1 norm [3]. The concept of *stabilized* methods was introduced in order to combat such oscillations, the most popular of which is the Streamline Upwind Petrov-Galerkin (SUPG) method [4]. The method can be interpreted as adding a sufficient amount of artificial viscosity in the streamline direction in order to restore discrete coercivity with respect to a new "streamline-diffusion" norm [5]. SUPG is also an example of a residual-based stabilization, where the stabilization mechanism disappears as the strong residual of the equation is satisfied.

A connection can be drawn between residual-based stabilized methods and Petrov-Galerkin schemes, where the trial (approximating) functions and test (weighting) functions are allowed to differ. Specifically, the SUPG method can be interpreted as a modification of standard test functions[1], biasing them in the upwind direction based on mesh size, order of polynomial approximation, the magnitude of convection, and the diffusion parameter. More recently, the concept of Petrov-Galerkin methods has been connected to a novel *minimum residual* framework – it is shown that the minimization of a specific residual corresponding to a variational formulation naturally leads to the concept of optimal test functions [6]. Additionally, this framework has been exploited as an alternative method of proving stability for more standard methods in the least squares finite element community [7].

Optimal test functions resulting from residual minimization were first implemented by Demkowicz and Gopalakrishnan in [1, 8]. The connection between stabilization and least squares/minimum residual methods has been observed previously [9]; however, the Discontinuous Petrov-Galerkin method distinguishes itself by measuring the residual of the operator form of the equation, which is posed in the dual space. Independently, Cohen, Dahmen and Welper introduced an alternative saddle-point formulation of such a minimum residual method in [2], which alluded to a Variational Multiscale (VMS) perspective [10, 11, 12]. We refer to the method of Cohen, Dahmen, and Welper as a Dual Petrov-Galerkin method, which distinguishes itself from the Discontinuous Petrov-Galerkin method in that it does not use a broken test space.

The goal of this paper is three-fold. The first is to present a method which borrows concepts from each of these recent works and to derive a more detailed connection between the Petrov-Galerkin, stabilized, and VMS perspectives of minimum-residual methods. The second is to demonstrate that the proposed method is robust in both convection and diffusion-dominated regimes provided the dual norm used to measure the residual is chosen intelligently. The last goal is to show that the method is stable for arbitrary higher order and adaptive meshes and easily implemented using existing finite element codes and technologies.

[1] We note that one cannot recover the SUPG formulation beginning with the strong form of the equations and the SUPG test functions; the interpretation of SUPG as a Petrov-Galerkin scheme holds only locally, on element interiors.

2. A minimum-residual method. Our starting point is the minimization of some measure of error over a finite-dimensional space. We begin by first introducing an abstract variational formulation

$$\begin{cases} \text{Given } l \in V^*, \text{ find } u \in U \text{ such that} \\ b(u,v) = l(v), \quad \forall v \in V, \end{cases} \tag{2.1}$$

where $b(\cdot,\cdot) : U \times V \to \mathbb{R}$ is a continuous bilinear form. Throughout the paper, we assume that the trial space U and test space V are real Hilbert spaces, and denote U^* and V^* as the respective topological dual spaces. Throughout the paper, we suppose the variational problem (2.1) to be well-posed in the inf-sup sense. We can then identify a unique operator $B : U \to V^*$ such that

$$\langle Bu, v\rangle_{V^*\times V} \coloneqq b(u,v), \quad u \in U, v \in V$$

with $\langle\cdot,\cdot\rangle_{V^*\times V}$ denoting the duality pairing between V^* and V, to obtain the operator form of the variational problem

$$Bu = l \quad \text{in } V^*. \tag{2.2}$$

We are interested in minimizing the residual over the discrete approximating subspace $U_h \subset U$

$$u_h = \underset{u_h \in U_h}{\arg\min}\, J(u_h) \coloneqq \frac{1}{2}\|l - Bu_h\|_{V^*}^2 \coloneqq \frac{1}{2}\sup_{v\in V\setminus\{0\}} \frac{|l(v) - b(u_h,v)|^2}{\|v\|_V^2}.$$

For convenience in writing, we will abuse the notation $\sup_{v\in V}$ to denote $\sup_{v\in V\setminus\{0\}}$ for the remainder of the paper. If we define the problem-dependent *energy norm*

$$\|u\|_E \coloneqq \|Bu\|_{V^*},$$

then we can equate the minimization of $J(u_h)$ with the minimization of error in $\|\cdot\|_E$.

The first order optimality condition for minimization of the quadratic functional $J(u_h)$ requires the Gâteaux derivative to be zero in all directions $\delta u \in U_h$,

$$(l - Bu_h, B\delta u)_{V^*} = 0, \quad \forall \delta u \in U, \tag{2.3}$$

which is nothing more than the least-squares condition enforcing orthogonality of error with respect to the inner product on V.

The difficulty in working with the first-order optimality condition (2.3) is that the inner product $(\cdot,\cdot)_{V^*}$ cannot be evaluated explicitly. However, we have that

$$(l - Bu_h, B\delta u)_{V^*} = \left(R_V^{-1}(l - Bu_h), R_V^{-1}B\delta u\right)_V = 0, \tag{2.4}$$

where $R_V : V \to V^*$ is the Riesz map mapping elements of a Hilbert space V to elements of the dual V^* defined by

$$\langle R_V v, \delta v\rangle_{V^*\times V} \coloneqq (v, \delta v)_V.$$

Furthermore, the Riesz operator is an isometry, such that $J(u_h) = \frac{1}{2}\|l - Bu_h\|_{V^*}^2 = \frac{1}{2}\left\|R_V^{-1}(l - Bu_h)\right\|_V^2$. Thus, satisfaction of (2.4) is exactly equivalent to satisfaction of the original optimality conditions (2.3).

2.1. Saddle point formulation. We can translate the optimality conditions (2.4) into a more standard variational problem. First, given some finite dimensional solution $u_h \in U_h$, we define the *error representation function* e such that

$$(e, v)_V = l(v) - b(u_h, v), \quad \forall v \in V.$$

We recognize this condition as defining $R_V e = l - Bu_h$; inversion of the Riesz map gives us

$$e = R_V^{-1}\left(l - Bu_h\right) = R_V^{-1} B\left(u - u_h\right),$$

from which we can see that $\|e\|_V = \|u - u_h\|_E$.

Secondly, we need to enforce the orthogonality of the error in equation (2.4), which becomes

$$\begin{aligned}\left(e, R_V^{-1} B\delta u\right)_V &= 0, \quad \forall \delta u \in U_h \\ \Rightarrow \left\langle e, B\delta u\right\rangle_{V\times V^*} &= b(\delta u, e) = 0.\end{aligned}$$

Combining these two conditions together returns the symmetric saddle-point problem for $(e, u_h) \in V \times U_h$

$$\begin{aligned}(e, v)_V + b(u_h, v) &= l(v), \quad \forall v \in V \\ b(\delta u, e) &= 0, \quad \forall \delta u \in U_h.\end{aligned}$$

We note that, under the perspective of the minimization of $J(u_h)$, this system can be viewed as a constrained optimization problem for e, where u_h are Lagrange multipliers enforcing satisfaction of the variational problem.

At the moment, V remains an infinite-dimensional space, and in practice, will have to be replaced by some $V_h \subset V$. Under this finite dimensional setting, the choice of discretization V_h will determine how effectively the solution of the discrete problem will minimize $J(u_h)$ – inadequately enriched choices of V_h may result in solutions that are far from the true minimizer.

This formulation was first introduced by Cohen, Dahmen, and Welper in [2], where V_h was determined adaptively. The method proceeded in two steps – first, V_h was set equal to U_h on what we will define as the *coarse mesh*. Next, *a posteriori* error estimators for e (based on a specific choice of inner product $(\cdot,\cdot)_V$) were used to drive adaptivity to determine a fine-scale mesh on which V_h was supported. We choose a simpler representation – if U_h is represented by piecewise polynomials of order p, we set V_h to be piecewise polynomials of order $p+\triangle p$. We note that these two choices of discretization for V are not mutually exclusive, and that novel choices for V_h are likely the key to yielding computationally effective methods under this framework.

2.2. DPG optimal test function framework. The DPG method was introduced by Demkowicz and Gopalakrishnan in 2009, and rapidly applied to a series of problems in computational mechanics [1, 13, 14, 15, 16]. Though the DPG method shares the same variational setting as Cohen et al. each method followed very different approaches in implementation. Recall the orthogonality of error (2.4); by defining the error representation function e and undoing the Riesz map acting on $B\delta u$, we recovered a saddle point formulation. However, it is possible instead to undo the Riesz map acting on the residual, to recover

$$\begin{aligned}\left(R_V^{-1}\left(l - Bu_h\right), R_V^{-1} B\delta u\right)_V &= 0, \quad \forall \delta u \in U_h \\ \Rightarrow \left\langle \left(l - Bu_h\right), R_V^{-1} B\delta u\right\rangle_{V^*\times V} &= 0 \\ \Rightarrow l\left(R_V^{-1} B\delta u\right) - b\left(u_h, R_V^{-1} B\delta u\right) &= 0\end{aligned}$$

If we define $v_{\delta u} \coloneqq R_V^{-1} B\delta u$, then we recover the original variational formulation with a non-standard test space

$$b(u_h, v_{\delta u}) = l(v_{\delta u}), \quad \forall \delta u \in U_h.$$

The DPG method with optimal test functions was constructed under this perspective of minimization of $J(u_h)$. While $v_{\delta u} \in V$ must still be solved for approximately, under the assumption of discontinuous test functions, this solve can be localized to a single element.[2]

[2] In practice, if $U_h(K) = P^p(K)$ is the space of polynomials of order p over the element K, then the enriched space $V_h \subset V$ is chosen to be $V_h(K) = P^{p+\triangle p}(K)$, where $\triangle p \geq n$, and d is the spatial dimension.

We can draw yet another connection to the DPG method by considering the discrete problem which arises from the saddle point formulation. Defining the vectors of coefficients $\boldsymbol{u} = \{u_j\}_{j=1}^{\dim U_h}$ and $\boldsymbol{e} = \{e_j\}_{j=1}^{\dim V_h}$, our discrete saddle point system is

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{e} \\ \boldsymbol{u} \end{bmatrix} = \begin{bmatrix} \boldsymbol{f} \\ 0 \end{bmatrix},$$

where, for basis functions $\psi_j \in U_h$ and $\phi_i, \phi_j \in V_h$,

$$\boldsymbol{A}_{ij} \coloneqq (\phi_i, \phi_j)_V, \quad \boldsymbol{B}_{ij} \coloneqq b(\psi_j, \phi_i),$$

such that $\boldsymbol{A}$ is a square matrix and $\boldsymbol{B}$ is an overdetermined rectangular matrix. Expressing $\boldsymbol{e} = \boldsymbol{A}^{-1}(\boldsymbol{f} - \boldsymbol{B}\boldsymbol{u})$ allows us to statically condense out the degrees of freedom $\boldsymbol{e}$ to form the Schur complement

$$\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{u} = \boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{f}.$$

The above system is essentially an algebraic least squares system, but preconditioned on the inside with the positive-definite operator $\boldsymbol{A}$, which is precisely the discrete system that arises under the optimal test function framework of DPG. We can see this further by noting that

$$\left(\boldsymbol{A}^{-1}\boldsymbol{B}\right)^T \boldsymbol{B}\boldsymbol{u} = \left(\boldsymbol{A}^{-1}\boldsymbol{B}\right)^T \boldsymbol{f}$$

is equivalent to

$$b\left(u_h, \sum_k^{\dim V_h} v_{ik}\phi_k\right) = l\left(\sum_k^{\dim V_h} v_{ik}\phi_k\right), \quad i = 1, \ldots, \dim(U_h)$$

where $v_{ik} = \left(\boldsymbol{A}^{-1}\boldsymbol{B}\right)_{ik}$ are the degrees of freedom associated with the ith approximated optimal test function.

The efficiency of the DPG method lies in that, under the assumption of a localizable norm and optimal test functions, $\boldsymbol{A}$ is block diagonal and can be inverted locally [6]. Under the ultra-weak formulation of DPG, one can show that this corresponds exactly to choosing a non-conforming DG space for V, where continuity is enforced weakly at element interfaces by requiring that the jump be orthogonal to polynomials of order p [17].[3]

2.3. Variational Multiscale Framework. The above saddle point system can also be directly derived from variational multiscale (VMS) principles [10, 11]. The key difference between this and standard multiscale methods is the way in which the fine scales influence the coarse; approximating the fine-scale problem is as difficult as approximating the original equation, so analytic approximations or assumptions must typically be made. However, through the proper change of variables, this method converts the fine-scale problem into a symmetric-positive definite one, allowing for a well-behaved subgrid model of fine scale behavior.

We begin again with the variational problem

$$\langle Bu, v\rangle_{V^* \times V} = \langle l, v\rangle_{V^* \times V},$$

where, as before, $B : U \to V^*$ is the operator form of $b(u, v)$, and $l \in V^*$ is the load. Standard VMS techniques then decompose U into a *coarse scale* space U_h and *fine scale* space U', such that the coarse and fine scales $u_h \in U_h \subset U$ and $u' \in U'$ are related through some yet-unspecified orthogonality condition. Decomposing $u = u_h + u'$ gives

$$\begin{aligned} u_h \in U_h, u' \in U', v \in V \\ \langle Bu_h, v\rangle_{V^* \times V} + \langle Bu', v\rangle_{V^* \times V} = \langle l, v\rangle_{V^* \times V} \end{aligned}$$

[3]Such a weakly conforming method was first introduced in [18], where optimal rates of convergence were proved under the assumption that the DG space is at most of order $p + 1$.

where $U = U_h \oplus U'$. We can now explicitly choose our orthogonality condition defining the fine scales and rewrite the above problem as

$$\begin{aligned} u_h \in U_h, u' \in U, v \in V \\ \langle Bu_h, v\rangle_{V^* \times V} + \langle Bu', v\rangle_{V^* \times V} &= \langle l, v\rangle_{V^* \times V} \\ (u', \bar{w})_E &= 0, \quad \forall \bar{w} \in U_h, \end{aligned}$$

where $(u', \bar{w})_E$ is defined through

$$\begin{aligned} (u', \bar{w})_E &= (Tu', T\bar{w})_V \\ T : U \to V, \quad T &\coloneqq R_V^{-1} B \end{aligned}$$

where $(v, \delta v)_V$ is an inner product on V and R_V is the Riesz map associated with this inner product. DPG literature refers to T as the *trial-to-test* operator, which, given a basis function $\phi_i \in U_h$, returns the optimal test function for ϕ_i [1].

If we define our fine-scale using the above orthogonality condition, we have that

$$\begin{aligned} (Tu', v)_V &= \left(R_V^{-1} Bu', v\right)_V = \langle Bu', v\rangle_{V^* \times V} \\ (u', \bar{w})_E &= \left(Tu', R_V^{-1} B\bar{w}\right)_V = \langle Tu', B\bar{w}\rangle_{V \times V^*} \end{aligned}$$

and can now rewrite our above VMS formulation as

$$\begin{aligned} u_h \in U_h, u' \in U, v \in V \\ \langle Bu_h, v\rangle_{V^* \times V} + (Tu', v)_V &= \langle l, v\rangle_{V^* \times V} \\ (Tu', B\bar{w})_{V \times V^*} &= 0, \quad \forall \bar{w} \in U_h. \end{aligned}$$

The saddle point formulation of Section 2.1 can be interpreted as solving not for u', but for $e \coloneqq Tu'$. The discrete approximation then follows directly from setting $V \coloneqq V_h$, some finite-dimensional enriched space, which impacts the method through the approximation of Tu' via approximation of the Riesz map R_V^{-1}.

To summarize, the minimum-residual framework motivating both Cohen, Dahmen, and Welper's variational stabilization and Demkowicz and Gopalakrishnan's DPG method can be characterized as a VMS method under which

- Orthogonality of U_h and U' is defined through a nonstandard *energy inner product*

$$(u', \delta u)_E = 0, \quad \forall \delta u \in U_h,$$

- A change of variables (i. e. solving for Tu' instead of u') symmetrizes the fine-scale contribution and overall system.

3. Choices for test norms. As with the DPG method, the primary theoretical difficulty is in defining the norm $\|v\|_V$ such that the method returns good results. This choice of norm defines the measure of the residual which we minimize or the orthogonality condition defining the separation of coarse and fine scales, depending upon whether you adopt the minimum-residual or VMS perspective.

From this point on, we will focus specifically on H^1 standard Galerkin formulation for the scalar convection-diffusion problem on domain Ω

$$\nabla \cdot (\beta u) - \epsilon \Delta u = f,$$

which represents the convection of some quantity u subject to a convection field $\beta \in R^n$ and diffusion of magnitude $\epsilon \in \mathbb{R}$, where $\epsilon > 0$. We impose inflow and outflow boundary conditions

$$\begin{aligned} u|_{\Gamma_{\text{in}}} &= u_{\text{in}} \\ u|_{\Gamma_{\text{out}}} &= u_{\text{out}} \end{aligned}$$

where the following boundaries are defined as

$$\Gamma_{\text{in}} \coloneqq \{x \in \partial\Omega : \beta \cdot \vec{n}(x) \leq 0\}$$
$$\Gamma_{\text{out}} \coloneqq \{x \in \partial\Omega : \beta \cdot \vec{n}(x) > 0\}$$

where $\vec{n}(x)$ is the outward normal at a point on the boundary $\partial\Omega$. If it is non-empty, we may also define $\Gamma_0 \coloneqq \{x \in \partial\Omega s.\ t.\ \beta \cdot \vec{n}(x) = 0\}$ as the boundary with outward normal orthogonal to the streamline. For the remainder of this paper, we treat non-homogeneous boundary conditions as lifted forcing data, and consider only the homogeneous case.

For $u \in H_0^1(\Omega)$, our weak formulation becomes

$$(\nabla u, \beta v + \epsilon \nabla v)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)}, \quad \forall v \in H_0^1.$$

A norm based on the symmetric part of the bilinear form was explored in [2], but yielded approximations which degnerated in ϵ. We present here two different methods based on alternative choices for test norms on V for the convection-diffusion equation that both yield good results as $\epsilon \to 0$.

We note that the first-order term $(\nabla u, \beta v)$ may also be integrated by parts. The two are equivalent for $\nabla \cdot \beta = 0$ and give nearly identical results.

3.1. Weighted streamline diffusion norm. Our first method is motivated by previous results for the DPG method in [19], which achieved good results through a modification of the test norm. We set

$$\|v\|_{V,w}^2 \coloneqq \frac{L}{|\beta|} \left\| \sqrt{w}\beta \cdot \nabla v \right\|_{L^2(\Omega)}^2 + \epsilon \|\nabla v\|_{L^2(\Omega)}^2 + \alpha \|v\|_{L^2(\Omega)}^2,$$

where $L = |\Omega|^{\frac{1}{d}}$ is the length measure of Ω in d spatial dimensions, $|\beta| = \max_\Omega \|\beta\|_{l^2}$, and we have introduced the scaling $\frac{L}{|\beta|}$ on the streamline diffusion term for unit consistency[4]. $\alpha \geq 0$ is a user-specified constant with units of inverse time[5], and $w(x)$ is some smoothly varying positive weighting function. Specifically, we require $|w(x)|$ to be of magnitude $O(\epsilon)$ for $x \in \Gamma_{\text{in}}$, but $O(1)$ away from the inflow boundary.[6]

A second motivation behind the weighting function can also be found in [20, 3, 5], which refer to $w(x)$ as a cut-off function. A theoretical issue with the convection-diffusion problem is that $\|u\|_{H_0^1}$ and $\|u - u_h\|_{H_0^1}$ do not remain bounded as $\epsilon \to 0$, due to the fact that $u' = O(\epsilon^{-1})$ in the region of boundary or internal layers. These cut-off functions were introduced in order to yield ϵ-independent bounds on the error by localizing error estimates away from layers.

We can demonstrate the need for the weight/cut-off function $w(x)$ in the test norm – Figure 3.1 shows the result of two simulations on a uniform 16 element piecewise-linear mesh. The error representation e is approximated by quadratics over the same mesh. Without the weight, the solution degenerates near the inflow; a similar phenomenon was observed in [2, 1, 17].

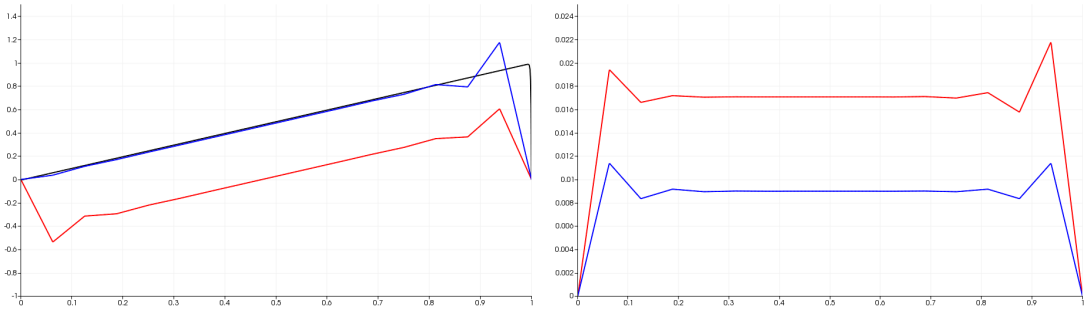
Physically speaking, the weight can be interpreted as a cutoff function for the adjoint solution [21], which displays a boundary layer at the inflow due to the fact that the direction of convection is reversed for the adjoint equation. Mathematically speaking, it can also be interpreted as restoring a sense of directionality to the test norm (which is sign of β).

3.2. Streamline norm with weak boundary conditions on e. While the above norm appears to give improved results, the issue of constructing a weighting function for complex geometries in higher dimensions can be fairly nontrivial. One possibility would be to solve a Laplace's equation for w on the domain, enforcing that $w(x) = \epsilon$ at inflow boundaries; however,

[4] In all problems considered, $\frac{L}{|\beta|} = O(1)$.

[5] In the case where $\alpha = 0$, the seminorm remains a norm due to equivalence with the H^1 seminorm, which is norm on H_0^1. Specifying $\alpha > 0$ does not change the results dramatically; however, it does appear to improve conditioning of $\boldsymbol{A}$.

[6] Specific restrictions on w are derived in [19] in context of the DPG method.



(a) Solution u (b) Fine-scales e

FIG. 3.1. *Comparisons of u and e for $\epsilon = 10^{-3}$, $f = 1$ and $\alpha = 0$. The solution when $w = 1$ is in red, and the solution for when $w = x + \epsilon$ is given in blue. The exact solution is given in black.*

the behavior of such a computationally determined weight has not been examined, and incurs additional complexity that is avoided by other methods.

Under this saddle-point framework, however, it is possible to avoid the use of a weight altogether. Currently, we seek $v, e \in H_0^1(\Omega)$; however, if we seek instead

$$v, e \in H^1_{\text{out}}(\Omega) \coloneqq \left\{ v \in H^1(\Omega) s.\ t.\ \ v|_{\Gamma_{out}} = 0 \right\},$$

we modify our variational formulation to be

$$(\nabla u, \beta v + \epsilon \nabla v)_{L^2(\Omega)} - \epsilon \int_{\Gamma_{\text{in}}} \frac{\partial u}{\partial n} v = (f, v)_{L^2(\Omega)}.$$

Under this modification of the variational formulation, we can set $w = 1$ and use the test norm

$$\|v\|_V^2 \coloneqq \frac{L}{|\beta|} \|\beta \cdot \nabla v\|^2_{L^2(\Omega)} + \epsilon \|\nabla v\|^2_{L^2(\Omega)} + \alpha \|v\|^2_{L^2(\Omega)}$$

without experiencing the issues displayed for $e \in H_0^1(\Omega)$ in Figure 3.1.

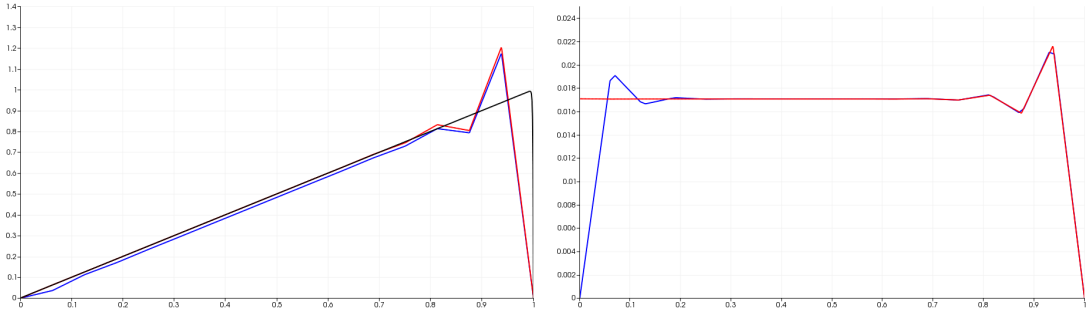
Physically speaking, the replacement of $H_0^1(\Omega)$ with $H^1_{\text{out}}(\Omega)$ can be viewed as releasing the strong imposition of a boundary condition at the inflow. The addition of the term $\epsilon \int_{\Gamma_{\text{in}}} \frac{\partial u}{\partial n} v$ is akin to a penalty term that weakly enforces a homogeneous boundary condition on e at the inflow; the orthogonality condition for e under this additional term is

$$(\nabla \delta u, \beta e + \epsilon \nabla e)_{L^2(\Omega)} - \epsilon \int_{\Gamma_{\text{in}}} \frac{\partial \delta u}{\partial n} e = 0, \quad \forall \delta u \in U_h$$

If U_h is taken to be a polynomial space of order p, the first term enforces orthogonality of $\beta e + \epsilon \nabla e$ to polynomials of order $p-1$ on the interior, while the second term enforces orthogonality of e to polynomials of order $p-1$ on the subset of the boundary Γ_{in}. This boundary condition is released in the limit as $\epsilon \to 0$; as a result, for small ϵ, the contribution $(e, v)_V$ no longer provides extraneous stabilization at the inflow for under-resolved meshes. However, for large ϵ, the boundary condition activates and e recovers homogeneous boundary conditions over the entire space V. The choice of weight $w(x)$ in the previous section can be seen as similarly removing the contribution of e near the inflow for small ϵ, but allowing it to return for large ϵ.

A comparison between the weighted norm $\|v\|_{V,w}$ with $e \in H_0^1(\Omega)$ and unweighted norm with $e \in H^1_{\text{out}}$ in one dimension is given in Figure 3.2. We note that Broersen and Stevenson have also achieved good results by making use of the space $H^1_{\text{out}}(\Omega)$ in their application of the DPG method to the mixed mild-weak formulation for convection-diffusion [22].

REMARK 1. *We remark that this inclusion of $\epsilon \int_{\Gamma_{\text{in}}} \frac{\partial u}{\partial n} v$ is a variational crime under the assumption that $u \in H_0^1(\Omega)$; the proper way to include such a term in a conforming manner (without introducing additional regularity assumptions on u) would be to introduce $\epsilon \frac{\partial u}{\partial n}$ as an*



(a) Solution u (b) Fine-scales e

FIG. 3.2. *Comparisons of u and e for $\epsilon = 10^{-3}$, $f = 1$ and $\alpha = 0$. The solution when $w = 1$ and $e \in H^1_{\text{out}}(\Omega)$ is in red, and the solution for when $w = x + \epsilon$ is given in blue. The exact solution is given in black.*

additional unknown on the boundary, similar to the work done using DPG under the ultra-weak formulation [1, 23]. However, we have found that doing so does not remove the boundary layer in the residual at the inflow, and numerical results appear to degenerate in quality as ϵ grows smaller and smaller. We believe this is due to the fact that, by identifying $\epsilon \frac{\partial u}{\partial n}$ as an additional unknown, the orthogonality boundary condition $\int_{\Gamma_{\text{in}}} \frac{\partial u}{\partial n} e = 0$ is decoupled from the interior integrals and enforced strongly. The same situation occurs when $\frac{\partial u}{\partial n}$ is declared as an additional unknown and when $\epsilon > 0$.

REMARK 2. *Various other methods of implementing a weak boundary condition on e at the inflow Γ_{in} were implemented; for example, the inner product $(e, v)_V$ was modified by adding the a weak penalization term*

$$\epsilon C(h, p + \triangle p) \int_{\Gamma_{\text{in}}} ev$$

to enforce the boundary condition at $e|_{\Gamma_{\text{in}}}$, where $C(h, p+\triangle p) = p^2/h$, which is a common choice of penalization parameter in high order finite element literature [24, 25]. However, the inclusion of the nonconforming term $\epsilon \int_{\Gamma_{\text{in}}} \frac{\partial u}{\partial n} v$ in the bilinear form still resulted in the best solutions for all our numerical tests.

Thus, the full formal characterization of the second method can be given as follows - solve for $u \in U_h \subset H^1_0(\Omega) \cap H^{\frac{3}{2}+\delta}(\Omega)$ (where δ is an positive real number), $e \in V_h \subset H^1_{\text{out}}(\Omega)$ such that

$$(e, v)_V + (\nabla u, \beta v + \epsilon \nabla v)_{L^2(\Omega)} - \epsilon \int_{\Gamma_{\text{in}}} \frac{\partial u}{\partial n} v = (f, v)_{L^2(\Omega)}, \quad \forall v \in V_h$$

$$(\nabla \delta u, \beta e + \epsilon \nabla e)_{L^2(\Omega)} - \epsilon \int_{\Gamma_{\text{in}}} \frac{\partial \delta u}{\partial n} e = 0, \quad \forall \delta u \in U_h$$

where

$$(e, v)_V \coloneqq (\beta \cdot \nabla e, \beta \cdot \nabla v)_{L^2(\Omega)} + \epsilon \, (\nabla e, \nabla v)_{L^2(\Omega)} \, .$$

We will utilize this second version of the method for all following numerical examples.

REMARK 3. *We note that, when the Peclet number $\text{Pe} = \frac{|\beta| L}{\epsilon} \gg 1$, the boundary term $\epsilon \int_{\Gamma_{\text{in}}} \frac{\partial u}{\partial n} v$ effectively disappears, and numerically speaking, we can use $e, v \in H^1_{\text{out}}(\Omega)$ without including the additional boundary term in the variational formulation.*

3.3. Stability theory. While a full proof of stability has not yet been developed, Brezzi theory for mixed methods [26] allows us to demonstrate stability on quasi-uniform meshes for under-resolved regimes.

LEMMA 3.1. *Assume the following conditions hold*

1. *V_h is sufficiently "rich", such that, for some $\gamma > 0$,*

$$\|\Pi_{V_h}(\beta \cdot \nabla \delta u)\|_{L^2(\Omega)} \geq \gamma \|\beta \cdot \nabla \delta u\|_{L^2(\Omega)}, \quad \forall \delta u \in U_h,$$

where Π_{V_h} is the L^2-orthogonal projection onto V_h.
2. *The convection field β satisfies $\nabla \cdot \beta = 0$ and*

$$\beta \cdot \nabla \psi \geq b_0 > 0, \quad \text{in}\Omega$$

for some $\psi \in C^1(\bar{\Omega})$.
3. *$\sqrt{\epsilon} \ll h$, where h is the mesh size.*

Under these assumptions, we can prove quasi-optimality of the method in the L^2-norm.

$$\|u - u_h\|_{L^2(\Omega)} + \|e - e_h\|_{L^2(\Omega)} \leq C_h \inf_{w_u \in U_h, w_e \in V_h} \left(\|u - w_u\|_{L^2(\Omega)} + \|e - w_e\|_{L^2(\Omega)} \right)$$

Proof. The result follows from proving two conditions in Brezzi theory; the first condition (inf-sup in the kernel) reduces to the coercivity condition

$$(e, v)_V = \epsilon (\nabla e, \nabla v)_{L^2(\Omega)} + (\beta \cdot \nabla e, \beta \cdot \nabla v)_{L^2(\Omega)} \geq C_1 \|v\|_{L^2(\Omega)}, \quad \forall v \in V_h, \tag{3.1}$$

while the second condition is

$$\sup_{v \in V_h \setminus \{0\}} \frac{b(u, v)}{\|v\|_{L^2}} = \frac{(\nabla u, \beta v + \epsilon \nabla v)_{L^2(\Omega)} - \epsilon \int_{\Gamma_{\text{in}}} \frac{\partial u}{\partial n} v}{\|v\|_{L^2}} \geq C_2 \|u\|_{L^2}. \tag{3.2}$$

The first condition 3.1 can be proven by noticing that

$$\|\beta \cdot \nabla v\|_{L^2(\Omega)} = \sup_{w \in L^2(\Omega) \setminus \{0\}} \frac{(\beta \cdot \nabla v, w)_{L^2(\Omega)}}{\|w\|_{L^2(\Omega)}}.$$

We take $w = -e^{\psi} v$, noting that $\nabla \cdot (\beta w) \geq \nabla \cdot \beta e^{\psi} v + b_0 e^{\psi} v + e^{\psi} \beta \cdot \nabla v$. Integrating by parts and rearranging, we recover

$$\begin{aligned} 2 (\beta \cdot \nabla v, w)_{L^2(\Omega)} &\geq \int_{\Gamma_{\text{in}}} -\beta_n v^2 e^{\psi} + \left(v, \nabla \cdot \beta e^{\psi} v + b_0 \psi e^{\psi} v \right)_{L^2(\Omega)} \\ &\geq \left(v, b_0 e^{\psi} v \right)_{L^2(\Omega)}. \end{aligned}$$

by $\beta_n < 0$ on Γ_{in} and Assumption 2. Since ψ is determined up to a constant, we set the constant such that $e^{\psi} = O(1)$ to arrive at the result.

To prove 3.2, we first take $v = \Pi_{V_h}(\beta \cdot \nabla u_h)$ for arbitrary $u_h \in U_h$. Then, by Assumption 1,

$$(\beta \nabla u_h, v)_{L^2(\Omega)} = \|\Pi_{V_h} \beta \cdot \nabla u_h\|^2_{L^2(\Omega)} \geq \gamma \|\beta \cdot \nabla u_h\|^2_{L^2(\Omega)}.$$

Since $u_h \in H^1_0$, by a similar argument in the proof of 3.1, we can show

$$(\beta \nabla u_h, v)_{L^2(\Omega)} \geq \gamma \|\beta \cdot \nabla u_h\|^2_{L^2(\Omega)} \geq \bar{C}_2 \left(\|u_h\|^2_{L^2(\Omega)} + \|\beta \cdot \nabla u_h\|^2_{L^2(\Omega)} \right).$$

Finally, note that

$$\begin{aligned} \left| \epsilon (\nabla u_h, \nabla v)_{L^2(\Omega)} \right| &\leq \epsilon \|\nabla u_h\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} \leq \epsilon h^{-2} \|u_h\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\ &\leq c_1 \|u_h\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}, \\ \left| \epsilon \int_{\Gamma_{\text{in}}} \frac{\partial u_h}{\partial n} v \right| &\leq \epsilon h^{-2} \|u_h\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \leq c_2 \|u_h\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}, \end{aligned}$$

using Assumption 3 and standard trace and inverse inequalities on quasi-uniform meshes. Note that $\|v\|_{L^2(\Omega)} = \|\beta \cdot \nabla u_h\|_{L^2(\Omega)}$; assuming c_1 and c_2 are sufficiently small relative to $\bar{C}_2$, an application of discrete Cauchy Schwarz gives

$$(\nabla u_h, \beta v + \epsilon \nabla v)_{L^2(\Omega)} - \epsilon \int_{\Gamma_{\text{in}}} \frac{\partial u_h}{\partial n} v \geq C_2 \left(\|u_h\|^2_{L^2(\Omega)} + \|\beta \cdot \nabla u_h\|^2_{L^2(\Omega)} \right)$$

which proves 3.2. □

REMARK 4. *Ideally, we would prove the above result under the stronger norm $\|v\|_V$ for v and a stronger norm on u. Under the test norm on V, 3.1 would be satisfied trivially; however, we have not yet been able to prove 3.2 yet under a stronger norm on v.*

4. Numerical examples. We present several numerical examples, demonstrating rates of convergence and performance of the method on several benchmark problems in 2D. All numerical examples were executed using the FEniCS codebase [27] to demonstrate that, unlike previous methods based upon the minimization of dual residuals, this method is easily implemented in most standard high-order finite element codes.

A remaining choice to be specified is the discretization of V_h: if U_h is given to be polynomials of order p, the enriched space V_h under which the fine-scales are approximated is set to be polynomials of order $p+\triangle p$ – in other words, we use a simple polynomial enrichment scheme over a fixed mesh to approximate the fine-scale error representation function e. In all experiments, $\triangle p = 1$ has been found to be sufficient, and larger $\triangle p$ has not been found to have a significant effect on the behavior of the method. Table 4.1 shows the effect of increased $\triangle p$ on a uniform mesh; there is less than .5% change in L^2 error between $\triangle p = 1$ and $\triangle p = 2$, and less than .1% change between $\triangle p = 2$ and $\triangle p = 3$.[7]

Order	Energy, $\epsilon = 1$	L^2, $\epsilon = 1$	Energy, $\epsilon = 10^{-4}$	L^2, $\epsilon = 10^{-4}$
$\triangle p = 1$	0.228580909	0.012003627	0.130406706	0.126763142
$\triangle p = 2$	0.230296524	0.011963812	0.130408851	0.126763338
$\triangle p = 3$	0.230781435	0.011953146	0.130410190	0.126763389
$\triangle p = 4$	0.230900562	0.011950356	0.130410964	0.126763422

FIG. 4.1. *Effect of $\triangle p$ on L^2 and energy error.*

We report only L^2 errors for each example, due to the fact that H^1 error is not a well-defined quantity as $\epsilon \to 0$. As a consequence, during adaptive refinement, the error does not always monotonically decrease or behave as expected, since adaptive refinement is driven by the the quantity $\|u - u_h\|_E = \|e\|_V$, which we expect to be equivalent (possibly with equivalence constants depending on ϵ) to the H^1 norm. We do expect $\|u - u_h\|_E$ to decrease monotonically for all ϵ (though not uniformly in ϵ) due to the minimum-residual nature of our method.

4.1. Erikkson-Johnson model problem. We adopt a problem first proposed by Eriksson and Johnson in [29] and later used in [19, 21] to determine the robustness of the Discontinuous Petrov-Galerkin method with respect to the diffusion parameter ϵ. For the choice of $\Omega = (0,1)^2$, $f = 0$, and $\beta = (1,0)^T$, the convection diffusion equation reduces to

$$\frac{\partial u}{\partial x} - \epsilon \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) = 0,$$

which has an exact solution by separation of variables, allowing us to analyze convergence of DPG for a wide range of ϵ. For boundary conditions, we impose

$$u = u_0, \quad x = 0,$$
$$u = 0, \quad x = 1, y = 0, 1$$

[7] This is in contrast to the original DPG method, which also adopts a polynomial enriched space (where enrichment is done locally over each element) – results by Gopalakrishnan and Qiu in [28] suggest that only $\triangle p \geq d$, where d is spatial dimension, is guaranteed to yield optimal convergence rates for Laplace's equation. However, it is unknown if these results carry over to this new formulation.

In this case, our exact solution is the series

$$u(x,y) = \sum_{n=0}^{\infty} C_n \frac{e^{r_{1,n}(x-1)} - e^{r_{2,n}(x-1)}}{e^{-r_{1,n}} - e^{-r_{2,n}}} \sin(n\pi y)$$

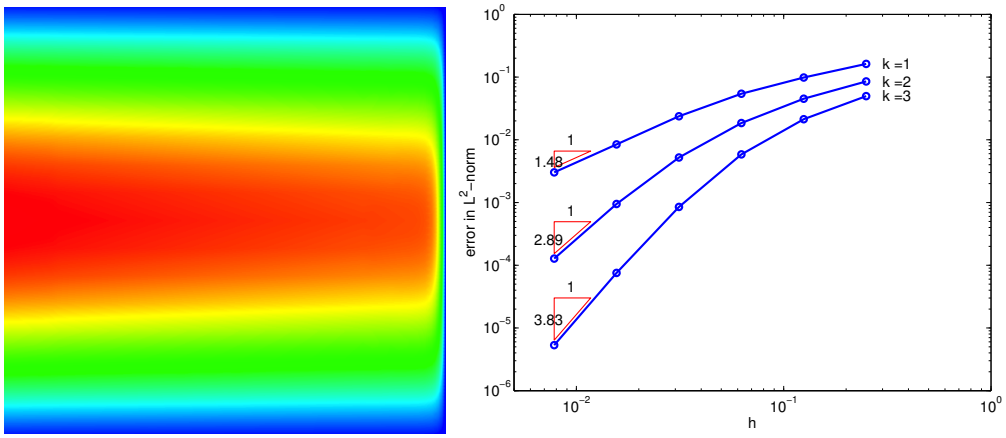
where

$$r_{1,2,n} = \frac{1 \pm \sqrt{1+4\lambda_n}}{2\epsilon},$$
$$\lambda_n = n^2\pi^2\epsilon.$$

We choose boundary conditions such that the exact solution is given with $C_0 = 1$ and $C_n = 0$ for $n \neq 1$.

$$u(x,y) = \frac{e^{r_1(x-1)} - e^{r_2(x-1)}}{e^{-r_1} - e^{-r_2}} \sin(\pi y)$$
$$r_{1,2} = \frac{1 \pm \sqrt{1+4\epsilon^2\pi^2}}{2\epsilon},$$

The problem is driven solely by inflow boundary conditions, and develops a boundary layer of width ϵ at the outflow $x = 1$. The resulting solution is shown in Figure 4.2 for $\epsilon = 10^{-2}$, along with convergence rates of $L^2(\Omega)$ error under uniform mesh refinement for various p.



(a) Exact solution for u

(b) Convergence rates for $\epsilon = 10^{-2}$

FIG. 4.2. *Erikkson-Johnson model problem for $\epsilon = 10^{-2}$, along with convergence rates under uniform refinement.*

Optimal rates of convergence are not expected for small ϵ due to the pre-asymptotic nature of the mesh relative to the solution – typically, h must be $O(\epsilon)$ in order to observe optimal rates of convergence [30].

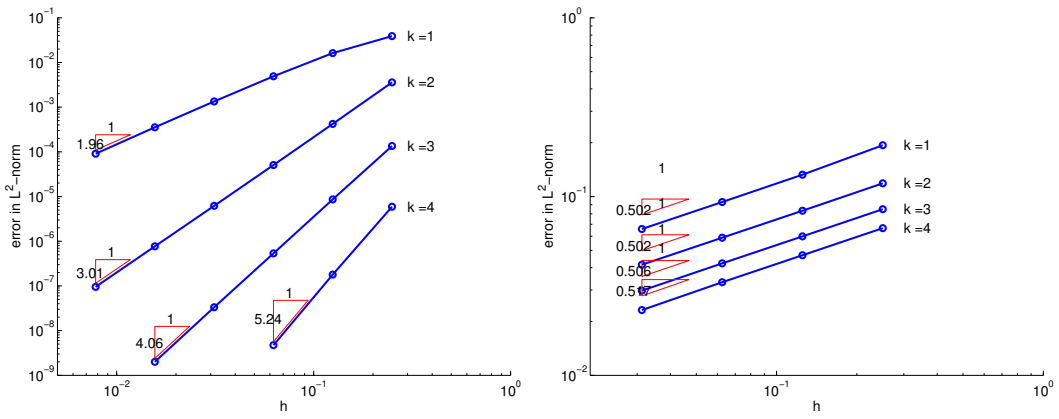
For large diffusion – $\epsilon = 1.0$, we do observe optimal $p+1$ rates of convergence in Figure 4.3. Likewise, when $\epsilon \ll h$, we observe a sub-optimal rate of $\frac{1}{2}$, the same as the rate suggested by theory and observed numerically in [22] due to the strong boundary layer present under small diffusion.

5. Adaptive mesh refinement. We experimented also with an adaptive refinement scheme. Since $\|e\|_V$ is localizable[8], we can evaluate it element-wise to get element error indicators

[8] A localizable norm $\|v\|_{V(\Omega)}$ can be written in the form

$$\|v\|^2_{V(\Omega)} = \sum_{K \in \Omega_h} \|v\|^2_{V(K)},$$

where $\|v\|_{V(K)}$ is a norm over each element K in the mesh Ω_h.



(a) $\epsilon = 1$ (b) $\epsilon = 10^{-4}$

FIG. 4.3. *Erikkson-Johnson model problem for $\epsilon = 1$ and $\epsilon = 10^{-4}$, along with convergence rates under uniform refinement.*

$e_K \coloneqq \|e\|_{V(K)}$. We implemented a bulk-chasing refinement strategy, where, given some factor $\theta \in [0,1]$, we refine the top $\lceil \theta N \rceil$ elements with the largest error indicators e_K. A greedy refinement scheme was also implemented, where all elements with $e_K > \theta \max_K e_K$ are refined; however, this refinement scheme tended to place refinements more conservatively, requiring many more refinement iterations to achieve a qualitatively good resolution. In all following experiments, θ is set to be .25.

5.1. Erikkson-Johnson model problem. We continue to verify our method using the Erikkson-Johnson with $\epsilon = 10^{-3}$. The solution u and the convergence/rates of $L^2(\Omega)$ error under adaptive mesh refinement for $p = 2$ are given in Figure 5.1. The energy error $\|u - u_h\|_E$ is also displayed for comparison.

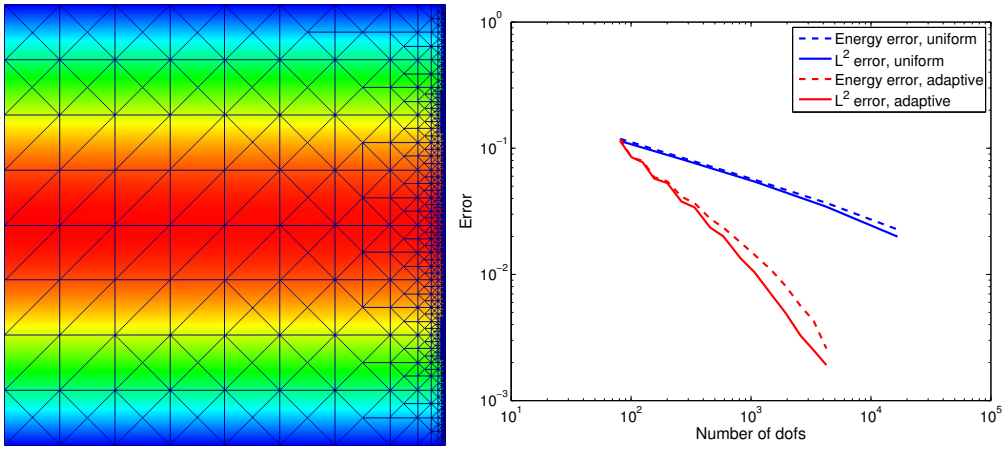


FIG. 5.1. *Erikkson-Johnson problem with $p = 2, \epsilon = 10^{-3}$ under 16 refinements.*

We can also examine the fine-scale error representation function e. Figure 5.2 displays a comparison between the error representation function after the first iteration of adaptive mesh refinement and the 16th iteration. Recall that the main contribution to the energy error is the streamline derivative of e; $\|u - u_h\|_E^2 = \|e\|_V^2 = \|\beta \cdot \nabla e\|_{L^2(\Omega)}^2 + \epsilon \|\nabla e\|_{L^2(\Omega)}^2$. Thus, variation in the streamline direction is picked up by $\|\beta \cdot \nabla e\|_{L^2(\Omega)}^2$ and is the primary contribution to the total error. After sufficient refinement and resolution of the outflow layer, the error representation function becomes smooth in the region of the boundary layer, and refinements will be placed

according to the viscous error $\epsilon \|\nabla e\|^2_{L^2(\Omega)}$.

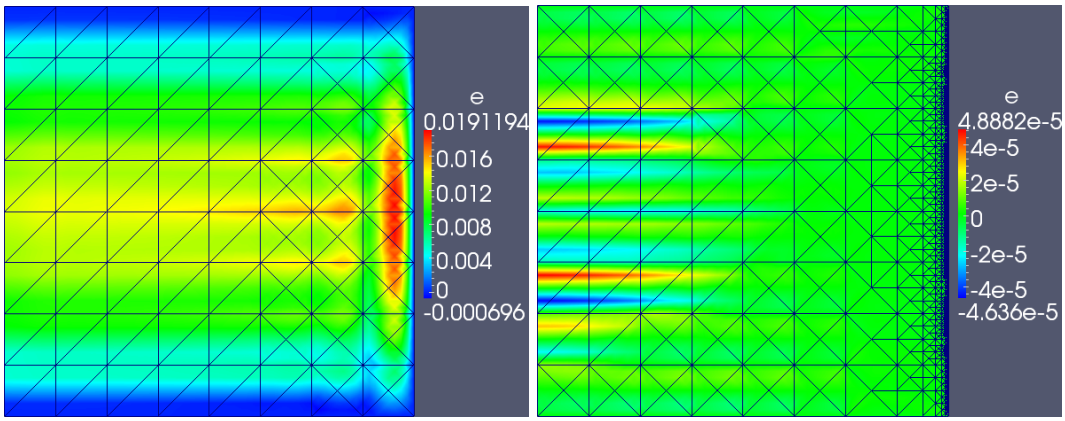


FIG. 5.2. *Comparison of the fine scale error representation e under the Erikkson-Johnson problem with $\epsilon = 10^{-3}$ for the mesh at the first and 16th refinement steps.*

One open question concerning this method is the energy norm $\|u\|_E$ which is induced by our choice of bilinear form and inner product on our choice of test space $V = H^1_{\text{out}}$. Figure 5.3 displays both L^2 and energy errors over a range of ϵ. We observe that the energy error appears to converge to the L^2 error as ϵ vanishes, however, for larger ϵ, an explicit expression for the energy error has not yet been derived.

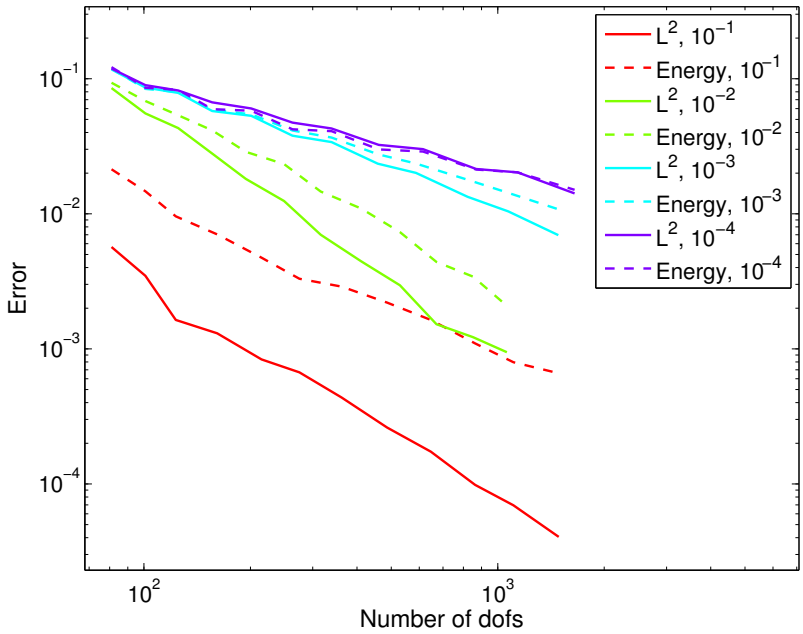
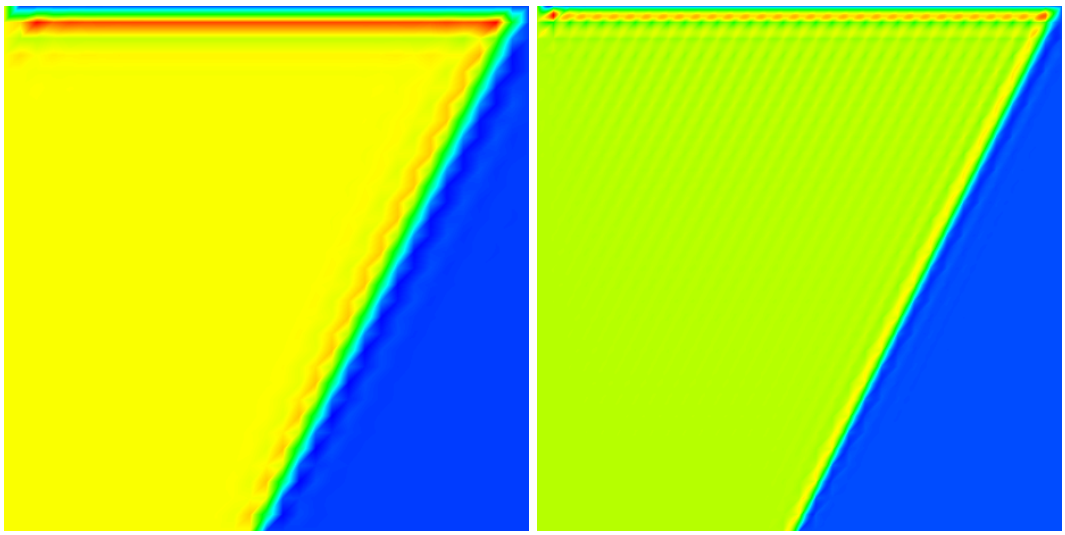


FIG. 5.3. *Comparison of $L^2(\Omega)$ and energy error convergences under adaptive refinement for the Erikkson-Johnson problem for $\epsilon = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$ and $p = 2$.*

5.2. Discontinuous forcing and advection skew to mesh. We examine first advection skew to the mesh, a common benchmark problem for convection-diffusion.

$$u = \begin{cases} 1 \text{ if } x = 0 \text{ or } y = 0 \text{ and } x < .5 \\ 0 \text{ otherwise,} \end{cases}$$



(a) $p = 1$ (b) $p = 2$

FIG. 5.4. *Solution u for advection skew to mesh with the upper-left hand corner discontinuity under both linear and quadratic uniform meshes.*

Under such an example, a boundary layer develops at the outflow boundaries $x, y = 1$, and an internal layer develops due to the fact that the discontinuity in f is parallel to the streamline. We note that under the uniform meshes in Figure 5.4 the method appears to deliver similar results to SUPG, and the discontinuity is convected across the domain in a minimially diffusive manner. We avoid the use of adaptive meshes for this problem; since there exists a discontinuity in the outflow boundary conditions at the top left corner of the domain, adaptive mesh refinement schemes will tend to place a large number of refinements near this discontinuity, while not improving resolution of the salient features of the solutions (boundary and interior layers) or decreasing error.[9]

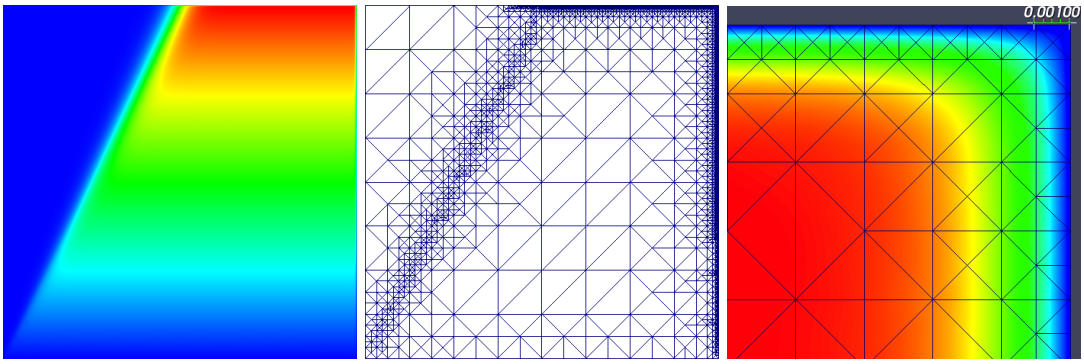


FIG. 5.5. *Model problem with $\epsilon = 10^{-3}$ and discontinuous forcing f.*

We consider next a slight variant of a model problem introduced by Hughes and Sangalli in [12]. The domain Ω is again taken to be the unit square, β is taken to be $\beta = [.5, 1]$, and $\epsilon = 10^{-3}$. Homogeneous Dirichlet boundary conditions are taken over the entire boundary, and

[9] For solutions with strong discontinuities, the placement of extraneous refinements is related to the fact that, for Sobolev-like residual norms, mesh refinement at discontinuities can actually increase residual error (this is also referred to as the Lavrentiev phenomenon; see [31] for related discussion). Thus, the use of adaptive mesh refinement for error reduction under such a problem is ineffective.

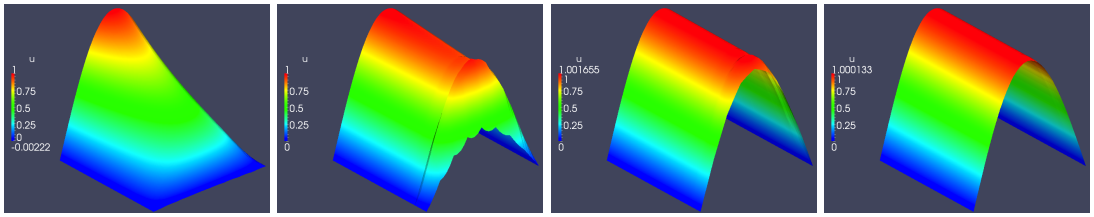


FIG. 5.6. *Weak boundary imposition for* $\epsilon = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$.

the problem is driven by a discontinuous forcing term $f(x, y)$, where

$$f(x,y) = \begin{cases} 1 \text{ if } y > 2x \\ 0 \text{ otherwise.} \end{cases}$$

Under such an example, a boundary layer develops at the outflow boundaries $x, y = 1$, and an internal layer develops due to the fact that the discontinuity in f is parallel to the streamline. Unlike the example of advection skew-to-mesh, the solution lies in H^1, implying that a good adaptive mesh refinement scheme should effectively resolve both interior and boundary layers. Figure 5.5 displays the solution and mesh resulting from 16 automatic mesh refinement iterations.

5.3. Weak boundary conditions and comparison to other methods. We close by introducing a method for the weak imposition of boundary conditions. For stability reasons, we modified the variational formulation with an additional boundary term on the inflow boundary. If we take instead

$$\begin{aligned} \delta u, u &\in H^{\frac{3}{2}+\delta}(\Omega) \cap H^1_{\text{in}} \\ e, v &\in H^{\frac{3}{2}+\delta}(\Omega) \cap H^1_{\text{out}} \end{aligned}$$

where $H^1_{\text{in}} \coloneqq \left\{ v \in H^1(\Omega) s.\ t.\ \ v|_{\Gamma_{\text{in}}} = 0 \right\}$, we can weakly impose the outflow boundary by modifying the variational formulation used in the saddle point system in Section 2.1. The resulting variational formulation is

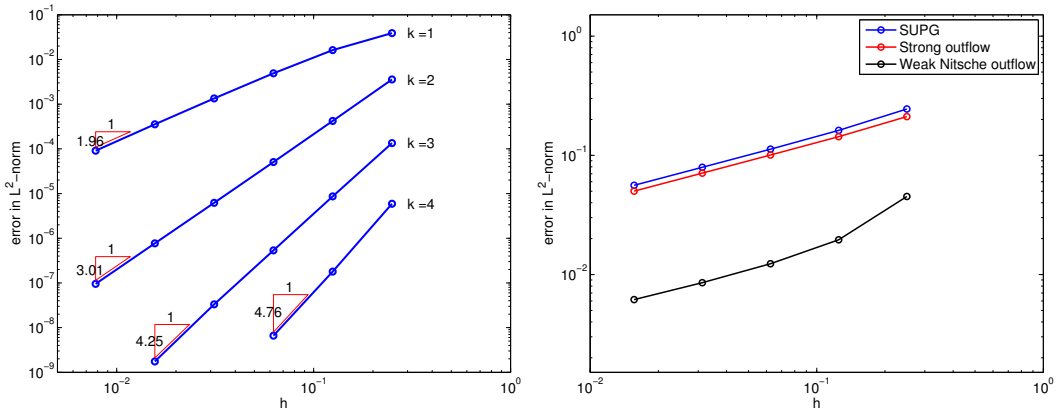
$$b(u,v) = (\nabla u, \beta v + \epsilon \nabla v)_{L^2(\Omega)} - \epsilon \int_{\Gamma_{\text{in}}} \frac{\partial u}{\partial n} v - \epsilon \int_{\Gamma_{\text{out}}} \frac{\partial v}{\partial n} u = (f, v)_{L^2(\Omega)}, \quad \forall v \in V_h.$$

Figure 5.6 demonstrates the release of the outflow boundary condition as ϵ vanishes relative to the mesh size: for a sufficiently resolved mesh with $\epsilon \approx h$, the boundary condition activates, but releases as the ratio $\epsilon/h \to 0$. Taking $\epsilon = 0$ convergences to a stable method for the pure convection equation.

Figure 5.7 compares the convergence of L^2 error with SUPG[10] to this method under both the strong and weak imposition of boundary conditions. While the rate of convergence does not change, the imposition of a weak outflow boundary condition greatly decreases the L^2 error. We note that the asymptotic rate of convergence is less than the optimal rate for $p = 4$; however, increasing to $\triangle p = 2$ recovers the optimal rate.

6. Conclusions and future direction. We have presented a higher order adaptive H^1-conforming method for convection-diffusion problems for very small ϵ, and have made connections to both the DPG method as well as the method of Cohen, Dahmen and Welper. The method has also been shown to be derivable in the Variational Multiscale framework as well, and is distinguished from traditional VMS methods in its definition of the fine-scale space and problem. Unlike standard stabilized methods, there are no stabilization parameters to specify.

[10]The stabilization parameter is taken to be $\tau = \frac{h}{|\beta|(p+1)}$ in this comparison.



(a) $\epsilon = 1.0$ (b) $\epsilon = 10^{-4}$, $p = 1$

FIG. 5.7. *Convergence rates for the diffusive regime, and comparison between methods (for $p = 1$) under uniform refinement for convective regime.*

We note that, while the method displays similarities to SUPG, we do not observe nodal exactness in 1D. The two methods differ in that we do not work directly with coarse scales defined by the H^1 inner product; unless we can induce an energy inner product such that $(u_h, u')_E = (u_h, u')_{H_0^1}$, then we should not expect nodal exactness in 1D. However, in higher space dimensions, the behavior of the method is qualitatively very similar to SUPG.

Since this method is derived from the same variational principles as DPG, it should likewise be automatically inf-sup stable for arbitrary *hp* meshes as well. *hp*-refinement is of interest due to the fact that, under proper choices between h and p refinement, exponential convergence rates are possible, and have been observed for several model problems [32].

6.1. Computational feasibility. The main obstacle to making this method computationally competitive is the doubling of the number of fully coupled degrees of freedom in solving the problem in saddle point system formulation. While a larger number of degrees of freedom is often cited as a detriment, it is often accompanied by an advantageous underlying structure – for example, discontinuous Galerkin methods are also often criticized for their "explosion of degrees of freedom" compared to continuous Galerkin methods [33], but their local nature aids significantly in parallelization, adaptivity, and explicit time integration, as opposed to methods based on continuous discretizations.

The cost of explicitly discretizing the fine scale error representation $e \in V_h$ roughly amounts to the cost of discretizing an additional coupled PDE under a higher resolution. The idea of solving an additional PDE for the purposes of stabilization is not unheard of [34]; however, the additional discretization cost must be minimized, or at least offset by additional positive properties of the resulting system. We propose several preliminary ideas for doing so:

- Dahmen, Huang, Schwab, and Welper utilize an iterative Uzawa method for the pure convection equation, which allows for the separate solution of two symmetric, coercive problems for u and e [35]. Dahmen et al. showed that the trial norm induced by their choice of test norm is precisely the L^2 norm, and utilized this in their Uzawa iteration. However, to use an Uzawa iteration for the convection-diffusion equation for $\epsilon > 0$ would require an explicit expression of the induced trial norm, which has not yet been derived in closed form.
- We hope to explore alternative discretizations for both U_h and V_h in order to minimize degrees of freedom or solution for in the overall system. A few examples of nonstandard discretizations include higher order continuity basis functions (splines and NURBS [36]), and discontinuous functions (DG). Preliminary experiments have yielded promising results in 1D under the combination of higher-continuity and C_0 bases for U_h and V_h, respectively. We hope to further explore intelligent mixing of discretizations for

trial and test spaces in future experiments.

- The DPG method achieves computational efficiency by block-diagonalizing $\boldsymbol{A}$ using discontinuous test functions and trace/flux degrees of freedom on element boundaries. The same method could be applied to the H^1 setting for convection-diffusion problems at the cost of introducing inter-element fluxes representing the trace on the boundary of an element resulting from integration by parts of the viscous term, similarly to the primal DPG method [37]. An alternative along the same lines would be to consider block diagonalization of $\boldsymbol{A}$ under a patch-based decomposition, such as FETI or one of its variants [38, 39]. Both of these methods would allow for more efficient inversion of $\boldsymbol{A}$ in the computation of the symmetric positive-definite Schur complement $\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B}$ at the cost of introducing additional hybrid unknowns on element faces. This method is akin to approximating the error representation e using a nonconforming primal hybrid finite element method [40].

7. Acknowledgements. The authors would like to acknowledge the help of Truman Ellis and Dr. Leszek Demkowicz for fruitful discussions and perspective. Jesse Chan was supported by the Department of Energy [National Nuclear Security Administration] under Award Number [DE- FC52-08NA28615]. John A. Evans was partially supported by the ICES Postdoctoral Fellowship at the Institute for Computational Engineering and Sciences, and by a grant from the Office of Naval Research [N00014-08-1-0992]. Weifeng Qiu was supported by the City University of Hong Kong under start-up grant No. 7200324.

REFERENCES

[1] L. Demkowicz and J. Gopalakrishnan. A class of discontinuous Petrov-Galerkin methods. ii. Optimal test functions. *Num. Meth. for Partial Diff. Eq*, 27:70–105, 2011.

[2] A. Cohen, W. Dahmen, and G. Welper. Adaptivity and variational stabilization for convection-diffusion equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46(5):1247–1273, 2012.

[3] H. Roos, M. Stynes, and L. Tobiska. *Robust numerical methods for singularly perturbed differential equations: convection-diffusion-reaction and flow problems*. Springer series in computational mathematics. Springer, 2008.

[4] A. Brooks and T. J. R. Hughes. Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Comp. Meth. Appl. Mech. Engr*, 32:199–259, 1982.

[5] C. Johnson, A. H. Schatz, and L. B. Wahlbin. Crosswind Smear and Pointwise Errors in Streamline Diffusion Finite Element Methods. *Mathematics of Computation*, 49, 1987.

[6] L. Demkowicz and J. Gopalakrishnan. An Overview of the DPG Method. Technical Report 13-02, ICES, January 2013.

[7] H. Chen, G. Fu, J. Li, and W. Qiu. First order least square method with ultra-weakly imposed boundary condition for convection dominated diffusion problems. *ArXiv e-prints*, September 2013.

[8] L. Demkowicz and J. Gopalakrishnan. A class of discontinuous Petrov-Galerkin methods. Part I: The transport equation. *Computer Methods in Applied Mechanics and Engineering*, 199(23-24):1558 – 1572, 2010.

[9] T. J. R. Hughes, L. Franca, and G. Hulbert. A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/least-squares method for advective-diffusive equations. *Comp. Meth. Appl. Mech. Engr*, 73:173–189, 1989.

[10] T. J. R. Hughes. Multiscale phenomena: Green's functions, the dirichlet-to-neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Computer Methods in Applied Mechanics and Engineering*, 127(1-4):387 – 401, 1995.

[11] T. J. R. Hughes, G. R. Feijoo, L. Mazzei, and J. Quincy. The variational multiscale method – a paradigm for computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 166(1-2):3 – 24, 1998. Advances in Stabilized Methods in Computational Mechanics.

[12] T. J. R. Hughes and G. Sangalli. Variational Multiscale Analysis: the Fine-scale Green's Function, Projection, Optimization, Localization, and Stabilized Methods. *SIAM J. Numer. Anal.*, 45(2):539–557, February 2007.

[13] L. Demkowicz, J. Gopalakrishnan, and A. Niemi. A class of discontinuous petrov-galerkin methods. part iii: Adaptivity. *Appl. Numer. Math.*, 62(4):396–427, April 2012.

[14] J. Zitelli, I. Muga, L. Demkowicz, J. Gopalakrishnan, D. Pardo, and V.M. Calo. A class of discontinuous Petrov–Galerkin methods. Part IV: The optimal test norm and time-harmonic wave propagation in 1D. *Journal of Computational Physics*, 230(7):2406 – 2432, 2011.

[15] L. Demkowicz and J. Li. Numerical simulations of cloaking problems using a DPG method. Technical Report 11-31, ICES, October 2011.

[16] N. Roberts, T. Bui Thanh, and L. Demkowicz. The DPG method for the Stokes problem. Technical Report 12-22, ICES, June 2012.
[17] J. Chan, J. Gopalakrishnan, and L. Demkowicz. Global properties of DPG test spaces for convection-diffusion problems. Technical Report 13-05, ICES, February 2013.
[18] S. C. Brenner and R. Scott. *The Mathematical Theory of Finite Element Methods*. Texts in Applied Mathematics. Springer, 2008.
[19] L. Demkowicz and N. Heuer. Robust DPG method for convection-dominated diffusion problems. *SIAM Journal on Numerical Analysis*, 51(5):2514–2537, 2013.
[20] M. Stynes. Convection-diffusion problems, SDFEM/SUPG and a priori meshes. *Int. J. Comput. Sci. Math.*, 1(2-4):412–431, January 2007.
[21] J. Chan, N. Heuer, T. Bui-Thanh, and L. Demkowicz. A robust {DPG} method for convection-dominated diffusion problems ii: Adjoint boundary conditions and mesh-dependent test norms. *Computers and Mathematics with Applications*, 2013. In Press, Corrected Proof.
[22] D. Broerson and R. Stevenson. A Petrov-Galerkin discretization with optimal test space of a mild-weak formulation of convection-diffusion equations in mixed form. Technical report, Korteweg-de Vries Institute for Mathematics, November 2012. Submitted.
[23] L. Demkowicz and J. Gopalakrishnan. Analysis of the DPG method for the Poisson equation. *SIAM J. Numer. Anal.*, 49(5):1788–1809, September 2011.
[24] J. A. Evans and T. J. R. Hughes. Explicit trace inequalities for isogeometric analysis and parametric hexahedral finite elements. *Numerische Mathematik*, 123(2):259–290, 2013.
[25] T. Warburton and J.S. Hesthaven. On the constants in hp-finite element trace inverse inequalities. *Computer Methods in Applied Mechanics and Engineering*, 192(25):2765 – 2773, 2003.
[26] L. Demkowicz. Babuska <=> Brezzi? Technical Report 06-08, ICES, 2006.
[27] A. Logg, K. A. Mardal, G. N. Wells, et al. *Automated Solution of Differential Equations by the Finite Element Method*. Springer, 2012.
[28] J. Gopalakrishnan and W. Qiu. An analysis of the practical DPG method. *Mathematics of Computation*, 230(286):537–552, 2013.
[29] K. Eriksson and C. Johnson. Adaptive streamline diffusion finite element methods for stationary convection-diffusion problems. *Mathematics of Computation*, 60(201):pp. 167–188, 1993.
[30] C. Schwab and M. Suri. The p and hp versions of the finite element method for problems with boundary layers. *Math. Comput.*, 65(216):1403–1429, October 1996.
[31] M. Gerritsma, R. van der Bas, B. De Maerschalck, B. Koren, and H. Deconinck. Least-squares spectral element method applied to the euler equations. *International Journal for Numerical Methods in Fluids*, 57(9):1371–1395, 2008.
[32] L. Demkowicz. *Computing With hp-adaptive Finite Elements: One and two dimensional elliptic and Maxwell problems*. Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. Chapman & Hall/CRC, 2006.
[33] J.A. Cottrell. *Isogeometric analysis and numerical modeling of the fine scales within the variational multiscale method*. ProQuest, 2007.
[34] G. Barter. *Shock Capturing with PDE-Based Artificial Viscosity for an Adaptive, Higher-Order Discontinuous Galerkin Finite Element Method*. PhD thesis in Aeronautics and Astronautics, Massachusetts Institute of Technology, 2008.
[35] W. Dahmen, C. Huang, C. Schwab, and G. Welper. Adaptive Petrov-Galerkin methods for first order transport equations. *SIAM J. Numerical Analysis*, 50(5):2420–2445, 2012.
[36] T. J. R. Hughes, J.A. Cottrell, and Y. Bazilevs. Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Computer Methods in Applied Mechanics and Engineering*, 194(39–41):4135 – 4195, 2005.
[37] L. Demkowicz and J. Gopalakrishnan. A primal DPG method without a first-order reformulation. *Computers and Mathematics with Applications*, 66(6):1058 – 1064, 2013.
[38] C. Farhat and F. Roux. A method of finite element tearing and interconnecting and its parallel solution algorithm. *International Journal for Numerical Methods in Engineering*, 32(6):1205–1227, 1991.
[39] C. Farhat, J. Mandel, and F. Roux. Optimal convergence properties of the FETI domain decomposition method. *Computer Methods in Applied Mechanics and Engineering*, 115(3–4):365 – 385, 1994.
[40] P. A. Raviart and J. M. Thomas. Primal hybrid finite element methods for 2nd order elliptic equations. *Mathematics of computation*, 31(138):391–413, 1977.