Space-Time Robustness

Truman Ellis, Jesse Chan, Leszek Demkowicz

1 Introduction

The discontinuous Petrov-Galerkin finite element method presents a promising new framework for developing robust numerical methods for computational mechanics. In a sense, DPG contains the promise of being an automated scientific computing technology; it provides stability for any variational formulation, optimal conversion rates in user-defined norm, no pre-asymptotic stability issues on coarse meshes, a measure of the error residual which can be used to robustly drive adaptivity, Hermitian positive definite stiffness matrices for any problem, weak enforcement of boundary conditions, and several other attractive properties. We start with an abstract derivation of the DPG framework then define the concept of a robust test norm, specialize to transient convection-diffusion, derive a new robust norm, then provide some numerical verifications of the theory.

1.1 Overview of DPG

1.1.1 A Generalized Minimum Residual Method

We begin with any well posed bilinear variational problem: find $u \in U$ such that

$$b(u, v) = l(v) \quad \forall v \in V$$

with operator $B : U \to V'$ $\quad$ (V' is the dual space to V) defined by $b(u, v) = \langle Bu, v\rangle_{V'\times V}$. This gives the operator equation:

$$Bu = l \in V' .$$

We wish to find the u_h in a finite dimensional subspace that minimizes the residual $Bu - l$ in V':

$$u_h = \underset{w_h\in U_h}{\arg\min} \frac{1}{2} \left\| Bu - l \right\|^2_{V'} .$$

This mathematical framework is very natural, but it is not yet practical as the V' norm is not especially translatable to computations. With the assumption that we are working with Hilbert spaces, we can use the Riesz representation theorem to find a complementary object in V rather than V'. Let $R_V : V \ni v \to (v, \cdot) \in V'$ be the Riesz map. Then the inverse Riesz map (which is an isometry) lets us represent our residual in V:

$$u_h = \underset{w_h\in U_h}{\arg\min} \frac{1}{2} \left\| R_V^{-1}(Bu - l) \right\|^2_{V} .$$

Since this is a minimization problem, we need to find the critical points, so taking the Gâteaux derivative to be zero in all directions $\delta u \in U_h$ gives,

$$\left(R_V^{-1}(Bu_h - l), R_V^{-1}B\delta u\right)_V = 0, \quad \forall \delta u \in U,$$

which by definition of the Riesz map is equivalent to the duality pairing

$$\left\langle Bu_h - l, R_V^{-1}B\delta u_h\right\rangle = 0 \quad \forall \delta u_h \in U_h .$$

We can define an optimal test function $v_{\delta u_h} := R_V^{-1}B\delta u_h$ for each trial function δu_h. This allows us to revert back to our original bilinear form with a finite dimensional set of trial and test functions:

$$b(u_h, v_{\delta u_h}) = l(v_{\delta u_h}).$$

Note that $v_{\delta u_h} \in V$ comes from the auxiliary problem

$$(v_{\delta u_h}, \delta v)_V = \langle R_V v_{\delta u_h}, \delta v \rangle = \langle B \delta u_h, \delta v \rangle = b(\delta u_h, \delta v) \quad \forall \delta v \in V.$$

We might call this an *optimal Petrov-Galerkin* method. We arrive at the same method by realizing the supremum in the inf-sup condition (see [1], motivating the *optimal* nomenclature. As a minimum residual method, optimal Petrov-Galerkin methods produce Hermitian, positive-definite stiffness matrices since

$$b(u_h, v_{\delta \boldsymbol{u}_h}) = (v_{u_h}, v_{\delta \boldsymbol{u}_h})_V = \overline{(v_{\delta \boldsymbol{u}_h}, v_{u_h})} = \overline{b(\delta u_h, v_{u_h})}\,.$$

The energy norm (defined by $\|u\|_E := \|Bu\|_{V'}$) of the residual comes from a straightforward post-process:

$$\|u_h - u\|_E = \|B(u_h - u)\|_{V'} = \|Bu_h - l\|_{V'} = \left\|R_V^{-1}(Bu_h - l)\right\|_V \,,$$

where we designate $R_V^{-1}(Bu_h - l)$ the *error representation function*. This has proven to be a very robust *a-posteriori* error estimator for driving adaptivity.

1.1.2 Optimal Stability

Babuška's theorem [2] states that discrete stability and approximability imply convergence. That is, if M is the continuity constant for $b(u, v)$ which satisfies the discrete inf-sup condition with constant γ_h,

$$\sup_{v_h \in V_h} \frac{|b(u, v)|}{\|v_h\|_V} \geq \gamma_h \|u_h\|_U \,,$$

then the Galerkin error satisfies the bound

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

Optimal test functions realize the supremum in the discrete discrete inf-sup condition such that $\gamma_h \geq \gamma$, the infinite-dimensional inf-sup constant. If we then use the energy norm for $\|\cdot\|_U$, then $M = \gamma = 1$ and Babuška's estimate implies that the optimal Petrov-Galerkin method is the most stable Petrov-Galerkin method possible.

1.2 A Concrete Example

1.2.1 Problem Description

In order to better illustrate choice of the U and V spaces, we introduce the transient convection-diffusion problem. Consider spatial domain domain Ω and corresponding space-time domain $Q = \Omega \times [0, T]$ with boundary $\Gamma = \Gamma_- \cup \Gamma_+ \cup \Gamma_0 \cup \Gamma_T$ where Γ_- is the spatial inflow boundary, Γ_+ is the spatial outflow boundary, Γ_0 is the initial time boundary, and Γ_T is the final time boundary. Let Γ_h denote the entire mesh skeleton.

The transient convection-diffusion equation is

$$\frac{\partial u}{\partial t} + \nabla \cdot (\boldsymbol{\beta} u) - \epsilon \Delta u = f \,,$$

where u is the quantity of interest, often interpreted to be a concentration of some quantity, $\boldsymbol{\beta}$ is the convection vector, ϵ is the diffusion scale, and f is the source term.

Based on previous experience [**?**], we apply flux boundary conditions on the inflow and trace boundary conditions on the outflow

$$\begin{aligned} \operatorname{tr}(\boldsymbol{\beta} \cdot u - \epsilon \nabla u) \cdot \boldsymbol{n}_x = t_- \quad &\text{on} \quad \Gamma_- \\ \operatorname{tr}(u) = u_+ \quad &\text{on} \quad \Gamma_+ \\ \operatorname{tr}(u) = u_0 \quad &\text{on} \quad \Gamma_0 \,. \end{aligned}$$

1.2.2 Relevant Sobolev Spaces

Begin by defining operators $\nabla_{xt} u := \begin{pmatrix} \nabla u \\ \frac{\partial u}{\partial t} \end{pmatrix}$ and $\nabla_{xt} \cdot \boldsymbol{u} := \nabla \cdot \boldsymbol{u}_x + \frac{\partial u_t}{\partial t}$. We will need the following Sobolev spaces defined on our space-time domain.

$$
\begin{aligned}
H^1(Q) &= \left\{ u \in L^2(Q) \,:\, \nabla u \in \mathbf{L}^2(Q) \right\} \\
H^1_{xt}(Q) &= \left\{ u \in L^2(Q) \,:\, \nabla_{xt} u \in \mathbf{L}^2(Q) \right\} \\
\boldsymbol{H}(\mathrm{div}, Q) &= \left\{ \boldsymbol{\sigma} \in \mathbf{L}^2(Q) \,:\, \nabla \cdot \boldsymbol{\sigma} \in L^2(Q) \right\} \\
\boldsymbol{H}(\mathrm{div}_{xt}, Q) &= \left\{ \boldsymbol{\sigma} \in \mathbf{L}^2(Q) \,:\, \nabla_{xt} \cdot \boldsymbol{\sigma} \in L^2(Q) \right\}
\end{aligned}
$$

We will also need the following broken Sobolev spaces.

$$
\begin{aligned}
H^1(Q_h) &= \left\{ u \in L^2(Q) \,:\, u|_K \in H^1(K),\, K \in Q_h \right\} &&= \prod_{K \in Q_h} H^1(K) \\
H^1_{xt}(Q_h) &= \left\{ u \in L^2(Q) \,:\, u|_K \in H^1_{xt}(K),\, K \in Q_h \right\} &&= \prod_{K \in Q_h} H^1_{xt}(K) \\
\boldsymbol{H}(\mathrm{div}, Q_h) &= \left\{ \boldsymbol{\sigma} \in \mathbf{L}^2(Q) \,:\, u|_K \in \boldsymbol{H}(\mathrm{div}, K),\, K \in Q_h \right\} &&= \prod_{K \in Q_h} \boldsymbol{H}(\mathrm{div}, K) \\
\boldsymbol{H}(\mathrm{div}_{xt}, Q_h) &= \left\{ \boldsymbol{\sigma} \in \mathbf{L}^2(Q) \,:\, u|_K \in \boldsymbol{H}(\mathrm{div}_{xt}, K),\, K \in Q_h \right\} &&= \prod_{K \in Q_h} \boldsymbol{H}(\mathrm{div}_{xt}, K)
\end{aligned}
$$

Consider the following trace operators:

$$
\begin{aligned}
\mathrm{tr}^K_{\mathrm{grad}} u &= u|_{\partial K_x} && u \in H^1(K) \\
\mathrm{tr}^K_{\mathrm{div}_{xt}} \boldsymbol{\sigma} &= \boldsymbol{\sigma}|_{\partial K_{xt}} \cdot \boldsymbol{n}_{K_{xt}} && \boldsymbol{\sigma} \in \boldsymbol{H}(\mathrm{div}_{xt}, K)
\end{aligned}
$$

where ∂K_x refers to spatial faces of element K, ∂K_{xt} to the full space-time boundary, and $\boldsymbol{n}_{K_{xt}}$ is the unit outward normal on ∂K_{xt}. The operators tr_{grad} and $\mathrm{tr}_{\mathrm{div}_{xt}}$ perform the same operation element by element producing the linear maps

$$
\begin{aligned}
\mathrm{tr}_{\mathrm{grad}} &: H^1(Q_h) \to \prod_{K \in Q_h} H^{1/2}(\partial K_x) \\
\mathrm{tr}_{\mathrm{div}_{xt}} &: \boldsymbol{H}(\mathrm{div}_{xt}, Q_h) \to \prod_{K \in Q_h} H^{-1/2}(\partial K_{xt})
\end{aligned}
$$

Finally, we define spaces of interface functions. In order that our functions be single valued, we use the following definitions.

$$
\begin{aligned}
H^{1/2}(\partial Q_h) &= \mathrm{tr}_{\mathrm{grad}} H^1(Q)\,, \\
H^{-1/2}_{xt}(\partial Q_h) &= \mathrm{tr}_{\mathrm{div}_{xt}} H(\mathrm{div}_{xt}, Q)\,.
\end{aligned}
$$

For more details on broken and trace Sobolev spaces, see [3].

1.2.3 Variational Formulations

There are many possible manipulations that could be performed before arriving at a variational formulation. Breaking this into a system of first order equations allows us to consider four related formulations. The equivalent first order system is

$$
\begin{aligned}
\frac{1}{\epsilon}\boldsymbol{\sigma} - \nabla u &= 0 \\
\nabla_{xt} \cdot \begin{pmatrix} \boldsymbol{\beta} u - \boldsymbol{\sigma} \\ u \end{pmatrix} &= f\,,
\end{aligned}
\tag{1}
$$

where $\nabla_{xt}\cdot$ denotes a divergence operator in the space-time domain Q. Multiplying (1) by test functions $\boldsymbol{\tau} \in \mathbf{L}^2(Q)$ and $v \in L^2(Q)$, we obtain the following bilinear form:

$$
\begin{aligned}
&u \in H^1_{xt}(Q) & u &= u_+ \quad \text{on } \Gamma_+ \\
&\boldsymbol{\sigma} \in \boldsymbol{H}(\mathrm{div}, Q) & (\boldsymbol{\beta} u - \epsilon \nabla u) \cdot \boldsymbol{n} &= t_- \quad \text{on } \Gamma_- \\
&\left(\frac{1}{\epsilon}\boldsymbol{\sigma}, \boldsymbol{\tau}\right) - (\nabla u, \boldsymbol{\tau}) & &= 0 \quad \forall \boldsymbol{\tau} \in \mathbf{L}^2(Q) \\
&\left(\nabla_{xt} \cdot \begin{pmatrix} \boldsymbol{\beta} u - \boldsymbol{\sigma} \\ u \end{pmatrix}, v\right) & &= f \quad \forall v \in L^2(Q)\,,
\end{aligned}
\tag{2}
$$

which is L^2-equivalent to the strong form.

We can now choose either to relax (integrate by parts and build in the boundary conditions) or strongly enforce each equation. The four resulting options are explored and analyzed in further detail in [4] and are termed the trivial formulation (don't relax anything), the classical formulation (relax the second equation), the mixed formulation (relax the first equation), and the ultra-weak formulation (relax both equations). The stability constants for the four formulations are related, but the functional settings and norms of convergence change. Early DPG work emphasized the ultra-weak formulation since in many ways it was the easiest to analyze, though recently the classical formulation has been under very active consideration. In the interests of simpler analysis, we focus on the ultra-weak formulation in this paper.

$$
\begin{aligned}
&u \in L^2(Q)\,,\ \boldsymbol{\sigma} \in \mathbf{L}^2(Q) \\
&\left(\frac{1}{\epsilon}\boldsymbol{\sigma}, \boldsymbol{\tau}\right) + (u, \nabla \cdot \boldsymbol{\tau}) &&= 0 \quad \forall \boldsymbol{\tau} \in \boldsymbol{H}(\mathrm{div}, Q) : \boldsymbol{\tau} \cdot \boldsymbol{n}_x = 0 \text{ on } \Gamma_- \\
&-\left(\begin{pmatrix} \boldsymbol{\beta} u - \boldsymbol{\sigma} \\ u \end{pmatrix}, \nabla_{xt} v\right) &&= f \quad \forall v \in H^1_{xt}(Q) : v = 0 \text{ on } \Gamma_+\,,
\end{aligned}
\tag{3}
$$

We can remove the conditions on the test functions by introducing trace unknowns

$$
\begin{aligned}
\hat{u} &= \mathrm{tr}(u) && \text{on } \partial Q_x \\
\hat{t} &= \mathrm{tr}\begin{pmatrix} \boldsymbol{\beta} u - \boldsymbol{\sigma} \\ u \end{pmatrix} \cdot \boldsymbol{n}_{xt} && \text{on } \partial Q_{xt}\,.
\end{aligned}
$$

Our new ultra-weak formulation with continuous test functions is

$$
\begin{aligned}
&u \in L^2(Q)\,,\ \boldsymbol{\sigma} \in \mathbf{L}^2(Q) \\
&\hat{u} \in H^{1/2}(\partial Q)\,, && \hat{u} = u_+ \text{ on } \Gamma_+ \\
&\hat{t} \in H^{-1/2}_{xt}(\partial Q)\,, && \hat{t} = t_- \text{ on } \Gamma_-\,, \quad \hat{t} = -u_0 \text{ on } \Gamma_0 \\
&\left(\frac{1}{\epsilon}\boldsymbol{\sigma}, \boldsymbol{\tau}\right) + (u, \nabla \cdot \boldsymbol{\tau}) - \langle \hat{u}, \boldsymbol{\tau} \cdot \boldsymbol{n}_x \rangle &&= 0 \quad \forall \boldsymbol{\tau} \in \boldsymbol{H}(\mathrm{div}, Q) \\
&-\left(\begin{pmatrix} \boldsymbol{\beta} u - \boldsymbol{\sigma} \\ u \end{pmatrix}, \nabla_{xt} v\right) + \langle \hat{t}, v \rangle &&= f \quad \forall v \in H^1_{xt}(Q)\,.
\end{aligned}
\tag{4}
$$

1.2.4 Broken Test Functions

One of the key insights that led to the development of the DPG framework was the process of breaking test functions, that is testing with functions from larger broken Sobolev spaces, replacing $H^1_{xt}(Q)$ with $H^1_{xt}(Q_h)$ and $\boldsymbol{H}(\mathrm{div}, Q)$ with $\boldsymbol{H}(\mathrm{div}, Q_h)$. Discretizing such spaces is much simpler than standard spaces which require enforcement of global continuity conditions. The cost of introducing broken spaces is that we have to extend our interface unknowns $\hat{u}$ and $\hat{t}$ to live on the mesh skeleton. Our ultra-weak formulation with broken test functions looks like

$$
\begin{aligned}
&u \in L^2(Q)\,,\ \boldsymbol{\sigma} \in \mathbf{L}^2\,(Q) && \\
&\hat{u} \in H^{1/2}(Q_h)\,, && \hat{u} = u_+ \text{ on } \Gamma_+ \\
&\hat{t} \in H^{-1/2}_{xt}(Q_h)\,, && \hat{t} = t_- \text{ on } \Gamma_-\,, \quad \hat{t} = -u_0 \text{ on } \Gamma_0 \\
&\left(\frac{1}{\epsilon}\boldsymbol{\sigma}, \boldsymbol{\tau}\right) + (u, \nabla\cdot\boldsymbol{\tau}) - \langle \hat{u}, \boldsymbol{\tau}\cdot\boldsymbol{n}_x\rangle && = 0 \quad \forall \boldsymbol{\tau} \in \boldsymbol{H}(\mathrm{div}, Q) \\
&-\left(\begin{pmatrix} \boldsymbol{\beta} u - \boldsymbol{\sigma} \\ u \end{pmatrix}, \nabla_{xt} v\right) + \langle \hat{t}, v\rangle && = f \quad \forall v \in H^1_{xt}(Q)\,.
\end{aligned}
\tag{5}
$$

The main consequence of breaking test functions is that it reduces the cost of solving for optimal test functions from a global solve to an embarrassingly parallel local solve element by element. Now that we've derived a suitable variational formulation, we are left with the task of selecting a test norm from which to compute our optimal test functions.

1.3 Robust Test Norms

The final unresolved choice is what norm to apply to the V space. This is one of the most important factors in designing a robust DPG method as the corresponding Riesz operator needs to be inverted to solve for the optimal test functions. If the norm produces unresolved boundary layers in the auxiliary problem, then many of the attractive features of DPG may fall apart. This is in fact the primary emphasis of this paper. The problem of constructing stable test norms for steady convection-diffusion was addressed in [5, 6]. In this paper, we extend that work to transient convection-diffusion in space-time.

We define a robust test norm such that the L^2 norm of the solution is bounded by the energy norm of the solution up to a constant independent of ϵ. We can rewrite any ultra-weak formulation with broken test functions as the following bilinear form with group variables:

$$
b\left((u, \hat{u}), v\right) = (u, A^* v)_{L^2} + \langle \widehat{u}, [\![v]\!]\rangle_{\Gamma_h}
$$

where A^* represents the adjoint. In the case of convection-diffusion, $u := \{u, \boldsymbol{\sigma}\}$, $\hat{u} := \left\{\hat{u}, \hat{t}\right\}$, $v := \{v, \boldsymbol{\tau}\}$.

Note that for conforming v^* satisfying $A^* v^* = u$

$$
\begin{aligned}
\|u\|^2_{L^2} = b(u, v^*) &= \frac{b(u, v^*)}{\|v^*\|_V}\,\|v^*\|_V \\
&\le \sup_{v^* \neq 0} \frac{|b(u, v^*)|}{\|v^*\|}\,\|v^*\| = \|u\|_E\,\|v^*\|_V
\end{aligned}
$$

This defines a necessary condition for robustness, namely that

$$
\|v^*\|_V \lesssim \|u\|_{L^2}\ . \tag{6}
$$

If this condition is satisfied, then we get our final result:

$$
\|u\|_{L^2} \lesssim \|u\|_E\ .
$$

So far, we've assumed that our finite set of optimal test functions are assembled from an infinite dimensional space. In practice, we have found it to be sufficient to use an "enriched" space of higher polynomial dimension than the trial space [7]. This adds an addition requirement when assembling a robust test norm, namely that our optimal test functions should be adequately representable within this enriched space. We illustrate this point by considering two norms which satisfy the above conditions for 1D steady convection-diffusion. The graph norm is $\|A^* v\|_{L^2} + \|v\|_{L^2}$:

$$
\|(v, \boldsymbol{\tau})\|^2 = \|\nabla\cdot\boldsymbol{\tau} - \boldsymbol{\beta}\cdot\nabla v\|^2 + \left\|\frac{1}{\epsilon}\boldsymbol{\tau} + \nabla v\right\|^2 + \|v\|^2 + \|\boldsymbol{\tau}\|^2\ .
$$

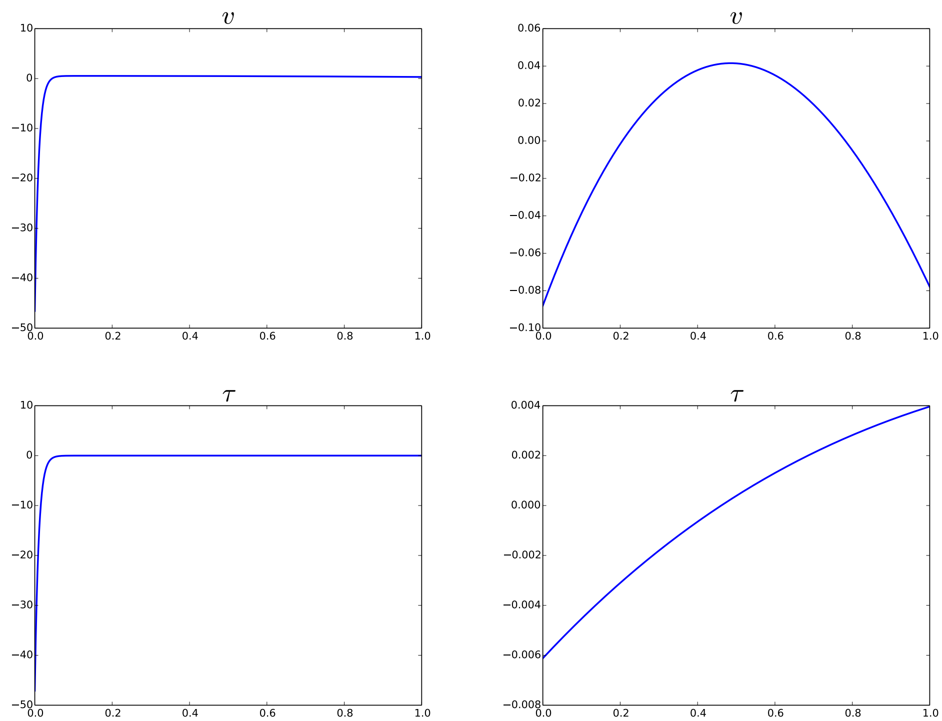
The coupled robust norm is a slight modification of the norm derived in [6] (for further details on the modification, see [8]):

$$\|(v, \boldsymbol{\tau})\|^2 = \|\nabla \cdot \boldsymbol{\tau} - \boldsymbol{\beta} \cdot \nabla v\|^2 + \min\left(\frac{1}{h^2}, \frac{1}{\epsilon}\right) \|\boldsymbol{\tau}\|^2 + \epsilon \|\nabla v\|^2 + \|\boldsymbol{\beta} \cdot \nabla v\|^2 + \|v\|^2 .$$

The bilinear form and test norm define a mapping from input trial functions to an optimal test function:

$$T = R_V^{-1} B : U \to V .$$

Below, we plot the optimal test functions produced given a representative trial function $u = x - \frac{1}{2}$, steady 1D ultra-weak convection-diffusion, and either the graph norm or the coupled robust norm. Note that the optimal test functions will be different for any other trial function. In the left column, we see the fully resolved *ideal* optimal test function that DPG theory relies on. On the right, we see the approximated optimal test function using a enriched cubic test space.

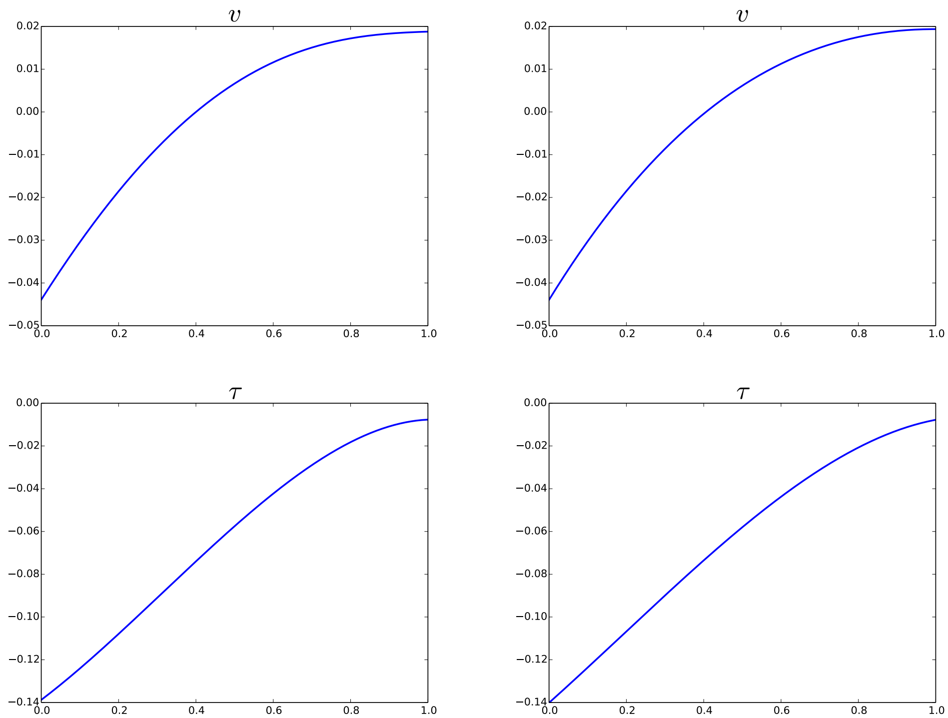


(a) Ideal

(b) Approximated

Figure 1: Graph norm optimal test functions for $u = x - \frac{1}{2}$

Mathematically, the graph norm satisfies the necessary condition to be a robust norm, but the ideal optimal test functions contain strong boundary layers which can not be realistically approximated with the provided enriched space. If the approximated optimal test functions can not come sufficiently close to the ideal, then the whole DPG theory falls apart. See [7] for more discussion. This provides an additional condition on a test norm before we can truly call it robust: the ideal test functions must be adequately representable within the provided enriched space. This ultimately comes down to an analysis of the relative magnitudes of individual terms within the test norm, usually attempting to bound reactive or convective terms by diffusive terms. The coupled robust norm satisfies 6 and also produces relatively smooth optimal test functions that can be sufficiently approximated with a cubic polynomial space.



(a) Ideal (b) Approximated

Figure 2: Coupled robust norm optimal test functions for $u = x - \frac{1}{2}$

2 A Robust Norm for Transient Convection-Diffusion

Consider transient convection-diffusion with homogeneous boundary conditions

$$\begin{aligned}
\frac{1}{\epsilon}\boldsymbol{\sigma} - \nabla u &= 0\\
\frac{\partial u}{\partial t} + \boldsymbol{\beta}\cdot\nabla u - \nabla\cdot\boldsymbol{\sigma} &= f\\
\beta_n u - \epsilon\frac{\partial u}{\partial n} &= 0 \text{ on } \Gamma_-\\
u &= 0 \text{ on } \Gamma_+\\
u &= u_0 \text{ on } \Gamma_0,
\end{aligned}$$

where $\nabla\cdot\boldsymbol{\beta} = 0$ and $\|\nabla\boldsymbol{\beta}\|_{L^\infty} \leq C_{\boldsymbol{\beta}}$. Let $\tilde{\boldsymbol{\beta}} := \begin{pmatrix}\boldsymbol{\beta}\\ 1\end{pmatrix}$, then we can rewrite this as

$$\begin{aligned}
\frac{1}{\epsilon}\boldsymbol{\sigma} - \nabla u &= 0\\
\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt} u - \nabla\cdot\boldsymbol{\sigma} &= f\\
\beta_n u - \epsilon\frac{\partial u}{\partial n} &= 0 \text{ on } \Gamma_-\\
u &= 0 \text{ on } \Gamma_+\\
u &= u_0 \text{ on } \Gamma_0.
\end{aligned}$$

We decompose the adjoint into three parts: a discontinuous part (Dr. Demkowicz, what kind of argument do we want to use to avoid dealing with this part? I assume it's something to do with breaking test functions.)

$$
\begin{aligned}
\frac{1}{\epsilon}\boldsymbol{\tau}_0+\nabla v_0&=0\\
-\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v_0+\nabla\cdot\boldsymbol{\tau}_0&=0\\
\boldsymbol{\tau}_0\cdot\boldsymbol{n}_x&=\boldsymbol{\tau}\cdot\boldsymbol{n}_x\text{ on }\Gamma_-\cup\Gamma_0\\
v_0&=v\text{ on }\Gamma_+\\
v_0&=v\text{ on }\Gamma_T\\
[\![v_0]\!]&=[\![v]\!]\text{ on }\Gamma_h^0\\
[\![\boldsymbol{\tau}_0\cdot\boldsymbol{n}_x]\!]&=[\![\boldsymbol{\tau}_0\cdot\boldsymbol{n}_x]\!]\text{ on }\Gamma_{hx}^0\,,
\end{aligned}
$$

a continuous part with forcing term g

$$
\begin{aligned}
\frac{1}{\epsilon}\boldsymbol{\tau}_1+\nabla v_1&=0\\
-\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v_1+\nabla\cdot\boldsymbol{\tau}_1&=g\\
\boldsymbol{\tau}_1\cdot\boldsymbol{n}_x&=0\text{ on }\Gamma_-\\
v_1&=0\text{ on }\Gamma_+\\
v_1&=0\text{ on }\Gamma_T\,,
\end{aligned}
$$

and a continuous part with forcing $\boldsymbol{f}$

$$
\begin{aligned}
\frac{1}{\epsilon}\boldsymbol{\tau}_2+\nabla v_2&=\boldsymbol{f}\\
-\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v_2+\nabla\cdot\boldsymbol{\tau}_2&=0\\
\boldsymbol{\tau}_2\cdot\boldsymbol{n}_x&=0\text{ on }\Gamma_-\\
v_2&=0\text{ on }\Gamma_+\\
v_2&=0\text{ on }\Gamma_T\,.
\end{aligned}
$$

(The boundary conditions can be derived by taking the ultra-weak formulation and choosing boundary conditions such that the temporal flux and spatial flux terms $\langle\widehat{u},[\![\tau_n]\!]\rangle_{\Gamma_{out}}$ and $\langle\hat{t}_n,[\![v]\!]\rangle_{\Gamma_{in}}$ are zero.)

We can then derive that the test norm

$$
\begin{aligned}
\|(v,\boldsymbol{\tau})\|_{V,K}^2\coloneqq\frac{1}{\epsilon}\|\boldsymbol{\tau}\|_K^2+&\left\|\nabla\cdot\boldsymbol{\tau}-\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\right\|_K^2\\
&+\|\boldsymbol{\beta}\cdot\nabla v\|_K^2+\epsilon\|\nabla v\|_K^2+\|v\|_K^2\,,
\end{aligned}
\tag{7}
$$

provides the necessary bound $\|v^*\|_V\lesssim\|u\|_{L^2(Q)}$.

In the following lemmas we establish the following bounds:

- Bound on $\|(v_0,\boldsymbol{\tau}_0)\|_V$. What do we do for this?
- Bound on $\|(v_1,\boldsymbol{\tau}_1)\|_V$. Lemma 2.1 gives $\left\|\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v_1\right\|\leq\|g\|$. Since $\nabla\cdot\boldsymbol{\tau}_1=g+\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v_1$,

  $$\|\nabla\cdot\boldsymbol{\tau}_1\|\leq\|g\|+\left\|\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v_1\right\|\leq 2\|g\|\,.$$

  Or, the fact that $\nabla\cdot\boldsymbol{\tau}-\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v_1=g$ clearly gives

  $$\left\|\nabla\cdot\boldsymbol{\tau}-\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v_1\right\|=\|g\|\,.$$

  Also, clearly

  $$\|\boldsymbol{\beta}\cdot\nabla v_1\|\leq\left\|\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v_1\right\|\leq\|g\|\,.$$

Lemma 2.2 gives $\|v_1\|^2 + \epsilon \|\nabla v_1\|^2 \le \|g\|^2$. Since $\epsilon^{1/2}\nabla v_1 = -\epsilon^{-1/2}\boldsymbol{\tau}_1$,

$$\frac{1}{\epsilon}\|\boldsymbol{\tau}_1\|^2 \le \|g\|^2 .$$

Thus, all $(v_1, \boldsymbol{\tau}_1)$ terms in (7) are accounted for.

- Bound on $\|(v_2, \boldsymbol{\tau}_2)\|_V$. The fact that $\nabla \cdot \boldsymbol{\tau} - \tilde{\boldsymbol{\beta}} \cdot \nabla_{xt} v = 0$ clearly gives

  $$\left\|\nabla \cdot \boldsymbol{\tau} - \tilde{\boldsymbol{\beta}} \cdot \nabla_{xt} v_2\right\| = 0 \le \|\boldsymbol{f}\| .$$

  Lemma 2.2 gives $\|v_2\|^2 + \epsilon \|\nabla v_2\|^2 \le \epsilon \|\boldsymbol{f}\|^2$. Since $\epsilon^{1/2}\nabla v_2 = \boldsymbol{f} - \epsilon^{-1/2}\boldsymbol{\tau}_2$,

  $$\frac{1}{\epsilon}\|\boldsymbol{\tau}_2\|^2 \le (1+\epsilon)\|\boldsymbol{f}\|^2 .$$

  Finally,

  $$\|\boldsymbol{\beta} \cdot \nabla v_2\| \le \|\boldsymbol{\beta}\|_\infty \|\nabla v_2\| \le \|\boldsymbol{\beta}\|_\infty \|\boldsymbol{f}\| .$$

  Thus, all $(v_2, \boldsymbol{\tau}_2)$ terms in (7) are accounted for.

Our goal is to analyze the stability properties of the adjoint equations by deriving bounds of the form $\|(v_1, \boldsymbol{\tau}_1)\|_V \le \|g\|^2_L(Q)$ and $\|(v_2, \boldsymbol{\tau}_2)\|_V \le \|f\|^2_L(Q)$.

Lemma 2.1. *We can bound*

$$\left\|\tilde{\boldsymbol{\beta}} \cdot \nabla_{xt} v_1\right\| \le \|g\| .$$

Proof. Multiply $-\tilde{\boldsymbol{\beta}} \cdot \nabla_{xt} v = g - \nabla \cdot \boldsymbol{\tau}$ by $-\tilde{\boldsymbol{\beta}} \cdot \nabla_{xt} v$ and integrate over Q to get

$$\left\|\tilde{\boldsymbol{\beta}} \cdot \nabla_{xt} v\right\|^2 = -\int_Q g\tilde{\boldsymbol{\beta}} \cdot \nabla_{xt} v + \int_Q \tilde{\boldsymbol{\beta}} \cdot \nabla_{xt} v \nabla \cdot \boldsymbol{\tau} . \tag{8}$$

Note that

$$
\begin{aligned}
\frac{1}{\epsilon}\int_Q \tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\nabla\cdot\boldsymbol{\tau} &= -\int_Q \tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\nabla\cdot\nabla v \\
&= -\int_{\Gamma_x} \tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\nabla v\cdot\boldsymbol{n}_x + \int_Q \nabla(\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v)\cdot\nabla v \\
&= -\int_{\Gamma_x} \tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\nabla v\cdot\boldsymbol{n}_x + \int_Q (\nabla\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v)\cdot\nabla v \\
&\quad + \int_Q \tilde{\boldsymbol{\beta}}\cdot\nabla\nabla_{xt}v\cdot\nabla v \\
&= -\int_{\Gamma_x} \tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\nabla v\cdot\boldsymbol{n}_x + \int_Q (\nabla\boldsymbol{\beta}\cdot\nabla v)\cdot\nabla v \\
&\quad + \frac{1}{2}\int_Q \tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}(\nabla v\cdot\nabla v) \\
&= -\int_{\Gamma_x} \tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\nabla v\cdot\boldsymbol{n}_x + \int_Q (\nabla\boldsymbol{\beta}\cdot\nabla v)\cdot\nabla v \\
&\quad + \frac{1}{2}\int_\Gamma \tilde{\boldsymbol{\beta}}\cdot\boldsymbol{n}(\nabla v\cdot\nabla v) - \frac{1}{2}\int_Q \nabla_{xt}\cdot\tilde{\boldsymbol{\beta}}(\nabla v\cdot\nabla v) \\
&= -\int_{\Gamma_x} \tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\nabla v\cdot\boldsymbol{n}_x + \int_Q (\nabla\boldsymbol{\beta}\cdot\nabla v)\cdot\nabla v \\
&\quad + \frac{1}{2}\int_\Gamma \tilde{\boldsymbol{\beta}}\cdot\boldsymbol{n}(\nabla v\cdot\nabla v) - \frac{1}{2}\int_Q \nabla\cdot\boldsymbol{\beta}(\nabla v\cdot\nabla v) \\
&= -\int_{\Gamma_x} \tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\nabla v\cdot\boldsymbol{n}_x + \frac{1}{2}\int_\Gamma \tilde{\boldsymbol{\beta}}\cdot\boldsymbol{n}(\nabla v\cdot\nabla v) \\
&\quad + \int_Q \nabla v(\nabla\boldsymbol{\beta} - \frac{1}{2}\nabla\cdot\boldsymbol{\beta}\boldsymbol{I})\nabla v
\end{aligned}
$$

Plugging this into (8), we get

$$
\begin{aligned}
\left\|\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\right\|^2 &= -\int_Q g\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v+\epsilon\int_Q\nabla v(\nabla\boldsymbol{\beta}-\frac{1}{2}\nabla\cdot\boldsymbol{\beta}\boldsymbol{I})\nabla v\\
&\quad-\epsilon\int_{\Gamma_x}\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\nabla v\cdot\boldsymbol{n}_x+\frac{\epsilon}{2}\int_\Gamma\tilde{\boldsymbol{\beta}}\cdot\boldsymbol{n}(\nabla v\cdot\nabla v)\\
&= -\int_Q g\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v+\epsilon\int_Q\nabla v(\nabla\boldsymbol{\beta}-\frac{1}{2}\nabla\cdot\boldsymbol{\beta}\boldsymbol{I})\nabla v\\
&\quad-\epsilon\int_{\Gamma_-}\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\underbrace{\nabla v\cdot\boldsymbol{n}_x}_{=0}-\epsilon\int_{\Gamma_+}\left(\underbrace{\frac{\partial v}{\partial t}}_{=0}+\boldsymbol{\beta}\cdot\nabla v\right)\nabla v\cdot\boldsymbol{n}_x\\
&\quad+\frac{\epsilon}{2}\int_{\Gamma_-}\underbrace{\boldsymbol{\beta}\cdot\boldsymbol{n}_x}_{<0}(\nabla v\cdot\nabla v)+\frac{\epsilon}{2}\int_{\Gamma_+}\boldsymbol{\beta}\cdot\boldsymbol{n}_x(\nabla v\cdot\nabla v)\\
&\quad+\frac{\epsilon}{2}\int_{\Gamma_0}\underbrace{n_t}_{<0}(\nabla v\cdot\nabla v)+\frac{\epsilon}{2}\int_{\Gamma_T}n_t\underbrace{(\nabla v\cdot\nabla v)}_{=0}\\
&\le -\int_Q g\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v+\epsilon\int_Q\nabla v(\nabla\boldsymbol{\beta}-\frac{1}{2}\nabla\cdot\boldsymbol{\beta}\boldsymbol{I})\nabla v\\
&\quad+\epsilon\int_{\Gamma_+}\left(-\frac{\partial v}{\partial\boldsymbol{n}_x}\boldsymbol{\beta}+\frac{1}{2}\boldsymbol{\beta}\cdot\boldsymbol{n}_x\nabla v\right)\cdot\nabla v\\
&= -\int_Q g\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v+\epsilon\int_Q\nabla v(\nabla\boldsymbol{\beta}-\frac{1}{2}\nabla\cdot\boldsymbol{\beta}\boldsymbol{I})\nabla v\\
&\quad+\epsilon\int_{\Gamma_+}\left(-\frac{\partial v}{\partial\boldsymbol{n}_x}\boldsymbol{\beta}+\frac{1}{2}\boldsymbol{\beta}\cdot\boldsymbol{n}_x\frac{\partial v}{\partial\boldsymbol{n}_x}\boldsymbol{n}_x\right)\cdot\frac{\partial v}{\partial\boldsymbol{n}_x}\boldsymbol{n}_x\\
&= -\int_Q g\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v+\epsilon\int_Q\nabla v(\nabla\boldsymbol{\beta}-\frac{1}{2}\nabla\cdot\boldsymbol{\beta}\boldsymbol{I})\nabla v\\
&\quad\underbrace{-\frac{\epsilon}{2}\int_{\Gamma_+}\left(\frac{\partial v}{\partial\boldsymbol{n}_x}\right)^2\boldsymbol{\beta}\cdot\boldsymbol{n}_x}_{<0}\\
&\le -\int_Q g\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v+\epsilon\int_Q\nabla v(\nabla\boldsymbol{\beta}-\frac{1}{2}\nabla\cdot\boldsymbol{\beta}\boldsymbol{I})\nabla v\\
&\le\frac{\|g\|^2}{2}+\frac{\left\|\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\right\|^2}{2}+\epsilon\int_Q\nabla v(\nabla\boldsymbol{\beta}-\frac{1}{2}\nabla\cdot\boldsymbol{\beta}\boldsymbol{I})\nabla v\\
&\le\frac{\|g\|^2}{2}+\frac{\left\|\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\right\|^2}{2}+\epsilon C_{\boldsymbol{\beta}}\|\nabla v\|^2
\end{aligned}
$$

Seems to additionally require that $\epsilon C_{\boldsymbol{\beta}}\|\nabla v\|^2\le\left\|\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}v\right\|^2$. □

Lemma 2.2. *For the duration of this lemma, let $v := v_1+v_2$. Then, for the above conditions on $\boldsymbol{\beta}$,*

$$\|v\|^2+\epsilon\|\nabla v\|^2\le\|g\|^2+\epsilon\|\boldsymbol{f}\|^2\,.$$

Proof. Define $w=e^t v$ and note that $\frac{\partial w}{\partial t}=\left(\frac{\partial v}{\partial t}+v\right)e^t$ while all spatial derivatives go through. Multiplying the adjoint by w and integrating over Q gives

$$-\int_Q\tilde{\boldsymbol{\beta}}\cdot\nabla_{xt}vw-\epsilon\Delta vw=\int_Q gw-\epsilon\int_Q\nabla\cdot\boldsymbol{f}w$$

or

$$-\int_Q e^t v\tilde{\boldsymbol{\beta}} \cdot \nabla_{xt} v - \epsilon \int_Q e^t v \Delta v = \int_Q e^t g v - \epsilon \int_Q e^t v \nabla \cdot \boldsymbol{f}$$

Integrating by parts:

$$\begin{aligned}\int_Q \nabla_{xt} \cdot \left(e^t \tilde{\boldsymbol{\beta}} v\right) v - \int_\Gamma e^t \tilde{\boldsymbol{\beta}} \cdot \boldsymbol{n} v^2 + \epsilon \int_Q e^t \nabla v \cdot \nabla v - \epsilon \int_{\Gamma_x} e^t v \cdot \nabla v \cdot \boldsymbol{n}_x \\ = \int_Q e^t g v + \epsilon \int_Q e^t \nabla v \cdot \boldsymbol{f} - \epsilon \int_{\Gamma_x} e^t v \boldsymbol{f} \cdot \boldsymbol{n}_x\end{aligned}$$

Note that $\nabla_{xt} \cdot e^t v \tilde{\boldsymbol{\beta}} = e^t(\tilde{\boldsymbol{\beta}} \cdot \nabla_{xt} v + v)$ if $\nabla \cdot \boldsymbol{\beta} = 0$. Moving some terms to the right hand side, we get

$$\begin{aligned}&\int_Q e^t v^2 + \int_Q \epsilon e^t \nabla v \cdot \nabla v \\ &\quad = \int_Q e^t g v + \epsilon \int_Q e^t \nabla v \cdot \boldsymbol{f} - \epsilon \int_{\Gamma_x} e^t v \boldsymbol{f} \cdot \boldsymbol{n}_x \\ &\qquad - \int_Q e^t \tilde{\boldsymbol{\beta}} \cdot \nabla_{xt} v v \quad + \int_\Gamma e^t \tilde{\boldsymbol{\beta}} \cdot \boldsymbol{n} v^2 \quad + \epsilon \int_{\Gamma_x} e^t v \cdot \nabla v \cdot \boldsymbol{n}_x\end{aligned}$$

Note that $1 \leq \|e^t\|_\infty = e^T$, then

$$\begin{aligned}&\|v\|^2 + \epsilon \|\nabla v\|^2 \\ &\leq e^T \Bigg(\int_Q g v + \epsilon \int_Q \nabla v \cdot \boldsymbol{f} - \epsilon \int_{\Gamma_-} v \underbrace{\boldsymbol{f} \cdot \boldsymbol{n}_x}_{= \cancelto{0}{\tau_n} + \frac{\partial v}{\partial \boldsymbol{n}_x}} - \epsilon \int_{\Gamma_+} \underbrace{v}_{=0} \boldsymbol{f} \cdot \boldsymbol{n}_x \\ &\quad - \int_Q \tilde{\boldsymbol{\beta}} \cdot \nabla_{xt} v v \quad + \int_\Gamma \tilde{\boldsymbol{\beta}} \cdot \boldsymbol{n} v^2 \quad + \epsilon \int_{\Gamma_-} v \cdot \nabla v \cdot \boldsymbol{n}_x \quad + \epsilon \int_{\Gamma_+} \underbrace{v}_{=0} \frac{\partial v}{\partial \boldsymbol{n}_x} \Bigg) \\ &= e^T \Bigg(\int_Q g v + \epsilon \int_Q \nabla v \cdot \boldsymbol{f} \cancel{- \epsilon \int_{\Gamma_-} v \frac{\partial v}{\partial \boldsymbol{n}_x} \quad + \epsilon \int_{\Gamma_x} v \frac{\partial v}{\partial \boldsymbol{n}_x}} \\ &\quad - \frac{1}{2} \int_Q \tilde{\boldsymbol{\beta}} \cdot \nabla_{xt} v^2 \quad + \int_\Gamma \tilde{\boldsymbol{\beta}} \cdot \boldsymbol{n} v^2 \Bigg) \\ &= e^T \left(\int_Q g v + \epsilon \int_Q \nabla v \cdot \boldsymbol{f} \right) \\ &\quad + \frac{1}{2} \int_Q \cancelto{0}{\nabla_{xt} \cdot \tilde{\boldsymbol{\beta}} v^2} \quad - \frac{1}{2} \int_\Gamma \tilde{\boldsymbol{\beta}} \cdot \boldsymbol{n} v^2 \quad + \int_\Gamma \tilde{\boldsymbol{\beta}} \cdot \boldsymbol{n} v^2 \Bigg) \\ &= e^T \Bigg(\int_Q g v + \epsilon \int_Q \nabla v \cdot \boldsymbol{f} \\ &\quad + \frac{1}{2} \left(\int_{\Gamma_0} \underbrace{-v^2}_{\leq 0} \quad + \int_{\Gamma_T} \cancelto{0}{v^2} + \int_{\Gamma_-} \underbrace{\boldsymbol{\beta} \cdot \boldsymbol{n}_x v^2}_{\leq 0} \quad + \int_{\Gamma_+} \boldsymbol{\beta} \cdot \boldsymbol{n}_x \cancelto{0}{v^2} \right) \Bigg) \\ &\leq e^T \left(\int_Q g v + \epsilon \int_Q \nabla v \cdot \boldsymbol{f} \right) \\ &\leq e^T \left(\frac{\|g\|^2}{2} + \epsilon \frac{\|\boldsymbol{f}\|^2}{2} + \frac{\|v\|^2}{2} + \epsilon \frac{\|\nabla v\|^2}{2} \right)\end{aligned}$$

$\square$

References

[1] L. Demkowicz, J. Gopalakrishnan, Recent Developments in Discontinuous Galerkin Finite Element Methods for Partial Differential Equations (eds. X. Feng, O. Karakashian, Y. Xing), Vol. 157, IMA Volumes in Mathematics and its Applications, 2014, Ch. An Overview of the DPG Method, pp. 149–180.

[2] I. Babuška, Error-bounds for finite element method, Numer. Math 16.

[3] C. Carstensen, L. D. J. Gopalakrishnan, Breaking spaces and forms for the DPG method and applications including Maxwell equations, Tech. rep., ICES (2015).

[4] L. Demkowicz, Various variational formulations and closed range theorem, Tech. rep., ICES (Jan. 2015).

[5] L. Demkowicz, N. Heuer, Robust DPG method for convection-dominated diffusion problems, SIAM J. Numer. Anal. 51 (5) (2013) 1514–2537.

[6] J. Chan, N. Heuer, T. Bui-Thanh, L. Demkowicz, A robust DPG method for convection-dominated diffusion problems II: Adjoint boundary conditions and mesh-dependent test norms, Comp. Math. Appl. 67 (4) (2014) 771 – 795, high-order Finite Element Approximation for Partial Differential Equations.

[7] J. Gopalakrishnan, W. Qiu, An analysis of the practical DPG method, Tech. rep., IMA, submitted (2011).

[8] J. Chan, A DPG method for convection-diffusion problems, Ph.D. thesis, University of Texas at Austin (2013).