ICES REPORT 10-01

January 2010

A Class of Discontinuous P Etrov-Galerkin Methods. Part III: Adaptivity

by

L. Demkowicz, J. Gopalakrishnan, A. Niemi

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

A CLASS OF DISCONTINUOUS PETROV-GALERKIN METHODS. Part III: ADAPTIVITY

L. Demkowicz[a], J. Gopalakrishnan[b], A. Niemi[a]

[a]**Institute for Computational Engineering and Sciences**
The University of Texas at Austin, Austin, TX 78712, USA
[b]**Department of Mathematics**
University of Florida, FL 32611, USA

Abstract

We continue our theoretical and numerical study on the Discontinuous Petrov Galerkin method with optimal test functions in context of 1D and 2D convection-dominated problems and hp-adaptivity. With a proper choice of norm for the test space, we prove robustness (uniform stability with respect to diffusion constant) and mesh-independence of the energy norm of the FE error for the 1D problem. With hp-adaptivity and a proper scaling of the norms for the test functions, we establish new limits for solving the convection-dominated problems numerically: $\epsilon = 10^{-11}$ for 1D and $\epsilon = 10^{-7}$ for 2D problems. The adaptive process is fully automatic and it starts with a mesh consisting of few elements only.

Key words: convection-dominated diffusion, hp-adaptivity, Discontinuous Petrov Galerkin

AMS subject classification: 65N30, 35L15

Acknowledgment

Demkowicz was supported in part by the Department of Energy [National Nuclear Security Administration] under Award Number [DE-FC52-08NA28615], and by a research contract with Boeing. Gopalakrishnan was supported in part by the National Science Foundation under grant DMS-0713833. Niemi was supported in part by KAUST. We thank Bob Moser and David Young for encouragement and stimulating discussions on the project.

1 Introduction

Petrov-Galerkin Method with Optimal Test Functions

Variational problem and energy norm. Consider an arbitrary abstract variational problem,

$$\begin{cases} u \in U \\ b(u,v) = l(v) \quad \forall v \in V \end{cases} \tag{1.1}$$

Here U, V are two real Hilbert spaces, and $b(u,v)$ is a continuous bilinear form on $U \times V$,

$$|b(u,v)| \leq M \|u\|_U \, \|v\|_V \tag{1.2}$$

that satisfies the inf-sup condition:

$$\inf_{\|u\|_U=1} \sup_{\|v\|_V=1} |b(u,v)| =: \gamma > 0 \tag{1.3}$$

Continuous functional $l \in V'$ represents the load. If the null space of the adjoint operator,

$$V_0 := \{v \in V \, : \, b(u,v) = 0 \quad \forall u \in U\} \tag{1.4}$$

is non-trivial, functional l is assumed to satisfy the compatibility condition:

$$l(v) = 0 \quad \forall v \in V_0 \tag{1.5}$$

By the Banach Closed-Range Theorem (see e.g. [11]), problem (1.1) possesses a unique solution u that depends continuously upon the data - functional l. More precisely,

$$\|u\|_U \leq \frac{1}{\gamma} \|l\|_{V'} \tag{1.6}$$

If we define an alternative, *energy norm*, equivalent to the original norm on U,

$$\|u\|_E := \sup_{\|v\|=1} |b(u,v)| \tag{1.7}$$

both continuity and inf-sup constants become equal one. Recalling that the Riesz operator,

$$R \, : \, V \ni v \rightarrow (v, \cdot) \in V' \tag{1.8}$$

is an isometry from V into its dual V', we may characterize the energy norm in an equivalent way as

$$\|u\|_E = \|v_u\|_V \tag{1.9}$$

where v_u is the solution of the variational problem,

$$\begin{cases} v_u \in V \\ (v_u, \delta v)_V = b(u, \delta v) \quad \forall \delta v \in V \end{cases} \tag{1.10}$$

with $(\cdot, \cdot)_V$ denoting the inner product in test space V. If the original norm in U is replaced with the (problem-dependent) energy norm, both continuity and inf-sup constants are equal one.

Optimal test functions. Let $U_{hp} \subset U$ be a finite-dimensional space with a basis $e_j = e_{j,hp}, j = 1, \ldots, N = N_{hp}$. For each basis *trial* function e_j, we introduce a corresponding *optimal test (basis) function* $\bar{e}_j \in V$ that realizes the supremum,

$$|b(e_j, \bar{e}_j)| = \sup_{\|v\|_V = 1} |b(e_j, v)| \tag{1.11}$$

i.e. it solves the variational problem,

$$\begin{cases} \bar{e}_j \in V \\ (\bar{e}_j, \delta v)_V = b(e_j, \delta v) \quad \forall \delta v \in V \end{cases} \tag{1.12}$$

The test space is now defined as the span of the optimal test functions, $\bar{V}_{hp} := \text{span}\{\bar{e}_j, j = 1, \ldots, N\} \subset V$. The Petrov-Galerkin discretization of problem (1.1) with the optimal test functions takes the form:

$$\begin{cases} u_{hp} \in U_{hp} \\ b(u_{hp}, v_{hp}) = l(v_{hp}) \quad \forall v_{hp} \in \bar{V}_{hp} \end{cases} \tag{1.13}$$

It follows from the construction of the optimal test functions that the *discrete* inf-sup constant

$$\inf_{\|u_{hp}\|_E = 1} \sup_{\|v_{hp}\| = 1} |b(u_{hp}, v_{hp})| \tag{1.14}$$

is also equal one. Consequently, Babuška's Theorem [1] implies that

$$\|u - u_{hp}\|_E \leq \inf_{w_{hp} \in U_{hp}} \|u - w_{hp}\|_E \tag{1.15}$$

i.e., the method delivers the *best approximation error* in the energy norm.

The construction of optimal test functions implies also that the global stiffness matrix is *symmetric and positive-definite*. Indeed,

$$b(e_i, \bar{e}_j) = (\bar{e}_i, \bar{e}_j)_V = (\bar{e}_j, \bar{e}_i)_V = b(e_j, \bar{e}_i) \tag{1.16}$$

Once the approximate solution has been determined, the energy norm of the FE error $e_{hp} := u - u_{hp}$ is computed by solving again the variational problem:

$$\begin{cases} v_{e_{hp}} \in V \\ (v_{e_{hp}}, \delta v)_V = b(u - u_{hp}, \delta v) = l(\delta v) - b(u_{hp}, \delta v) \quad \forall \delta v \in V \end{cases} \tag{1.17}$$

We have then

$$\|e_{hp}\|_E = \|v_{e_{hp}}\|_V \tag{1.18}$$

We shall call solution $v_{e_{hp}}$ to problem (1.17), *the error representation function*. Notice that we can compute the energy norm of the error *without knowing* the exact solution. Indeed, the energy norm of the error is nothing else than a properly defined norm of the residual.

Approximate optimal test functions. In practice, the optimal test functions $\bar{e}_j$ and the error representation function $v_{e_{hp}}$ are computed approximately using a sufficiently large *"enriched"* subspace $V_{hp} \subset V$. The optimal test space $\bar{V}_{hp}$ is then a subspace of the enriched test space V_{hp}. For instance, if elements of order p are used, the enriched test space may involve polynomials of order $p + \Delta p$. The approximate optimal test space forms a proper subspace of the enriched space and it should not be confused with the enriched space itself, though.

Discontinuous Petrov Galerkin Method

For conforming discretizations, approximate solution of variational problem (1.12),

$$\begin{cases} \bar{e}_j \in V_{hp} \\ (\bar{e}_j, \delta v)_V = b(e_j, \delta v) \quad \forall \delta v \in V_{hp} \end{cases} \tag{1.19}$$

leads to a global system of equations and, consequently, the whole discussed concept has little practical value. The situation changes, if the methodology is applied in the framework of *Discontinuous Petrov Galerkin* (DPG) method introduced by Bottaso, Micheletti, Sacco and Causin [2, 3, 4, 5]. The starting point of the method is a system of first order differential equations. The equations are multiplied with test functions, integrated over the domain and integrated by parts. The resulting boundary flux terms are *not* related to the field variables supported over the elements (concept of the *numerical flux*) but treated rather as *independent unknowns*. Fluxes known from boundary conditions are moved to the right-hand side and they contribute to the linear functional $l(v)$. Contrary to classical variational formulations where some of the equations are relaxed and other are treated in a strong form, in the DPG method *all equations* are treated in a weak sense. Formulations like this are sometimes termed as *ultra-weak variational formulations*.

The DPG framework of Bottaso, Micheletti et al. has been combined with the concept of optimal test functions in the first two papers of this series [7, 8]. [1] The approximate optimal test functions (and the energy error) are determined element-wise using higher order approximation. The global error (1.18) (squared) is equal to the sum of corresponding element contributions,

$$\|e_{hp}\|_E^2 = \|v_{e_{hp}}\|_V^2 = \sum_K \underbrace{\|v_{e_{hp}}\|_{V(K)}^2}_{e_K} \tag{1.20}$$

where the element error representation function $v_{e_{hp}}$ is computed by solving a local counterpart of (1.17) using the element enriched space. Element contributions e_K serve as element error indicators, and provide a basis for adaptivity.

Scope of this paper. The presented work is a continuation of [8] and focuses on convection-dominated diffusion problems. In Section 2 we present a new stability analysis for the 1D problem and show how

[1] Contrary to our work, Bottaso et al. have used approximate polynomial test spaces defined a-priori. Optimal test functions are non-polynomial.

a proper choice of the norm for the test functions leads to a *robust* method, i.e. a method with stability properties uniform in diffusion constant ϵ. We also prove a fundamental property of the error representation function - *global continuity*. The main focus of the paper though is automatic adaptivity. Section 3 presents numerical results for the 1D problem, and in Section 4 we share our 2D experience. With a clever choice of norms for the test functions we achieve not only the robustness but also have been able to push the existing limits for solving the convection-dominated diffusion problems with double precision arithmetic. The reported "records" are $\epsilon = 10^{-11}$ for 1D, and $\epsilon = 10^{-7}$ for 2D problems. Most importantly, the whole optimization process is completely automatic, and it starts with an initial mesh consisting of few elements only.

2 Convection-Dominated Diffusion in 1D

In this section, we continue our study on a 1D convection-dominated diffusion model problem initiated in [8]. We demonstrate that, with a proper choice of the norm for the test functions, we obtain a *robust* DPG method, i.e. a method whose stability properties are *uniform* in diffusion constant ϵ. More precisely, the standard L^2-norms of velocity u and stress $\sigma = \epsilon u'$ are bounded by the DPG energy norm times an ϵ-independent, order one constant.

In general, the DPG energy norm is mesh dependent. If we h-refine the mesh, the norm changes and there is no guarantee that the energy norm of the error decreases. However, if we work with the standard (possibly weighted) H^1-norm for the test functions,

$$\|\boldsymbol{v}\|_K^2 = \int_K (|\boldsymbol{v}'|^2 + |\boldsymbol{v}|^2)\alpha(x)\, dx \tag{2.21}$$

the corresponding energy norm of the FE error turns out to be *mesh-independent*. We will show that the DPG energy norm of the FE error for an arbitrary hp mesh *coincides* with the spectral (one element) DPG norm of the error.

1D model problem. Given a right-hand side $f(x), x \in (0,1)$ and inflow data u_0, we consider the problem:

$$\begin{cases} u(0) = u_0, & u(1) = 0 \\ \dfrac{1}{\epsilon}\sigma - u' & = 0 \\ -\sigma' + u' & = f \end{cases} \tag{2.22}$$

Introducing an arbitrary partition,

$$0 = x_0 < x_1 < \ldots < x_{k-1} < x_k < \ldots < x_N = 1 \tag{2.23}$$

we formulate the DPG method as follows. The unknowns include field variables σ_k, u_k defined over element (x_{k-1}, x_k) and fluxes $\hat{\sigma}(x_k), \hat{u}(x_k),\ k = 0, \ldots, N$. Fluxes $\hat{u}(0) = u_0$, $\hat{u}(1) = 0$ are known from the boundary conditions. For each element $K = (x_{k-1}, x_k)$, we have test functions $(\tau, v) = (\tau_k, v_k)$.

For each $k = 1, \ldots, N$, we satisfy the following variational equations,

$$\left\{ \begin{array}{llll} \dfrac{1}{\epsilon}\displaystyle\int_{x_{k-1}}^{x_k} \sigma_k \tau & & + \displaystyle\int_{x_{k-1}}^{x_k} u_k \tau' & -(\hat{u}\tau)|_{x_{k-1}}^{x_k} = 0 \\ \displaystyle\int_{x_{k-1}}^{x_k} \sigma_k v' & -(\hat{\sigma} v)|_{x_{k-1}}^{x_k} & - \displaystyle\int_{x_{k-1}}^{x_k} u_k v' & +(\hat{u} v)|_{x_{k-1}}^{x_k} = \displaystyle\int_{x_{k-1}}^{x_k} f v \end{array} \right. \tag{2.24}$$

for every optimal test function τ, v. Again, for $k = 1$, $\hat{u}(0) = u_0$ is known and is moved to the right-hand side. Similarly, $\hat{u}(1) = 0$ in the last equation for $k = N$.

Norm for the test functions. The norm for the test functions is defined element-wise:

$$\|(\boldsymbol{\tau}, \boldsymbol{v})\| = \left(\sum_{k=1}^{N} \|\tau_k\|_k^2 + \|v_k\|_k^2 \right)^{\frac{1}{2}} \tag{2.25}$$

We will work with two choices for the element norms,

- A weighted H^1-norm,

$$\begin{aligned} \|\tau_k\|_k^2 &= \int_{x_{k-1}}^{x_k} \alpha(|\tau_k'|^2 + |\tau_k|^2)\, dx \\ \|v_k\|_k^2 &= \int_{x_{k-1}}^{x_k} \alpha(|v_k'|^2 + |v_k)|^2)\, dx \end{aligned} \tag{2.26}$$

- A mesh-dependent norm,

$$\begin{aligned} \|\tau_k\|_k^2 &= \int_{x_{k-1}}^{x_k} \alpha|\tau_k'|^2 + \beta_k|\tau_k(x_k)|^2 \\ \|v_k\|_k^2 &= \int_{x_{k-1}}^{x_k} \alpha|v_k'|^2 + \beta_k|v_k(x_k)|^2 \end{aligned} \tag{2.27}$$

Notice that, for a globally continuous function, the norm corresponding to element norm (2.26) reduces to a *mesh-independent* global (spectral) norm, whereas in the second case, it does not. Weights $\alpha = \alpha(x),\ \beta_k,\ k = 1, \ldots, N$ are to be selected.

Global Continuity of the Error Representation Function

We start with a notation. Let $\mathcal{U} = (\boldsymbol{\sigma}, \boldsymbol{u}, \hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{u}})$ denote the exact solution,

$$\begin{aligned} \boldsymbol{\sigma} &:= (\sigma_1, \ldots, \sigma_N) \\ \boldsymbol{u} &:= (u_1, \ldots, u_N) \\ \hat{\boldsymbol{\sigma}} &:= (\hat{\sigma}(0), \ldots, \hat{\sigma}(1)) \\ \hat{\boldsymbol{u}} &:= (\hat{u}(x_1), \ldots, \hat{u}(x_{N-1})) \end{aligned} \tag{2.28}$$

and $\mathcal{U}_{hp}$ the corresponding approximate counterpart. Consider two neighboring elements $(x_{k-1}, x_k), (x_k, x_{k+1})$. Let $(\tau_{\hat{\sigma}_k}, v_{\hat{\sigma}_k})$ be the optimal, vector-valued test function corresponding to flux $\hat{\sigma}_k$, spanned over the two elements. We have the standard orthogonality condition for the error function $\mathcal{E}_{hp} := \mathcal{U} - \mathcal{U}_{hp}$,

$$b(\mathcal{U} - \mathcal{U}_{hp}, (\tau_{\hat{\sigma}_k}, v_{\hat{\sigma}_k})) = b_k(\mathcal{U} - \mathcal{U}_{hp}, (\tau_{\hat{\sigma}_k}, v_{\hat{\sigma}_k})) + b_{k+1}(\mathcal{U} - \mathcal{U}_{hp}, (\tau_{\hat{\sigma}_k}, v_{\hat{\sigma}_k})) = 0 \tag{2.29}$$

Here b_k denotes the k-th element contribution to the bilinear form defined by (2.24). Let (ϕ_k, ψ_k) be now the error representation function for the k-th element,

$$\begin{cases} (\phi_k, \psi_k) \in V((x_{k-1}, x_k)) \\ ((\phi_k, \psi_k), (\delta\phi, \delta\psi))_k = b_k(\mathcal{E}_{hp}, (\delta\phi, \delta\psi)), \quad \forall (\delta\phi, \delta\psi) \in V((x_{k-1}, x_k)) \end{cases} \tag{2.30}$$

Here $(\cdot, \cdot)_k$ denotes the inner product corresponding to element norm $\|\cdot\|_k$, and $V((x_{k-1}, x_k))$ is the element test space. For the exact optimal test functions $V((x_{k-1}, x_k)) = H^1((x_{k-1}, x_k))$, for the approximate test functions, $V((x_{k-1}, x_k)) = \mathcal{P}^{p+\Delta p}((x_{k-1}, x_k))$ where p is the element order of approximation that varies with element, and Δp is a global increment in p defining the *enriched space*. We have,

$$((\phi_k, \psi_k), (\tau_{\hat{\sigma}_k}, v_{\hat{\sigma}_k}))_k + ((\phi_{k+1}, \psi_{k+1}), (\tau_{\hat{\sigma}_k}, v_{\hat{\sigma}_k}))_{k+1} = b(\mathcal{U} - \mathcal{U}_{hp}, (\tau_{\hat{\sigma}_k}, v_{\hat{\sigma}_k})) = 0 \tag{2.31}$$

On the other side, the definition of optimal test functions implies that

$$((\tau_{\hat{\sigma}_k}, v_{\hat{\sigma}_k}), (\delta\phi, \delta\psi))_k = -\delta\psi(x_k), \quad \forall (\delta\phi, \delta\psi) \in V((x_{k-1}, x_k)) \tag{2.32}$$

and

$$((\tau_{\hat{\sigma}_k}, v_{\hat{\sigma}_k}), (\delta\phi, \delta\psi))_{k+1} = \delta\psi(x_k), \quad \forall (\delta\phi, \delta\psi) \in V((x_k, x_{k+1})) \tag{2.33}$$

Selecting the error representation functions (ϕ_k, ψ_k) and (ϕ_{k+1}, ψ_{k+1}) for $(\delta\phi, \delta\psi)$ above, and summing up the two equations, we get

$$-\psi_k(x_k) + \psi_{k+1}(x_k) = 0 \tag{2.34}$$

In a similar way, we use the optimal test function corresponding to flux $\hat{u}(x_k)$ to demonstrate that $\phi_k(x_k) - \phi_{k+1}(x_k) = 0$ as well.

Applying the same argument to fluxes at the endpoints $\hat{\sigma}(0)$ and $\hat{\sigma}(1)$, we show that the corresponding error representation function ϕ vanishes there.

The discussed result is very general and it holds for any choice of inner products for the test functions and other systems of differential equations as well.

Continuity of the error representation function has an important implication. Solution of the *global problem*:

$$\begin{cases} (\phi, \psi) \in V((0, 1)) \\ ((\phi, \psi), (\delta\phi, \delta\psi)) = b(\mathcal{E}_{hp}, (\delta\phi, \delta\psi)), \quad \forall (\delta\phi, \delta\psi) \in V((0, 1)) \end{cases} \tag{2.35}$$

leads to error representation function (ϕ, ψ) where ϕ, ψ coincide with unions of element error representation functions ϕ_k, ψ_k. In order to see it, consider globally continuous test functions $(\delta\tau, \delta v)$ and sum up the local

problems corresponding to element error representation functions (τ_k, v_k) to obtain:

$$\begin{aligned}
(\tau,\delta\tau) &= \sum_{k=1}^{N}(\tau_k,\delta\tau)_k &&= \sum_{k=1}^{N}\left\{\frac{1}{\epsilon}\int_{x_{k-1}}^{x_k}\sigma_k\delta\tau + \int_{x_{k-1}}^{x_k}u_k\delta\tau' - [\hat{u}\delta\tau]|_{x_{k-1}}^{x_k}\right\} \\
&&&= \frac{1}{\epsilon}\int_0^1\sigma\delta\tau + \int_0^1 u\delta\tau' \\
(v,\delta v) &= \sum_{k=1}^{N}(v_k,\delta v)_k &&= \sum_{k=1}^{N}\left\{\int_{x_{k-1}}^{x_k}\sigma_k\delta v' - [\hat{\sigma}\delta v]|_{x_{k-1}}^{x_k} - \int_{x_{k-1}}^{x_k}u_k\delta v' - [\hat{u}\delta v]|_{x_{k-1}}^{x_k}\right\} \\
&&&= \int_0^1\sigma\delta v' - \int_0^1 u\delta v'
\end{aligned} \tag{2.36}$$

where σ and u represent unions of σ_k and u_k, respectively. In other words, continuous test functions $\delta\tau, \delta v$ "do not see" the jump terms corresponding to inter-element fluxes. This has an important consequence. If the formula for the global norm obtained by summing up the formulas for the element norms coincides with the spectral norm, like in case of (2.26), the energy norm of the FE error *coincides with the value of the spectral energy norm*. In general, convergence analysis for DG methods is based on mesh-dependent norms and, usually, we have to work rather hard to demonstrate that those mesh-dependent norms are bounded below by standard L^2-type norms. For the DPG method with optimal test functions, at least in the 1D case, the analysis can be reduced to the spectral case only.

Another important consequence is the behavior of the energy norm of the error during refinements. As the DPG method delivers the best approximation error in the energy norm, the error cannot increase for a p-refined mesh. The discussed result implies that this is true for the h-refinements as well. Continuity of the error representation function and non-increase of the error are, of course, true only under the assumption of perfect integration and arithmetic and, for this reason, the lack of either in practical computation is a great indicator of round off error effects.

Spectral Stability Analysis

We begin with a slight generalization of the stability analysis for the spectral case presented in [8] and norm (2.27) with $\beta = 1$. In particular, we will find a suitable choice for weight α to guarantee the robustness of the method.

Following the reasoning in [8] (see also Appendix), we derive the explicit formula for the energy norm,

$$\|(\sigma, u, \hat{\sigma}(0), \hat{\sigma}(1))\|_E^2 =$$

$$\|\frac{1}{\epsilon}\int_0^x \sigma(s)\,ds - u(x)\|_{1/\alpha}^2 + \|-\sigma(x) + u(x) + \hat{\sigma}(0)\|_{1/\alpha}^2 + |\frac{1}{\epsilon}\int_0^1\sigma(s)\,ds|^2 + |-\hat{\sigma}(1) + \hat{\sigma}(0)|^2 \tag{2.37}$$

where $\|\cdot\|_{1/\alpha}$ denotes the weighted L^2-norm with weight $1/\alpha$. The use of norm (2.27) is essential in establishing the explicit formula for the energy norm.

Formula (2.37) and triangle inequality lead to:

$$\begin{aligned} \|\frac{1}{\epsilon}\int_0^x \sigma - \sigma(x) + \hat{\sigma}(0)\|^2_{1/\alpha} &\leq \|(\sigma, u, \hat{\sigma}(0), \hat{\sigma}(1))\|^2_E \\ |\frac{1}{\epsilon}\int_0^1 \sigma|^2 &\leq \|(\sigma, u, \hat{\sigma}(0), \hat{\sigma}(1))\|^2_E \end{aligned} \tag{2.38}$$

Denoting,

$$\frac{1}{\epsilon}\int_0^x \sigma - \sigma(x) + \hat{\sigma}(0) =: g(x), \quad \frac{1}{\epsilon}\int_0^1 \sigma =: A$$

we solve for σ to obtain,

$$\sigma(x) = \hat{\sigma}(0)e^{\frac{x}{\epsilon}} - g(x) - \frac{1}{\epsilon}\int_0^x e^{\frac{x-s}{\epsilon}} g(s)\, ds \tag{2.39}$$

Integrating both sides over $(0,1)$ and dividing by ϵ, we get

$$A = \hat{\sigma}(0)(e^{\frac{1}{\epsilon}} - 1) - \frac{1}{\epsilon}\int_0^1 e^{\frac{1-s}{\epsilon}} g(s)\, ds \tag{2.40}$$

Solving for $\hat{\sigma}(0)$, we obtain,

$$\hat{\sigma}(0) = A\frac{e^{-\frac{1}{\epsilon}}}{1 - e^{-\frac{1}{\epsilon}}} + \frac{1}{\epsilon(1 - e^{-\frac{1}{\epsilon}})}\int_0^1 e^{-\frac{s}{\epsilon}} g(s)\, ds \tag{2.41}$$

Now comes the main point. If we control function g only in the L^2-norm, then the stability constant reflecting dependence $\hat{\sigma}(0)$ upon g is of order $\epsilon^{-\frac{1}{2}}$, i.e. the stability is not uniform with respect to ϵ (the method is not *robust*). If, however, function g is controlled in a weighted norm, we can claim the robustness. Indeed, choosing for instance

$$\alpha(x) = x \tag{2.42}$$

we have,

$$\begin{aligned} |\int_0^1 e^{-\frac{s}{\epsilon}} g(s)\, ds| &= |\int_0^1 e^{-\frac{s}{\epsilon}} s^{\frac{1}{2}} s^{-\frac{1}{2}} g(s)\, ds| \\ &\leq \left(\int_0^1 e^{-\frac{2s}{\epsilon}} s\, ds\right)^{\frac{1}{2}} \left(\int_0^1 \frac{1}{s}|g(s)|^2\, ds\right)^{\frac{1}{2}} \end{aligned} \tag{2.43}$$

and

$$\int_0^1 e^{-\frac{2s}{\epsilon}} s\, ds = \frac{\epsilon^2}{4}(1 - e^{-\frac{2}{\epsilon}}) - \frac{\epsilon}{2}e^{-\frac{2}{\epsilon}} \leq \frac{\epsilon^2}{4} \tag{2.44}$$

which leads to the final estimate,

$$|\hat{\sigma}(0)| \leq \frac{e^{-\frac{1}{\epsilon}}}{1 - e^{-\frac{1}{\epsilon}}} A + \frac{1}{2(1 - e^{-\frac{1}{\epsilon}})}\|g\|_{1/s} \tag{2.45}$$

where $\|\cdot\|_{1/s}$ is the weighted L^2-norm with weight $1/s$.

L^2-stability result for σ. Substituting expression for $\hat{\sigma}(0)$ into (2.39), we get the final formula for the stress,

$$\begin{aligned}\sigma(x) &= \frac{A}{1-e^{-\frac{1}{\epsilon}}}e^{\frac{x-1}{\epsilon}} + \frac{1}{\epsilon(1-e^{-\frac{1}{\epsilon}})}\int_0^1 e^{\frac{x-s}{\epsilon}}g(s)\,ds - g(x) - \frac{1}{\epsilon}\int_0^x e^{\frac{x-s}{\epsilon}}g(s)\,ds \\ &= \frac{A}{1-e^{-\frac{1}{\epsilon}}}e^{\frac{x-1}{\epsilon}} + \frac{e^{-\frac{1}{\epsilon}}}{\epsilon(1-e^{-\frac{1}{\epsilon}})}\int_0^x e^{\frac{x-s}{\epsilon}}g(s)\,ds - g(x) + \frac{1}{\epsilon(1-e^{-\frac{1}{\epsilon}})}\int_x^1 e^{\frac{x-s}{\epsilon}}g(s)\,ds\end{aligned} \tag{2.46}$$

The L^2-norm of σ is bounded in terms of the constant A and the L^2-norm of g. We have shown in [8] that the corresponding constants are independent of ϵ, i.e. the method is robust in σ without using the weights. Since the standard L^2-norm of g is controlled by the weighted norm $\|\cdot\|_{1/x}$, we also can claim corresponding estimate using the weighted norm for g,

$$\|\sigma\|_{L^2} \leq C(|A| + \|g\|_{1/s}) \tag{2.47}$$

where the constant C is independent of ϵ.

L^2-stability result for u Once we can control $\hat{\sigma}(0)$ and σ in L^2-norm, formula for the energy norm implies that we also control u in L^2-norm. Indeed,

$$\|-\sigma+u+\hat{\sigma}(0)\|^2 \leq \|-\sigma+u+\hat{\sigma}(0)\|_{1/x}^2 \leq \|(\sigma,u,\hat{\sigma}(0),\hat{\sigma}(1))\|_E^2 \tag{2.48}$$

and the triangle inequality imply that

$$\|u\|^2 \leq 2(\|\sigma\|^2 + |\hat{\sigma}(0)|^2 + \|-\sigma+u+\hat{\sigma}(0)\|^2) \leq C\|(\sigma,u,\hat{\sigma}(0),\hat{\sigma}(1))\|_E^2 \tag{2.49}$$

where constant C is independent of ϵ.

Equivalence of norms (2.26) and (2.27) on the spectral level, implies that the discussed spectral stability results hold also for norm (2.26).

FE Stability Result

Global continuity of the error representation function and the spectral stability result imply that the results for the spectral case are automatically valid for an arbitrary FE mesh and norm (2.26). As proved, the relation between the L^2 and energy norms holds only for the FE error function. One can establish a slightly stronger algebraic result and show that, with an appropriate choice of parameters β_k in norm (2.27), the L^2 norms of both σ and u are bounded by the energy norm of $\mathcal{U}$ times an ϵ-independent constant, for *any function* $\mathcal{U}$ for which the energy is finite. The proof involves the use of "telescoping sums" and can be found in Appendix.

3 1D Numerical Experiments

In this section we present numerical results for the 1D model convection-dominated diffusion problem studied in the previous section. All 1D experiments are reported for data $f = 0$ and $u_0 = 1$. The corresponding solution develops a boundary layer at $x = 1$,

$$\sigma(x) = -\frac{1}{1-e^{-\frac{1}{\epsilon}}} e^{\frac{x-1}{\epsilon}}, \quad u(x) = \frac{1}{1-e^{-\frac{1}{\epsilon}}} \left(1 - e^{\frac{x-1}{\epsilon}}\right) \tag{3.50}$$

In all reported numerical experiments we start with a mesh of four elements of order $p = 0$ (piecewise constants). We have used norm (2.26) for computing the optimal test functions and error representation functions. In place of the weight $\alpha(x) = x$ used in the proof of robustness, we have used a simpler choice,

$$\alpha(x) = \begin{cases} 0.1 & x \in (0, \frac{1}{4}) \\ 1.0 & x \in (\frac{1}{4}, 1) \end{cases} \tag{3.51}$$

The increment Δp used to define the enriched space was set to $\Delta p = 4$, and the maximum order was set to $p_{max} = 4$. These choices are related to an attempt to determine the minimum diffusion constant ϵ for which we can solve the problem numerically using the double precision arithmetic.

We have used a "poor man greedy hp algorithm" summarized below.

```
Set δ = 0.5
do while δ > 0.1
  solve the problem on the current mesh
  for each element K in the mesh
    compute element error contribution e_K
  end of loop through elements
  for each element K in the mesh
    if e_K > δ² max_K e_K then
      if new h ≥ ε then
        h-refine the element
      elseif new p ≤ p_max then
        p-refine the element
      endif
    endif
  end of loop through elements
if (new N_dof = old N_dof) reset δ = δ/2
end of loop through mesh refinements
```

The strategy reflects our old experiments with boundary layers [9], and rigorous approximability results

of Schwab and Suri [12] and Melenk [10] on optimal hp discretizations of boundary layers: we proceed with h-refinements until the diffusion scale ϵ is reached and then continue with p-refinements.

We start with an easy case of $\epsilon = 10^{-3}$. Fig. 1 presents the corresponding convergence history on a log-log scale, energy norm of the error vs. number of degrees-of-freedom (d.o.f.). As predicted by the theory, the error goes monotonically down. The curve displays three different regions. The first one corresponds to a series of h-refinements "localizing" the boundary layer and it is characterized with relatively slow convergence. The second region, characterized with a rapid convergence, corresponds to the p-refinements in the boundary layer and illustrates the power of p-refinements on a properly refined mesh. Finally, the last part corresponds to the last few refinements when the mesh restrictions in terms of both minimum element size and maximum polynomial order, become active. Fig. 2 presents a $10\times$ zoom on the boundary layer

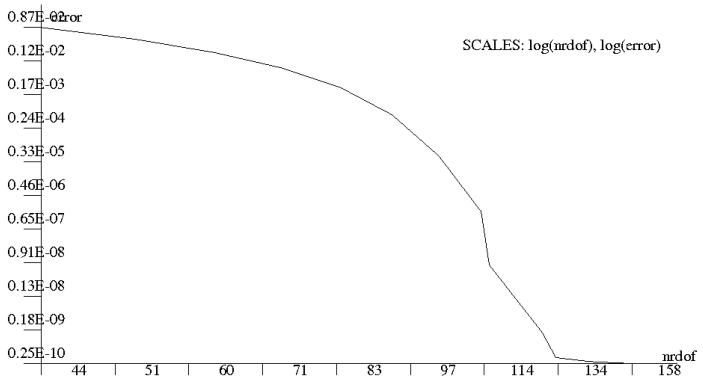

Figure 1: Diffusion $\epsilon = 10^{-3}$: convergence history.

revealing the optimal hp mesh and (an eye ball norm) perfect resolution of the boundary layer. Finally, Fig. 3 displays the corresponding error representation function ϕ. The function is continuous, consistently with the theory.

We continue now with a more challenging case of $\epsilon = 10^{-6}$. Fig. 4 presents the convergence history. The error goes monotonically down until the last few refinements where it begins to oscillate. Fig. 5 presents a $10^5\times$ zoom on the boundary layer revealing the optimal hp mesh and (still) a perfect resolution of the boundary layer. So far, everything looks OK. If we look however at Fig. 6 showing the error representation function ϕ, we see that, contrary to the theory, the function is no longer globally continuous. For $\epsilon = 10^{-7}$ the method falls apart and the code reports zero pivots in solving local problems for optimal test functions. The lost of continuity of the error representation functions (the same behavior is observed for function ψ) is visible, starting already with $\epsilon = 10^{-4}$. For elements resolving the boundary layer, element size $h = \epsilon$. The ratio of the stiffness and mass terms in formula (2.26) is of order p^2/h^2 and, for $h = 10^{-3}, p = 4$, is

Figure 2: Diffusion $\epsilon = 10^{-3}$: resolution of the boundary layer.

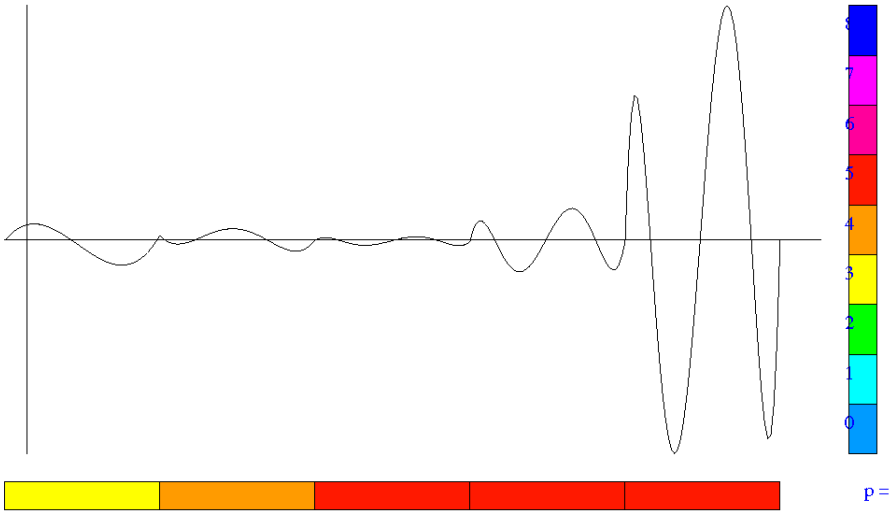

Figure 3: Diffusion $\epsilon = 10^{-3}$: error representation function ϕ.

of order 10^7, which corresponds to half of the 14-15 digits available with double precision. With $h = 10^{-4}$,

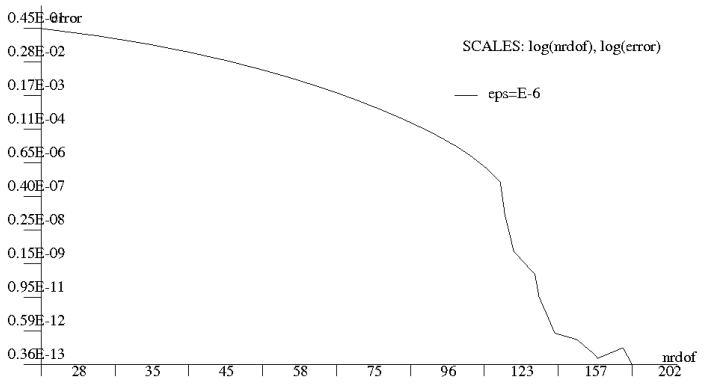

Figure 4: Diffusion $\epsilon = 10^{-6}$: convergence history.

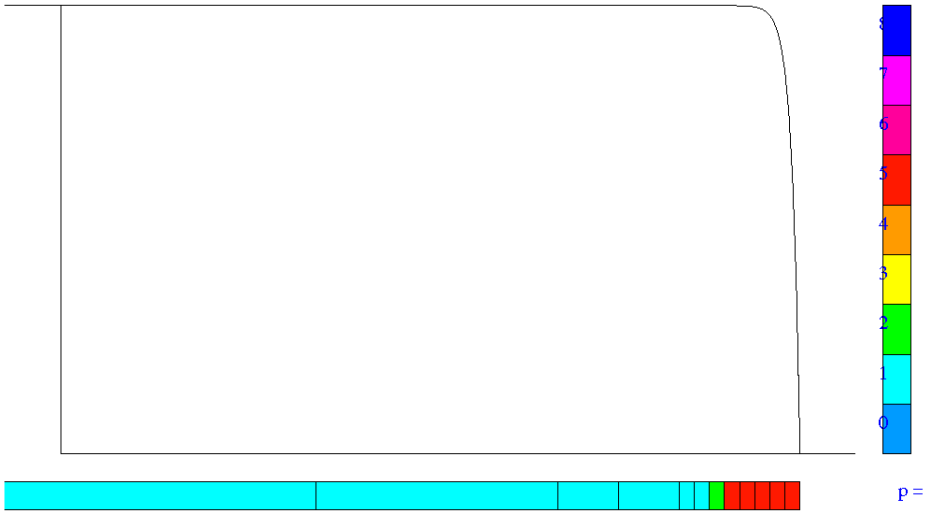

Figure 5: Diffusion $\epsilon = 10^{-6}$: resolution of the boundary layer.

the overlap is minimized to 5 digits and we begin to observe the effects of round off errors [2]. The problem apparently is not just with the solution of the element, local problems for the optimal test functions. To investigate the issue, we have employed for local shape functions (a basis for the enriched space used to determine optimal test functions and solve for the error representation functions) approximate eigenvectors

[2] The situation resembles experience with penalty methods.

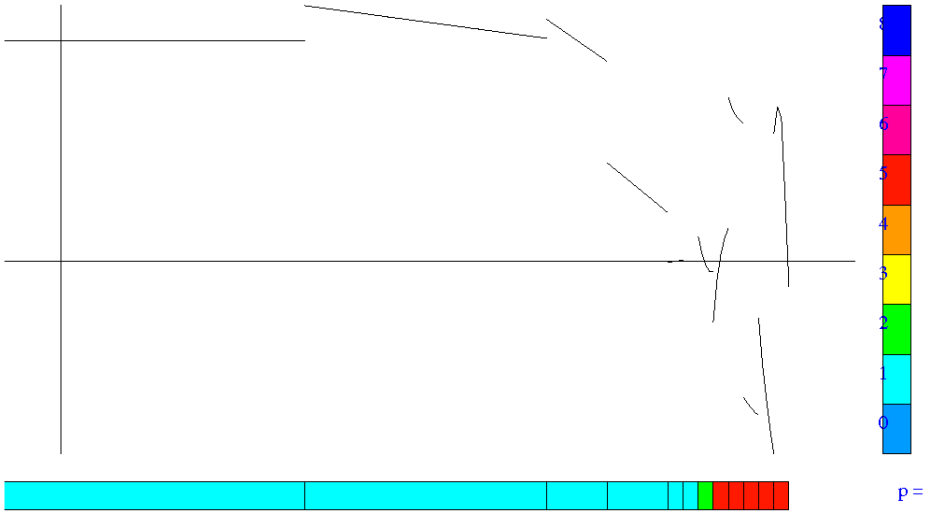

Figure 6: Diffusion $\epsilon = 10^{-6}$: error representation function ϕ.

of the 1D Laplacian (with Neumann boundary conditions),

$$\begin{cases} \lambda_n \in \mathbb{R}, \phi_n \in \mathcal{P}^{p+\Delta p}(0,1) \\ \displaystyle\int_0^1 \phi_n' \delta\phi' \, dxi = \lambda_n \int_0^1 \phi_n \delta\phi \, dxi, \quad \forall \delta\phi \in \mathcal{P}^{p+\Delta p}(0,1) \end{cases} \tag{3.52}$$

The eigenproblem above was discretized in turn with Legendre polynomials. With the use of the eigenvectors, the stiffness matrix for the local, element problems was reduced to a diagonal matrix (the off-diagonal terms were set to "clean" zeros). Still, we have not observed any significant improvement in the global results which indicates conditioning problems with the global and not just the local stiffness matrix.

A remedy. A simple idea that came to mind was to work with a rescaled inner product:

$$(v, \delta v) = \int_{x_{k-1}}^{x_k} (h_k v' \delta v' + v \delta v) \alpha(x) \, dx \tag{3.53}$$

where h_k is the element size. The norm is now mesh-dependent, and it changes with every h-refinement (a "moving target" strategy). The global energy error is no longer guaranteed to behave monotonically with h-refinements. Indeed, looking at the convergence history shown in Fig. 7, the period corresponding to h-refinements is characterized with practically constant error (the numerical values display a modest decrease..). Fig. 8 showing a $10^5\times$ zoom on the boundary layer reveals a perfect resolution of the boundary layer. Most importantly, the error representation function ϕ shown in Fig. 9, is again globally continuous.

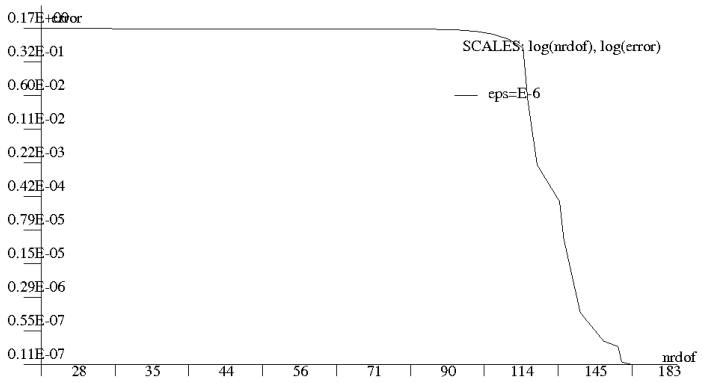

Figure 7: Diffusion $\epsilon = 10^{-6}$, rescaled inner product: convergence history.

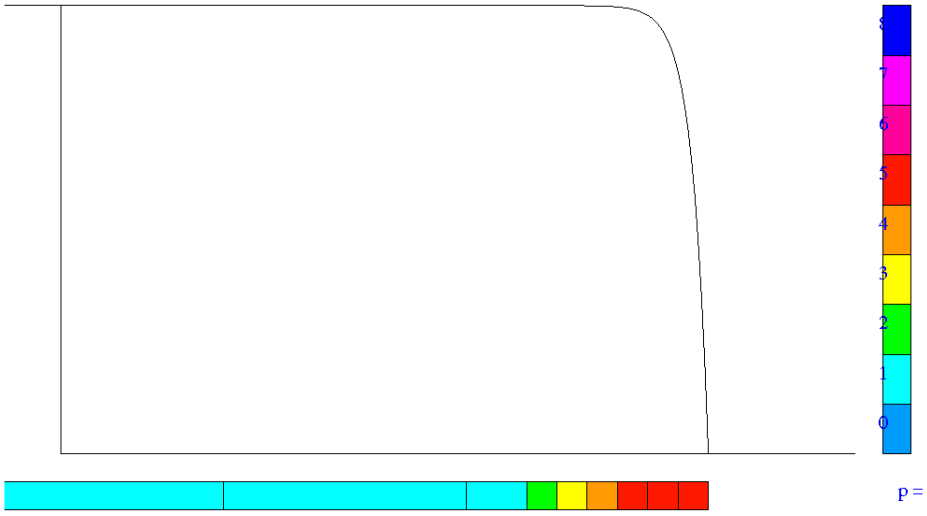

Figure 8: Diffusion $\epsilon = 10^{-6}$, rescaled inner product: resolution of the boundary layer.

The minimum value of ϵ, we have been able to solve for with the rescaled inner product, is $\epsilon = 10^{-11}$. We have tried to increase the power of element size h in (3.53) beyond one. It looks like the code does work with the exponent greater than one (e.g. $1.1 - 1.3$) but the method falls apart with the exponent equal 2. As practical range of interest is $\epsilon \geq 10^{-7}$, we have been content with the use of (3.53).

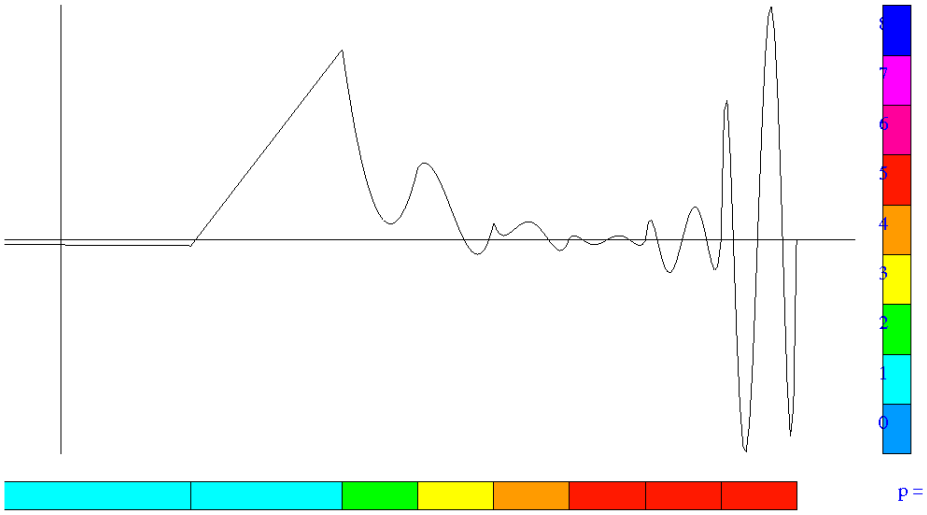

Figure 9: Diffusion $\epsilon = 10^{-6}$, rescaled inner product: error representation function ϕ.

Remark 1 All presented numerical results were done within hp FE codes supporting the whole exact sequence. The color code for displayed meshes, both in 1D and 2D, corresponds to H^1-conforming elements. The actual order of approximation used for the L^2-conforming elements, is one order less for 1D and quads in 2D, and two orders less for triangles in 2D. ∎

4 2D Numerical Experiments

We consider the 2D convection-dominated diffusion problem.

$$\left\{ \begin{array}{rcll} \dfrac{1}{\epsilon}\boldsymbol{\sigma} - \nabla u & = 0 & \text{in } \Omega \\ -\text{div}\boldsymbol{\sigma} + \text{div}(\boldsymbol{\beta} u) & = f & \text{in } \Omega \end{array} \right. \tag{4.54}$$

accompanied with boundary condition for velocity u,

$$u = u_0 \text{ on } \partial\Omega \tag{4.55}$$

where $\Omega \subset \mathbb{R}^2$, right-hand side f and boundary data u_0 are given, and $\boldsymbol{\beta}$ is a prescribed advection field.

Given an element K, we multiply equations (4.54) with test functions $\boldsymbol{\tau}, v$, integrate over the element,

and integrate by parts to arrive at the variational formulation:

$$\left\{\begin{array}{llllll} \frac{1}{\epsilon}\int_K \boldsymbol{\sigma}\boldsymbol{\tau} & & + \int_K u\,\mathrm{div}\boldsymbol{\tau} & - \int_{\partial K}\hat{u}\tau_n & = 0 & \forall \boldsymbol{\tau} \\ \int_K \boldsymbol{\sigma}\nabla v & - \int_{\partial K}\hat{\sigma}_n \mathrm{sgn}(\boldsymbol{n})v & - \int_K u\,\boldsymbol{\beta}\cdot\nabla v & + \int_{\partial K}\hat{u}\beta_n v & = \int_K fv & \forall v \end{array}\right. \tag{4.56}$$

where normal stress $\hat{\sigma}_n$ and velocity $\hat{u}$ have become independent unknowns on the boundary of the element. The normal stress is defined using a globally fixed normal $\boldsymbol{n}_e$ for element edge e, with function $\mathrm{sgn}(\boldsymbol{n})$ indicating consistency of global and local, element outward normal unit vectors,

$$\mathrm{sgn}(\boldsymbol{n}) = \left\{\begin{array}{ll} 1 & \text{if } \boldsymbol{n} = \boldsymbol{n}_e \\ -1 & \text{if } \boldsymbol{n} = -\boldsymbol{n}_e \end{array}\right. \tag{4.57}$$

Energy setting. Presence of grad and div operators acting on test functions and Cauchy-Schwarz inequality clearly suggests the choice of norm for the test functions - space $\boldsymbol{H}(\mathrm{div}, K)$ for $\boldsymbol{\tau}$, and $H^1(K)$ for v , as well as for the field unknowns $\boldsymbol{\sigma} \in \boldsymbol{L}^2(K), u \in L^2(K)$. In 1D, the fluxes are just numbers so the functional setting for them is trivial. In multi-dimensions, we only can say that the fluxes should live in the space dual to the test functions:

$$\prod_K (\boldsymbol{H}(\mathrm{div}, K) \times H^1(K)) \tag{4.58}$$

A more concrete characterization of this space is non-trivial.

Discretization. We will compute with both triangular and quadrilateral elements, restricting ourselves to 1-irregular meshes only. Each element K is supporting two field variables: vector values stress $\boldsymbol{\sigma}$ and velocity u. Each element edge is supporting two fluxes: normal stress $\hat{\sigma}_n$ and velocity $\hat{u}$. For triangles, we use spaces $\mathcal{P}^p$, for quads we use spaces $\mathcal{Q}^{(p,q)} = \mathcal{P}^p \otimes \mathcal{P}^q$ with the standard relation between the master and physical element shape functions,

$$\phi(\boldsymbol{x}) := \hat{\phi}(\boldsymbol{\xi}), \quad \boldsymbol{x} = \boldsymbol{x}_K(\boldsymbol{\xi}) \tag{4.59}$$

where $\boldsymbol{x}_K$ stands for the element map transforming the corresponding master triangle or square into the physical element. The logic of exact sequence and L^2-conforming discretizations suggests a slightly different relation between physical element shape functions $\phi(\boldsymbol{x})$ and master element shape functions $\hat{\phi}(\boldsymbol{\xi})$ [6],

$$\phi(\boldsymbol{x}) := \frac{1}{j}\hat{\phi}(\boldsymbol{\xi}), \quad \boldsymbol{x} = \boldsymbol{x}_K(\boldsymbol{\xi}) \tag{4.60}$$

where j denotes the jacobian for the element map, $j = \det(\nabla \boldsymbol{x}_K)$. However, all presented numerical examples will involve affine element only, for which the difference between the two relations reduces to a local scaling with the constant jacobian only. Element order of approximation p for triangles and (p, q) for quads can vary locally, from element to element.

Discretization for fluxes is fixed using a *maximum rule*. For two neighboring triangles of order p_1 and p_2, the order for the shared edge is set to

$$p_e = \max\{p_1, p_2\} + 1 \tag{4.61}$$

There is no particular strong reason, why we have raised the order for the edge by one. It is mostly a legacy issue reflecting our initial experience with the DPG method for pure convection problems [7]. In the case of hanging nodes and an edge with two small neighbors on one side and equal size neighbor on the other, the discretization for the flux is chosen to be *piecewise polynomial* with the order set to the maximum of orders for all three neighboring elements,

$$p_e = \max\{p_1, p_2, p_3\} + 1 \tag{4.62}$$

In that sense the *maximum rule* applies here to both element order and size.

We use the same principle for quadrilateral meshes accounting for the directionality of approximation. This time, again for no deep reason, we chose to set the order of fluxes just to the maximum of the directional orders for the neighboring elements,

$$p_e = \max\{q_1, q_2(, q_3)\} \tag{4.63}$$

Here q_i is the order of discretization of the i-th neighbor in the direction of the edge.

For each element trial shape function, and for each edge trial shape functions, we determine the optimal test function by solving the local problem,

$$\begin{cases} \tau_K \in \boldsymbol{V}(K), v_K \in V(K) \\ (\boldsymbol{\tau}_K, \boldsymbol{\delta\tau})_{H(div)} + (v_K, \delta v)_{H^1} = b_K((\boldsymbol{\sigma}, u, \hat{\sigma}_n, \hat{u}), (\boldsymbol{\delta\tau}, \delta v)) \quad \forall \boldsymbol{\delta}\tau \in \boldsymbol{V}(K),\ \delta v \in V(K) \end{cases} \tag{4.64}$$

where $\boldsymbol{V}(K), V(K)$ are the *enriched spaces*, $\mathcal{P}^{p+\Delta p}$ for triangular elements, and $\mathcal{P}^{p+\Delta p} \otimes \mathcal{P}^{q+\Delta p}$ for quads. At this point, we have ignored $H(\mathrm{div})$-conforming elements and approximated components of τ and $\boldsymbol{\delta}\tau$ with polynomials of equal order. Despite the piecewise polynomial fluxes for 1-irregular meshes, we also have not experimented yet with piecewise polynomials for the enriched spaces corresponding to a local h-refinement[3]. The optimal test functions corresponding to field variables $\boldsymbol{\sigma}, u$ span over one element only, the optimal test functions corresponding to fluxes $\hat{\sigma}_n, \hat{u}$ span over the elements neighboring the particular edge.

Fluxes $\hat{u}$ on the boundary of the domain are determined by performing L^2-projections of the corresponding boundary data onto the corresponding trial space.

Finally, we start our experiments with weighted inner products for the $\boldsymbol{H}(\mathrm{div})$ and H^1 spaces,

$$\begin{aligned} (\boldsymbol{\tau}, \boldsymbol{\delta\tau})_{H(div)} &= \int_K \{\mathrm{div}\boldsymbol{\tau}\ \mathrm{div}\boldsymbol{\delta\tau} + \boldsymbol{\tau}\ \boldsymbol{\delta\tau}\} w(\boldsymbol{x})\ d\boldsymbol{x} \\ (v, \delta v)_{H^1} &= \int_K \{\nabla v \nabla \delta v + v\ \delta v\} w(\boldsymbol{x})\ d\boldsymbol{x} \end{aligned} \tag{4.65}$$

[3] A must in the future work...

The same setting is always used to determine the element error representation functions $(\boldsymbol{\psi}_K, \phi_K)$.

All numerical experiments reported in this paper were done for a unit square domain $\Omega = (0,1)^2$, constant advection vector $\boldsymbol{\beta}$, and boundary conditions defined in Fig. 10. The weight $w(\boldsymbol{x})$ for the inner

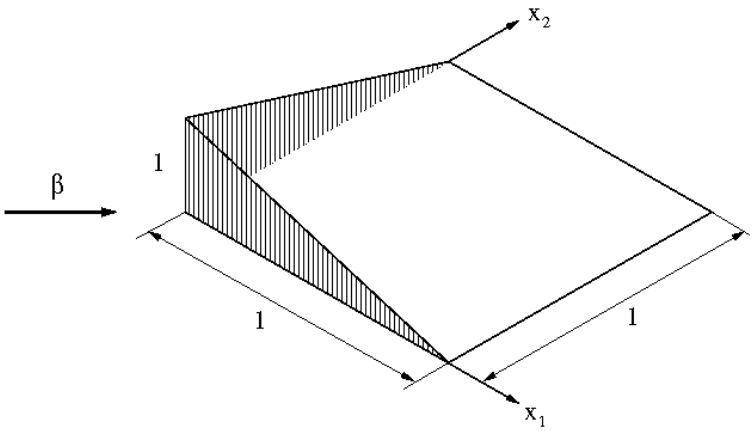

Figure 10: Convection-dominated diffusion in 2D: boundary conditions

products (4.65) was selected following our 1D experience and it is shown in Fig. 11. The 2D situation is a bit more difficult than in 1D as the inflow boundary meets the no-flow boundary at north-west and south-east corners. The choice of weight reflects the intention of emphasizing the inflow boundary over the no-flow boundary in the resulting DPG energy norm. Our numerical experiments focus on determining the minimum

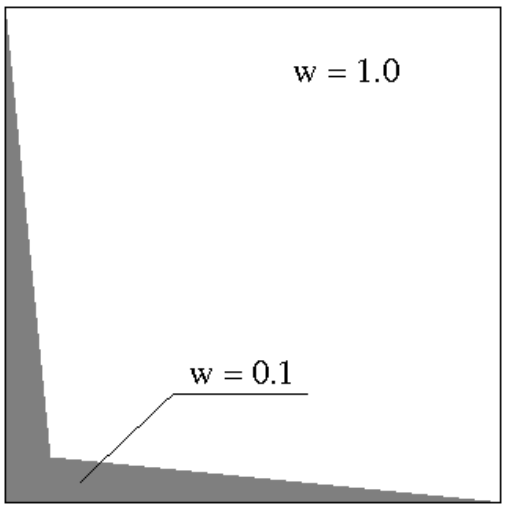

Figure 11: Convection-dominated diffusion in 2D: definition of the weight.

values for the diffusion constant ϵ for which we can solve the problem. The global system of equations was solved using a frontal solver for symmetric problems (with no pivoting). The element d.o.f were ordered

in the order: field variables first, fluxes next. This effectively results in a static condensation of the interior d.o.f. and improves the conditioning of the global matrix.

Triangular Meshes

The advection vector was set to $\boldsymbol{\beta} = (1, 2)$ to avoid an alignment with the mesh.[4]

We have used the same adaptive algorithm as for the 1D problems. The increment in polynomial order was set to $\Delta p = 4$, and maximum order was limited to $p_{max} = 4$. The "working horse" are elements of second order and we always start with a mesh of eight elements with $p = 2$. We begin with a modest case $\epsilon = 10^{-3}$. Fig.12 presents convergence history on the log-log scale. The maximum number of d.o.f. after 21 iterations reached a total of 284,287, and the energy norm was reduced by two orders of magnitude. Convergence is monotone throughout the whole process. Fig. 13 presents the final mesh. The inflow conditions

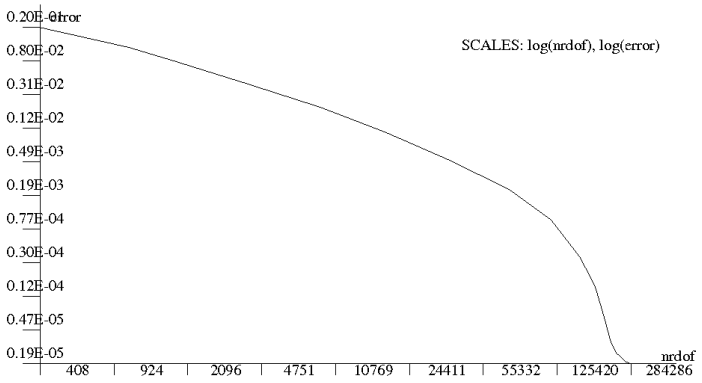

Figure 12: $\epsilon = 10^{-3}$: convergence history for triangular meshes.

are non-smooth at the south-west corner which causes a visible diffusion along the streamline picked by the refinements. Figures 14 and 15 present $10^2\times$ zooms on the mid-point of the north edge and the north-east corner showing the hp meshes generated by the algorithm. Figures 16 through 16 present the corresponding resolution of the flow and the boundary layers, with and without the mesh. The visible mesh structure on the contour plots without the mesh indicates still a non-perfect resolution. The computed solution range is $(-0.81043E-01, 0.10783E+01)$.

Solution for $\epsilon = 10^{-4}$ was possible but at the limit of the four year old IBM ThinkPad on which the presented experiments were made. The final mesh reached almost 1M d.o.f. and, in order to avoid the round-off breakdown, we had to rescale the mass terms in the inner product by a factor of 10. Clearly, the

[4]The problem then would have been easier.

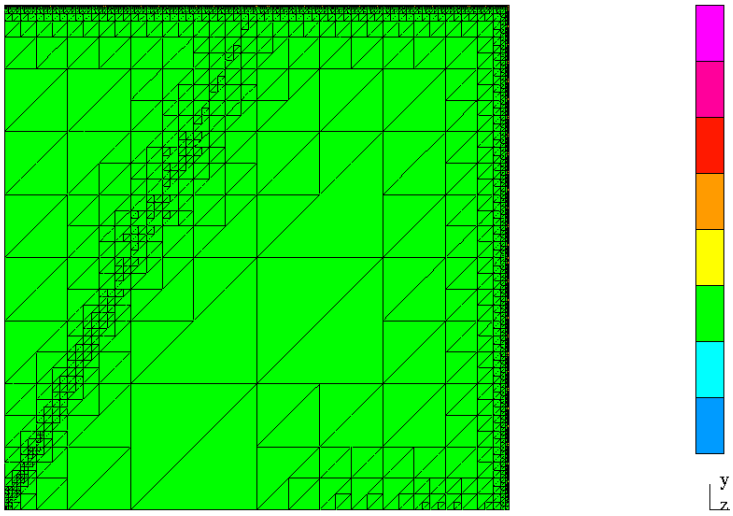

Figure 13: $\epsilon = 10^{-3}$: final mesh after 21 iterations.

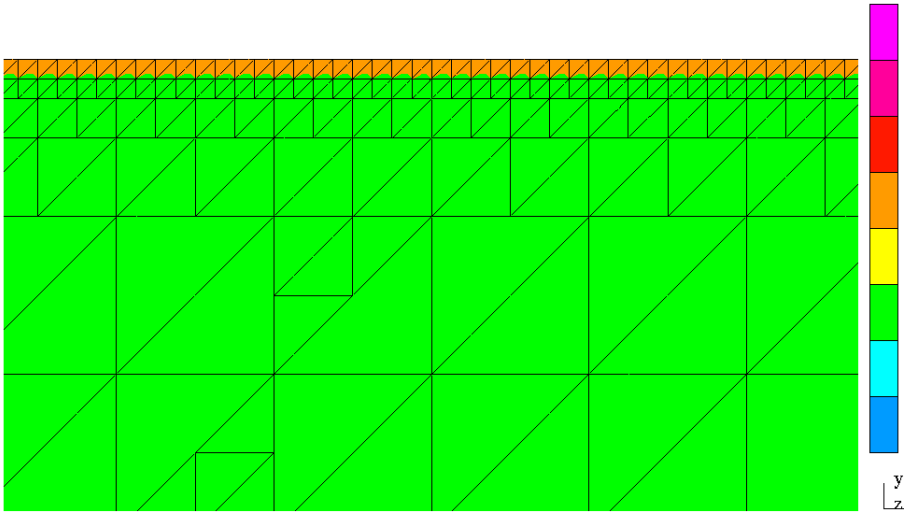

Figure 14: $\epsilon = 10^{-3}$: final mesh after 21 iterations, $10^{-2}\times$ zoom on upper boundary.

triangular meshes and isotropic refinements is not the way to go when it comes to boundary layers.

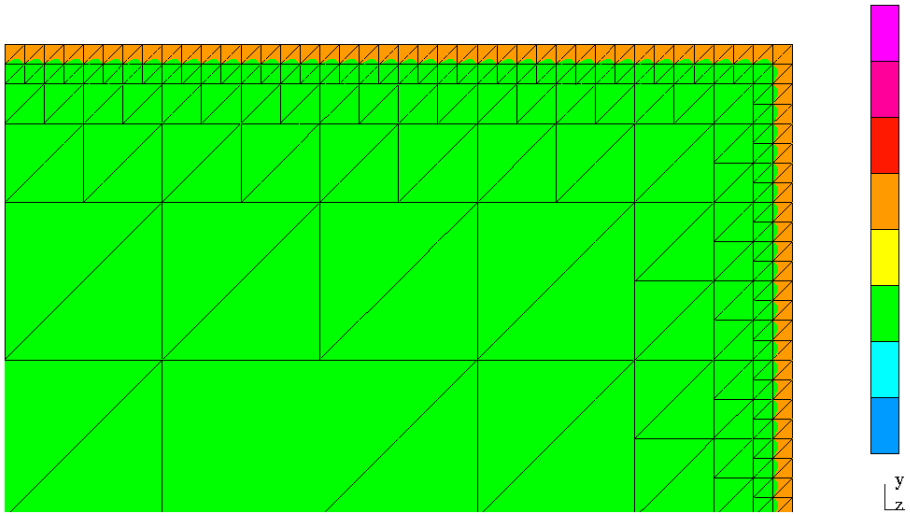

Figure 15: $\epsilon = 10^{-3}$: final mesh after 21 iterations. $10^{-2}\times$ zoom on north-east corner.

Quadrilateral Meshes

Each quadrilateral element is assigned an *isotropy flag* which is fixed upon solving the local problem for determining the error representation functions $(\boldsymbol{\psi}_K, \phi_K)$,

$$\begin{cases} \boldsymbol{\psi}_K \in \boldsymbol{V}(K), \phi_K \in V(K) \\ ((\boldsymbol{\psi}_K, \phi_K), (\delta\boldsymbol{\tau}, \delta v))_K = b_K((\boldsymbol{\sigma}_{hp}, u_{hp}, \hat{\sigma}_n, \hat{u}), (\delta\boldsymbol{\tau}, \delta v)) - l_K((\delta\boldsymbol{\tau}, \delta v)) \\ \qquad\qquad \forall \delta\boldsymbol{\tau} \in \boldsymbol{V}(K), \delta v \in V(K) \end{cases} \tag{4.66}$$

Here $(\boldsymbol{\sigma}_{hp}, u_{hp}, \hat{\sigma}_n, \hat{u})$ is the FE solution in the element, b_K is the element bilinear form, $(\cdot,\cdot)_K$ is the scalar product for the test functions, and $\boldsymbol{V}(K), V(K)$ are the enriched spaces for determining the approximate optimal test functions and the error representation functions. Once functions $(\boldsymbol{\psi}_K, \phi_K)$ are known, we define two directional contributions to the element error:

$$c_1 := \int_K (|\tau_1|^2 + |\frac{\partial v}{\partial x_1}|^2) w(x)\, dx \quad c_2 := \int_K (|\tau_2|^2 + |\frac{\partial v}{\partial x_2}|^2) w(x)\, dx \tag{4.67}$$

The anisotropy flag is defined now as follows.

$$\text{Anisotropy flag} = \begin{cases} 10 & \text{if } c_1 \geq 10 c_2 \\ 01 & \text{if } c_2 \geq 10 c_1 \\ 11 & \text{otherwise} \end{cases} \tag{4.68}$$

The choice of factor 10 is arbitrary.

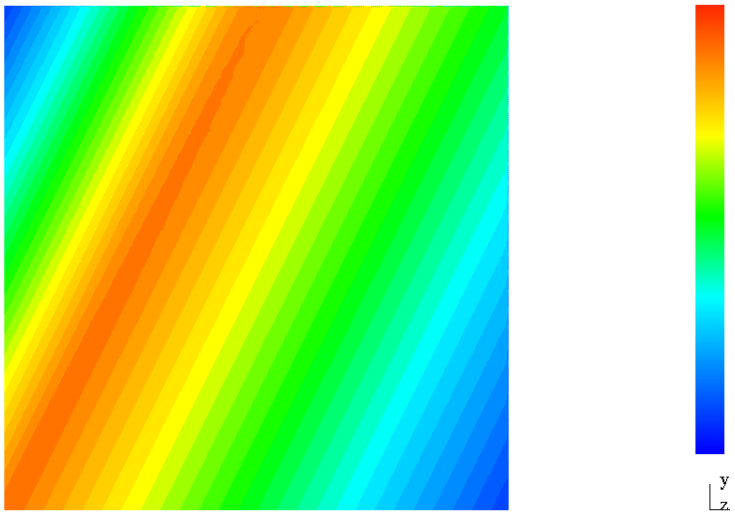

Figure 16: $\epsilon = 10^{-3}$: solution u.

The elements are refined according to their isotropy flags. If the flag points in one direction only, the element is h- or p-refined only in this direction. The formal algorithm is a slight modification of the algorithm for the triangular meshes and it looks as follows.

```
Set δ = 0.5
do while δ > 0.1
  solve the problem on the current mesh
  for each element K in the mesh
    compute element error contribution e_K and isotropy flag
  end of loop through elements
  for each element K in the mesh
    if e_K > δ² max_K e_K then
      if new h_1, h_2 ≥ ε then
        h-refine the element
      elseif new p, q ≤ p_max then
        p-refine the element
      endif
    endif
  end of loop through elements
if (new N_dof = old N_dof) reset δ = δ/2
```

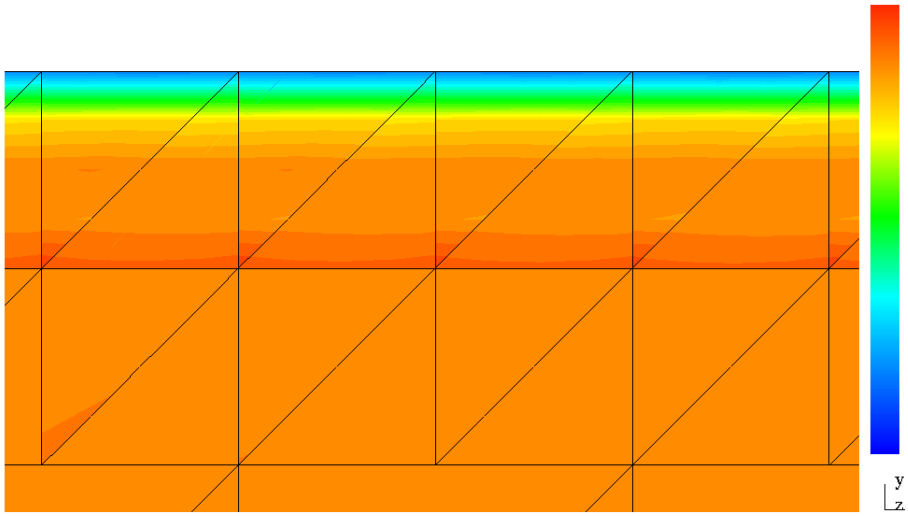

Figure 17: $\epsilon = 10^{-3}$: solution u, $10^2\times$ zoom on top boundary with the mesh.

```
end of loop through mesh refinements
```

The ratio of maximum to minimum element size was limited to $10,000$. The number was determined experimentally in process of determining the minimum diffusion constant ϵ for which we can solve the problem despite the conditioning problems. In the experiments with rectangular meshes, the advection vector was changed to $\boldsymbol{\beta} = (1, 1)$. The corresponding solution is then symmetric along the diagonal. Any loss of symmetry points to either bugs in the code or round off error effects.

Use of anisotropic refinements results in dramatic savings in terms of number of d.o.f.. Fig. 21 presents convergence history for $\epsilon = 10^{-4}$. In place of nearly 1M d.o.f. for triangles, the maximum number of d.o.f. reached only 134k. After 27 refinements, the energy error is reduced by four orders of magnitude (displayed is the error squared). As in 1D examples, the error decreases monotonically down. Slightly non-monotone behavior of the error in the last three iterations indicates that we are at the limit of machine precision. Quality of the solution obtained with the hp refinements is illustrated in Fig. 22 displaying the solution in the north-east corner of the square domain. The range for the computed solution u is $(-0.19044E-04, 0.99980E+00)$.

We have managed to solve the problem for $\epsilon = 10^{-5}$ with additional restrictions on element aspect ratio (to 100) and rescaling the mass terms in the norm for the test functions by a factor of 10. Limiting the aspect ratio to 100 resulted in a blow up in number of d.o.f. (over 700k).

It is not just the local problem that is affected by the round off error effects. We have implemented the

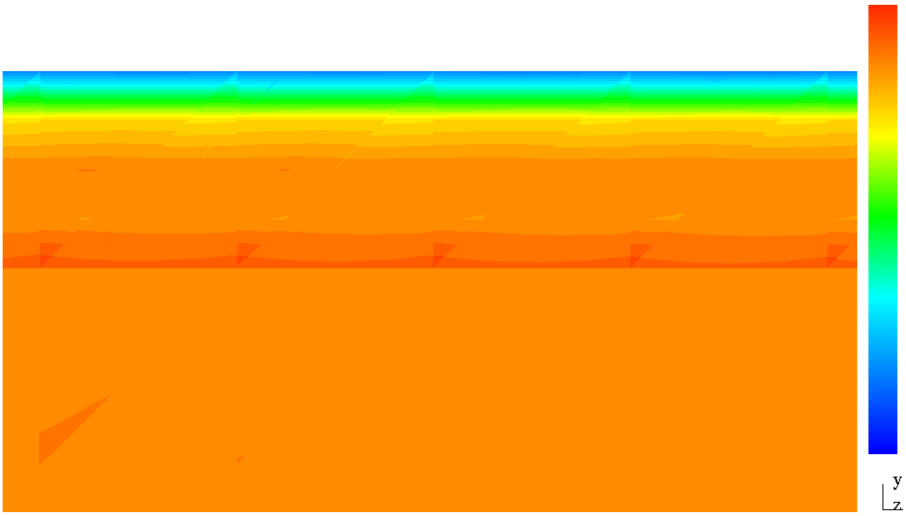

Figure 18: $\epsilon = 10^{-3}$: solution u, $10^2\times$ zoom on top boundary without the mesh.

solution of local problems using a quadruple arithmetic [5] , with little effect on the final results.

Similarly to 1D experiments, the key point in "pushing the envelope" turned out to be the rescaling of inner product. The new, mesh-dependent, rescaled norm for the test functions was defined as:

$$\begin{aligned} \|(\boldsymbol{\tau}, v)\|_K^2 &= \int_K \Big\{ |\sqrt{h_1}\frac{\partial \tau_1}{\partial x_1} + \sqrt{h_2}\frac{\partial \tau_2}{\partial x_2}|^2 + |\tau_1|^2 + |\tau_2|^2 \\ &\quad + h_1|\frac{\partial v}{\partial x_1}|^2 + h_2|\frac{\partial v}{\partial x_2}|^2 + |v|^2 \Big\} w(x)\, dx \end{aligned} \tag{4.69}$$

With the rescaled inner product, we were able to solve the problem for $\epsilon = 10^{-7}$. Fig. 23 shows the corresponding convergence history. The convergence is no longer monotone, as the error slightly increases during the h-refinements and decreases only when the p-refinements begin. The results for $\epsilon = 10^{-7}$ are illustrated in Figures 24-29. The last picture shows still some room for improvement as one can detect the trace of the mesh on the contour plot. The range for the computed solution u is $(-0.65259E-06, 0.99980E+00)$.

5 Conclusions

A summary. Construction of stable discretizations has been at the heart of the numerical analysis since the inception of the field. The phrase "*discrete* stability and approximability imply convergence" is known to

[5] The code slowed down from minutes to hours. With the current hardware, the quadruple precision does not present a practical solution to round off error related problems.

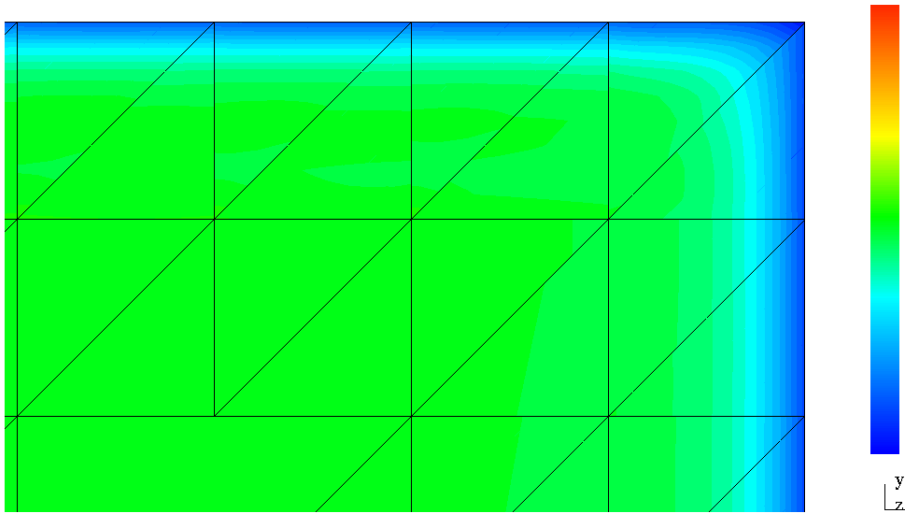

Figure 19: $\epsilon = 10^{-3}$: solution u, $10^2\times$ zoom on north-east corner with the mesh.

every numerical analyst. Equally known fact is that, except for a nice but small class of coercive problems, stability at the continuous level *does not* imply the discrete stability of Bubnov-Galerkin approximations. The DPG method with optimal test functions breaks through this barrier - continuous stability implies automatically discrete stability, for any system of linear PDE's. In context of hp-discretizations, this implies automatically the uniform stability in terms of polynomial order. The difficulty is, however, shifted to proving uniform stability with respect to mesh. The "natural" DPG norm is mesh-dependent, as it is for all DG methods and one of the first tasks in analyzing those methods is to establish a lower bound for the mesh-dependent norm in terms of a global, mesh-independent norm.

We have established such a result for the 1D convection-dominated diffusion problem. With a proper choice of a weight for the norm in test space, we have managed also to construct a "robust" version of the method, with stability properties uniform in diffusion constant ϵ. Convinced with numerical experience, we hope to extend our 1D theoretical results to multidimensional setting.

The choice of the test space norm turned out also to be a crucial tool in fighting the round off error effects. With the rescaled norms, we have been able to extend the current limit [6] for diffusion constant ϵ for which the 2D problem can be solved on a double precision platform from $\epsilon = 10^{-4}$ to $\epsilon = 10^{-7}$.

Most importantly, the whole adaptive process is fully automatic and its starts with few elements only. The method does not exhibit any preasymptotic range behavior where stability properties do not hold. In other words, one does not need to start with a mesh that reflects an expertise on the problem.

[6] Dominik Schotzau, private communication.

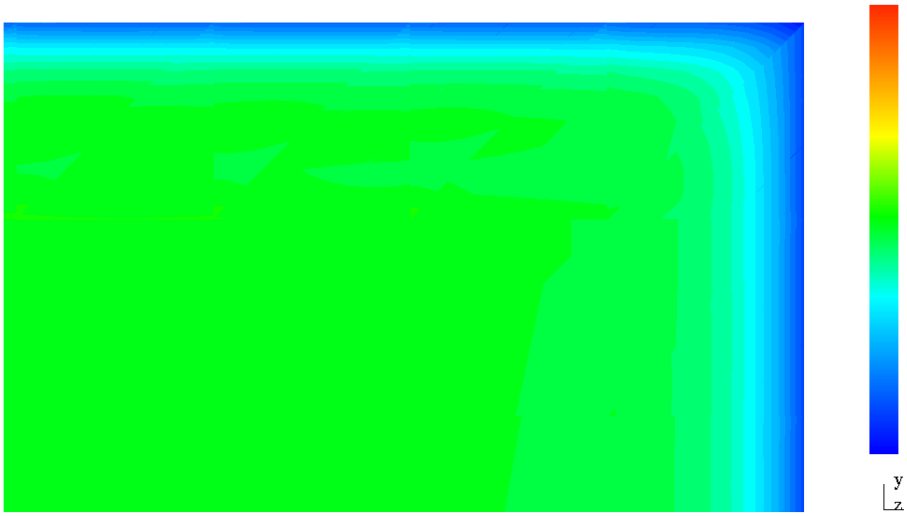

Figure 20: $\epsilon = 10^{-3}$: solution u, $10^2\times$ zoom on north-east corner without the mesh.

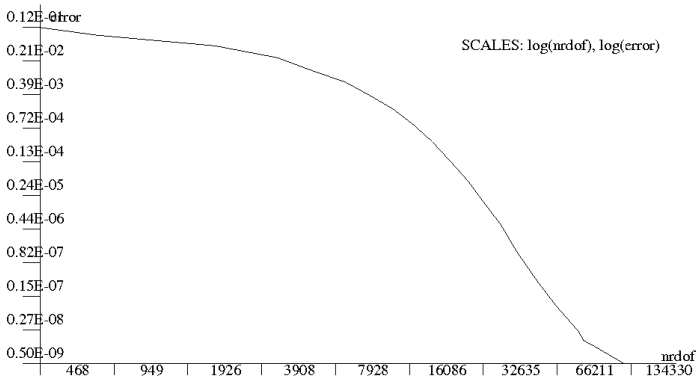

Figure 21: $\epsilon = 10^{-4}$: convergence history.

Convergence analysis in multi-dimensions. A careful examination of arguments used in 1D to prove the global continuity of the error representation function reveals that we cannot count on the same result in 2D or 3D. However, repeating the arguments, we learn that in place of continuity we have an orthogonality result: jump in error representation functions must be orthogonal to the trial space for fluxes. We have for

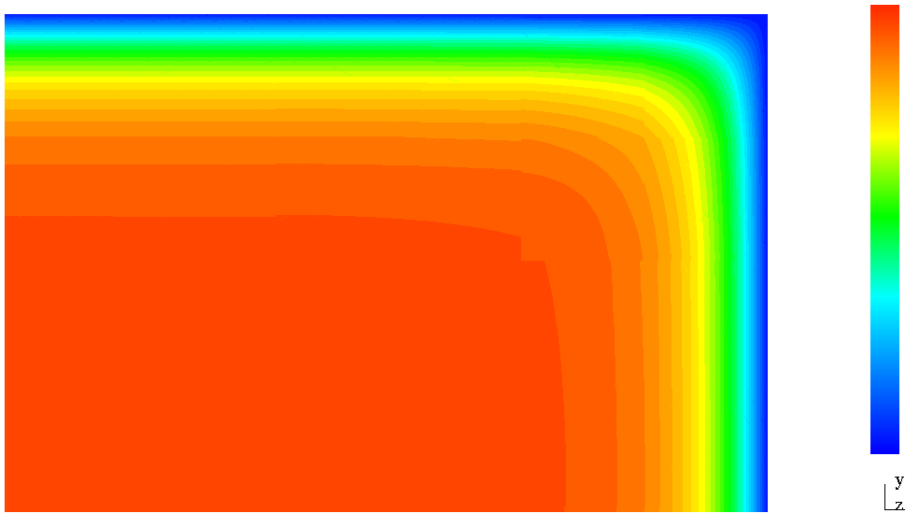

Figure 22: $\epsilon = 10^{-4}$: solution u, $10^3\times$ zoom on the north-east corner.

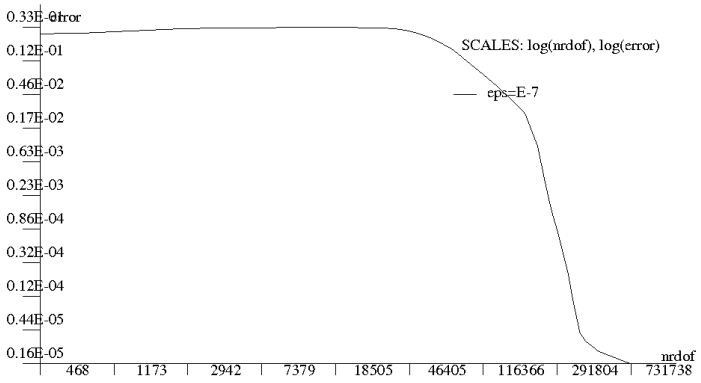

Figure 23: $\epsilon = 10^{-7}$: convergence history.

instance,

$$\int_e [\phi_K]\delta\hat{u}_{hp}\, ds = 0 \tag{5.70}$$

for every function $\delta\hat{u}_{hp}$ from the trial space for edge e. Here $[\phi_K]$ denotes the jump of the element error representation functions ϕ_K across the edge.

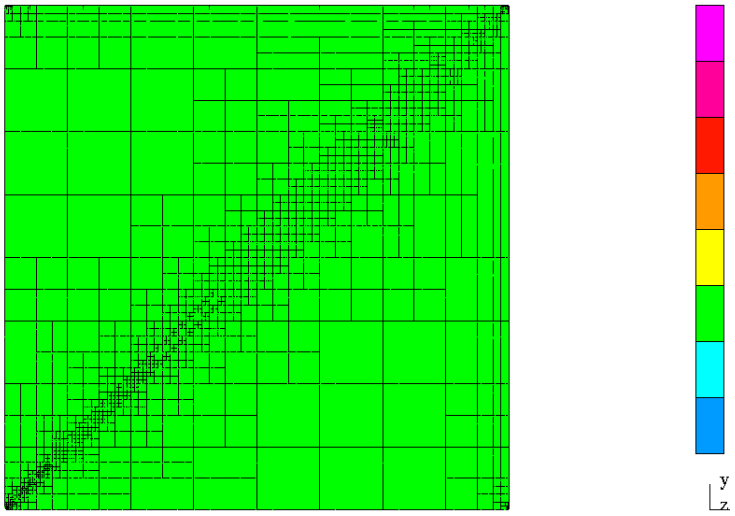

Figure 24: $\epsilon = 10^{-7}$: hp mesh after 45 mesh refinements.

An explicit computation of the energy norm in multi-dimensions does not seem possible, and we will have to resort to a qualitative analysis only. This, along with the orthogonality property above should suffice for proving L^2-error estimates and, hopefully, robustness in ϵ clearly indicated by the numerical results.

Current and future work. Besides the theoretical work, our current effort focuses on applying the DPG method to nonlinear problems - Burgers' and compressible Navier-Stokes equations. We hope to present new exciting results in a forthcoming paper.

References

[1] I. Babuška. Error-bounds for finite element method. *Numer. Math*, 16, 1970/1971.

[2] C.L. Bottasso, S. Micheletti, and R. Sacco. The discontinuous Petrov-Galerkin method for elliptic problems. *Comput. Methods Appl. Mech. Engrg.*, 191:3391–3409, 2002.

[3] C.L. Bottasso, S. Micheletti, and R. Sacco. A multiscale formulation of the discontinuous Petrov-Galerkin method for advective-diffusive problems. *Comput. Methods Appl. Mech. Engrg.*, 194:2819–2838, 2005.

[4] P. Causin and R. Sacco. A discontinuous Petrov-Galerkin method with Lagrangian multipliers for second order elliptic problems. *SIAM J. Numer. Anal.*, 43, 2005.

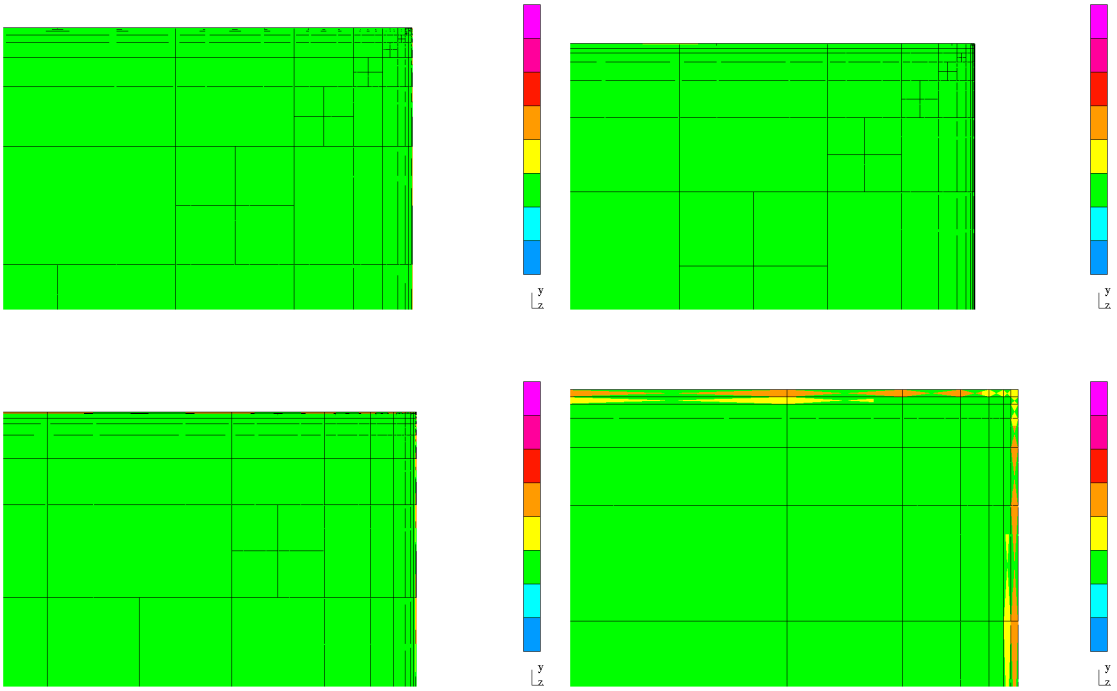

Figure 25: $\epsilon = 10^{-7}$: hp mesh after 45 mesh refinements, $10^2 - 10^5\times$ zooms on the north-east corner.

[5] P. Causin, R. Sacco, and C.L. Bottasso. Flux-upwind stabilization of the discontinuous Petrov-Galerkin formulation with Lagrange multipliers for advection-diffusion problems. *M2AN Math. Model. Numer. Anal.*, 39:1087–1114, 2005.

[6] L. Demkowicz. *Computing with hp Finite Elements. I.One- and Two-Dimensional Elliptic and Maxwell Problems*. Chapman & Hall/CRC Press, Taylor and Francis, October 2006.

[7] L. Demkowicz and J. Gopalakrishnan. A class of discontinuous Petrov-Galerkin methods. Part I: The transport equation. *Comput. Methods Appl. Mech. Engrg.*, 2009. in review, see also ICES Report 2009-12.

[8] L. Demkowicz and J. Gopalakrishnan. A class of discontinuous Petrov-Galerkin methods. Part II: Optimal test functions. Technical Report 16, ICES, 2009.

[9] L. Demkowicz, J.T. Oden, and R. Rachowicz. A new finite element method for solving compressible Navier-Stokes equations based on an operator splitting method and hp adaptivity. *Comput. Methods Appl. Mech. Engrg.*, 84:275–326, 1990.

[10] J.M. Melenk. *hp-Finite Element Methods for Singular Perturbations*. Springer, Berlin, 2002.

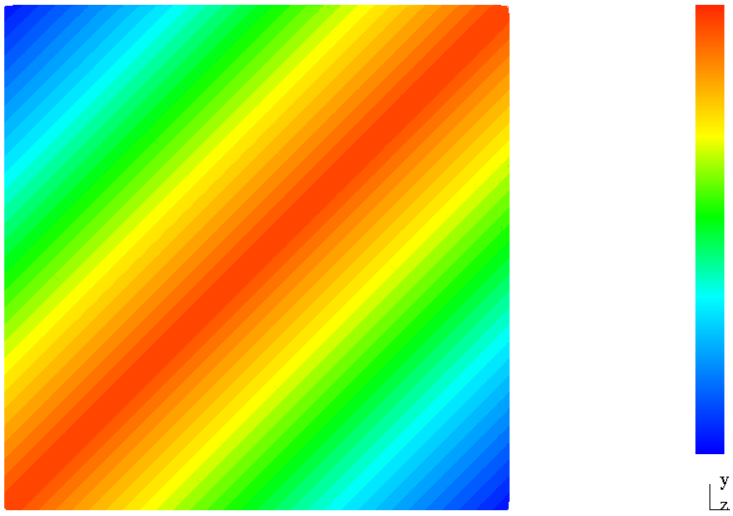

Figure 26: $\epsilon = 10^{-7}$: velocity u.

Figure 27: $\epsilon = 10^{-7}$: velocity u, 10^5 zoom on the north-east corner.

[11] J.T. Oden and L.F. Demkowicz. *Applied Functional Analysis for Science and Engineering*. Chapman & Hall/CRC Press, Boca Raton, 2010. Second edition.

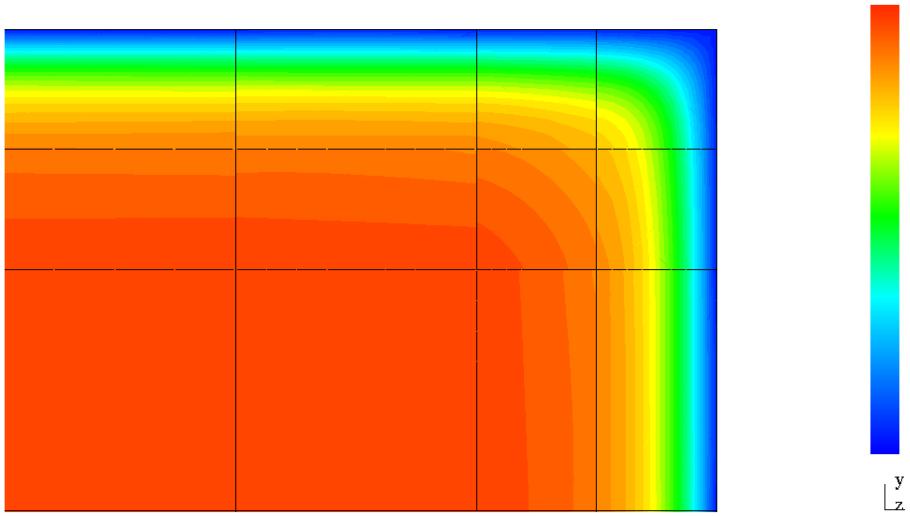

Figure 28: $\epsilon = 10^{-7}$: velocity u, 10^6 zoom on the north-east corner with the mesh.

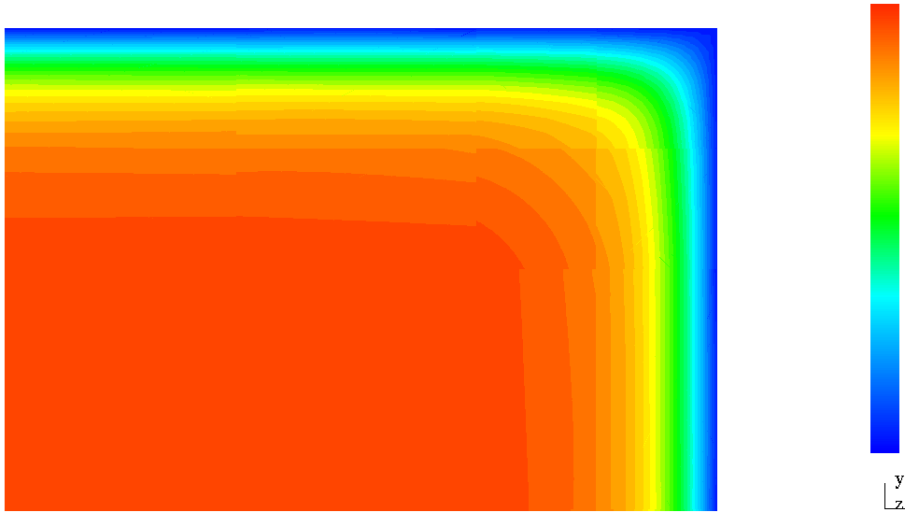

Figure 29: $\epsilon = 10^{-7}$: velocity u, 10^6 zoom on the north-east corner without the mesh.

[12] 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, 1996.

A Relation Between 1D Spectral and FE Energy Norms

In this section, we show that the global energy norm corresponding to norm (2.27), with a proper choice of constants β_k, is bounded below with the corresponding spectral norm of the solution premultiplied with order 1 constant.

The local variational problems for determining optimal test functions look as follows.

$$\begin{cases} \displaystyle\int_{x_{k-1}}^{x_k} \alpha\tau'\delta\tau' + \tau(x_k)\delta\tau(x_k) \\ \displaystyle= \frac{1}{\epsilon}\int_{x_{k-1}}^{x_k} \sigma_k\delta\tau + \int_{x_{k-1}}^{x_k} u_k\delta\tau' - (\hat{u}\delta\tau)|_{x_{k-1}}^{x_k} \qquad \forall\delta\tau \\ \displaystyle\int_{x_{k-1}}^{x_k} \alpha v'\delta v' + \beta_k v(x_k)\delta v(x_k) \\ \displaystyle= \int_{x_{k-1}}^{x_k} \sigma_k v' - (\hat{\sigma}\delta v)|_{x_{k-1}}^{x_k} - \int_{x_{k-1}}^{x_k} u_k v' + (\hat{u}\delta v)|_{x_{k-1}}^{x_k} \quad \forall\delta v \end{cases} \tag{A.71}$$

For each flux unknown $\hat{\sigma}(x_k)$, we have an optimal test function which spans across neighboring elements (x_{k-1}, x_k) and (x_k, x_{k+1}). For the first flux $\hat{\sigma}(0)$, the corresponding test function spans over the first element and, similarly, for the last flux $\hat{\sigma}(1)$, the corresponding test function spans over the last element only. Variational problem (A.71) leads to the following differential equations and boundary conditions for the optimal test functions.

$$\begin{cases} -(\alpha\tau')' &= \dfrac{1}{\epsilon}\sigma_k - u_k' \\ \alpha(x_k)\tau'(x_k) + \beta_k\tau(x_k) &= u_k(x_k) - \hat{u}(x_k) \\ -\alpha(x_{k-1})\tau'(x_{k-1}) &= -u_k(x_{k-1}) + \hat{u}(x_{k-1}) \end{cases} \tag{A.72}$$

and

$$\begin{cases} -(\alpha v')' &= -\sigma_k' + u_k' \\ v'(x_k) + \beta_k v(x_k) &= \sigma_k(x_k) - \hat{\sigma}(x_k) - u_k(x_k) + \hat{u}(x_k) \\ -v'(x_{k-1}) &= -\sigma_k(x_{k-1}) + \hat{\sigma}(x_{k-1}) + u_k(x_{k-1}) - \hat{u}(x_{k-1}) \end{cases} \tag{A.73}$$

Solving for the optimal test functions, we obtain the following formula for the energy norm

$$\begin{aligned} \|(\boldsymbol{\sigma}, \boldsymbol{u}, \hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{u}})\|_E^2 = \\ \sum_{k=1}^{N}\left[\|\frac{1}{\epsilon}\int_{x_{k-1}}^{x} \sigma_k(s)\,ds - u_k(x) + \hat{u}(x_{k-1})\|_{1/\alpha}^2 + \|-\sigma_k(x) + u_k(x) + \hat{\sigma}(x_{k-1}) - \hat{u}(x_{k-1})\|_{1/\alpha}^2\right] \\ + \sum_{k=1}^{N}\frac{1}{\beta_k}\left[|\frac{1}{\epsilon}\int_{x_{k-1}}^{x_k} \sigma_k(s)\,ds - \hat{u}|_{x_{k-1}}^{x_k}|^2 + |-\hat{\sigma}|_{x_{k-1}}^{x_k} + \hat{u}|_{x_{k-1}}^{x_k}|^2\right] \end{aligned} \tag{A.74}$$

with $\hat{u}(0) = \hat{u}(1) = 0$. More precisely,

$$\begin{aligned} \boldsymbol{\sigma} &= (\sigma_1, \ldots, \sigma_N) \\ \boldsymbol{u} &= (u_1, \ldots, u_N) \\ \hat{\boldsymbol{\sigma}} &= (\hat{\sigma}(0), \hat{\sigma}(x_1), \ldots, \hat{\sigma}(1)) \\ \hat{\boldsymbol{u}} &= (\hat{u}(x_1), \ldots, \hat{u}(x_{N-1})) \end{aligned} \tag{A.75}$$

In order to simplify the notation, we will drop indices in field unknowns $\sigma_k(x) = \sigma$ and $u_k(x) = u$ and some of the independent and integration variables,

$$\begin{aligned} \|(\boldsymbol{\sigma}, \boldsymbol{u}, \hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{u}})\|_E^2 = \\ \sum_{k=1}^{N} \left[\| \frac{1}{\epsilon} \int_{x_{k-1}}^{x} \sigma - u + \hat{u}(x_{k-1}) \|_{1/\alpha}^2 + \| - \sigma + u + \hat{\sigma}(x_{k-1}) - \hat{u}(x_{k-1}) \|_{1/\alpha}^2 \right] \\ + \sum_{k=1}^{N} \frac{1}{\beta_k} \left[| \frac{1}{\epsilon} \int_{x_{k-1}}^{x_k} \sigma - \hat{u}|_{x_{k-1}}^{x_k}|^2 + | - \hat{\sigma}|_{x_{k-1}}^{x_k} + \hat{u}|_{x_{k-1}}^{x_k}|^2 \right] \end{aligned} \tag{A.76}$$

The element norms are *weighted* L^2-norms with weight $1/\alpha$ (inverse of the weight used for the norm of the test functions).

We start with a simple estimate,

$$\begin{aligned} |\frac{1}{\epsilon} \int_0^1 \sigma| &= | \sum_{k=1}^{N} \left(\frac{1}{\epsilon} \int_{x_{k-1}}^{x_k} \sigma - \hat{u}|_{x_{k-1}}^{x_k} \right) | \\ &\leq \sum_{k=1}^{N} \sqrt{\beta_k} \frac{1}{\sqrt{\beta_k}} \left| \frac{1}{\epsilon} \int_{x_{k-1}}^{x_k} \sigma - \hat{u}|_{x_{k-1}}^{x_k} \right| \\ &\leq (\sum_{k=1}^{N} \beta_k)^{\frac{1}{2}} \left(\sum_{k=1}^{N} \frac{1}{\beta_k} \left| \frac{1}{\epsilon} \int_{x_{k-1}}^{x_k} \sigma - \hat{u}|_{x_{k-1}}^{x_k} \right| \right)^{\frac{1}{2}} \end{aligned} \tag{A.77}$$

The estimate suggests choosing $\beta_k = h_k$. In the same way,

$$\begin{aligned} |\hat{\sigma}(1) - \hat{\sigma}(0)| &= | \sum_{k=1}^{N} (\hat{\sigma}|_{x_{k-1}}^{x_k} - \hat{u}|_{x_{k-1}}^{x_k}) | \\ &\leq \sum_{k=1}^{N} \sqrt{\beta_k} \frac{1}{\sqrt{\beta_k}} |\hat{\sigma}|_{x_{k-1}}^{x_k} - \hat{u}|_{x_{k-1}}^{x_k}| \\ &\leq (\sum_{k=1}^{N} \beta_k)^{\frac{1}{2}} \left(\sum_{k=1}^{N} \frac{1}{\beta_k} |\hat{\sigma}|_{x_{k-1}}^{x_k} - \hat{u}|_{x_{k-1}}^{x_k}|^2 \right)^{\frac{1}{2}} \end{aligned} \tag{A.78}$$

implies a similar estimate for $|\hat{\sigma}(1) - \hat{\sigma}(0)|$.

Assume now that $x \in (x_{k-1}, x_k), k > 1$. We have,

$$-\sigma(x) + u(x) + \hat{\sigma}(0) = -\sigma(x) + u(x) + \hat{\sigma}(x_{k-1}) - \hat{u}(x_{k-1}) + \sum_{l-1}^{k-1} \left[-\hat{\sigma}|_{x_{l-1}}^{x_l} + \hat{u}|_{x_{l-1}}^{x_l} \right] \tag{A.79}$$

which implies

$$\int_{x_{k-1}}^{x_k} \frac{1}{\alpha} | -\sigma + u + \hat{\sigma}(0)|^2 \leq 2 \int_{x_{k-1}}^{x_k} \frac{1}{\alpha} | -\sigma + u + \hat{\sigma}(x_{k-1}) + \hat{u}(x_{k-1})|^2 + 2 \frac{h_k}{\alpha_k} \left| \sum_{l-1}^{k-1} \left[-\hat{\sigma}|_{x_{l-1}}^{x_l} + \hat{u}|_{x_{l-1}}^{x_l} \right] \right|^2 \tag{A.80}$$

where $\alpha_k = \min_{x \in (x_{k-1}, x_k)} \alpha(x), k > 1$ is assumed to be finite.

We use the discrete Cauchy-Schwarz inequality to estimate the second term on the right-hand side,

$$\frac{h_k}{\alpha_k} \left| \sum_{l-1}^{k-1} \left[-\hat{\sigma}|_{x_{l-1}}^{x_l} + \hat{u}|_{x_{l-1}}^{x_l} \right] \right|^2 \leq \frac{h_k}{\alpha_k} \left(\sum_{l=1}^{k-1} \beta_l \right) \left(\sum_{l-1}^{k-1} \frac{1}{\beta_l} \left| -\hat{\sigma}|_{x_{l-1}}^{x_l} + \hat{u}|_{x_{l-1}}^{x_l} \right|^2 \right) \tag{A.81}$$

Taking into account that $\sum_{l=1}^{k-1} \beta_l = x_{k-1}$ and summing up the estimate in k, we get,

$$\begin{aligned} \int_0^1 \frac{1}{\alpha} | -\sigma + u + \hat{\sigma}(0)|^2 \quad & \leq 2 \sum_{k=1}^{N} \| -\sigma + u + \hat{\sigma}(x_{k-1}) + \hat{u}(x_{k-1}) \|_{1/\alpha}^2 \\ & + 2 \max_{k>1} \left| \frac{x_{k-1}}{\alpha_k} \right| \left(\sum_{k=1}^{N} \frac{1}{\beta_k} \left[-\hat{\sigma}|_{x_{k-1}}^{x_k} + \hat{u}|_{x_{k-1}}^{x_k} \right] \right) \end{aligned} \tag{A.82}$$

Notice that for $\alpha(x) = x$, term $\frac{x_{k-1}}{\alpha_k} = 1$.

We follow the same strategy to estimate the last term. Starting with identity for $x \in (x_{k-1}, x_k),\ k > 1$,

$$\frac{1}{\epsilon} \int_0^x \sigma - u(x) = \frac{1}{\epsilon} \int_{x_{k-1}}^x \sigma - u(x) + \hat{u}(x_{k-1}) + \sum_{l=1}^{k-1} \left[\frac{1}{\epsilon} \int_{x_{l-1}}^{x_l} \sigma - \hat{u}|_{x_{l-1}}^{x_l} \right] \tag{A.83}$$

we obtain,

$$\begin{aligned} \int_0^1 \frac{1}{\alpha} | \frac{1}{\epsilon} \int_0^x \sigma - u(x)|^2 \quad & \leq 2 \sum_{k=1}^{N} \| \int_{x_{k-1}}^x \sigma - u + \hat{u}(x_{k-1}) \|_{1/\alpha}^2 \\ & + 2 \max_{k>1} \left| \frac{x_{k-1}}{\alpha_k} \right| \left(\sum_{k=1}^{N} \frac{1}{\beta_k} \left[\frac{1}{\epsilon} \int_{x_{k-1}}^{x_k} \sigma - \hat{u}|_{x_{k-1}}^{x_k} \right]^2 \right) \end{aligned} \tag{A.84}$$

We can formulate now our finite result.

THEOREM 1

Consider an arbitrary $(\boldsymbol{\sigma}, \boldsymbol{u}, \hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{u}})$ for which norm (A.74) is finite. Set constants $\beta_k = h_k := x_k - x_{k-1}$ and

assume that the weight $\alpha(x)$ has been selected in such a way that $\min_{x\in(x_{k-1},x_k)}\alpha(x) =: \alpha_k$ is finite, for $k = 2,\ldots,N$. Let σ and u denote the unions of element functions σ_k, u_k. Then the spectral energy norm for $(\sigma, u, \hat{\sigma}(0), \hat{\sigma}(1))$ is bounded by by the FE energy norm of $(\boldsymbol{\sigma}, \boldsymbol{u}, \hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{u}})$,

$$\|\frac{1}{\epsilon}\int_0^x \sigma - u\|_{1/\alpha}^2 + \|-\sigma + u + \hat{\sigma}(0)\|_{1/\alpha}^2 + |\frac{1}{\epsilon}\int_0^1 \sigma(s)ds|^2 + |-\hat{\sigma}(1) + \hat{\sigma}(0)|^2 \leq C\|(\boldsymbol{\sigma}, \boldsymbol{u}, \hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{u}})\|_E^2 \quad \text{(A.85)}$$

where $C = (1 + 2\max\{1, \max_{k>1}\frac{x_{k-1}}{\alpha_k}\})$. For $\alpha(x) = x$, $C = 3$. ∎