ICES REPORT 14-36

December 2014

Robust Discontinuous Petrov Galerkin (DPG) Methods for Reaction–Dominated Diffusion

by

Leszek Demkowicz and Isaac Harari

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

Reference: Leszek Demkowicz and Isaac Harari, "Robust Discontinuous Petrov Galerkin (DPG) Methods for Reaction–Dominated Diffusion," ICES REPORT 14-36, The Institute for Computational Engineering and Sciences, The University of Texas at Austin, December 2014.

Robust Discontinuous Petrov Galerkin (DPG) Methods for Reaction–Dominated Diffusion

Leszek Demkowicz[a] and Isaac Harari[b]

[a] Institute for Computational Engineering and Sciences (ICES), University of Texas at Austin, Austin, TX 78712, USA, *leszek@ices.utexas.edu*

[b] School of Mechanical Engineering, Tel Aviv University, 69978 Ramat Aviv, ISRAEL, *harari@eng.tau.ac.il*

Abstract

Using the reaction-dominated diffusion problem, we investigate various DPG formulations in an attempt to construct a scheme that delivers robust discretization in the norm of our choice. In particular, we generalize the strategy of Heuer and Demkowicz [10] to the primal DPG formulation.

Keywords: Discontinuous Petrov-Galerkin Method, robust discretization, reaction-dominated diffusion.

1 Introduction

This work is a continuation of [10, 4] aiming at working out a general methodology for constructing robust discretizations for Singularly Perturbed Boundary-Value Problems (SPBVP) within the framework of the Discontinuous Petrov-Galerkin (DPG) Method with Optimal Test Functions, see [7, 9] and the literature therein. The methodology presented in [10, 4] was based on the so-called *ultra-weak variational formulation* based on a first order system setting. In this paper, we extend it to the so-called *primal DPG method* [8] based on the standard variational formulation for second order problems, and apply it to reaction-dominated diffusion.

Reaction-dominated diffusion. We shall focus on perhaps the simplest singularly perturbed problem:

$$\left\{ \begin{array}{ll} u = 0 & \text{on } \Gamma \,, \\ -\epsilon^2 \Delta u + c(x) u = f(x) & \text{in } \Omega \,. \end{array} \right. \tag{1.1}$$

Here $\Omega \subset I\!\!R^N, N = 2, 3$, is a Lipschitz domain with boundary Γ, the reaction coefficient $c(x)$ is a sufficiently regular function satisfying lower and upper bounds,

$$0 < c_0 \le c(x) \le c_1 < \infty \,, \tag{1.2}$$

$f(x)$ is a sufficiently regular "load", and the diffusion coefficient is $\epsilon \le 1$.

The bilinear form

$$(u,v)_{\epsilon^k} := \epsilon^k(\boldsymbol{\nabla} u, \boldsymbol{\nabla} v) + (cu, v), \quad k = 1,2,3\,,$$

defines a scalar product on $H^1(\Omega)$ with the corresponding norm,

$$\|u\|^2_{\epsilon^k} = (u,u)_{\epsilon^k}\,.$$

We use the standard notation $(\cdot,\cdot)$ for the $L^2(\Omega)$-inner product, with the corresponding norm $\|u\|^2 = (u,u)$.

Under favorable regularity assumptions on the reaction coefficient $c(x)$ and domain regularity, the solution u is bounded in norm $\|u\|_\epsilon$, the so-called *balanced norm* [12], *uniformly* in diffusion coefficient ϵ,

$$\|u\|_\epsilon \lesssim |||f|||\,.$$

Above, symbol $\lesssim$ indicates a *robust bound*, i.e. a bound with a constant *independent of* ϵ, and $|||f|||$ is an appropriate norm involving higher order derivatives of f, see [12] for details.

Problem (1.1) admits the standard variational formulation:

$$\left\{\begin{array}{l} u \in H^1_0(\Omega)\,, \\ (u,v)_{\epsilon^2} = \epsilon^2(\boldsymbol{\nabla} u, \boldsymbol{\nabla} v) + (cu,v) = (f,v) \quad v \in H^1_0(\Omega)\,. \end{array}\right. \tag{1.3}$$

Galerkin discretization. The standard (Bubnov-) Galerkin discretization of (1.3) is obtained by replacing $H^1(\Omega)$ with a finite-dimensional subspace $U_h \subset H^1(\Omega)$,

$$\left\{\begin{array}{l} u_h \in U_h\,, \\ \epsilon^2(\boldsymbol{\nabla} u_h, \boldsymbol{\nabla} v_h) + (cu_h, v_h) = (f, v_h) \quad v_h \in U_h\,. \end{array}\right. \tag{1.4}$$

Testing in (1.3) with $v_h \in U_h$, and subtracting from it (1.4), we obtain the *Galerkin orthogonality condition* in the energy inner product:

$$(u - u_h, v_h)_{\epsilon^2} = 0 \quad \forall v_h \in U_h\,. \tag{1.5}$$

The Galerkin orthogonality condition in the energy inner product implies that the Galerkin method delivers the best approximation error in the corresponding *energy norm* $\|u\|^2_{\epsilon^2} := (u,u)_{\epsilon^2}$. Indeed,

$$\begin{aligned} \|u-u_h\|^2_{\epsilon^2} &= (u-u_h, u-u_h) \\ &= (u-u_h, u-w_h) + \underbrace{(u-u_h, w_h-u_h)}_{=0} \\ &\le \|u-u_h\|_{\epsilon^2}\,\|u-w_h\|_{\epsilon^2}\,, \end{aligned} \tag{1.6}$$

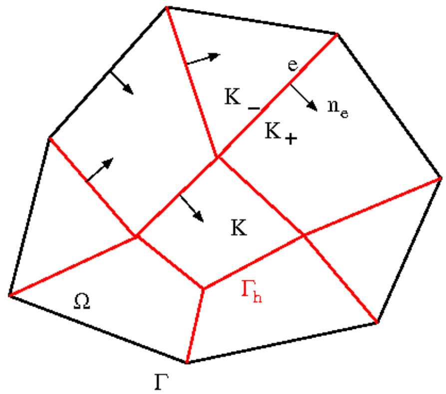
for an arbitrary $w_h \in U_h$. Dividing through by $\|u-u_h\|_{\epsilon^2}$, and taking infimum on the right-hand side with respect to $w_h \in U_h$, we obtain,

$$\|u-u_h\|_{\epsilon^2} \le \inf_{w_h \in U_h} \|u-w_h\|_{\epsilon^2}\,.$$

As u_h itself belongs to U_h, the inequality turns into equality,

$$\|u-u_h\|_{\epsilon^2} = \inf_{w_h \in U_h} \|u-w_h\|_{\epsilon^2}\,. \tag{1.7}$$

The Galerkin method delivers the orthogonal projection in the energy norm.



Elements:K
Edges:e
Skeleton:$\Gamma_h = \bigcup_K \partial K = \bigcup e$

Figure 1: Standard DG notation.

Robust discretization. A discretization of a singular perturbation problem is said to be *robust* in a norm U if

$$\|u - u_h\|_U \le C \inf_{w_h \in U_h} \|u - w_h\|_U \tag{1.8}$$

where the stability constant C is independent of the perturbation parameter. In other words, the actual approximation error is bounded by the best approximation error *uniformly* in the perturbation parameter ϵ. For the reaction-diffusion problem, in fact, for any hermitian positive-definite problem, the Galerkin method is automatically robust in the energy norm with $C = 1$.

Primal DPG formulation is obtained by "breaking" test functions in (1.3) and eliminating the zero boundary condition on test functions, see [8],

$$\left\{ \begin{array}{l} u \in H_0^1(\Omega),\, t \in H^{-1/2}(\Gamma_h)\,, \\ \epsilon^2(\boldsymbol{\nabla} u, \boldsymbol{\nabla}_h v) + (cu, v) - \langle t, v\rangle_{\Gamma_h} = (f, v) \quad v \in H^1(\Omega_h)\,. \end{array} \right. \tag{1.9}$$

Here $H^1(\Omega_h)$ denotes the *broken* Sobolev space of order one,

$$H^1(\Omega_h) := \{v = \{v_K\}_{K\in\mathcal{T}_h} \,:\, v_K \in H^1(K)\}\,, \tag{1.10}$$

and $\boldsymbol{\nabla}_h$ is the element-wise computed gradient. As usual, $\mathcal{T}_h$ denotes a triangulation of Ω with elements K. The price paid for breaking the test functions is the introduction of the new unknown - flux t defined on *mesh skeleton* Γ_h, see Fig. 1. For regular solutions,

$$t = \epsilon^2 \frac{\partial u}{\partial n_e}\,, \tag{1.11}$$

where n_e denotes the edge normal defining its orientation. The flux comes from an energy space defined on the mesh skeleton,

$$H^{-1/2}(\Gamma_h) := \{t = \{t_K\}_{K\in\mathcal{T}_h},\, t_K \in H^{1/2}(\partial K) \,:\, \exists w \in H(\mathrm{div}, \Omega) \text{ and } \mathrm{tr}_{\partial K} w = t_K\} \tag{1.12}$$

equipped with the quotient (minimum energy extension) norm, see [6, 8, **?**] for a detailed explanation. Finally, the bracket denotes the integral over the mesh skeleton,

$$\langle t, v\rangle_{\Gamma_h} := \sum_K \langle t_K, v_K\rangle_{\partial K} \,. \tag{1.13}$$

For sufficiently regular flux t, the term can be rewritten edge-wise,

$$\langle t, v\rangle_{\Gamma_h} = \sum_{e\subset\Gamma_h} \int_e t[v]_e \tag{1.14}$$

where we use the standard DG notation for the jump in test functions,

$$[v]_e := \begin{cases} v_{K_+} - v_{K_-} & e \subset \Omega \\ v & e \subset \Gamma \,. \end{cases} \tag{1.15}$$

It takes a slight modification of arguments presented in [8] to show that, with standard Sobolev norms, the problem is well-posed with mesh independent continuity and inf-sup constants, see also the general theory presented in [5].

DPG method. One of three possible characterizations of the DPG method is that of a *Minimum Residual Method*. Fixing a norm $\|v\|_V$ in the broken test space $V = H^1(\Omega_h)$, we introduce the corresponding DPG energy (residual) norm,

$$\|(u,t)\|_E := \sup_v \frac{|b((u,t),v)|}{\|v\|_V} \,. \tag{1.16}$$

The (ideal) DPG method delivers then the orthogonal projection in the DPG energy norm. Given finite-dimensional subspaces $U_h \subset H^1(\Omega), F_h \subset H^{-1/2}(\Gamma_h)$, we have:

$$(u_h, t_h) \in U_h \times F_h, \qquad \|(u-u_h, t-t_h)\|_E = \inf_{(w_h,r_h)\in U_h\times F_h} \|(u-w_h, t-r_h)\|_E \,. \tag{1.17}$$

Goal of this work. The purpose of this paper is to work out a systematic methodology for constructing a test norm $\|v\|_V$ in such a way that the corresponding DPG method is robust in a norm of our choice, and attempt to apply it to the model reaction-diffusion problem.

In this study we focus on the *balanced norm* $\|u\|_\epsilon$ and, an appropriately defined, corresponding norm $\|t\|_F$ for fluxes t. We strive thus for the result:

$$\|u-u_h\|_\epsilon + \|t-t_h\|_F \lesssim \inf_{w_h\in U_h} \|u-w_h\|_\epsilon + \inf_{r_h\in F_h} \|t-r_h\|_F \,. \tag{1.18}$$

The balanced norm is stronger in terms of derivatives then the standard energy norm, and we expect the corresponding approximate solutions to exhibit less oscillatory behavior. A special FE method that delivers convergence in the balanced norm was constructed by Lin and Stynes in [12].

2 Optimal Test Norm

Barret-Morton optimal test functions. We begin by recalling the concept of optimal test functions introduced by Barret and Morton in [2] and generalized by Demkowicz and Oden [11]. Related ideas were explored in [1, 13]. While Barret and Morton used the concept of the optimal test functions to symmetrize a problem, Demkowicz and Oden pushed it one step further, and used it to *change the norm* in which the solution converges or, more precisely, to execute an orthogonal projection in a desired norm.

We will explain the old idea using a general, abstract variational BVP:

$$\begin{cases} u \in U \,, \\ b(u,v) = l(v) \quad v \in V \,. \end{cases} \tag{2.19}$$

Here U, V are trial and test Hilbert spaces, and $b(u,v), l(v)$ are continuous bilinear and linear forms, resp.

Let $\|u\|_U^2 = (u,u)_U$ be the desired Euclidean norm in which we would like the solution to converge. For each trial function $u \in U$, we determine the corresponding optimal test function v_u that satisfies the variational problem:

$$\begin{cases} v_u \in V \,, \\ b(\delta u, v_u) = (\delta u, u)_U \quad \forall \delta u \in U \,. \end{cases} \tag{2.20}$$

Similarly to the DPG method, $u \to v_u$ executes a *trial-to-test operator* $T : U_h \ni u_h \to v_{u_h} \in U$. The optimal test space is then $V_h := TU_h$. If u_h is now the Petrov-Galerkin solution obtained with such optimal test functions, then, by the construction of the optimal test functions, the Galerkin orthogonality condition for form b translates into the orthogonality condition in terms of the desired inner product,

$$0 = b(u - u_h, v_{\delta u_h}) = (u - u_h, \delta u_h)_U \quad \forall \delta u_h \in U_h \,.$$

As explained in the Introduction, the orthogonality condition in inner product $(u, \delta u)_U$ implies then that the method delivers the best approximation error in the desired norm $\|u\|_U$.

Remark 1 Since the optimal test function v_u solves a global problem and does not have local support, it is unsuitable for FE computations. The original idea was to *localize* the optimal test space i.e. to find shape functions that span the optimal test space. This turned out to be possible in 1D but practically impossible in multidimensions. There is a strong logical link between the old optimal test functions and the DPG optimal test functions in context of ultra-weak formulation and graph norms, see [3]. We will return to this issue in Section 5.1. ∎

Optimal test norm. Given a trial norm $\|u\|_U$, we can define the corresponding optimal test norm[14] on the abstract level as:

$$\|v\|_{opt} := \sup_{u \in U} \frac{|b(u,v)|}{\|u\|_U} \,. \tag{2.21}$$

The Banach Closed Range Theorem implies that the optimal test norm and the original norm on u are then *in duality*, i.e. the DPG energy norm corresponding to the optimal test norm coincides with the original norm on u [14],

$$\|u\|_E = \|u\|_U \,.$$

Consequently, the DPG method delivers the orthogonal projection in $\|u\|_U$.

This can be also shown using an elementary argument. With the optimal test norm, both continuity and inf-sup constant for form $b(u,v)$ are unity. The argument for the continuity constant is straightforward,

$$|b(u,v)| = \|u\|_U \frac{|b(u,v)|}{\|u\|_U} \leq \|u\|_U \sup_u \frac{|b(u,v)|}{\|u\|_U} = \|u\|_U \, \|v\|_{opt} \,.$$

The Barret-Morton optimal test functions come handy to show the inf-sup condition. Indeed, the map $u \to v_u$ executes an isometry between $(U, \|u\|_U)$ and $(V, \|v\|_{opt})$,

$$\|v_u\|_{opt} = \sup_{\delta u} \frac{|b(\delta u, v_u)|}{\|\delta u\|_U} = \sup_{\delta u} \frac{|(\delta u, u)_U|}{\|\delta u\|_U} = \|u\|_U \,. \tag{2.22}$$

Consequently,

$$\sup_v \frac{|b(u,v)|}{\|v\|_{opt}} \geq \frac{|b(u,v_u)|}{\|v_u\|_{opt}} = \frac{|b(u,v_u)|}{\|u\|_U} = \frac{\|u\|_U^2}{\|u\|_U} = \|u\|_U \,.$$

Babuška's Theorem implies then that the Petrov-Galerkin method delivers the orthogonal projection in $\|u\|_U$.

Quasi-optimal test norms. Any norm that is robustly equivalent to the optimal test norm, will be called *quasi-optimal*,

$$\|v\|_{qopt} \approx \|v\|_{opt} \,, \tag{2.23}$$

meaning,

$$\|v\|_{qopt} \leq C_1 \|v\|_{opt} \quad \text{and} \quad \|v\|_{opt} \leq C_2 \|v\|_{qopt} \,, \tag{2.24}$$

with constants C_1, C_2 independent of the perturbation parameter ϵ. With a quasi-optimal test norm, the corresponding DPG method does not longer deliver an orthogonal projection in the desired norm $\|u\|$ but it is robust. Indeed, the inequalities above translate into corresponding inequalities in the DPG energy norms,

$$\|u\|_{E,opt} := \sup_v \frac{|b(u,v)|}{\|v\|_{opt}} \leq C_1 \sup_v \frac{|b(u,v)|}{\|v\|_{qopt}} = C_1 \|u\|_{E,qopt}$$

and, similarly,

$$\|u\|_{E,qopt} \leq C_2 \|u\|_{E,opt} \,.$$

Consequently,

$$\begin{aligned} \|u - u_h\|_U &= \|u - u_h\|_{E,opt} \\ &\leq C_1 \|u - u_h\|_{E,qopt} = C_1 \inf_{w_h} \|u - w_h\|_{E,qopt} \\ &\leq C_1 C_2 \inf_{w_h} \|u - w_h\|_{E,opt} = C_1 C_2 \inf_{w_h} \|u - w_h\|_U \,. \end{aligned}$$

Notice that equivalence constant C_1 affects the inf-sup stability condition, while constant C_2 affects the best approximation estimate.

A general strategy for determining quasi-optimal test norms. A systematic procedure for constructing quasi-optimal test norms in context of ultra-weak variational formulations has been proposed in [10]. We can generalize it to the primal DPG formulation as follows. We return to the definition (2.20) of Barret-Morton optimal test functions, and search for a test norm $\|v\|_{qopt}$ for which we can claim the robust stability estimate:

$$\|v_u\|_{qopt} \lesssim \|u\|_U \,. \tag{2.25}$$

We have then,

$$\begin{aligned} \|u-u_h\|_U^2 &= (u-u_h, u-u_h)_U = b(u-u_h, v_{u-u_h}) \\ &= \frac{b(u-u_h, v_{u-u_h})}{\|v_{u-u_h}\|_{qopt}} \|v_{u-u_h}\|_{qopt} \\ &\leq \sup_v \frac{|b(u-u_h, v)|}{\|v\|_{qopt}} \|v_{u-u_h}\|_{qopt} \\ &\lesssim \|u-u_h\|_E \, \|u-u_h\| \end{aligned} \tag{2.26}$$

Dividing through by $\|u-u_h\|$, we get the robust estimate:

$$\|u-u_h\| \lesssim \|u-u_h\|_E \tag{2.27}$$

where $\|\cdot\|_E$ is the energy norm corresponding to the quasi-optimal test norm, compare also (2.22). The practical point of this simple algebraic result is the relation of the stability of the adjoint problem (2.25) with the first of the inequalities (2.24). Contrary to the ultra-weak variational formulations and the first order systems setting though, we are dealing in (2.20) from the very beginning[1] with a distributional load.

An attempt to construct a quasi-optimal test norm for the reaction-diffusion problem. We return now to the primal DPG method for the reaction-diffusion problem. Let $b((u,t),v)$ represent the bilinear form corresponding to the primal DPG formulation,

$$\begin{aligned} &b((u,t),v) := \underbrace{\epsilon^2(\boldsymbol{\nabla} u, \boldsymbol{\nabla}_h v) + (u,v)}_{=:b_1(u,v)} \underbrace{-\langle t, v\rangle}_{=:b_2(t,v)} \,, \\ &u \in H_0^1(\Omega), t \in H^{-1/2}(\Gamma_h), v \in H^1(\Omega_h) \,. \end{aligned} \tag{2.28}$$

The test norm $\|v\|_{\epsilon^3}$ is naturally suggested by the continuity estimate for $b_1(u,v)$,

$$\begin{aligned} |\epsilon^2(\boldsymbol{\nabla} u, \boldsymbol{\nabla}_h v) + (u,v)| &\leq \epsilon^{1/2}\|\boldsymbol{\nabla} u\| \, \epsilon^{3/2}\|\boldsymbol{\nabla}_h v\| + \|c^{1/2}u\| \, \|c^{1/2}v\| \\ &\leq (\epsilon\|\boldsymbol{\nabla} u\|^2 + \|c^{1/2}u\|^2)^{1/2} \, (\epsilon^3\|\boldsymbol{\nabla}_h v\|^2 + \|c^{1/2}v\|^2)^{1/2} \end{aligned}$$

[1] Things are not that different in the end. After we reduced the first order system to a second order problem, we had to deal with distributional loads in [10] as well.

or,

$$|(u,v)_{\epsilon^2}| \le \|u\|_\epsilon \, \|v\|_{\epsilon^3} \, .$$

The continuity estimate implies immediately the second of inequalities (2.24). Indeed, the corresponding energy norm is bounded by the balanced norm,

$$\|u\|_E = \sup_v \frac{|(u,v)_{\epsilon^2}|}{\|v\|_{\epsilon^3}} \le \|u\|_\epsilon$$

which, by the duality of energy norm and the optimal test norm, is equivalent to the second inequality in (2.24).

Can we claim the first inequality as well, i.e. is this test norm quasi-optimal ?

The Barret-Morton optimal test function $v_u \in H^1(\Omega_h)$ corresponding to the balanced norm and a trial function $u \in H^1_0(\Omega)$ is defined by:

$$\left\{ \begin{array}{l} v_u \in H^1(\Omega_h) \, , \\ b((\delta u, \widehat{\delta t}), v_u) = (\delta u, u)_\epsilon \quad \forall \delta u \in H^1_0(\Omega), \, \forall \widehat{\delta t} \in H^{-1/2}(\Gamma_h) \, . \end{array} \right. \tag{2.29}$$

Orthogonality to fluxes $\widehat{\delta t}$ implies that $v_u \in H^1_0(\Omega)$, and we arrive at the Barret-Morton test function for the standard variational formulation,

$$\left\{ \begin{array}{l} v_u \in H^1_0(\Omega) \, , \\ (\delta u, v_u)_{\epsilon^2} = (\delta u, u)_\epsilon \quad \forall \delta u \in H^1_0(\Omega) \, . \end{array} \right. \tag{2.30}$$

In order to secure optimality of our stability estimates, we will rely on spectral calculus. Let $(\lambda_i, e_i), i = 1, 2, \ldots$, denote the eigenpairs of the Laplace operator with respected to the L^2 product weighted with coefficient $c(x)$,

$$\left\{ \begin{array}{l} e_i \in H^1(\Omega) \\ (\boldsymbol{\nabla} e_i, \boldsymbol{\nabla} v) = \lambda_i (c e_i, v) \quad \forall v \in H^1(\Omega) \end{array} \right.$$

where

$$0 < \lambda_1 \le \lambda_2 \le \ldots \le \lambda_n \to \infty \quad \text{and} \quad (c e_i, e_j) = \delta_{ij} \quad i, j = 1, \ldots, \infty \, .$$

With spectral decompositions of u and v_u,

$$u = \sum_{j=1}^{\infty} u_j e_j, \quad v_u = \sum_{j=1}^{\infty} v_j e_j \, ,$$

we can calculate coefficients v_j in terms of u_j,

$$v_j = \frac{\epsilon \lambda_j + 1}{\epsilon^2 \lambda_j + 1} u_j, \quad j = 1, \ldots, \infty \, .$$

The two norms in question are:

$$\|u\|_\epsilon^2 = \sum_{j=1}^{\infty} (\epsilon\lambda_j + 1)|u_j|^2\,, \qquad \|v_u\|_{\epsilon^3}^2 = \sum_{j=1}^{\infty} \frac{(\epsilon^3\lambda_j+1)(\epsilon\lambda_j+1)^2}{(\epsilon^2\lambda_j+1)^2}|u_j|^2\,.$$

Clearly,

$$\|v_u\|_{\epsilon^3} \lesssim \|u\|_\epsilon\,,$$

if and only if

$$\frac{(\epsilon^3\lambda+1)(\epsilon\lambda+1)^2}{(\epsilon^2\lambda+1)^2} \le C(\epsilon\lambda+1)\,,$$

for some $C > 0$, for every $\epsilon \le 1$ and $\lambda > 0$. Equivalently, setting $x = \epsilon\lambda$, we need,

$$\frac{(\epsilon^2 x+1)(x+1)^2}{(\epsilon x+1)^2} \le C(x+1) \quad x > 0\,,$$

or,

$$(\epsilon^2 x+1)(x+1) \le C(\epsilon x+1)^2 \quad x > 0\,.$$

Equivalently,

$$\epsilon^2 x^2 + (1+\epsilon^2)x + 1 \le C(\epsilon^2 x^2 + 2\epsilon x + 1)\,.$$

Whereas the first and the third term on the left-hand side are easily estimated robustly by the right-hand side, estimation of the second term would require a robust estimate

$$x \lesssim (\epsilon x+1)^2$$

or, equivalently (take $t = \sqrt{x}$),

$$1 \lesssim f(t) := \epsilon t + \frac{1}{t}\,.$$

Function $f(t)$ attains its minimum at $t = \sqrt{\epsilon}$, equal to $2\sqrt{\epsilon}$. Clearly, a robust estimate is thus impossible.

On the other side, $f(t) > 1$ in a relatively small interval:

$$\frac{1-\sqrt{1-4\epsilon}}{2\epsilon} < t < \frac{1+\sqrt{1-4\epsilon}}{2\epsilon}$$

which, for $\epsilon << 1$, translates into:

$$1 \le \epsilon\lambda \le \frac{1}{\epsilon}\,.$$

Thus, we lose the robust estimate only for relatively large frequencies $\lambda \ge 1/\epsilon$.

In order to guarantee a robust stability estimate for the adjoint problem, we need to decrease the coefficient in front of the 0-th order term,

$$\|v_u\|^2 = \epsilon^3\|\boldsymbol{\nabla} v\|^2 + \epsilon\|v\|^2\,.$$

The estimate can be established using the spectral techniques but we can also obtain it through elementary energy estimation. Substituting $\delta u = v_u$ in (2.30) we get,

$$\epsilon^2 \|\nabla v_u\|^2 + \|v_u\|^2 = \epsilon(\nabla v_u, \nabla u) + (v_u, u) \le \frac{1}{2}\epsilon^2 \|\nabla v_u\|^2 + \frac{1}{2}\|\nabla u\|^2 \le \frac{1}{2}\epsilon^2 \|v_u\|^2 + \frac{1}{2}\|u\|^2 .$$

This leads to:

$$\epsilon^2 \|\nabla v_u\|^2 + \|v_u\|^2 \le \|\nabla u\|^2 + \|u\|^2$$

or, upon multiplication by ϵ,

$$\epsilon(\epsilon^2 \|\nabla v_u\|^2 + \|v_u\|^2) \le \epsilon(\|\nabla u\|^2 + \|u\|^2) \le \epsilon\|\nabla u\|^2 + \|u\|^2 .$$

Remark 2 The quasi-optimal test norm that guarantees the robust inf-sup constant is somehow disappointing as it is simply obtained by scaling the original energy norm with ϵ. For the energy test norm, the corresponding DPG method reproduces simply the Galerkin method, i.e. the optimal test functions coincide with the corresponding trial functions. Scaling the test norm with ϵ does not change the test space. ▮

3 Practical Test Norms

Unfortunately, in the formula for the $\|v\|_{\epsilon^3}$ test norm:

$$\|v\|^2_{\epsilon^3} = \epsilon^3 \|\nabla v\|^2 + \|c^{1/2} v\|^2 , \tag{3.31}$$

the reaction term dominates the diffusion term, and resolving the optimal test functions corresponding to that norm is even more difficult than solving the original problem. We follow the idea from [10] and rescale the reaction term with a mesh dependent coefficient to arrive at:

$$\|v\|^2_{ropt} := \epsilon^3 \|\nabla v\|^2 + \min\{1, \frac{\epsilon^3}{h^2}\} \|c^{1/2} v\|^2 . \tag{3.32}$$

In other words, for large elements, the O(1) term is replaced with the smaller term ϵ^3/h^2 which makes the diffusion and reaction terms comparable and eliminates boundary layers in the optimal test functions. However, for smaller elements that are necessary to resolve boundary layers, we recover the original coefficient of order 1.

The rescaled test norm (3.32) is weaker than the $\|v\|_{\epsilon^3}$ norm, i.e.,

$$\|v\|_{ropt} \le \|v\|_{\epsilon^3}$$

which in turn implies that the corresponding DPG energy norm is *stronger* than the one corresponding to $\|v\|_{\epsilon^3}$ test norm. We improve thus on the stability side (we cannot show though that we regain the robustness) but we lose on the side of the best approximation error estimate. The loss happens where the elements are large, i.e. the solution is smooth. In places of large gradients (boundary layers), we maintain the optimality.

4 Numerical experiments

We begin with the manufactured solution example from [12]. In the unit square domain, we have a reaction coefficient

$$c = \left(1 + x^2y^2e^{xy/2}\right)$$

The source f and Dirichlet boundary conditions are chosen such that the solution is

$$u(x) = x^3\left(1+y^2\right) + \sin\left(\pi x^2\right) + \cos\left(\pi y/2\right) + (x+y)\left(e^{-2x/\epsilon} + e^{-2(1-x)/\epsilon} + e^{-3y/\epsilon} + e^{-3(1-y)/\epsilon}\right)$$

The solution is depicted in Fig. 2. The problem has been solved using rectangular elements of second order

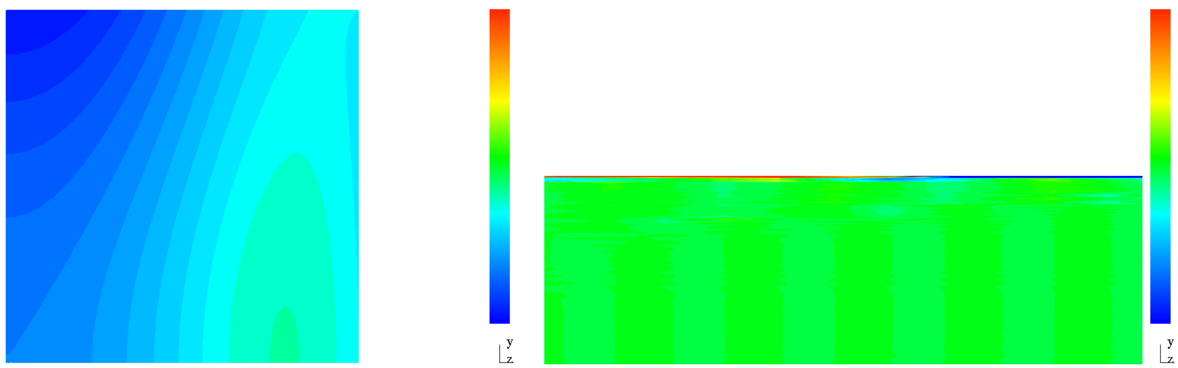


Figure 2: Left: Manufactured solution of Lin and Stynes for $\epsilon = 10^{-4}$. The function exhibits strong boundary layers invisible with this scale. Range: $0 - 0.6$. Right: zoom on the north boundary layer.

and linear fluxes. The optimal test functions and error representation function have been approximated with quartic elements, i.e. $\Delta p = 2$. We used the standard greedy strategy and h-adaptivity. A sample adaptive mesh is shown in Fig. 3. Fig. 4 shows the convergence history for the residual (DPG energy norm) and the solution u measured in the balanced norm $\epsilon = 10^0, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$. With the decreasing values of diffusion, the residual and error depart from each other on coarse meshes but converge to each other as the boundary layers are resolved. The residual always overestimates the error. Most importantly, the value of the residual can be used as a stopping criterion for terminating the adaptive process.

For $\epsilon = 10^{-5}$, the computation aborted by the end of the solution due to the round off error[2] in resolving the optimal test functions (error representation function).

Study of different test norms. In order to help realize the importance of the selection of the test norm, we have run the case for $\epsilon = 10^{-2}$ using four different test norms:

[2]We used double precision.

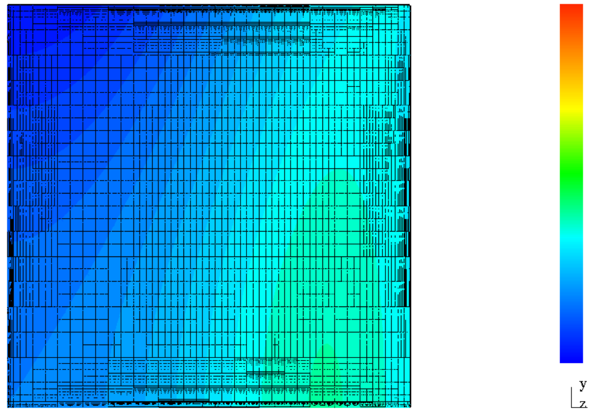


Figure 3: Optimal h-adaptive mesh and numerical solution for $\epsilon = 10^{-1}$.

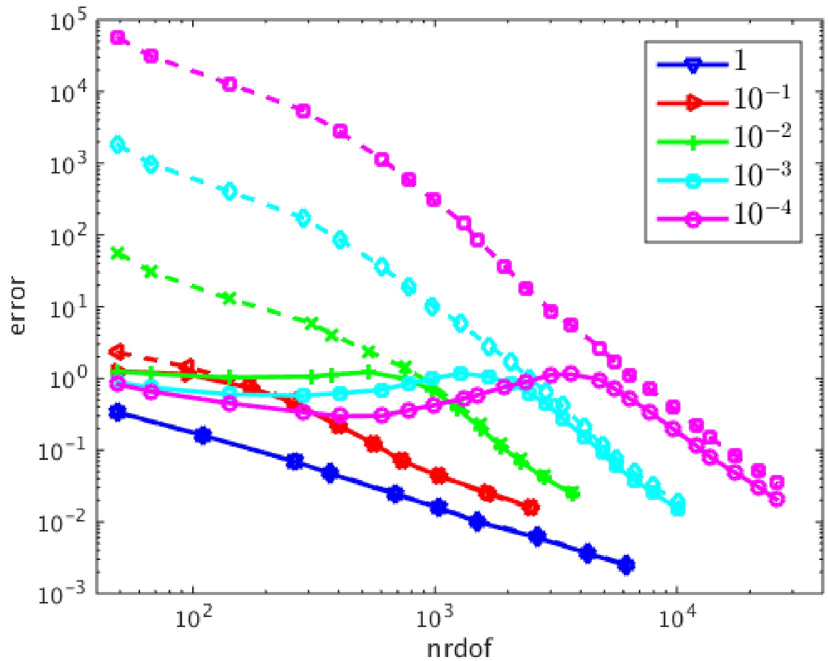


Figure 4: Convergence history for residual (dashed lines) and "balanced" error (solid lines) of u for different values of ϵ.

- standard H^1 norm:

$$\|v\|_1^2 = \|\boldsymbol{\nabla} v\|^2 + \|v\|^2 \tag{4.33}$$

- standard H^1 norm scaled with ϵ^3:

$$\|v\|_2^2 = \epsilon^3(\|\boldsymbol{\nabla} v\|^2 + \|v\|^2) \tag{4.34}$$

- standard energy inner product scaled with ϵ:

$$\|v\|_3^2 = \epsilon(\epsilon^2\|\boldsymbol{\nabla} v\|^2 + \|v\|^2) \tag{4.35}$$

- $\|v\|_{\epsilon^3}$ norm:

$$\|v\|_4^2 = \epsilon^3\|\boldsymbol{\nabla} v\|^2 + \|v\|^2 \tag{4.36}$$

In all cases, the error has been measured in the balanced norm. The results are reported in Fig. 5. The first two norms differ only by scaling and deliver the same error. For the first norm, as expected, the residual and error differ by two orders of magnitude. The scaling in the second norm is significant. Asymptotically, the energy norm (residual) and the balanced norm are very close to each other. For the third norm, the only one which satisfies robustly the inf-sup condition, the residual and error asymptotically coincide with each other. The fourth test norm performs worse but in the limit the residual also approaches the error.

Finally, we compare the convergence for the different test norms in Fig. 6. The rescaled $\|v\|_{\epsilon^3}$ test norm and the standard H^1 norm (first two test norms) deliver practically the same convergence. The rescaled energy norm performs significantly worse, and the under-resolved $\|v\|_{\epsilon^3}$ test norm (w/o rescaling) is the worst. Clearly, the under-resolved test functions make little sense. The optimal relation between the residual and error is lost and the overall convergence is the worst. The best convergence is obtained with the rescaled $\|v\|_{\epsilon^3}$ test norm and standard H^1 norm (the first and second norms deliver the same convergence). Both the rescaled $\|v\|_{\epsilon^3}$ test norm and the rescaled H^1 norm secure the residual and error to coincide with each other by the end of the convergence process.

A 3D example .

For illustration, we present the solution to the common test problem in 3D:

$$\left\{\begin{array}{rll} u & = 0 & \text{on } \Gamma \\ -\epsilon^2\Delta u + u & = 1 & \text{in } \Omega\,. \end{array}\right.$$

With $\epsilon \to 0$, the exact solution converges pointwise to unity and experiences a boundary layer around the whole boundary. Fig.7 shows the numerical solution for $\epsilon = 0.1$ and the mesh after seven mesh refinements with the corresponding convergence history for the residual.

Other tricks we can play with the DPG method. With the use of weights, we can emphasize the residual for a preselected subdomain. For reaction-diffusion problems, the relation between residual and error is very

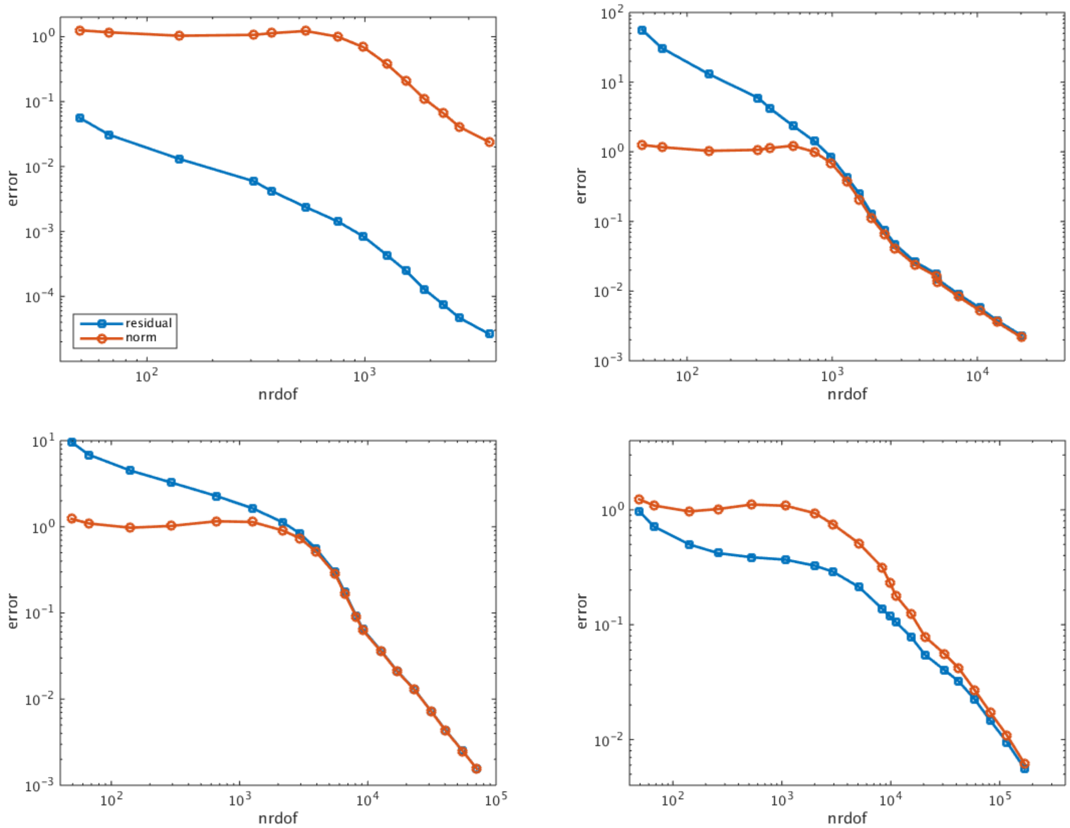


Figure 5: Lin-Stynes example. Convergence history for residual and “balanced” error of u for different test norms and $\epsilon = 10^{-2}$.

local which effectively enables a high accuracy solution in the preselected region. By weighting the rescaled optimal test norm with the weight;

$$w(x) = \begin{cases} 0.001 & x \in (0, \frac{1}{2})^2 \\ 1 & \text{elsewhere} \end{cases}$$

we “motivate” the DPG method to minimize the residual (and, therefore, the error) in the first quadrant of the domain. The concept is illustrated in Fig. 8. Majority of the refinements have taken place in the first quadrant delivering there a numerical solution with at least three digits of accuracy. The error elsewhere is large.

5 Discussion

We have generalized the strategy of Heuer and Demkowicz [10] to the primal DPG method. The strategy leads naturally to the old concept of Barret-Morton optimal test functions [2]. Contrary to the old ideas, the Barret-Morton functions are not used in computations but only analytically in the stability analysis for the

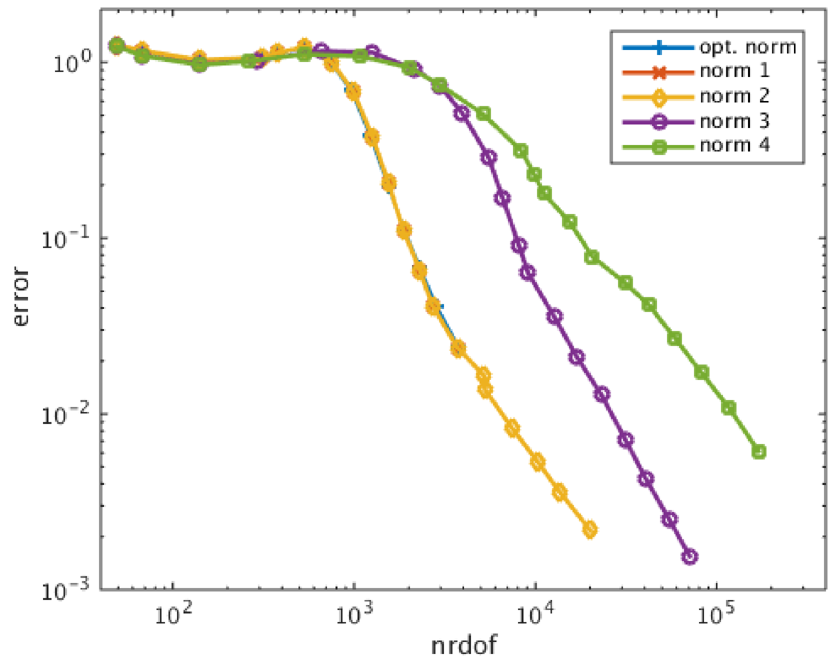


Figure 6: Lin-Stynes example. Convergence history for for different test norms and $\epsilon = 10^{-2}$.

dual problem.

Compared with ideas and results for convection-dominated diffusion [10], the results obtained for the reaction-diffusion problem are somehow surprising. For convection-diffusion problems we managed to construct test norms for which the inf-sup constant was robust. A subsequent rescaling of 0-th order terms to avoid unresolvable (by simple means) test functions with boundary layers, led to loss of robustness on the best approximation error side. There is a significant difference between the theoretical analysis of inf-sup constant (stability) and approximability. The stability analysis is always global. In contrary, the best approximation error in energy norm can be estimated locally, and local modifications of quasi-optimal test norm (rescaling) result in a local loss of robustness only. This remains true in the case of the primal DPG method as well.

We can illustrate this point with our model problem. We have,

$$
\begin{aligned}
(u - w_h, v)_{\epsilon^2} &= \sum_K \{\epsilon^2(\boldsymbol{\nabla}(u - w_h), \boldsymbol{\nabla} v)_K + (c(u - w_h), v)_K\} \\
&\leq \left(\sum_K \epsilon \|\boldsymbol{\nabla}(u - w_h)\|_K^2 + \gamma_K^{-1} \|c^{1/2}(u - w_h)\|_K^2 \right)^{1/2} \left(\sum_K \epsilon^3 \|\boldsymbol{\nabla} v\|_K^2 + \gamma_K \|c^{1/2} v\|_K^2 \right) .
\end{aligned}
$$

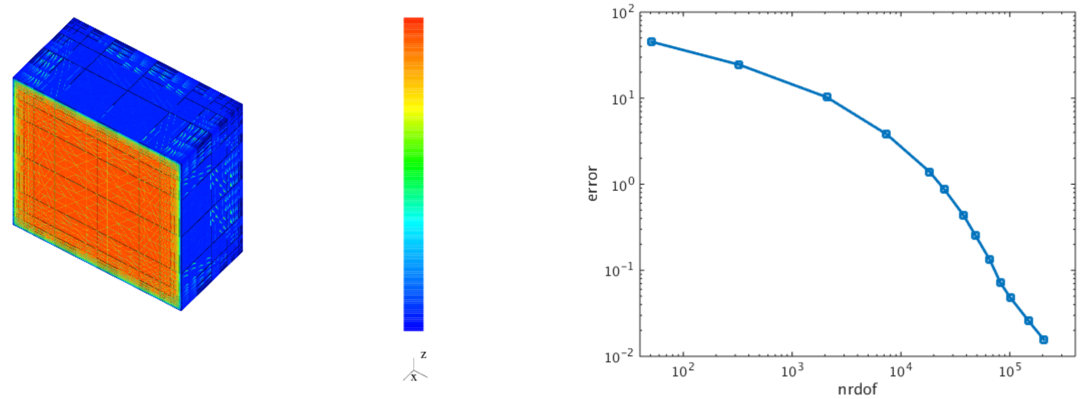


Figure 7: 3D example. Left: solution and mesh after 13 iterations. Right: convergence history for the residual.

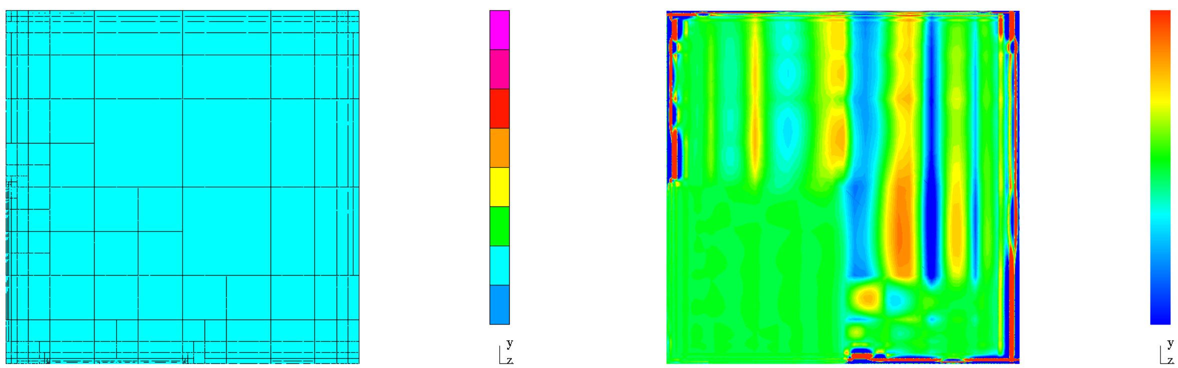


Figure 8: Optimal mesh and the corresponding pointwise error (range $(-0.001 - 0.001)$) with the "zoom" on the first quadrant of the domain.

where $\gamma_K =:= \min\{1, \frac{\epsilon^3}{h^2}\}$ is the rescaling term. This leads to the best approximation energy error estimate:

$$\inf_{w_h} \|u - w_h\|_E \le \inf_{w_h} \left(\sum_K \epsilon \|\boldsymbol{\nabla}(u - w_h)\|_K^2 + \gamma_K^{-1} \|c^{1/2}(u - w_h)\|_K^2 \right)^{1/2} .$$

The scaling constant γ_K^{-1} is large for large elements only. If the adaptive procedure works, these are employed only where the solution is smooth, and the corresponding best approximation error $\|c^{1/2}(u - w_h)\|^2$ is small. The conjecture is thus that the rescaling does not destroy the quasi-optimality of the test norm.

Contrary to convection-diffusion problems though, an attempt to construct a (provably) quasi-optimal test norm has led as back, modulo rescaling, to the original energy test norm. Means for scaling the 0-th order term in test norm are very limited (different powers of ϵ only) so we must conclude (whether we like it or not...) that we *cannot* construct a test norm with which we could prove the stability of the DPG in the balanced norm of Lin and Stynes who *were able* to construct a robust FE method in such a norm by different means, see [12].

5.1 Ultraweak variational formulation.

Our last attempt was to study the reaction-diffusion problem within the ultraweak setting. For simplicity of notation, we will assume now $c(x) = 1$. Introducing an extra variable $\sigma = \epsilon^\alpha \boldsymbol{\nabla} u$, we convert the original second order problem into a system of first order equations,

$$\left\{ \begin{aligned} \sigma - \epsilon^\alpha \boldsymbol{\nabla} u &= 0\,, \\ -\epsilon^{2-\alpha} \operatorname{div} \sigma + u &= f\,. \end{aligned} \right. \tag{5.37}$$

Here $\alpha \in [0,2]$. In context of the balanced norm, we are interested in case $\alpha = \frac{1}{2}$. Multiplying the equations with test functions τ and v, integrating by parts and incorporating boundary condition $u = 0$, we arrive at the so-called ultraweak variational formulation [7].

$$\left\{ \begin{array}{l} \sigma \in (L^2(\Omega))^2,\ u \in L^2(\Omega), \\ (\sigma, \tau + \epsilon^{2-\alpha} \boldsymbol{\nabla} v) + (u, \epsilon^\alpha \operatorname{div} \tau + v) = (f, v) \quad \forall \tau \in H(\operatorname{div}, \Omega),\ v \in H^1(\Omega)\,. \end{array} \right. \tag{5.38}$$

The construction of quasi-optimal test norm for the balanced norm leads to the stability analysis of the adjoint system:

$$\left\{ \begin{aligned} \tau + \epsilon^{2-\alpha} \boldsymbol{\nabla} v &= \sigma\,, \\ \epsilon^\alpha \operatorname{div} \tau + v &= u\,. \end{aligned} \right. \tag{5.39}$$

We can perform the analysis in two steps.

Case 1: $\sigma = 0, u \in L^2(\Omega)$.

Eliminating σ, we reduce the system to the original reaction-diffusion equation,

$$-\epsilon^2 \Delta v + v = u\,.$$

Multiplying by v, integrating over Ω, and integrating by parts, we obtain the standard energy estimate,

$$\epsilon^2 \|\boldsymbol{\nabla} v\|^2 + \|v\|^2 = (u, v) \leq \frac{1}{2}\|u\|^2 + \frac{1}{2}\|v\|^2$$

which leads to robust estimates for v,

$$\epsilon \|\boldsymbol{\nabla} v\|, \|v\| \lesssim \|u\|\,, \tag{5.40}$$

and in turn to estimates for τ,

$$\epsilon^{\alpha-1} \|\tau\|,\ \epsilon^\alpha \|\operatorname{div} \tau\| \lesssim \|u\|\,. \tag{5.41}$$

Case 2: $\sigma \in (L^2(\Omega))^2, u = 0$.

Eliminating v, we reduce the system to

$$\tau + \epsilon^2 \boldsymbol{\nabla}(\mathrm{div}\, \tau) = \sigma\,,$$

with homogeneous BC div $\tau = 0$ on Γ. Multiplication by τ and the same energy estimate as above, lead to the robust estimates:

$$\|\tau\|,\, \epsilon\|\mathrm{div}\, \tau\|,\, \epsilon^{2-\alpha}\|\boldsymbol{\nabla} v\|,\, \epsilon^{1-\alpha}\|v\| \lesssim \|\sigma\|\,. \tag{5.42}$$

One can use again the spectral analysis to demonstrate that all estimates above are sharp.

What have we learned?

First of all, we *cannot* be simultaneously robust in u and σ unless $\alpha = 1$, i.e. we are back to the Galerkin energy scaling. Secondly, no matter if we combine terms (5.42) or (5.40) and (5.41) into test norms, we end up with the same test spaces as for $\alpha = 1$. In other words, the extra relaxation coming from the integration by parts in the ultraweak formulation does not help. Standard Galerkin is the only provably robust method.

Last attempt. The last resource that we would like to mention is the scaled graph norm,

$$\|v\|_V^2 := \|\tau + \epsilon^{2-\alpha}\boldsymbol{\nabla} v\|^2 + \|\epsilon^{\alpha}\mathrm{div}\, \tau + v\|^2 + \beta(\|\tau\|^2 + \|v\|^2)\,. \tag{5.43}$$

It was shown in [3] that, with $\beta \to 0$, the DPG method delivers nothing else than the Barret-Morton optimal test functions corresponding to the L^2 trial norm,

$$\|(\sigma, u)\|_U^2 = \|\sigma\|^2 + \|v\|^2\,.$$

In other words, with $\beta \to 0$, the DPG method will converge to the L^2 projection. In particular, for $\alpha = \frac{1}{2}$, we recover the projection in the balanced norm. The main trouble with this approach is that the optimal test functions corresponding to this test norm develop strong boundary layers whose resolution requires special means like Shishkin meshes. We are back to techniques of Lin and Stynes.

Acknowledgments The work of L. Demkowicz was partially supported with the AFOSR under grant FA9550-12-1-0484. The work of I. Harari was partially supported by the Thomas J.R. Hughes Chair in Computational and Applied Mathematics and the J. Tinsley Oden Faculty Fellowship Research Program, while visiting the Institute for Computational Engineering and Sciences at the University of Texas at Austin, as well as the Diane and Arthur B. Belfer Chair in Mechanics and Biomechanics.

References

[1] P. E. Barbone and I. Harari. Nearly H^1-optimal finite element methods. *Comput. Methods Appl. Mech. Engrg.*, 190(43-44):5679–5690, 2001.

[2] J.W. Barret and K.W. Morton. Approximate symmetrization and petrov-galerkin methods for diffusion-convection problems. *Comput. Methods Appl. Mech. Engrg.*, 46:97–122, 1984.

[3] J. Chan, J. Gopalakrishnan, and L. Demkowicz. Global properties of DPG test spaces for convection-diffusion problems. Technical Report 5, ICES, 2013.

[4] J. Chan, N. Heuer, T. Bui-Thanh, and L. Demkowicz. A Robust DPG Method for Convection-dominated Diffusion Problems II: Adjoint Boundary Conditions and Mesh-dependent Test Norms. *Comput. Math. Appl.*, 67:771–795, 2014.

[5] L. Demkowicz, C. Carstensen, and J. Gopalakrishnan. The paradigm of broken test functions in DPG discretizations of elliptic second–order PDEs. 2014. in preparation.

[6] L. Demkowicz and J. Gopalakrishnan. Analysis of the DPG method for the Poisson problem. *SIAM J. Num. Anal.*, 49(5):1788–1809, 2011.

[7] L. Demkowicz and J. Gopalakrishnan. A class of discontinuous Petrov-Galerkin methods. Part II: Optimal test functions. *Numer. Meth. Part. D. E.*, 27:70–105, 2011. see also ICES Report 9/16.

[8] L. Demkowicz and J. Gopalakrishnan. A primal DPG method without a first order reformulation. *Comput. Math. Appl.*, 66(6):1058–1064, 2013.

[9] L. Demkowicz and J. Gopalakrishnan. *Recent Developments in Discontinuous Galerkin Finite Element Methods for Partial Differential Equations (eds. Xiaobing Feng, Ohannes Karakashian, Yulong Xing)*, volume 157, chapter An Overview of the DPG Method, pages 149–180. IMA Volumes in Mathematics and its Applications, 2014.

[10] L. Demkowicz and N. Heuer. Robust DPG method for convection-dominated diffusion problems. *SIAM J. Num. Anal*, 51:2514–2537, 2013. see also ICES Report 2011/13.

[11] L. Demkowicz and J. T. Oden. An adaptive characteristic Petrov-Galerkin finite element method for convection-dominated linear and nonlinear parabolic problems in one space variable. *Journal of Computational Physics*, 68(1):188–273, 1986.

[12] R. Lin and M. Stynes. A balanced finite element method for singularly perturbed reaction-diffusion problems. *SIAM J. Numer. Anal.*, 50(5):2729–2743, 2012.

[13] A. Nesliturk and I. Harari. The nearly-optimal Petrov-Galerkin method for convection-diffusion problems. *Comput. Methods Appl. Mech. Engrg.*, 192(22-24):2501–2519, 2003.

[14] J. Zitelli, I. Muga, L. Demkowicz, J. Gopalakrishnan, D. Pardo, and V. Calo. A class of discontinuous Petrov-Galerkin methods. Part IV: Wave propagation problems. *J. Comp. Phys.*, 230:2406–2432, 2011.