REDUCED TEST SPACES FOR DPG METHODS USING RECTANGULAR ELEMENTS

DOW DRAKE, JAY GOPALAKRISHNAN, AND AMMAR HARB

Abstract. This paper shows that the test spaces in discontinuous Petrov Galerkin (DPG) methods can be reduced on rectangular elements without affecting unisolvency or rates of convergences. One reduced case is obtained by decreasing the polynomial degree of a standard test space in both coordinate directions by one. A further reduction of test space by almost another full degree is possible if one is willing to implement a nonstandard test space. The error analysis of such cases is based on an extension of the second Strang lemma and an interpretation of the DPG method as a nonconforming method. The key technical ingredient in obtaining unisolvency is the identification of a discontinous piecewise polynomial on the element boundary that is orthogonal to all continuous piecewise polynomials of one degree higher.

1. Introduction

The development of discontinuous Petrov Galerkin (DPG) methods [6, 7] opened an avenue to construct methods whose stability is inherited from a residual minimization in a dual norm. When the minimization is carried out in a discrete (or approximate) dual norm, discrete stability depends on the choice of the finite dimensional spaces used. A few possible options for such choices on rectangular meshes are the subject of this short note.

We consider the primal DPG method introduced in [8] applied to a model Dirichlet problem. For triangular elements, it yields *(a)* an approximation u_h to the solution u of the Laplace equation, which is of degree k_u on each mesh element, and *(b)* an approximation to the flux q which is of degree k_q on each mesh edge. The degrees k_u and k_q form the so-called "trial space" degrees and the method's convergence rates depend on them. In addition, the method uses a discrete "test space," on which the stability of the method depends. The variational equations are imposed using these test functions, which are in a space $Y_h(K)$ on each mesh element K. In the DPG method, these test functions have no interelement continuity constraints.

We are interested in making $Y_h(K)$ as small as we can, while maintaining stability and rates of convergence. This has been the theme of at least two previous works. In the context of the so-called "ultraweak" DPG formulation on simplicial meshes, this issue was studied in [5]. In the context of the primal DPG method, which we shall be concerned with in this paper, the same issue was studied in [1] for triangular meshes. In fact, this study can be viewed as a continuation of [1] for rectangular element shapes. For the reader's convenience, the comparable findings of [1] are reproduced in Table 1. The table shows a comparison of a reference case (Case 1) with a reduced test space case (Case 3). We see that while maintaining the degrees k_u and k_q (and thus maintaining the same rate of convergence), it is possible to reduce the dimension of the element test space $Y_h(K)$.

Key words and phrases. Discontinuous Petrov Galerkin, DPG method, Strang lemma, unisolvency, quadrilateral.

This work was partially supported by the NSF grant DMS-1318916 and AFOSR grant FA9550-17-1-0090.

	k_u	k_q	dim $Y_h(K)$	$\|u - u_h\|_{H^1(\Omega)}$
Case 1 of [1]:	$k+1$	k	$(k+3)(k+4)/2$	$O(h^{k+1})$
Case 3 of [1], even k only:	$k+1$	k	$(k+2)(k+3)/2$	$O(h^{k+1})$

TABLE 1. Known results in the triangular case ($k \geqslant 0$)

Results of this paper are in the same spirit, but for rectangular elements. In this case, k_u represents the polynomial degree of u within each element in *each coordinate*. Precise definitions of the involved spaces appear later, but to provide an immediate summary of the results proved in later sections, we present Table 2.

	k_u	k_q	dim $Y_h(K)$	$\|u - u_h\|_{H^1(\Omega)}$
Case A	$k+1$	$k+1$	$(k+4)^2$	$O(h^{k+1})$
Case B	$k+1$	k	$(k+3)^2$	$O(h^{k+1})$
Case C	$k+1$	k	$(k+2)^2+1$	$O(h^{k+1})$

TABLE 2. Summary of the cases studied in this paper ($k \geqslant 0$) and the convergence rate findings for rectangular elements

Comparing Tables 1 and 2, we observe a few interesting differences between the triangular and rectangular cases. First, in the rectangular case, we are able to reduce the dimension of the space $Y_k(K)$ more than in the triangular case without altering the provable convergence rates. This requires the construction of a nonstandard polynomial test space, as we shall see. Second, unlike the triangular case, there are no even-odd discrepancies in the results of the rectangular case. The main new technical ingredient in this paper is motivated by the classical work of [12] and involves the identification of a null space on the boundary of a square element. This gives insight into a smaller test space that can be used without losing solvability.

We begin by describing the DPG method and the three cases of Table 2 precisely in Section 2. The motivation for the reduced degree cases becomes clearer in Section 3 where unisolvency of the various cases is discussed. Section 4 proves that each of the three cases produces a numerical approximation u_h that converges at the same rate. We conclude in Section 5 by reporting results of numerical experiments that match the theory.

2. Three choices of spaces on rectangular meshes

We study the DPG approximation to the Dirichlet problem

$$-\Delta u = f \qquad \text{on } \Omega, \tag{1a}$$

$$u = 0 \qquad \text{on } \partial\Omega. \tag{1b}$$

Here Ω is a bounded open Lipschitz polygon in $\mathbb{R}^2$ that is meshed by Ω_h, a geometrically conforming, finite element mesh of rectangles. Note that since the boundary $\partial\Omega$ is Lipschitz, Ω is not on both sides of $\partial\Omega$ at any point of $\partial\Omega$. Let $h = \max_{K\in\Omega_h} \operatorname{diam}(K)$. We assume for simplicity that all element sizes are comparable, i.e., there are fixed constants c_1 and c_2 such that $c_1 h \leqslant \operatorname{diam}(K) \leqslant c_2 h$ for all $K \in \Omega_h$. Finally, let $\partial\Omega_h$ denote the collection of all element boundaries ∂K for all elements K in Ω_h, 2.1. **The formulation and the method.** It is based on the following variational formulation [4, 8]: Find $(u, \hat{q}_n) \in X_0 \times \hat{X}$ satisfying

$$(\operatorname{grad} u, \operatorname{grad} v)_{\Omega_h} - \langle \hat{q}_n, v\rangle_{\partial\Omega_h} = (f, v)_{\Omega_h}, \quad \forall v \in Y, \tag{2}$$

where

$$\begin{aligned} X_0 &= H_0^1(\Omega), \\ Y &= H^1(\Omega_h) = \{v : \ v|_K \in H^1(K), \ \forall K \in \Omega_h\} \equiv \prod_{K \in \Omega_h} H^1(K), \\ \hat{X} &= H^{-1/2}(\partial\Omega_h) = \Big\{ \eta \in \prod_K H^{-1/2}(\partial K) : \ \exists\, r \in H(\operatorname{div}, \Omega) \text{ such that} \\ &\qquad\qquad \eta|_{\partial K} = r \cdot n|_{\partial K}, \quad \forall\, K \in \Omega_h \Big\}. \end{aligned}$$

Here and elsewhere n denotes the unit outward normals on the boundary of mesh elements. Above, the derivatives are calculated element by element, and we have used the following abbreviated notations for summing over mesh elements:

$$(r, s)_{\Omega_h} = \sum_{K \in \Omega_h} (r, s)_K, \qquad \langle \ell, w\rangle_{\partial\Omega_h} = \sum_{K \in \Omega_h} \langle \ell, w\rangle_{1/2, \partial K},$$

where $(\cdot, \cdot)_K$ denotes the $L^2(K)$-inner product and $\langle \ell, \cdot\rangle_{1/2,\partial K}$ denotes the action of a functional ℓ in $H^{-1/2}(\partial K)$. The space $\hat{X} = H^{-1/2}(\partial\Omega_h)$ is normed, as in [12], by

$$\|\hat{r}_n\|_{\hat{X}} = \inf\big\{ \|r\|_{H(\operatorname{div},\Omega)} : \ r \in H(\operatorname{div}, \Omega) \text{ such that } \hat{r}_n|_{\partial K} = r \cdot n|_{\partial K}\ \forall\, K \in \Omega_h \big\}. \tag{3}$$

Throughout, all function spaces are over $\mathbb{R}$. The wellposedness of this formulation was first proved in [8]. The proof was later considerably simplified in [4].

The DPG discretization uses finite element subspaces $X_{h,0} \subset X_0$, $\hat{X}_h \subset \hat{X}$, and $Y_h \subset Y$. Setting

$$b(\,(w, \hat{r}_n),\ v\,) = (\operatorname{grad} w, \operatorname{grad} v)_{\Omega_h} - \langle \hat{r}_n, v\rangle_{\partial\Omega_h}, \tag{4}$$

we can formulate the equations of the DPG method as follows: Find $e_h \in Y_h$, $u_h \in X_{h,0}$, and $\hat{q}_{n,h} \in \hat{X}_h$ satisfying

$$(e_h, v)_Y + b(\,(u_h, \hat{q}_{n,h}), v) = (f, v)_{\Omega_h} \qquad \forall v \in Y_h, \tag{5a}$$

$$b(\,(w, \hat{r}_n), e_h) = 0 \qquad \forall w \in X_{h,0},\ \hat{r}_n \in \hat{X}_h. \tag{5b}$$

Here $(r, s)_Y = (\operatorname{grad} r, \operatorname{grad} s)_{\Omega_h} + (r, s)_{\Omega_h}$ denotes the Y-inner product. We will also often write

$$X = X_0 \times \hat{X}, \qquad X_h = X_{h,0} \times \hat{X}_h.$$

2.2. **Three cases.** We want to study the effect of various finite element space choices in (5). To this end, we define certain standard spaces of polynomials [11] and one non-standard polynomial space [12]. Let $P_k(D)$ to be the set of polynomials of total degree at most k, restricted to the domain D, whenever $k \geqslant 0$, and the empty set when $k < 0$. Let $Q_{k,k}(D)$ be the space of polynomials of the form

$$p(x, y) = \sum_{0 \leqslant i,j \leqslant k} c_{ij} x^i y^j$$

on D. For the interface variables we need local polynomial spaces defined on the element boundary ∂K:

$$P_r(\partial K) = \{p \in L^2(\partial K) : p|_E \in P_r(E) \text{ for all edges } E \text{ of } K\},$$

Also set $Q_{k,k}(\Omega_h) = \{v : v|_K \in Q_{k,k}(K) \text{ for all } K \in \Omega_h\}$ and set $P_k(\partial\Omega_h)$ to the set of functions v on $\cup_{K\in\Omega_h}\partial K$ having the property that $v|_E \in P_k(E)$ for all edges of ∂K and for all $K \in \Omega_h$. Using these, the discrete spaces $X_h = X_{h,0} \times \hat{X}_h$ are set below for various cases. In all cases,

$$Y_h = \prod_{K\in\Omega_h} Y_h(K)$$

but the element space $Y_h(K)$ is not the same for all cases.

For any integer $k \geqslant 0$, we consider these three cases:

Case A	Case B	Case C
$X_{h,0} = Q_{k+1,k+1}(\Omega_h) \cap X_0$	$X_{h,0} = Q_{k+1,k+1}(\Omega_h) \cap X_0$	$X_{h,0} = Q_{k+1,k+1}(\Omega_h) \cap X_0,$
$\hat{X}_h = P_{k+1}(\partial\Omega_h) \cap \hat{X}$	$\hat{X}_h = P_k(\partial\Omega_h) \cap \hat{X}$	$\hat{X}_h = P_k(\partial\Omega_h) \cap \hat{X},$
$Y_h(K) = Q_{k+3,k+3}(K)$	$Y_h(K) = Q_{k+2,k+2}(K)$	$Y_h(K) = Q^{+1}_{k+1,k+1}(K),$

where $Q^{+1}_{k,k}(K)$ is the pullback of the following space defined on the reference element $\hat{K} = [0,1] \times [0,1]$.

Definition 2.1. If k is odd, set

$$\rho(x,y) = [x(1-x) - y(1-y)]\left[(x(1-x))^{(k-1)/2} + (y(1-y))^{(k-1)/2}\right],$$

and if k is even, set

$$\rho(x,y) = [x(1-x) - y(1-y)](2x-1)(2y-1)\left[(x(1-x))^{(k-2)/2} + (y(1-y))^{(k-2)/2}\right].$$

Then define

$$Q^{+1}_{k,k}(\hat{K}) = Q_{k,k}(\hat{K}) + \text{span}(\rho).$$

Definition 2.1 is motivated by a similar space used in [12] which they use for different purposes (to construct a hybrid method that is not a DPG method). When there is potential for confusion between the cases, we will explicitly indicate the case as a superscript in the notation for underlying spaces and operators, e.g., X_h^{C} and Y_h^{C} denote the spaces of Case C. Note that the spaces in the Cases A and B are standard. Case C however requires implementation of the nonstandard space $Q^{+1}_{k,k,}(\hat{K})$. This is not difficult since the space is obtained by adding to the standard $Q_{k,k}(\hat{K})$-space a single interior basis function (ρ). Case C is the most attractive practically, because it has the lowest number of degrees of freedom among the three cases. The rationale behind Definition 2.1 and the choice of spaces will become clear in the next section.

3. Unisolvency

In this section, we will show the unisolvency of the DPG method (5) in each of the abovementioned three cases. Although unisolvency alone does not give us a complete error analysis, it serves to motivate our choices of the spaces in the three cases. Let $B_h : X_h \to Y_h^*$ be the operator generated by the form in (4), i.e.,

$$(B_h x_h)(y) = b(x_h, y), \qquad \forall x_h \in X_h,\ y \in Y_h.$$

Similarly, let $\hat{B}_h : \hat{X}_h \to Y_h^*$ be defined by

$$(\hat{B}_h \hat{z}_h)(y) = -\langle \hat{z}_h, y\rangle_{\partial\Omega_h}, \qquad \forall \hat{z}_h \in \hat{X}_h,\ y \in Y_h. \tag{6}$$

It is obvious from the structure of the mixed method (5) that

$$(5) \text{ is uniquely solvable} \iff B_h \text{ is injective.} \tag{7}$$

Furthermore, by [1, Theorem 2.8],

$$B_h \text{ is injective} \iff \hat{B}_h \text{ is injective.} \tag{8}$$

Using these facts, we will show below that B_h, in each of the three cases (namely $B_h^{\mathrm{A}}, B_h^{\mathrm{B}}$, and B_h^{C}), is injective.

3.1. **The standard case (Case A).** Among the previously presented cases of DPG methods, we consider Case A to be the standard case for rectangular elements, because it is suggested by the analysis of [2].

Specifically, in [2, Lemma 4.3], they exhibited a bounded Fortin operator for rectangular meshes $\varPi_k : H^1(\varOmega_h) \to Q_{k+2,k+2}(\varOmega_h)$ satisfying

$$(\varPi_k z, v)_K = (z, v)_K \qquad \forall v \in Q_{k,k}(K), \tag{9}$$

$$\langle \varPi_k z, w\rangle_{\partial K} = \langle z, w\rangle_{\partial K} \qquad \forall w \in P_k(\partial K) \tag{10}$$

for all $K \in \varOmega_h$. This implies, by integration by parts element by element, that

$$b((w_h, r_{n,h}), \varPi_{k+1} v - v) = 0, \qquad \forall\, (w_h, r_{n,h}) \in Q_{k,k}(K) \times P_k(\partial K), \tag{11}$$

for all $v \in H^1(\varOmega_h)$. Equation (11) verifies a basic assumption in one of the known techniques [3, 9] of error analysis.

The injectivity of B_h^{A} in Case A follows from (11). Indeed, if $B_h^{\mathrm{A}} x_h = 0$, then $b(x_h, y_h) = 0$ for all $y_h \in Y_h$, so in particular, $b(x_h, \varPi_{k+1} y) = 0$. But this implies, due to (11), that $b(x_h, y) = 0$ for all $y \in Y$. Hence the wellposedness of the undiscretized problem implies $x_h = 0$.

As we shall see next (in the remaining cases), it is possible to reduce the space Y_h without losing the injectivity of B_h.

3.2. **Unisolvency of Case B.** In this case, we reduce the degree of $Y_h(K)$ by one. To prove that B_h remains injective, first observe from Case A that (11), with $k-1$ in place of k, implies

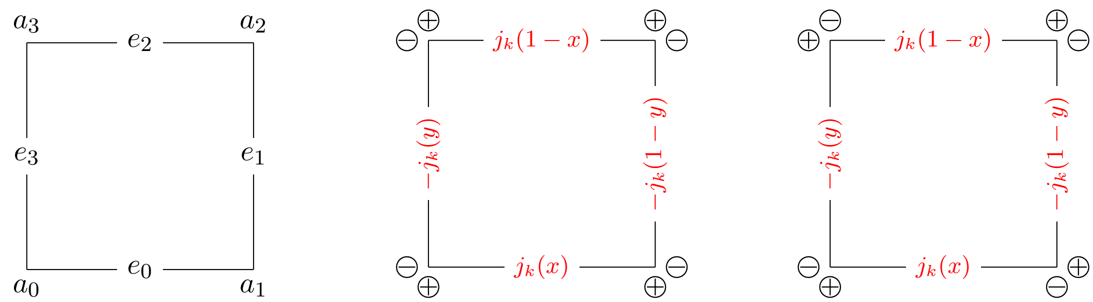
$$b((w_h, r_{n,h}), \varPi_k v - v) = 0, \tag{12}$$

for all $w_h \in Q_{k,k}(\varOmega_h) \cap X_0$, $r_{n,h} \in P_k(\partial\varOmega_h) \cap \hat{X}$ and $v \in H^1(\varOmega_h)$. Since the corresponding B_h is injective, by (8), the associated $\hat{B}_h$ is also injective, namely if $r_{n,h} \in P_k(\partial\varOmega_h)$ satisfies

$$\langle r_{n,h}, v_h\rangle_{\partial\varOmega_h} = 0, \qquad \forall v_h \in Q_{k+2,k+2}(\varOmega_h), \tag{13}$$

then $r_{n,h} = 0$. But this is exactly the same as the injectivity of $\hat{B}_h^{\mathrm{B}}$. Using (8) again, we conclude that B_h^{B} is injective.

3.3. **Unisolvency of Case C.** This case is more involved. The degree of $Y_h(K)$ is further reduced by one, but an extra dimension is added. To understand the need for this extra dimension, consider an $r_{n,h} \in P_k(\partial K)$, which as in (13), satisfies $\langle r_{n,h}, v_h\rangle_{\partial K} = 0$, but now only for all $v_h \in Q_{k+1,k+1}(K)$. This is a square system of equations for $r_{n,h}$ because $\dim P_k(\partial K)$ and the number of nontrivial equations both equal $4(k+1)$. Unfortunately, this system has a one-dimensional null space as seen in Lemma 3.1 below. The addition of an appropriate extra equation – which corresponds to the addition of an extra dimension to $Y_h(K)$ – will remove this null space, as we shall see below.



(A) Notations (B) Even k (C) Odd k

FIGURE 1. The unit square element and sign patterns of ν_k on its boundary

Let $J_k^{\alpha,\beta}(\hat{s})$ denote the family of Jacobi polynomials, indexed by degree k, orthogonal with respect to the weight function $\hat{\omega}(\hat{s}) = (1-\hat{s})^{\alpha}(1+\hat{s})^{\beta}$ on the interval $[-1,1]$. Let us recall the properties specific to the case $\alpha = \beta = 1$. Mapping to $[0,1]$, set

$$j_k(s) = J_k^{1,1}(2s-1), \qquad 0 \leqslant s \leqslant 1.$$

Then, with respect to the mapped weight $\omega(s) = s(1-s)$, we have

$$\int_0^1 \omega(s)\, j_k(s)\, p(s)\; ds = 0, \qquad \forall p \in P_{k-1}([0,1]). \tag{14}$$

Let $s_1, \ldots, s_{k-1}$ denote the roots of $j_k(s)$. Its well known that the nodes of the Gauss-Lobatto quadrature on $[0,1]$ are $0 = s_0 < s_1 < \cdots < s_{k-1} < s_k = 1$. If the corresponding weights are w_i, then the quadrature is exact for polynomials of degree at most $2k-1$, i.e.,

$$\int_0^1 \omega(s) p(s)\; ds = \sum_{i=0}^{k} w_i p(s_i), \qquad p \in P_{2k-1}([0,1]). \tag{15}$$

Next, consider the unit square $\hat{K} = [0,1] \times [0,1]$ in the (x,y)-plane with vertices $a_0 = (0,0)$, $a_1 = (1,0)$, $a_2 = (1,1)$, $a_3 = (0,1)$ and edges $e_0, \ldots, e_3$ as shown in Figure 1. By mapping j_k to the four edges of the unit square $\hat{K}$, we define a function ν_k on $\partial\hat{K}$ as follows.

$$\nu_k(x,y) = \begin{cases} j_k(x), & 0 < x < 1, \quad y = 0, \\ j_k(1-x), & 0 < x < 1, \quad y = 1, \\ -\,j_k(y), & x = 0, \quad 0 < y < 1, \\ -\,j_k(1-y), & x = 1, \quad 0 < y < 1. \end{cases}$$

The next result is essentially contained in [12, Lemma 8], but we state and prove it in a different form more suitable for our purposes.

Lemma 3.1. *Suppose $\mu \in P_{k-1}(\partial\hat{K})$ satisfies*

$$\int_{\partial\hat{K}} \mu v = 0, \qquad \forall v \in Q_{k,k}(K). \tag{16}$$

Then, there is a $c \in \mathbb{R}$ such that

$$\mu = c\,\nu_{k-1}.$$

Proof. Let $w \in Q_{k-2,k-1}(\hat{K})$. In (16), put $v = x(1-x)(1-y)w(x,y)$. Then

$$0 = \int_{\partial K} \mu v = \int_{e_0} x(1-x)(1-y)\mu(x,0)w(x,0)\,dx.$$

Hence, there is a scalar $c_0 \in \mathbb{R}$ such that $\mu(x,0) = c_0 j_{k-1}(x)$ – see (14). Arguing similarly on remaining edges, we find that there are constants $c_i \in \mathbb{R}$ such that $\mu|_{e_i} = c_i j_{k-1}$ for each edge e_i.

Next, let $v_0 \in P_k([0,1])$ satisfy $v_0(s_i) = 0$ for all $i = 1,\ldots,k$ and $v_0(s_0) = 1$. Then, putting $v(x,y) = v_0(x)v_0(y)$ in (16), and using (15), we obtain

$$\begin{aligned} 0 &= \int_{e_0 \cup e_3} \mu v = \int_0^1 c_0 j_{k-1}(x)v_0(x)v_0(0)\,dx + \int_0^1 c_3 j_{k-1}(y)v_0(0)v_0(y)\,dy \\ &= (c_0 + c_3)j_{k-1}(0)v_0(0)^2 w_0. \end{aligned}$$

Thus we obtain a condition at the vertex a_0, namely $c_0 + c_3 = 0$, i.e., the two limits of μ at a_0 are equal in magnitude and opposite in sign. Using a smilar argument at the remaining vertices, we find that sum of the two limits of μ at each vertex a_i must be zero.

Recall that $J_k^{1,1}$ is even or odd, depending on whether k is even or odd, respectively. Hence, fitting together the mapped orthogonal polynomials on each edge, reflecting about edge midpoints as needed to satisfy the sign change constraint at each vertex (see Figure 1), we find that μ must coincide with a scalar multiple of ν_{k-1}. □

Lemma 3.2. *Suppose ρ is any smooth scalar function on $\hat{K}$ that is linearly independent from $Q_{k,k}(\hat{K})$ and satisfies*

$$\int_{\partial\hat{K}} \nu_{k-1}\rho \neq 0. \tag{17}$$

Then the only $\mu \in P_{k-1}(\partial\hat{K})$ that satisfies

$$\int_{\partial\hat{K}} \mu v = 0 \qquad \forall v \in Q_{k,k}(\hat{K}) \oplus \mathrm{span}(\rho) \tag{18}$$

is the zero function.

Proof. By Lemma 3.1, equation (18) for all $v \in Q_{k,k}(\hat{K})$ implies that $\mu = c\nu_{k-1}$ for some $c \in \mathbb{R}$. Hence, putting $v = \rho$ in (18), we find that

$$c\int_{\partial\hat{K}} \nu_{k-1}\rho = 0.$$

Due to (17), this implies that $c = 0$, so $\mu \equiv 0$. □

Clearly, one can design many functions ρ that satisfy the conditions of Lemma 3.2. One possible choice is the one we have settled on in Definition 2.1. We conclude this section by summarizing the results in the next theorem.

Theorem 3.3. *In Cases A, B, and C, the DPG method* (5) *is uniquely solvable.*

Proof. We have already indicated the proofs in Cases A and B. To complete the proof for Case C, we verify that the ρ set in Definition 2.1 satisfies the conditions of Lemma 3.2. Indeed, for that ρ, the integral of $\nu_{k-1}\rho$ over all edges of $\partial\hat{K}$ are equal, and

$$\int_{\partial\hat{K}} \nu_{k-1}\rho = \begin{cases} 4\displaystyle\int_0^1 x(1-x)\, j_{k-1}(x)\, [x(1-x)]^{(k-1)/2}\,dx, & \text{if } k \text{ is odd,} \\ 4\displaystyle\int_0^1 x(1-x)\, j_{k-1}(x)\,(1-2x)\,[x(1-x)]^{(k-2)/2}\,dx, & \text{if } k \text{ is even.} \end{cases}$$

The functions $[x(1-x)]^{(k-1)/2}$ and $(1-2x)\,[x(1-x)]^{(k-2)/2}$, for odd and even k respectively, are polynomials of strict degree $k-1$, hence the integrals on the right hand side above cannot vanish. Thus having verified the conditions of Lemma 3.2, its conclusion implies that $\hat{B}_h^{\mathrm{C}}$ is injective. By (8), we conclude that the DPG method is uniquely solvable in Case C. $\square$

3.4. **A necessary condition for unisolvency.** DPG methods are usually designed with a rich enough space Y_h that makes B_h injective. We quantify this as a condition on the degrees of freedom that is necessary for unisolvency of (5):

Proposition 3.4. *The system* (5) *is uniquely solvable only if*

$$\dim(X_h) \leqslant \dim(Y_h). \tag{19}$$

Proof. Let $n_X = \dim(X_h)$ and $n_Y = \dim(Y_h)$. By (7), unisolvency of (5) implies that B_h is injective. Therefore, the $n_Y \times n_X$ stiffness matrix of B_h, denoted by $\mathtt{B}$, has the trivial nullspace. By the rank-nullity theorem, $\operatorname{rank}(\mathtt{B}) = n_X$. Now the result is obvious from the fact that $\operatorname{rank}(\mathtt{B}) \leqslant \min(n_X, n_Y)$. $\square$

Since we have already verified unisolvency of all the cases previously, the necessary condition (19) must hold in each case. To provide more insight into this condition, we now count the degrees of freedom to provide an alternate verification that the necessary condition holds in the two nonstandard cases we studied.

Proposition 3.5. *In Cases B and C on a mesh of n_K rectangular elements, we have*

$$\dim(Y_h) - \dim(X_h) \geqslant 2n_K - 1.$$

Proof. Decomposing Ω into disjoint connected components, it suffices to prove the stated inequality on each component. Hence without loss of generality, we assume that Ω is a multiply connected domain with m holes. Furthermore, we may restrict ourselves to the Y_h of Case C since it the smaller of the two cases. We will prove that in Case C

$$\dim(Y_h) - \dim(X_h) = 2n_K + m - 1. \tag{20}$$

Let n_V, n_E, n_V^∂, and n_E^∂ denote the number of vertices, edges, boundary vertices, and boundary edges, respectively. Removing the dimensions corresponding to the Dirichlet boundary conditions, we have $\dim(X_{0,h}) = n_V + n_E k + n_K k^2 - n_E^\partial k - n_V^\partial$. Using $n_V^\partial = n_E^\partial$ together with $\dim(\hat{X}_h) = n_E(k+1)$, we obtain $\dim(X_h) = n_V + n_E(2k+1) + n_K k^2 - n_E^\partial(k+1)$. Simplifying using the Euler relations (see e.g., [11, Lemma 1.57]) $n_V = n_E - n_K + 1 - m$ and $2n_E - n_E^\partial = 4n_K$,

$$\dim(X_h) = n_K(k+3)(k+1) + 1 - m.$$

Subtracting this from $\dim(Y_h) = n_K((k+2)^2 + 1)$, the proof of (20) is complete. $\square$

4. Error analysis

In this section we show that all the three cases admit the same convergence rate for u_h. The method of analysis for Cases B and C is different from that proposed in [9], but is the same as the one found in [1] and is akin to the second Strang lemma using the non-conforming space

$$Y_{h,0} = \{v \in Y_h : \langle r_{n,h}, v\rangle_{\partial\Omega_h} = 0 \text{ for all } r_{n,h} \in \hat{X}_h\}.$$

We use c to denote a generic constant, independent of h, whose value at different occurrences may vary.

Theorem 4.1. *In Cases A, B, and C, the solution u_h of the DPG method satisfies*

$$\|u - u_h\|_{H^1(\Omega)} \leqslant ch^{k+1}|u|_{H^{k+2}(\Omega)}. \tag{21}$$

Proof. For Case A, the Fortin operator of [2] satisfies (11), thus verifying the assumptions of [9, Theorem 2.1], which together with the Bramble-Hilbert lemma, gives the stated error estimate.

For Cases B and C, we adopt the approach in the proof of [1, Theorem 3.5]. We use a characterization of the DPG solution u_h using the so-called weakly conforming optimal test space $Y_{h,0}^{\text{opt}} = T_h X_{h,0}$, where $T_h : X_{h,0} \to Y_{h,0}$ is defined as the solution of $(T_h w_h, z_h)_Y = (\operatorname{grad} w_h, \operatorname{grad} z_h)_{\Omega_h}$, for all $z_h \in Y_{h,0}$. It is easy to see that for all $w_h \in X_{h,0}$,

$$\sup_{v_h \in Y_{h,0}} \frac{(\operatorname{grad} w_h, \operatorname{grad} v_h)_{\Omega_h}}{\|v_h\|_Y} = \sup_{v_h \in Y_{h,0}^{\text{opt}}} \frac{(\operatorname{grad} w_h, \operatorname{grad} v_h)_{\Omega_h}}{\|v_h\|_Y}. \tag{22}$$

Also recall that by [1, Theorem 2.6], u_h satisfies

$$(\operatorname{grad} u_h, \operatorname{grad} v_h)_{\Omega_h} = (f, v_h)_{\Omega_h}, \qquad \forall v_h \in Y_{h,0}^{\text{opt}}. \tag{23}$$

We use these facts in the argument below.

If c_p is the Poincaré constant such that $\|w\|_{H^1(\Omega)} \leqslant c_p \|\operatorname{grad} w\|_{L^2(\Omega)}$ for all $w \in H_0^1(\Omega)$, then, for any $w_h \in X_{h,0}$,

$$\begin{aligned}
\|u_h - w_h\|_{H^1(\Omega)} &\leqslant c_p^2 \sup_{z_h \in X_{h,0}} \frac{(\operatorname{grad}(u_h - w_h), \operatorname{grad} z_h)_{\Omega_h}}{\|z_h\|_{H^1(\Omega)}} \\
&\leqslant c_p^2 \sup_{v_h \in Y_{h,0}} \frac{(\operatorname{grad}(u_h - w_h), \operatorname{grad} v_h)_{\Omega_h}}{\|v_h\|_Y}, && \text{as } X_{h,0} \subseteq Y_{h,0}, \\
&= c_p^2 \sup_{v_h \in Y_{h,0}^{\text{opt}}} \frac{(\operatorname{grad}(u_h - u), \operatorname{grad} v_h)_{\Omega_h} + (\operatorname{grad}(u - w_h), v_h)_{\Omega_h}}{\|v_h\|_Y},
\end{aligned}$$

due to (22). Focusing on the first term in numerator, for any $v_h \in Y_{h,0}^{\text{opt}}$,

$$\begin{aligned}
(\operatorname{grad}(u_h - u), \operatorname{grad} v_h)_{\Omega_h} &= (f, v_h)_{\Omega_h} - (\operatorname{grad} u, \operatorname{grad} v_h)_{\Omega_h} && \text{by (23)} \\
&= -\langle \hat{q}_n, v_h \rangle_{\partial\Omega_h} && \text{by (2)} \\
&= -\langle \hat{q}_n - \hat{r}_{n,h}, v_h \rangle_{\partial\Omega_h}
\end{aligned}$$

for any $\hat{r}_{n,h} \in \hat{X}_h$ (since $v_h \in Y_{h,0}$). Therefore,

$$\|u_h - w_h\|_{H^1(\Omega)} \leqslant c_p^2 \sup_{v_h \in Y_{h,0}^r} \frac{b((u - w_h, \hat{q}_n - \hat{r}_{n,h}), v_h)}{\|v_h\|_Y} \leqslant c \left(\|\hat{q}_n - \hat{r}_{n,h}\|_{\hat{X}} + \|u - w_h\|_{H^1(\Omega)} \right).$$

Thus, by triangle inequality,

$$\|u - u_h\|_{H^1(\Omega)} \leqslant c \inf_{(w_h, \hat{r}_{n,h}) \in X_h} \Bigg(\|\hat{q}_n - \hat{r}_{n,h}\|_{\hat{X}} + \|u - w_h\|_{H^1(\Omega)} \Bigg).$$

It only remains to estimate the right hand side. The argument is standard: Let R_k denote the Raviart-Thomas subspace [12] of $H(\operatorname{div}, \Omega)$ consisting of all vector functions which when restricted to an element takes the form $xp_1 + p_2$ for some $p_1 \in P_k(K)$ and some $p_2 \in P_k(K)^2$. Since $\hat{r}_{n,h}$ and $\hat{q}_n$ are element-by-element traces of an $r_h \in R_k$ and the exact flux $q = \operatorname{grad} u$, respectively, by (3),

$$\|\hat{q}_n - \hat{r}_{n,h}\|_{\hat{X}} \leqslant \|q - r_h\|_{H(\operatorname{div},\Omega)} \leqslant ch^{k+1} \Big(|u|_{H^{k+1}(\Omega)} + |\operatorname{div} q|_{H^{k+1}(\Omega)} \Big).$$

Together with the well-known bounds for $H^1(\Omega)$-best approximation, the estimate is proved. $\square$

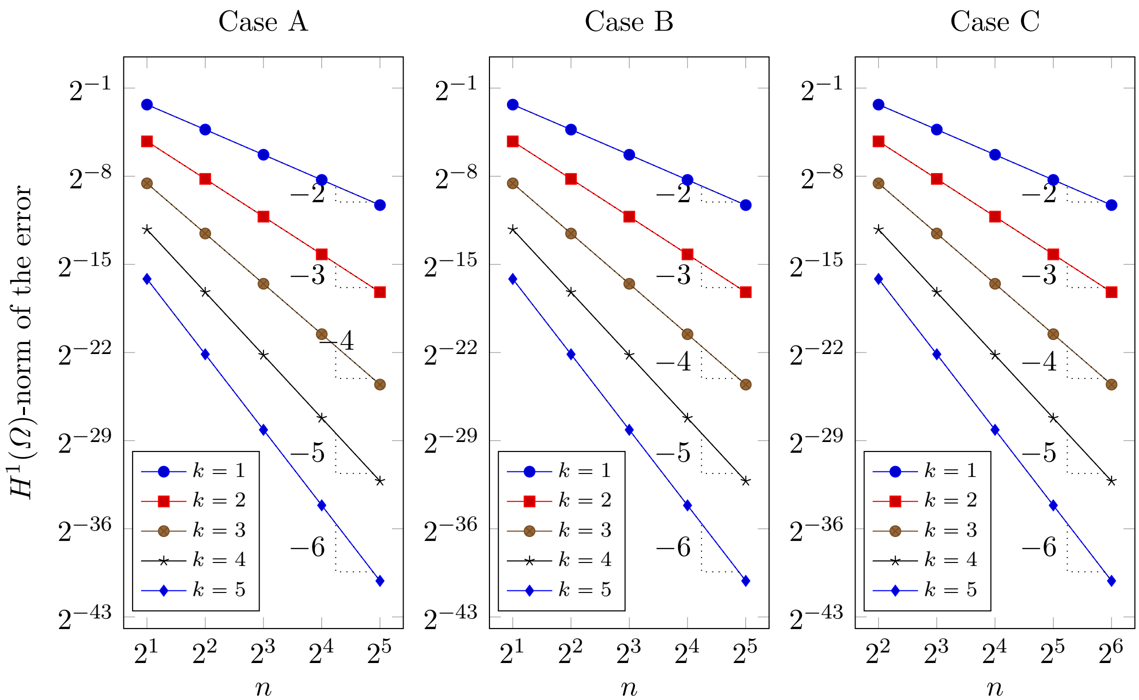


FIGURE 2. Observed convergence rates of the three cases

5. NUMERICAL RESULTS

For numerical illustration of the previously established theoretical results, we solve the Poisson equation with Dirichlet boundary condition using the three cases of the DPG method.

The domain Ω was set to the unit square. The function f was chosen so that the exact solution is $u = \sin(\pi x)\sin(\pi y)$. We construct an $n \times n$ uniform mesh by dividing Ω into n^2 congruent squares. Its mesh size is $h = \sqrt{2}/n$. The method is applied on a sequence of such meshes with geometrically increasing n.

All three cases were implemented using the open source NGSolve finite element library [14]. Since the test space of Case C is not a standard space, it is not available in NGSolve. So we have implemented it within a DPG shared library freely available in an online repository [10]. This code uses existing NGSolve classes to make a new finite element class representing $Q^+_{k,k}$ (as well as the corresponding finite element space it generates). Upon compilation it provides a shared library that can be loaded into NGSolve's (python or other) interfaces at run time.

Figure 2 shows plots of the error $\|u - u_h\|_{H^1(\Omega)}$ as a function of $n = \sqrt{2}/h$ in all three cases. The error decreases at the rate $O(n^\alpha)$. Slopes of regression fits to each data set determine an approximate observed convergence rate α, which is also marked in the figure. They are in good agreement with the theoretically predicted rates for all orders except $k = 2$. The observed convergence rate for $k = 2$ is inexplicably one order higher than what is expected from our theory.

References

[1] T. Bouma, J. Gopalakrishnan, and A. Harb, *Convergence rates of the DPG method with reduced test space degree*, Computers and Mathematics with Applications, 68 (2014), pp. 1550–1561.

[2] V. M. Calo, N. O. Collier, and A. H. Niemi, *Analysis of the discontinuous Petrov-Galerkin method with optimal test functions for the Reissner-Mindlin plate bending model*, Computers and Mathematics with Applications, 66 (2014), pp. 2570–2586.

[3] C. Carstensen, L. Demkowicz, and J. Gopalakrishnan, *A posteriori error control for DPG methods*, SIAM J Numer. Anal., 52 (2014), pp. 1335–1353.

[4] C. Carstensen, L. Demkowicz, and J. Gopalakrishnan, *Breaking spaces and forms for the dpg method and applications including maxwell equations*, Computers and Mathematics with Applications, 72 (2016), pp. 494–522.

[5] C. Carstensen, D. Gallistl, F. Hellwig, and L. Weggler, *Low-order dPG-FEM for an elliptic PDE*, Computers & Mathematics with Applications, 68 (2014), pp. 1503 – 1512. Minimum Residual and Least Squares Finite Element Methods.

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

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

[8] ———, *A primal DPG method without a first-order reformulation*, Computers and Mathematics with Applications, 66 (2013), pp. 1058–1064.

[9] J. Gopalakrishnan and W. Qiu, *An analysis of the practical DPG method*, Mathematics of Computation, 83 (2014), pp. 537–552.

[10] J. Gopalakrishnan, N. Olivares, L. Kogler and J. Schöberl, *DPG Methods in NGSolve*, GitHub Repository: github.com/jayggg/DPG, See commits pushed August 2016.

[11] A. Ern and J.-L. Guermond, *Theory and practice of finite elements*, vol. 159 of Applied Mathematical Sciences, Springer-Verlag, New York, 2004.

[12] P.-A. Raviart and J. M. Thomas, *Primal hybrid finite element methods for 2nd order elliptic equations*, Math. Comp., 31 (1977), pp. 391–413.

[13] M. Schedensack, A new generalization of the P1 non-conforming FEM to higher polynomial degrees. Comput. Methods Appl. Math., 2016. Published Online; also available as INS Preprint No. 1507 and arXiv e-print 1505.02044.

[14] J. Schöberl, *C++11 implementation of finite elements in NGSolve*, To appear in ACM Transactions on Mathematical Software, (2016).

Address for all authors: PO Box 751, Portland State University, Portland, OR 97207-0751.