Low-Order dPG-FEM for an Elliptic PDE

C. Carstensen[a], D. Gallistl[b], F. Hellwig[a], L. Weggler[a]

[a] *Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany*
[b] *Institut für Numerische Simulation, Universität Bonn, Wegelerstraße 6, 53115 Bonn, Germany*

Abstract

This paper introduces a novel lowest-order discontinuous Petrov Galerkin (dPG) finite element method (FEM) for the Poisson model problem. The ultra-weak formulation allows for piecewise constant and affine ansatz functions and for piecewise affine and lowest-order Raviart-Thomas test functions. This lowest-order discretization for the Poisson model problem allows for a direct proof of the discrete inf-sup condition and a complete a priori and a posteriori error analysis. Numerical experiments investigate the performance of the method and underline the quasi-optimal convergence.

Keywords: FEM, discontinuous, Petrov-Galerkin, DPG, Poisson
2010 MSC: 65N30

1. Introduction

The discontinuous Petrov Galerkin (dPG) finite element method (FEM) was introduced by [1, 2, 3, 4] for particular choices of polynomial degrees, namely the lowest-order case with piecewise constant and affine trial functions and piecewise first-order Raviart-Thomas and quadratic test functions in 2D. A refined analysis is presented in [5]. The novel low-order variant of this paper reduces the test-function space dramatically to piecewise zero-order Raviart-Thomas and affine functions and still allows for a stable and quasi-optimal scheme.

The dPG finite element method uses a mixture of ideas from the least-squares and mixed finite element methodology and started with the search of an optimal test-function space.

Given $F \in Y^*$, seek $x \in X$ with

$$b(x, y) = F(y) \quad \text{for all } y \in Y. \tag{1.1}$$

For a quite general functional setting, i.e., X, Y reflexive Banach spaces, well-posedness requires the inf-sup condition

$$0 < \beta := \inf_{x \in S(X)} \sup_{y \in S(Y)} b(x, y). \tag{1.2}$$

Here and throughout the paper, the unit sphere of a normed linear space $(W, \|\cdot\|_W)$ is denoted by $S(W) = \{w \in W | \|w\|_W = 1\}$. A closer look at (1.2) reveals that

1. The larger Y, the better chances we have to guarantee (1.2).

2. The inf-sup condition gives rise to an a posteriori error estimate for conforming discretizations. Provided (1.2), any $x_h \in X_h \subset X$ satisfies
$$\|x - x_h\|_X \leq \frac{1}{\beta} \|F - b(x_h, \cdot)\|_{Y^*}. \tag{1.3}$$

The dPG method was originally thought of as a method to exploit those two aspects by a minimization of the right-hand side of (1.3) for $X_h \subset X$. The idealized dPG (idPG) method seeks

$$x_h \in \operatorname*{arg\,min}_{x_h \in X_h} \|F - b(x_h, \cdot)\|_{Y^*}$$

and is not practical because the norm of the dual space Y^* is, in general, not computable as a supremum over an infinite-dimensional space Y. The (practical) dPG method is the least squares minimization with the finite-dimensional dual space Y_h^* and seeks $x_h \in X_h$ such that

$$x_h \in \operatorname*{arg\,min}_{x_h \in X_h} \|F - b(x_h, \cdot)\|_{Y_h^*}.$$

The finite-dimensional spaces $X_h \subset X$ and $Y_h \subset Y$ form conforming approximations of the Banach spaces X and Y where $X \neq Y$ in general. In the discrete setting, this may lead to an unbalancy of dimension meaning that $\dim X_h < \dim Y_h$. The dPG method can be, therefore, regarded as either a finite element method with nonstandard test spaces, or, as least-squares finite element method minimizing the residual in a nonstandard norm.

This paper proposes a lowest-order dPG finite element discretization for an elliptic model problem and proves its stability and optimal approximation properties. The notation is fixed for two space dimensions in order to simplify the presentation. Nevertheless, all the analysis and the proofs directly carry over to three dimensions.

The paper is organized as follows. In Section 2 the ultraweak formulation of the model problem along with necessary notation of regular triangulations and nonstandard function spaces is introduced. Section 3 is concerned with the dPG method for the Hilbert space case. The discussion includes results about a priori as well as a posteriori properties of the dPG solution. The paper is concluded by numerical experiments of the lowest-order dPG discretization presented in Section 4.

Throughout the paper standard notation on Lebesgue and Sobolev spaces is employed and $\|\bullet\|$ abbreviates $\|\bullet\|_{L^2(\Omega)}$. The dot denotes the scalar product of two one-dimensional lists of reals of the same length. The notation $a \lesssim b$ abbreviates $a \leq Cb$ for a positive generic constant C that may depend on the domain Ω but not on the mesh-size. The notation $a \approx b$ stands for $a \lesssim b \lesssim a$. The measure $|\cdot|$ is context-sensitive and refers to the number of elements of some finite set or the length of an edge or the area of some domain and not just the modulus of a real number or the Euclidean length of a vector.

2. Ultraweak formulation of the Poisson model problem

The idea behind the ultraweak formulation is to enlarge the test function space. This means that one gives up on global continuity properties of the test functions at the expense of explicit discretization of traces on the interfaces of simplices. Different from conventional methods, the triangulation is inextricably linked with the functional setting even on the continuous level.

2.1. Triangulation of the domain

Given a regular triangulation $\mathcal{T}$ of the polygonal Lipschitz domain $\Omega \subset \mathbb{R}^2$ into simplices, the set of faces of a simplex T is denoted $\mathcal{F}(T)$; $\mathcal{F}$ (resp. $\mathcal{F}(\Omega)$) denotes the set of all faces in $\mathcal{T}$ (resp. interior faces of $\mathcal{T}$). The union of all faces $\mathcal{F}$ is called the skeleton $\partial\mathcal{T}$ regarding it as the set of all the boundaries ∂T of the simplices $T \in \mathcal{T}$, i.e.,

$$\partial\mathcal{T} := \bigcup_{T\in\mathcal{T}} \bigcup_{F\in\mathcal{F}(T)} F.$$

We further use some notational convention for the unit normal vector fields. The unit normal vector field along the boundary ∂T of a fixed simplex $T \in \mathcal{T}$ is denoted by ν_T or, when it is clear from the context, simply by ν.

Each interior face is attributed a unique unit normal vector ν_F by fixing one of the two possible orientations. On the faces which lie on the boundary $\partial\Omega$, the exterior direction is selected.

Let Π_0 denote the L^2 projection onto piecewise constant functions with respect to $\mathcal{T}$; the same notation applies component-wise to vector-valued functions. Let for any $T \in \mathcal{T}$ its barycenter be denoted by $\operatorname{mid}(T)$ and define the piecewise constant function $\operatorname{mid}(\mathcal{T}) : \Omega \to \mathbb{R}^2$ via $\operatorname{mid}(\mathcal{T})|_K := \operatorname{mid}(K)$ for any $K \in \mathcal{T}$.

2.2. Function spaces

Besides Sobolev spaces H^1 and $H(\mathrm{div})$, we consider functions which exhibit certain (piecewise) regularity properties only on the isolated simplices in $\mathcal{T}$. Those function spaces are

$$H(\nabla_{NC}, \mathcal{T}) := \{v \in L^2(\Omega) : \ \forall T \in \mathcal{T}, \ v|_T \in H^1(T)\},$$
$$H(\mathrm{div}_{NC}, \mathcal{T}) := \{\tau \in L^2(\Omega; \mathbb{R}^2) : \ \forall T \in \mathcal{T}, \ \tau|_T \in H(\mathrm{div}, T)\}.$$

Throughout this paper, div_{NC} and ∇_{NC} denote the piecewise application of the differential operators div and ∇.

Let us consider a simplex $T \in \mathcal{T}$ with outer unit normal vector ν and recall that the trace operators

$$\gamma_0 : H^1(T) \to H^{1/2}(\partial T), \qquad \gamma_0 v := v|_{\partial T} \quad \text{for all } v \in H^1(T),$$
$$\gamma_\nu : H(\mathrm{div}, T) \to H^{-1/2}(\partial T), \quad \gamma_\nu q := q \cdot \nu_T|_{\partial T} \text{ for all } q \in H(\mathrm{div}, T)$$

are surjective with continuous right inverses. Thus, the local regularity of functions in the above spaces allows for the definition of traces on the skeleton $\partial\mathcal{T}$ and we set

$$H_0^{1/2}(\partial\mathcal{T}) := \{w \in \prod_{T\in\mathcal{T}} H^{1/2}(\partial T) : \exists\, \tilde{u} \in H_0^1(\Omega) \forall T \in \mathcal{T}, \ w|_{\partial T} = \gamma_0(\tilde{u}|_T)\},$$
$$H^{-1/2}(\partial\mathcal{T}) := \{t \in \prod_{T\in\mathcal{T}} H^{-1/2}(\partial T) : \exists\, \tilde{\sigma} \in H(\mathrm{div}, \Omega) \forall T \in \mathcal{T}, \ t|_{\partial T} = \gamma_\nu(\tilde{\sigma}|_T)\}.$$

We conclude this section with two remarks on the spaces $H_0^{1/2}(\partial\mathcal{T})$ and $H^{-1/2}(\partial\mathcal{T})$ as well as their duality relation.

Remark 2.1. *Each $w \in H_0^{1/2}(\partial\mathcal{T})$ may be identified with some $\tilde{u} \in H_0^1(\Omega)$. Such $\tilde{u}$ is an extension of the skeleton trace w on $\partial\mathcal{T}$. In a finite element discretization, $\tilde{u}$ must be discretized by globally $H^1(\Omega)$-conforming finite element functions to render γ_0 well defined.*

Similarly, each $t \in H^{-1/2}(\partial\mathcal{T})$ may be identified with some $\tilde{\sigma} \in H(\mathrm{div}, \Omega)$. Such $\tilde{\sigma}$ is an extension of the skeleton trace $t := \sigma \cdot \nu$ on $\partial\mathcal{T}$. In a finite element discretization, $\tilde{\sigma}$ must be discretized by globally $H(\mathrm{div}, \Omega)$-conforming finite element functions to render γ_ν well defined.

Remark 2.2. *Locally we know that $H^{-1/2}(\partial T)$ is the realization of the dual space $H^{1/2}(\partial T)^\star$ with $L^2(T)$ as pivot space. This means that the duality $\langle \bullet, \bullet \rangle_{\partial T}$ is an extension of the $L^2(\partial T)$ scalar product taking into account the boundary ∂T as a whole and* not *the single faces $F \in \mathcal{F}(T)$. In conclusion, the skeleton spaces $H_0^{1/2}(\partial\mathcal{T})$ and $H^{-1/2}(\partial\mathcal{T})$ are related by duality with*

$$\forall w \in H_0^{1/2}(\partial\mathcal{T})\, \forall t \in H^{-1/2}(\partial\mathcal{T}) \quad \langle t, w \rangle_{\partial\mathcal{T}} := \sum_{T\in\mathcal{T}} \left\langle w|_{\partial T}, t|_{\partial T} \right\rangle_{\partial T}. \tag{2.1}$$

2.3. Model problem

In this section, the ultraweak formulation of a second-order model problem is presented. Let Ω be a polygonal bounded Lipschitz domain in $\mathbb{R}^2$ with closed boundary $\partial\Omega$. The starting point is the boundary value problem

$$-\mathrm{div}\,\nabla u = f \quad \text{in } \Omega \quad \text{and} \quad \gamma_0 u = 0 \quad \text{on } \partial\Omega$$

and the equivalent system of first-order equations

$$\begin{aligned} -\mathrm{div}\,\sigma &= f && \text{in } \Omega, \\ \sigma - \nabla u &= 0 && \text{in } \Omega, \\ \gamma_0 u &= 0 && \text{on } \partial\Omega. \end{aligned} \tag{2.2}$$

For any right-hand side $f \in L^2(\Omega)$ there exist some unique $u \in H_0^1(\Omega)$ and $\sigma \in H(\mathrm{div}, \Omega)$ with (2.2). The weak formulation of *both* first order equations in (2.2) leads to a bilinear form $b : X \times Y \to \mathbb{R}$ where X and Y are the following Hilbert spaces

$$X := L^2(\Omega) \times L^2(\Omega; \mathbb{R}^2) \times H_0^{1/2}(\partial\mathcal{T}) \times H^{-1/2}(\partial\mathcal{T}),$$
$$Y := H(\nabla_{NC}, \mathcal{T}) \times H(\mathrm{div}_{NC}, \mathcal{T})$$

(with natural norms discussed below). For $x = (u, \sigma, w, t) \in X$ and $y = (v, \tau) \in Y$, recall (2.1) and define

$$\begin{aligned} b(x, y) := & \int_\Omega \sigma \cdot \nabla_{NC} v \,\mathrm{d}x - \langle t, \gamma_0 v \rangle_{\partial\mathcal{T}} \\ & + \int_\Omega \sigma \cdot \tau \,\mathrm{d}x + \int_\Omega u \,\mathrm{div}_{NC}\, \tau \,\mathrm{d}x - \langle \gamma_\nu \tau, w \rangle_{\partial\mathcal{T}} \,. \end{aligned} \tag{2.3}$$

For all $y = (v, \tau) \in Y$, the right-hand side in (2.2) defines the linear functional

$$F(y) := \int_\Omega f v \,\mathrm{d}x \quad \text{with} \quad |F(y)| \leq \|y\|_Y \|f\| \,. \tag{2.4}$$

Remark 2.3. *In contrast to the weak formulation based on the space $H^1_0(\Omega)$, the triangulation $\mathcal{T}$ enters explicitly in the definition of the infinite-dimensional spaces X and Y.*

2.4. Equivalent formulation

Note that in Remark 2.1 it has been proposed a strategy how to remove the duality pairings from the ultraweak formulation. Consider any $\tilde{u} \in H^1_0(\Omega)$ with $w = \gamma_0 \tilde{u}$ and any $\tilde{\sigma} \in H(\mathrm{div}, \Omega)$ with $t = \gamma_\nu \tilde{\sigma}$. Note that the right-hand sides do *not* depend on the values of $\tilde{u}$ and $\tilde{\sigma}$ on $\Omega \setminus \partial\mathcal{T}$ and, thus, (2.3) equivalently reads

$$\begin{aligned} b(x, y) = & \int_\Omega (\sigma - \tilde{\sigma}) \cdot \nabla_{NC} v \,\mathrm{d}x - \int_\Omega v \,\mathrm{div}\, \tilde{\sigma} \,\mathrm{d}x \\ & + \int_\Omega (\sigma - \nabla \tilde{u}) \cdot \tau \,\mathrm{d}x + \int_\Omega (u - \tilde{u}) \,\mathrm{div}_{NC}\, \tau \,\mathrm{d}x. \end{aligned} \tag{2.5}$$

The identifications of skeleton traces with globally defined functions as in Remark 2.1 specify the norm topologies of X and Y via

$$\begin{aligned} X &= L^2(\Omega) \times L^2(\Omega; \mathbb{R}^2) \times H^{1/2}_0(\partial\mathcal{T}) \times H^{-1/2}(\partial\mathcal{T}) \\ &\equiv L^2(\Omega) \times L^2(\Omega; \mathbb{R}^2) \times H^1_0(\Omega) \times H(\mathrm{div}, \Omega) \\ \|x\|^2_X &:= \|u\|^2 + \|\sigma\|^2 + \|\tilde{u}\|^2 + \|\nabla \tilde{u}\|^2 + \|\tilde{\sigma}\|^2 + \|\mathrm{div}\, \tilde{\sigma}\|^2, \\ Y &= H(\nabla_{NC}, \mathcal{T}) \times H(\mathrm{div}_{NC}, \mathcal{T}; \mathbb{R}^2) \\ \|(v, \tau)\|^2_Y &:= \|v\|^2 + \|\nabla_{NC} v\|^2 + \|\tau\|^2 + \|\mathrm{div}_{NC}\, \tau\|^2. \end{aligned}$$

Remark 2.4 (well-posedness)**.** *The bilinear form $b : X \times Y \to \mathbb{R}$ is bounded*

$$|b(x, y)| \lesssim \|x\|_X \|y\|_Y \,. \tag{2.6}$$

For any $x \in X$ it holds $b(x, \bullet) \in Y^*$ and, thus, the ultraweak problem is well-posed in the continuous case and it reads: Seek $x \in X$ such that

$$b(x, y) = F(y) \;\; \textit{for all } y \in Y \,.$$

The stability of the problem was proven in [2].

Lemma 2.5 (inf-sup condition, Theorem 4.2 of [2])**.** *The constant*

$$0 < \beta := \inf_{x \in S(X)} \sup_{y \in S(Y)} b(x, y)$$

depends on the domain Ω but is independent of the triangulation $\mathcal{T}$. □

3. Lowest-order dPG FEM

The following two subsections are concerned with the numerical analysis of the dPG finite element method. The starting point is the ultraweak formulation with the bilinear form $b : X \times Y \to \mathbb{R}$ from (2.5).

The introduction of the finite element spaces is followed by an a priori result stating that the dPG solution possesses best-approximation properties.

The dPG method is a mixed finite element formulation. This means that solvability and stability of the method is captured by the inf-sup condition. With this idea in mind the lowest-order dPG method for the Poisson problem is investigated and proved to be a stable numerical method.

3.1. Discretization and a priori error estimates

The finite-dimensional subspaces $X_h \subset X$ and $Y_h \subset Y$ are spaces of piecewise polynomial functions. For $k \geq 0$, the spaces of piecewise polynomial functions read as

$$\begin{aligned}
P_k(T) &:= \{v_k \in L^\infty(T) : \ v_k \text{ is polynomial on } T \text{ of degree } \leq k\},\\
P_k(\mathcal{T}) &:= \{v_k \in L^\infty(\Omega) : \forall T \in \mathcal{T},\ v_k|_T \in P_k(T)\},\\
P_k(T;\mathbb{R}^2) &:= \{q_k \in L^\infty(T;\mathbb{R}^2) : \ \text{each component of } q_k \text{ belongs to } P_k(T)\},\\
P_k(\mathcal{T};\mathbb{R}^2) &:= \{q_k \in L^\infty(\Omega;\mathbb{R}^2) : \forall T \in \mathcal{T},\ v_k|_T \in P_k(T;\mathbb{R}^2)\}.
\end{aligned}$$

The space of lowest-order discontinuous Raviart-Thomas vector fields is defined as

$$RT_0(\mathcal{T}) := \{q_0 \in L^\infty(\mathcal{T};\mathbb{R}^2) : \forall T \in \mathcal{T}\ \exists (a,b,c) \in \mathbb{R}^3, q_0|_T(x) = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) + cx\}.$$

Define $X_h \subset X$ and $Y_h \subset Y$ as

$$\begin{aligned}
X_h &:= P_0(\mathcal{T}) \times P_0(\mathcal{T};\mathbb{R}^2) \times (P_1(\mathcal{T}) \cap C_0(\Omega)) \times (RT_0(\mathcal{T}) \cap H(\mathrm{div},\Omega)),\\
Y_h &:= P_1(\mathcal{T}) \times RT_0(\mathcal{T}).
\end{aligned}$$

The discretization is based on the equivalent formulation described in Subsection 2.4. Since the normal traces of $(RT_0(\mathcal{T}) \cap H(\mathrm{div},\Omega))$ are exactly the piecewise constant functions on the edges $P_0(\mathcal{F})$ and since the traces of $(P_1(\mathcal{T}) \cap C_0(\Omega))$ are equivalent to the continuous piecewise affines on $\mathcal{F}$ that vanish on the boundary, the proposed scheme corresponds to a direct discretization of the formulation based on the bilinear form (2.3). The following two lemmas show that also the involved norms are equivalent.

Lemma 3.1. *Given $\tilde{u}_1 \in P_1(\mathcal{T}) \cap C_0(\Omega)$. The trace $\gamma_0\tilde{u}_1$ on the skeleton $\partial\mathcal{T}$ is a unique function in $H^{1/2}(\partial\mathcal{T})$. The norm of $w := \gamma_0\tilde{u}_1$ is defined as follows*

$$\|w\|_{H^{1/2}(\partial\mathcal{T})} := \inf_{\substack{v \in H^1_0(\Omega)\\ \gamma_0 v = w}} \|v\|_{H^1(\Omega)} = \inf_{\substack{v \in H^1_0(\Omega)\\ \gamma_0 v = w}} \sqrt{\|v\|^2 + \|\nabla v\|^2}.$$

It holds the equivalence of norms

$$\|w\|_{H^{1/2}(\partial\mathcal{T})} \approx \|\tilde{u}_1\|_{H^1(\Omega)}.$$

Proof. For the proof of the non-trivial direction $\|\tilde{u}_1\|^2_{H^1(\Omega)} \lesssim \|w\|^2_{H^{1/2}(\partial\mathcal{T})}$, note that any $v \in H^1_0(\Omega)$ with $\gamma_0 v = w$ satisfies on each $K \in \mathcal{T}$ that

$$\nabla\tilde{u}_1\big|_K = |K|^{-1} \int_K \nabla\tilde{u}_1 \,\mathrm{d}x = |K|^{-1} \int_{\partial K} vw \,\mathrm{d}x = |K|^{-1} \int_K \nabla v \,\mathrm{d}x\,.$$

Hence, $\nabla\tilde{u}_1 = \Pi_0\nabla v$ which implies

$$\|\nabla\tilde{u}_1\| = \|\Pi_0\nabla v\| \leq \|\nabla v\| \leq \|v\|_{H^1(\Omega)}.$$

The Friedrichs inequality states $\|\tilde{u}_1\|_{H^1(\Omega)} \lesssim \|\nabla\tilde{u}_1\|$. The claimed equivalence follows from infimizing over $v \in H^1_0(\Omega)$. □

Lemma 3.2. *Given $\tilde{\sigma}_0 \in RT_0(\mathcal{T}) \cap H(\mathrm{div}, \Omega)$. The trace $\gamma_\nu \tilde{\sigma}_0$ on the skeleton $\partial\mathcal{T}$ is a unique function in $H^{-1/2}(\partial\mathcal{T})$. The norm of $t = \gamma_\nu \tilde{\sigma}_0$ is defined as follows*

$$\|t\|_{H^{-1/2}(\partial\mathcal{T})} := \inf_{\substack{v \in H(\mathrm{div},\Omega) \\ \gamma_\nu v = t}} \|v\|_{H(\mathrm{div},\Omega)} = \inf_{\substack{v \in H(\mathrm{div},\Omega) \\ \gamma_\nu v = t}} \sqrt{\|v\|^2 + \|\mathrm{div}\, v\|^2}$$

It holds the equivalence of norms

$$\|t\|_{H^{-1/2}(\partial\mathcal{T})} \approx \|\tilde{\sigma}_0\|_{H(\mathrm{div},\Omega)}$$

Proof. To prove the non-trivial direction $\|\tilde{\sigma}_0\| \lesssim \|t\|_{H^{-1/2}(\partial\mathcal{T})}$, consider $v \in H(\mathrm{div}, \Omega)$ with $\gamma_\nu v = t$. Since v and $\tilde{\sigma}_0$ have the same normal trace on the boundary of each $K \in \mathcal{T}$ with $t \in P_0(\partial\mathcal{T})$ it follows

$$\mathrm{div}\, \tilde{\sigma}_0\big|_K = |K|^{-1} \int_K \mathrm{div}\, \tilde{\sigma}_0 \,\mathrm{d}x = |K|^{-1} \int_{\partial K} t \,\mathrm{d}x = |K|^{-1} \int_K \mathrm{div}\, v \,\mathrm{d}x\,.$$

Hence,

$$\|\mathrm{div}\, \tilde{\sigma}_0\| = \|\Pi_0 \,\mathrm{div}\, v\| \leq \|\mathrm{div}\, v\|.$$

Recall that $\tilde{\sigma}_0 \in RT_0(\mathcal{T}) \cap H(\mathrm{div}, \Omega)$, i.e., there exist $c_1 \in P_0(\mathcal{T}; \mathbb{R}^2)$ and $c_2 \in P_0(\mathcal{T})$ with

$$\tilde{\sigma}_0\big|_K(x) = c_1 + c_2(x - \mathrm{mid}(K)) = \nabla\left(c_1 \cdot x + \frac{c_2}{2}\|x - \mathrm{mid}(K)\|^2\right)$$

and, therefore, for any $K \in \mathcal{T}$, $\tilde{\sigma}_0|_K = \nabla\alpha_K$ is the gradient of some $\alpha_K \in H^1(K)$ which solves the local Neumann problem

$$\begin{aligned} \Delta\,\alpha_K &= \mathrm{div}\,\tilde{\sigma}_0 \quad \text{in } K \\ \gamma_\nu \alpha_K &= t \quad \text{on } \partial K \\ \int_K \alpha_K \,\mathrm{d}x &= 0. \end{aligned}$$

As α_K has integral mean zero, we can apply the Poincaré inequality with the Poincaré constant C_P and the diameter $h_K := \mathrm{diam}(K)$ and obtain

$$\begin{aligned} \|\nabla\alpha_K\|_K^2 &= \int_K \nabla\alpha_K \cdot \nabla\alpha_K \,\mathrm{d}x \\ &= -\int_K \alpha_K \,\mathrm{div}\, \tilde{\sigma}_0 \,\mathrm{d}x + \int_{\partial K} t\,\alpha_K \,\mathrm{d}s \\ &= -\int_K \alpha_K \,\mathrm{div}\, \tilde{\sigma}_0 \,\mathrm{d}x + \int_K \alpha_K \,\mathrm{div}\, v \,\mathrm{d}x + \int_K v \cdot \nabla\alpha_K \,\mathrm{d}x \\ &\leq h_K C_P \|\nabla\alpha_K\|_K \|\mathrm{div}\, \tilde{\sigma}_0\|_K + h_K C_P \|\nabla\alpha_K\|_K \|\mathrm{div}\, v\|_K + \|\nabla\alpha_K\|_K \|v\|_K. \end{aligned}$$

Recall the estimate from above, then it holds

$$\|\nabla\alpha_K\|_K \leq 2\, h_K C_P \|\mathrm{div}\, v\|_K + \|v\|_K$$

and summing up both local results we obtain after infimizing over v that

$$\|\tilde{\sigma}_0\|_{H(\mathrm{div},\Omega)} \lesssim \|t\|_{H^{-1/2}(\partial\mathcal{T})}.$$ □

Recall that the main issue for mixed finite element methods is the proof of the discrete inf-sup condition, that is the existence of $\beta_h > 0$ such that

$$0 < \beta_h = \inf_{x_h \in S(X_h)} \sup_{y_h \in S(Y_h)} b(x_h, y_h). \tag{3.1}$$

Note that for X_h and Y_h as above it holds $\mathrm{dim} X_h < \mathrm{dim} Y_h$. For the analysis of unique solvability of the dPG formulation, consider the closed subspace N_h where b degenerates, namely

$$N_h := \{y_h \in Y_h : \; b(x_h, y_h) = 0 \text{ for all } x_h \in X_h\} \subset Y_h.$$

Furthermore, define $M_h := N_h^\perp$ as the orthogonal complement of N_h with respect to the scalar product in Y_h.

Theorem 3.3 (abstract a priori error estimate)**.** *Provided the spaces X_h and Y_h satisfy the discrete inf-sup condition* (3.1)*, then for any $F \in Y^*$ there exists a unique solution $x_h \in X_h$ to the problem*

$$b(x_h, y_h) = F(y_h) \text{ for all } y_h \in M_h. \tag{3.2}$$

It holds quasi-optimal convergence

$$\|x - x_h\|_X \leq \frac{\|b\|}{\beta_h} \operatorname{dist}_X(x, X_h). \tag{3.3}$$

Proof. The orthogonal projection P_h onto M_h satisfies $P_h : Y_h \to M_h$, $\|P_h\| = 1$. This and the discrete inf-sup condition (3.1) reveal for any $\xi_h \in X_h$ that

$$\begin{aligned}\|b(\xi_h, \bullet)\|_{Y_h^*} &= \sup_{y_h \in S(Y_h)} b(\xi_h, y_h) = \sup_{y_h \in S(Y_h)} b(\xi_h, P_h y_h) \\ &\leq \sup_{y_h \in S(M_h)} b(\xi_h, y_h) = \|b(\xi_h, \bullet)\|_{M_h^*}.\end{aligned}$$

This implies the following discrete inf-sup condition for M_h

$$\beta_h \leq \inf_{x_h \in S(X_h)} \sup_{y_h \in S(M_h)} b(x_h, y_h).$$

Since b is non-degenerate on M_h, the classical Babuška-Brezzi theory proves that problem (3.2) has a unique solution $x_h \in X_h$ which satisfies (3.3). The constant $\|b\|/\beta_h$ in the error estimate can be obtained (in the present case of a Hilbert space) from the fact that the discrete solution operator $x \mapsto x_h$ is a nontrivial oblique projection. □

Remark 3.4 (convergence rate)**.** *The error estimate* (3.3) *shows that the proposed scheme is of first order: Provided the exact solution allows for the smoothness $u \in H^2(\Omega)$ and and the data f is piecewise smooth, standard approximation results prove that the error $\|x - x_h\|_X$ converges as $O(h)$.*

The remaining part of this section is devoted to the proof of the discrete inf-sup condition.

Theorem 3.5 (discrete inf-sup condition)**.** *Let $C_F \leq \operatorname{width}(\Omega)/\pi$ denote the Friedrichs constant and let $h_{max} := \max_{T \in \mathcal{T}} \operatorname{diam}(T) \leq \operatorname{diam}(\Omega)$ denote the maximum mesh-size in $\mathcal{T}$. The bilinear form b satisfies the discrete inf-sup condition* (3.1) *with the constant β_h equal to*

$$\begin{aligned}(1 + h_{max}^2)^{-1/2} \min \Big\{ &(1 + 2C_F^2(5 + 3C_F^2) + 2h_{max}^2(6 + 3C_F^2))^{-1/2}, \\ &(2 + 4(6 + 3C_F^2))^{-1/2} \Big\}.\end{aligned}$$

The inf-sup constant $\beta_h \approx 1$ exclusively depends on the diameter of the domain Ω.

Proof. Recall the L^2 projection Π_0 onto piecewise constants and the definition of $\operatorname{mid}(\mathcal{T})$ from Subsection 2.1. Given $x_h = (u_0, \sigma_0, \tilde{u}_1, \tilde{\sigma}_0) \in X_h$, define $y_h = (v_1, \tau_0) \in Y_h$ by

$$\begin{aligned}\tau_0 &:= (\sigma_0 - \nabla \tilde{u}_1) + \frac{1}{2}(u_0 - \Pi_0 \tilde{u}_1)(\bullet - \operatorname{mid}(\mathcal{T})), \\ v_1 &:= -\operatorname{div} \tilde{\sigma}_0 + (\sigma_0 - \Pi_0 \tilde{\sigma}_0) \cdot (\bullet - \operatorname{mid}(\mathcal{T})).\end{aligned}$$

The design of y_h shows that

$$\begin{aligned}b(x_h, y_h) &= \int_\Omega (\sigma_0 - \nabla \tilde{u}_1) \cdot \tau_0 \, \mathrm{d}x + \int_\Omega (u_0 - \tilde{u}_1) \operatorname{div}_{NC} \tau_0 \, \mathrm{d}x \\ &\quad + \int_\Omega (\sigma_0 - \tilde{\sigma}_0) \cdot \nabla_{NC} v_1 \, \mathrm{d}x - \int_\Omega \operatorname{div} \tilde{\sigma}_0 \, v_1 \, \mathrm{d}x \\ &= \|\sigma_0 - \nabla \tilde{u}_1\|^2 + \|u_0 - \Pi_0 \tilde{u}_1\|^2 + \|\sigma_0 - \Pi_0 \tilde{\sigma}_0\|^2 + \|\operatorname{div} \tilde{\sigma}_0\|^2.\end{aligned}$$

This and direct calculations with the orthogonality of $\bullet - \mathrm{mid}(\mathcal{T})$ onto the piecewise constants and the estimate $\|\bullet - \mathrm{mid}(\mathcal{T})\|_{L^\infty} \leq h_{max}$ show

$$\begin{aligned}\|y_h\|_Y^2 &\leq \|\sigma_0 - \nabla\tilde{u}_1\|^2 + h_{max}^2/4\,\|u_0 - \Pi_0\tilde{u}_1\|^2 + \|u_0 - \Pi_0\tilde{u}_1\|^2 \\ &\quad + \|\operatorname{div}\tilde{\sigma}_0\|^2 + h_{max}^2\|\sigma_0 - \Pi_0\tilde{\sigma}_0\|^2 + \|\sigma_0 - \Pi_0\tilde{\sigma}_0\|^2 \\ &\leq (1 + h_{max}^2)b(x_h, y_h).\end{aligned}$$

The Friedrichs, Young, and triangle inequalities imply

$$\begin{aligned}\|x_h\|_X^2 &\leq \|\operatorname{div}\tilde{\sigma}_0\|^2 + \|\tilde{\sigma}_0\|^2 + \|\sigma_0\|^2 + (1 + C_F^2)\|\nabla\tilde{u}_1\|^2 + \|u_0\|^2 \\ &\leq \|\operatorname{div}\tilde{\sigma}_0\|^2 + 2\|\tilde{\sigma}_0 - \nabla\tilde{u}_1\|^2 + 2\|\sigma_0 - \nabla\tilde{u}_1\|^2 \\ &\quad + (5 + 3C_F^2)\|\nabla\tilde{u}_1\|^2 + 2\|u_0 - \Pi_0\tilde{u}_1\|^2.\end{aligned} \tag{3.4}$$

Even for multiply connected domains, the Helmholtz decomposition shows that there exist unique $\alpha \in H_0^1(\Omega)$ and $r \in H(\operatorname{div}, \Omega)$ with $\operatorname{div} r = 0$ such that

$$\tilde{\sigma}_0 - \nabla\tilde{u}_1 = \nabla\alpha + r \quad \text{and} \quad \|\nabla\alpha\|^2 + \|r\|^2 = \|\tilde{\sigma}_0 - \nabla\tilde{u}_1\|^2.$$

Since $\tilde{\sigma}_0 \in H(\operatorname{div}, \Omega)$ and $-\Delta(\alpha + \tilde{u}_1) = \operatorname{div}\tilde{\sigma}_0 \in L^2(\Omega)$, the Friedrichs inequality implies that

$$\|\nabla(\alpha + \tilde{u}_1)\|^2 = \int_\Omega \operatorname{div}\tilde{\sigma}_0(\alpha + \tilde{u}_1)\mathrm{d}x \leq C_F\|\operatorname{div}\tilde{\sigma}_0\|\|\nabla(\alpha + \tilde{u}_1)\|\,.$$

In other words,

$$\|\nabla(\alpha + \tilde{u}_1)\| \leq C_F\|\operatorname{div}\tilde{\sigma}_0\|.$$

This plus the triangle inquality and the aforementioned stability of the Helmholtz decomposition reveal

$$\|\nabla\tilde{u}_1\| \leq \|\nabla\alpha\| + \|\nabla(\alpha + \tilde{u}_1)\| \leq \|\tilde{\sigma}_0 - \nabla\tilde{u}_1\| + C_F\|\operatorname{div}\tilde{\sigma}_0\|.$$

The combination with (3.4) and the triangle and Young inequalities lead to

$$\begin{aligned}\|x_h\|_X^2 &\leq (1 + 2C_F^2(5 + 3C_F^2))\|\operatorname{div}\tilde{\sigma}_0\|^2 + 2(6 + 3C_F^2)\|\tilde{\sigma}_0 - \nabla\tilde{u}_1\|^2 \\ &\quad + 2\|\sigma_0 - \nabla\tilde{u}_1\|^2 + 2\|u_0 - \Pi_0\tilde{u}_1\|^2.\end{aligned}$$

Since $\tilde{\sigma}_0|_T = \Pi_0\tilde{\sigma}_0 + \operatorname{div}\tilde{\sigma}_0(\bullet - \mathrm{mid}(T))/2$ on any $T \in \mathcal{T}$, the triangle inequality and Young inequalities lead to

$$\begin{aligned}\|\tilde{\sigma}_0 - \nabla\tilde{u}_1\|^2 &\leq h_{max}^2\|\operatorname{div}\tilde{\sigma}_0\|^2 + \|\Pi_0\tilde{\sigma}_0 - \nabla\tilde{u}_1\|^2 \\ &\leq h_{max}^2\|\operatorname{div}\tilde{\sigma}_0\|^2 + 2\|\sigma_0 - \Pi_0\tilde{\sigma}_0\|^2 + 2\|\sigma_0 - \nabla\tilde{u}_1\|^2.\end{aligned}$$

The combination of the foregoing two displayed inequalities results in

$$\begin{aligned}\|x_h\|_X^2 &\leq (1 + 2C_F^2(5 + 3C_F^2) + 2h_{max}^2(6 + 3C_F^2))\|\operatorname{div}\tilde{\sigma}_0\|^2 \\ &\quad + (2 + 4(6 + 3C_F^2))\|\sigma_0 - \nabla\tilde{u}_1\|^2 \\ &\quad + 2\|u_0 - \Pi_0\tilde{u}_1\|^2 + 4(6 + 3C_F^2)\|\sigma_0 - \Pi_0\tilde{\sigma}_0\|^2 \\ &\leq \max\left\{1 + 2C_F^2(5 + 3C_F^2) + 2h_{max}^2(6 + 3C_F^2), 2 + 4(6 + 3C_F^2)\right\} b(x_h, y_h).\end{aligned}$$

The product of the two resulting estimates $\|x_h\|_X^2 \lesssim b(x_h, y_h)$ and $\|y_h\|_Y^2 \lesssim b(x_h, y_h)$ concludes the proof. □

The validity of (3.1) implies that X_h and M_h define a stable pair for mixed finite elements and, by Theorem 3.3, the lowest-order dPG solution of the Poisson problem is quasi-optimal.

3.2. A posteriori analysis

Recall that P_h denotes the orthogonal projection onto M_h in the Hilbert space Y.

Lemma 3.6 (reliability)**.** *Let b be a continuous bilinear form on $X \times Y$ with* (1.2) *and let $F \in Y^*$ be a given linear functional. Suppose that finite-dimensional subspaces $X_h \subset X$ and $Y_h \subset Y$ are specified. Then, any $x_h \in X_h$ satisfies*

$$\beta\|x - x_h\|_X \leq \|F - b(x_h, \bullet)\|_{Y_h^*} + \|F \circ (1 - P_h)\|_{Y^*} .$$

Proof. The proof is contained in [6] and repeated here for convenient reading. The inf-sup condition for b guarantees

$$\beta\|x - x_h\|_X \leq \|F - b(x_h, \bullet)\|_{Y^*} .$$

The argument in the norm of the right hand side reads

$$F(y) - b(x_h, y) = F(P_h y) - b(x_h, P_h y) + F(y - P_h y) - b(x_h, y - P_h y).$$

Any $y \in S(Y)$ satisfies

$$|F(y) - b(x_h, y)| \leq \|P_h\|\|F - b(x_h, \bullet)\|_{Y_h^*} + \|F \circ (1 - P_h)\|_{Y^*}.$$

Thus,

$$\beta\|x - x_h\|_X \leq \|F - b(x_h, \bullet)\|_{Y_h^*} + \|F \circ (1 - P_h)\|_{Y^*}. \quad \square$$

Remark 3.7 (a posteriori error control)**.** *The estimate in the previous lemma makes clear that the error is controlled by two terms. The first is the residual which is minimized by the dPG method. The second term measures how well the right hand side was approximated. If the data approximation is efficient, the error behaves quantitatively as the first term. In this context, note that for any $x_h \in X_h$ and $y \in S(Y)$ it holds*

$$\begin{aligned}(F \circ (1 - P))(y) = F(y - Py) = b(x, y - Py) = b(x - x_h, y - Py)\\ \leq \|b\|\|x - x_h\|_X\|1 - P\| .\end{aligned}$$

4. Numerical results

This section presents numerical experiments on three domains with Dirichlet and Neumann boundary conditions and investigates the error in the norm $\|\cdot\|_{X_h}$ and the errors in the the single components of $x_h = (u_0, \sigma_0, \tilde{u}_1, \tilde{\sigma}_0)$ as well as the convergence behaviour of the a posteriori error estimator on uniform and adaptive meshes.

4.1. Adaptive Algorithm

The a posteriori error estimate from Subsection 3.2 motivates an adaptive algorithm based on local residuals. The fact that the test functions in Y_h are discontinuous allows to choose an element-by-element basis with local supports on one element only which leads to a decoupled mass matrix of block-diagonal type. The restriction of Y_h to a single element reads

$$Y_h(T) := \operatorname{span}\{y_h \in Y_h \mid \operatorname{supp} y_h \subseteq T\}.$$

The local error estimator contributions read

$$\eta^2(T) := \|F - b(x_h, \cdot)\|^2_{Y_h(T)^\star} \quad \text{for any } T \in \mathcal{T}$$

and sum up to the global error estimator $\eta := \big(\sum_{T\in\mathcal{T}} \eta^2(T)\big)^{1/2}$.

The adaptive algorithm with Dörfler marking runs the following loop.

Input: Initial triangulation $\mathcal{T}_0$, bulk parameter $0 < \theta \leq 1$.

for $\ell = 0, 1, 2, \ldots$

Solve. Compute discrete solution x_ℓ to (3.2) with respect to $\mathcal{T}_\ell$.

Estimate. Compute local contributions of the error estimator $(\eta_\ell^2(T))_{T\in\mathcal{T}_\ell}$.

Mark. Choose a minimal subset $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ such that $\theta\eta_\ell^2(\mathcal{T}_\ell) \leq \eta_\ell^2(\mathcal{M}_\ell)$.

Refine. Generate $\mathcal{T}_{\ell+1}$ from $\mathcal{T}_\ell$ and $\mathcal{M}_\ell$ with newest-vertex bisection.

end for

Output: Triangulations $(\mathcal{T}_\ell)_\ell$ and discrete solutions $(x_\ell)_\ell$.

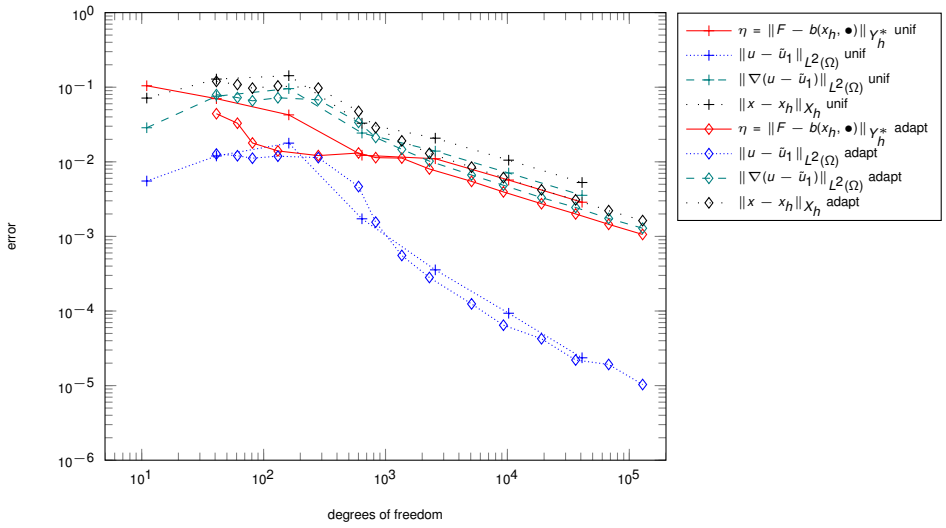

Figure 1: Convergence history for the square domain under uniform and adaptive mesh-refinement.

4.2. Square Domain

The exact solution of the Poisson model problem on the unit square $\Omega := (0,1)^2$ reads

$$u(x,y) = x(x-1)y(y-1)\exp(-100(x-1/2)^2 - (y-117)^2/10^4)$$

and specifies the right-hand side $f = -\Delta u$. Figure 1 displays the convergence history on uniform and adaptive meshes. The L^2 error $\|u - \tilde{u}_1\|_{L^2(\Omega)}$ converges at rate 1 with respect to the number of degrees of freedom, whereas the energy error $\|\nabla(u-\tilde{u}_1)\|_{L^2(\Omega)}$, the total error $\|x - x_h\|_{X_h}$ and the error estimator converge at rate $1/2$. These convergence rates are observable starting from 800 degrees of freedom. For coarser meshes, the oscillatory right-hand side f might not be resolved accurately by the finite element meshes. The error estimator $\eta = \|F - b(x_h, \cdot)\|_{Y_h^\star}$ is reliable and efficient in that it converges at the same rate as the $\|x - x_h\|_{X_h}$. Moreover, η seems to be a close approximation to $\|\nabla(u - \tilde{u}_1)\|_{L^2(\Omega)}$.

4.3. L-shaped domain

The L-shaped domain $\Omega = (-1,1)^2 \setminus [-1,0]^2$ is treated with mixed boundary conditions. On the Dirichlet boundary $\Gamma_D := \mathrm{conv}\{(0,-1),(0,0)\} \cup \mathrm{conv}\{(0,0),(1,0)\}$, zero boundary conditions are imposed. The Neumann condition on $\Gamma_N := \partial\Omega \setminus \Gamma_D$ is set according to the exact solution, which reads in polar coordinates (r,θ) as

$$u(r,\theta) = r^{2/3}\sin(2(\theta + \pi/2)/3).$$

Figure 2 displays the convergence history under uniform and adaptive mesh-refinement. The generic singularity leads to the suboptimal convergence rate of $1/3$ for the error estimator η, the gradient approximation $\|\nabla u - \sigma_0\|_{L^2(\Omega)}$ and the total error $\|x - x_h\|_{X_h}$ for uniform refinement.

As in the previous example, the L^2 error $\|u - \tilde{u}_1\|_{L^2(\Omega)}$ shows the double rate of convergence. For all error quantities, the convergence rate is observed for less than 100 degrees of freedom. As in the previous example, the a posteriori error estimator η appears to be a close approximation to $\|\nabla(u - \tilde{u}_1)\|_{L^2(\Omega)}$. Adaptive mesh-refinement can recover the optimal convergence rates.

Acknowledgements

The first author acknowledges support from the Deutsche Forschungsgemeinschaft through SPP 1748; the second author acknowledges support from the DFG Research Center Matheon via Project C22; the third author acknowledges support from the Berlin Mathematical School; the fourth author acknowledges support from the Humboldt School and Frauenförderung des Landes Berlin.

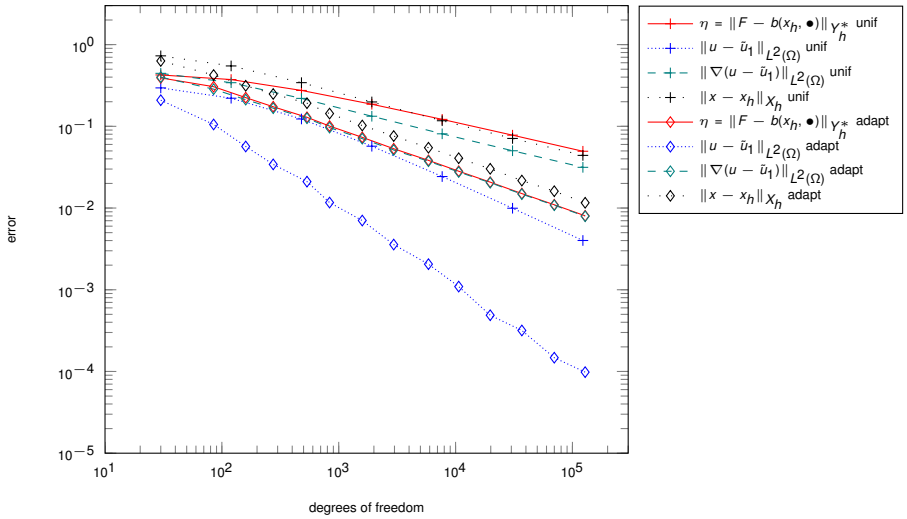

Figure 2: Convergence history for the L-shaped domain under uniform and adaptive mesh-refinement.

References

[1] L. Demkowicz, J. Gopalakrishnan, A class of discontinuous Petrov-Galerkin methods. Part I: the transport equation, Comput. Methods Appl. Mech. Engrg. 199 (23-24) (2010) 1558–1572.

[2] L. Demkowicz, J. Gopalakrishnan, Analysis of the DPG method for the Poisson equation, SIAM J. Numer. Anal. 49 (5) (2011) 1788–1809.

[3] L. Demkowicz, J. Gopalakrishnan, A class of discontinuous Petrov-Galerkin methods. II. Optimal test functions, Numer. Methods Partial Differential Equations 27 (1) (2011) 70–105.

[4] L. Demkowicz, J. Gopalakrishnan, A. H. Niemi, A class of discontinuous Petrov-Galerkin methods. Part III: Adaptivity, Appl. Numer. Math. 62 (4) (2012) 396–427.

[5] T. Bouma, J. Gopalakrishnan, A. Harb, Convergence rates of the DPG method with reduced test space degree, Comput. Math. Appl.To appear.

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